



(12) 发明专利申请

(10) 申请公布号 CN 103282377 A

(43) 申请公布日 2013. 09. 04

(21) 申请号 201180063093. 7

(22) 申请日 2011. 12. 28

(30) 优先权数据

2010-293783 2010. 12. 28 JP

(85) PCT申请进入国家阶段日

2013. 06. 27

(86) PCT申请的申请数据

PCT/JP2011/007337 2011. 12. 28

(87) PCT申请的公布数据

W02012/090499 EN 2012. 07. 05

(71) 申请人 丰田自动车株式会社

地址 日本爱知县丰田市

(72) 发明人 服部悦子 西村哲 伊藤和代

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

代理人 彭鲲鹏 卢蓓

(51) Int. Cl.

C07K 14/415 (2006. 01)

C12N 15/82 (2006. 01)

A01H 5/12 (2006. 01)

权利要求书1页 说明书10页

序列表6页 附图6页

(54) 发明名称

成熟叶特异性启动子

(57) 摘要

本发明提供了基因表达调控性DNA,其具有以成熟叶特异性方式启动基因表达的活性。提供了以下:具有以成熟叶特异性方式启动基因表达之活性的基因表达调控性DNA,其包括以下(a)至(d)中的任一种:(a)由SEQ ID NO:1或SEQ ID NO:4所示核苷酸序列组成的DNA;(b)由在SEQ ID NO:1或SEQ ID NO:4所示核苷酸序列中具有一个或多个核苷酸的缺失、替换、添加或插入的核苷酸序列组成并且具有以成熟叶特异性方式启动基因表达之活性的DNA;(c)由与SEQ ID NO:1或SEQ ID NO:4所示核苷酸序列有90%或更高序列同一性的核苷酸序列组成并且具有以成熟叶特异性方式启动基因表达之活性的DNA;和(d)在严格条件下与由互补于SEQ ID NO:1或SEQ ID NO:4所示核苷酸序列之一部分或全部的序列所组成的DNA杂交并且具有以成熟叶特异性方式启动基因表达之活性的DNA。

1. 一种基因表达调控性 DNA, 其具有以成熟叶特异性方式启动基因表达的活性, 其包括以下 (a) 至 (d) 中的任一种:

(a) 由 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列组成的 DNA ;

(b) 由在 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列中具有一个或多个核苷酸的缺失、替换、添加或插入的核苷酸序列组成并且具有以成熟叶特异性方式启动基因表达之活性的 DNA ;

(c) 由与 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列有 90% 或更高序列同一性的核苷酸序列组成并且具有以成熟叶特异性方式启动基因表达之活性的 DNA ; 和

(d) 在严格条件下与由互补于 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列之一部分或全部的序列所组成的 DNA 杂交并且具有以成熟叶特异性方式启动基因表达之活性的 DNA。

2. 一种重组载体, 其包含根据权利要求 1 所述的基因表达调控性 DNA。

3. 一种经转化植物, 其通过使用根据权利要求 2 所述的重组载体进行转化而获得。

4. 一种经转化植物, 其由经转化的植物细胞获得, 所述转化通过使用根据权利要求 2 所述的重组载体的植物细胞转化来进行。

成熟叶特异性启动子

技术领域

[0001] 本发明涉及具有以成熟叶特异性方式启动基因表达之活性的基因表达调控性 DNA 及其用途。

背景技术

[0002] 生物体的遗传信息通过被称为“中心法则 (central dogma)”的一系列过程进行传递,其中功能性基因的 DNA 信息转录为 mRNA,通过 mRNA 信息的翻译合成功能性蛋白质。从而表达生物学功能。基因组 DNA 是由亲本单倍体的组合而形成的遗传信息的聚集体。另外,对于植物,基因组 DNA 一般记载了包含所述 DNA 的细胞以及来源于此细胞的植物的特性。

[0003] 为了以正确方式通过细胞中所含基因组 DNA 的功能性遗传信息传递来诱导生物学功能的表达,必须在适当位点适当时间并且以适当强度诱导适当基因的表达。因此,需要严格调控特定基因的表达。

[0004] 对于功能性基因的表达,所述基因 5' 上游区中存在的基因表达调控性 DNA 调控表达的时间、位点和强度。

[0005] 近年来,由于使用拟南芥 (*Arabidopsis thaliana*)、稻米等的基因组对 DNA 进行的破译,所以这些植物各个功能性基因的基因表达调控性 DNA 已变得可以容易地获得 (非专利文献 1 和 2)。

[0006] 与此同时,甘蔗 (一种易于获得的作物) 的基因组 DNA 仍未被破译。因此,得到甘蔗各个功能性基因的基因表达调控性 DNA 并不容易。迄今为止,为了将基因引入甘蔗中,已使用了来自非甘蔗植物的基因表达调控性 DNA 来调控转基因的表达 (非专利文献 3 和 4)。

[0007] 但是,当使用非甘蔗植物的基因表达调控性 DNA 在甘蔗中诱导功能性基因表达时,观察到在一些情况下无法严格地调控表达的时间、位点、强度和其他条件。因此,相关领域期望获得来源于甘蔗的基因表达调控性 DNA,特别是组织特异性基因表达调控性 DNA。

[0008] 引用列表

[0009] 非专利文献

[0010] NPL1 :Nature,2000,408 :769-815

[0011] NPL2 :Science,2002,296 :92-100

[0012] NPL3 :Plant Mol Biol. 1992, Feb ;18(4) :675-89

[0013] NPL4 :Planta,1998,206 :20-27

发明内容

[0014] 技术问题

[0015] 因此,本发明的一个目的是提供基因表达调控性 DNA,其具有以成熟叶特异性方式启动基因表达的活性。

[0016] 问题的解决方案

[0017] 作为为了实现以上目的而进行的深入研究的结果,本发明人发现在甘蔗成熟叶中

所表达的基因的 5' 上游区中具有以成熟叶特异性方式启动基因表达之活性的基因表达调控性 DNA。这导致了本发明的完成。

[0018] 具体地,本发明涵盖以下特征 {1} 至 {4}。

[0019] {1} 具有以成熟叶特异性方式启动基因表达之活性的基因表达调控性 DNA,其包含以下 (a) 至 (d) 中的任何一种:

[0020] (a) 由 SEQ ID NO:1 或 SEQ ID NO:4 所示核苷酸序列组成的 DNA;

[0021] (b) 由在 SEQ ID NO:1 或 SEQ ID NO:4 所示核苷酸序列中具有一个或多个核苷酸的缺失、替换、添加或插入的核苷酸序列组成并且具有以成熟叶特异性方式启动基因表达之活性的 DNA;

[0022] (c) 由与 SEQ ID NO:1 或 SEQ ID NO:4 所示核苷酸序列有 90% 或更高序列同一性的核苷酸序列组成并且具有以成熟叶特异性方式启动基因表达之活性的 DNA;和

[0023] (d) 在严格条件下与由互补于 SEQ ID NO:1 或 SEQ ID NO:4 所示核苷酸序列之一部分或全部的序列所组成的 DNA 杂交并且具有以成熟叶特异性方式启动基因表达之活性的 DNA。

[0024] {2} 重组载体,其包含根据 {1} 所述的基因表达调控性 DNA。

[0025] {3} 经转化植物,其通过使用根据 {2} 所述的重组载体转化而得到。

[0026] {4} 经转化植物,其由经转化的植物细胞得到,所述转化通过使用根据 {2} 所述的重组载体的植物细胞转化进行。

[0027] 日本专利申请 No. 2010-293783 (本申请的优先权文件) 的说明书和 / 或附图所公开内容的一部分或全部通过引用并入本文。

[0028] 发明的有利效果

[0029] 根据本发明,可提供以下内容:具有以成熟叶特异性方式启动基因表达之活性的基因表达调控性 DNA;包含所述 DNA 的重组载体,其可引起功能性基因以成熟叶特异性方式表达;以及用所述重组载体转染的经转化植物。

[0030] 附图简述

[0031] [图 1] 图 1 示出来源于白甘蔗 (*Saccharum officinarum*) 的 ecc0001EST 的核苷酸序列 (使用 DFCI 甘蔗基因指数 (Sugarcane Gene Index) 检索)。

[0032] [图 2] 图 2 示出甘蔗属杂交变种 NiF8 (*Saccharum spp. hybrids cv. NiF8*) 各个组织的 ecc0001EST 表达水平分析的结果。将对于茎表皮所确认的最强表达水平指定为 100。

[0033] [图 3] 图 3 示出在 5' 和 3' 端分别插入 HindIII 限制酶识别序列和 BlnI 限制酶识别序列的 ecc0001 基因之基因表达调控区的核苷酸序列 (SEQ ID NO:1)。

[0034] [图 4-1] 图 4-1 示出使用启动子分析工具鉴定的 ecc0001 基因之基因表达调控区中的预测启动子区 (SEQ ID NO:4)。

[0035] [图 4-2] 图 4-2 (续接图 4-1)。

[0036] [图 5] 图 5 示意性地示出包含彼此连接的 ecc0001 基因之基因表达调控区和 β -葡糖醛酸糖苷酶基因的基因表达载体。

[0037] [图 6] 图 6 示出转基因甘蔗各个组织的 β -葡糖醛酸糖苷酶基因表达水平分析的结果,其中 β -葡糖醛酸糖苷酶基因受 ecc0001 基因的表达调控性 DNA 调控。将对于转基因甘蔗成熟叶所确认的最强表达水平指定为 100。

[0038] 实施方案的描述

[0039] 以下详细描述本发明。

[0040] 首先,描述了本发明的具有以成熟叶特异性方式启动基因表达之活性的基因表达调控性 DNA。本发明的基因表达调控性 DNA 包含以下 (a) 至 (d) 中的任何一种:

[0041] (a) 由 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列组成的 DNA ;

[0042] (b) 由在 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列中具有一个或多个核苷酸的缺失、替换、添加或插入的核苷酸序列组成并且具有以成熟叶特异性方式启动基因表达之活性的 DNA ;

[0043] (c) 由与 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列有 90% 或更高序列同一性的核苷酸序列组成并且具有以成熟叶特异性方式启动基因表达之活性的 DNA ;和

[0044] (d) 在严格条件下与由互补于 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列之一部分或全部的序列所组成的 DNA 杂交并且具有以成熟叶特异性方式启动基因表达之活性的 DNA。

[0045] 本发明的基因表达调控性 DNA 可通过以下方式得到:通过使用来自各个甘蔗组织(茎、成熟叶、幼叶等)的总 RNA 或来自这些 RNA 的 cDNA 的基因表达分析得到以成熟叶特异性方式表达的候选基因;评价所述候选基因的表达特征;基于评价结果指明被评价为以成熟叶特异性方式表达的基因;以及基于相关所指明基因的 cDNA 或基因组 DNA 鉴定每个候选基因 5' 上游区的核苷酸序列。本文中,可使用本领域技术人员众所周知的穷尽性(exhaustive)基因表达分析技术(例如 DNA 芯片法和差异显示法)进行基因表达分析。

[0046] 特别地,SEQ ID NO :1 或 SEQ ID NO :4 的核苷酸序列存在于基因(下文称为“ecc0001”)的 5' 上游区,其在成熟甘蔗叶中特异性表达。本文所述的“甘蔗”植物的实例包括(但不特别限于)属于甘蔗属的植物,例如白甘蔗(*Saccharum officinarum*)、中国竹蔗(*Saccharum sinense*)、细秆甘蔗(*Saccharum barberi*)、大茎野生种(*Saccharum robustum*)、割手密(*Saccharum spontaneum*)、肉质花穗野生种(*Saccharum edule*)和甘蔗属杂交栽培变种 NiF8;以及属于与甘蔗属/甘蔗种密切相关的属/种的植物,例如高粱(*Sorghum*)或蔗茅属(*Erianthus*)。其中,优选甘蔗属杂交栽培变种 NiF8。

[0047] 用于分离所述 5' 上游区中存在的 DNA 的方法无特别限制。DNA 分离可通过本领域技术人员众所周知的方法进行。例如,DNA 可通过常规已知方法分离,所述方法包括基于 ecc0001 基因的核苷酸序列(SEQ ID NO :2)克隆未知区域(即,上述情况下的 5' 上游区)。在这样的方法中,使包含 ecc0001 基因 5' 上游区的基因组 DNA 经历限制酶处理以使得由预定核苷酸序列组成的接受体(adopter)与所述 DNA 连接。为 ecc0001 基因和接受体的核苷酸序列指定引物,然后进行 PCR。因此可扩增毗连 ecc0001 基因核苷酸序列 5' 上游区的未知核苷酸序列。在确定扩增的核苷酸序列之后,基于确定的核苷酸序列设计另一对引物。因此,可以以类似方式扩增毗连所确定核苷酸序列的另外未知核苷酸序列。该方法可使用可商购克隆试剂盒如 RightWalk(注册商标)试剂盒(BEX Co., Ltd.)进行。或者,除以上之外可建议使用反向 PCR 的方法。在这种情况下,基于 ecc0001 基因的核苷酸序列信息设计一对引物。使用该引物对和通过用某些限制酶处理和自身连接所得到的基因组 DNA 片段进行 PCR。因此,可扩增 ecc0001 基因的上游区。此外,可建议另一种用于从基因组 DNA 文库中分离 ecc0001 基因上游区的方法。在这种情况下,使用包含 ecc0001 基因的 cDNA 作为探

针来筛选已通过标准方法制备的基因组 DNA 文库以获得包含 ecc0001 基因的基因组 DNA。然后,测定通过筛选得到的基因组 DNA 的核苷酸序列。因此,可指明 ecc0001 基因上游区中存在的 5' 上游区。此外,可通过 PCR 或其他方式扩增仅仅 5' 上游区。

[0048] 如上所述,依次扩增或筛选位于 ecc0001 基因上游的未知核苷酸序列用以通过标准方法测定核苷酸序列。因此,可指明 SEQ ID NO:1 或 SEQ ID NO:4 所示的核苷酸序列。测定 SEQ ID NO:1 或 SEQ ID NO:4 所示核苷酸序列后,使用提取自甘蔗的基因组 DNA 作为模板和基于 SEQ ID NO:1 或 SEQ ID NO:4 所示核苷酸序列设计的引物通过 PCR 得到 SEQ ID NO:1 或 SEQ ID NO:4 所示核苷酸序列就变得可能。

[0049] SEQ ID NO:1 或 SEQ ID NO:4 所示核苷酸序列作为能够以成熟叶特异性方式诱导基因表达的基因表达调控区起作用。基因表达调控区包含参与基因转录控制的核苷酸序列,例如启动子区、增强子区、TATA 盒和 / 或 CAT 盒 (但是所述区的内容不特别受限于此)。

[0050] 本文使用的术语“特异性”指以下情况:基因表达诱导型功能 (gene expression inducible function) 仅存在于作为构成植株的各个组织之一的成熟叶组织中;以及成熟叶组织中的基因表达诱导型功能显著的或统计上显著的 (例如约 2 倍、3 倍、4 倍、5 倍、6 倍、7 倍、8 倍、9 倍、10 倍或更多倍) 高于非成熟叶的组织 (例如茎、幼叶、根或分生组织) 中的基因表达诱导型功能。

[0051] 本文使用的术语“成熟叶 (组织)”指包含细胞中积聚的用于光合作用的叶绿体并因此带有绿色的叶 (组织)。它还表示除不含用于光合作用的叶绿体的幼叶 (组织) 之外的叶 (组织)。

[0052] 可通过本领域技术人员众所周知的报告基因测定 (reporter assay) 或类似测定来确定基因表达诱导型功能。基于报告基因测定制备载体,其中多种报告基因 (例如 β -葡糖醛酸糖苷酶基因 (beta-glucuronidase gene, GUS)、萤光素酶基因 (luciferase gene, LUC) 和绿色荧光蛋白基因 (green fluorescent protein, GFP)) 与待研究基因表达诱导型功能的核苷酸序列之下游区连接,使得所述报告基因受所述核苷酸序列调控。使用所述载体将基因引入 (或瞬时基因引入) 宿主基因组中。然后,测定每个报告基因的表达水平。由此可确定基因表达诱导型功能。只要报告基因的表达水平可检测,所述报告基因就没有具体限制。这样的报告基因的实例包括本领域技术人员常规使用的报告基因,例如 CAT 基因、lacZ 基因、萤光素酶 (下文表示为“LUC”) 基因、 β -葡糖醛酸糖苷酶 (下文表示为“GUS”) 基因和绿色荧光蛋白 (下文表示为“GFP”) 基因。

[0053] 可根据报告基因的类型通过本领域技术人员众所周知的方法测定报告基因的表达水平。例如,如果报告基因是 CAT 基因,则可通过检查基因产物对氯霉素的乙酰化来测定报告基因的表达水平。可通过以下技术测定各个报告基因的表达水平。对于报告基因是 lacZ 基因的情况,检测由基因表达产物催化作用诱导的染料化合物的显色。对于报告基因是 LUC 基因的情况,检测由基因表达产物催化作用诱导的荧光化合物的荧光发射。对于报告基因是 GFP 基因的情况,检测 GFP 蛋白质的荧光发射。例如,如果报告基因是 GUS,则 GUS 活性通过以下两种方法之一测定为宿主细胞中的启动子活性:(i) 涉及组织化学 GUS 染色的方法 (EMBO J. 6, 3901-3907(1987)) 和 / 或 (ii) 涉及使用荧光底物的 Castle&Morris 的方法 (Plant Molecular Biology Manual, B5, 1-16(1994); S. B. Gelvin&R. A. Schilperoord, Kluwer Academic Publishers)。此外,通过 Bradford 方

法测定蛋白质量 (Anal. Biochem. 72, 248-254 (1976))。基于蛋白质量将 GUS 活性转化为单位为 $\text{nmol}^{14}\text{-MU}/\text{min}/\text{mg}$ 蛋白质。因此,对于每种情况都可确认基因表达诱导型功能。

[0054] 此外,如果将除以上之外的基因用作报告基因,则通过 Northern 杂交、RT-PCR、DNA 阵列技术等来测定基因转录水平。或者,通过电泳例如 SDS-PAGE、Western 印迹等来测定由基因编码的蛋白质的表达水平。因此,可确认基因表达诱导型功能。

[0055] 本发明的基因表达调控性 DNA 不限于 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列。如以上 (b) 中所述,它可以是在 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列中具有一个或多个核苷酸缺失、替换、添加或插入的核苷酸序列,只要它具有以成熟叶特异性方式启动基因表达的活性。

[0056] 例如,本发明的基因表达调控性 DNA 甚至包括在 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列中具有 1 至 100 个核苷酸,优选 1 至 50 个核苷酸,并且更优选 1 至 10 个核苷酸缺失、替换、添加或插入的核苷酸序列,只要它显示出以成熟叶特异性方式启动基因表达的活性。

[0057] 此外,本发明的基因表达调控性 DNA 不限于 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列。如以上 (c) 所述,它可以是与 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列具有 80% 或更多,更优选 90% 或更多,更优选 95% 或更多,并且最优选 99% 或更多序列同一性的核苷酸序列,只要它表现出以成熟叶特异性方式启动基因表达的活性。核苷酸序列可通过众所周知的方法进行比较。可使用例如 BLAST (美国国家生物学信息中心 (National Center for Biological Information) 的局部序列排列检索基本工具 (Basic Local Alignment Search Tool)) 基于默认设置进行比较。

[0058] 此外,本发明的基因表达调控性 DNA 不限于 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列。如以上 (d) 所述,它可以是在严格条件下与由互补于 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列一部分或全部的序列所组成的 DNA 杂交的核苷酸序列,只要它显示出以成熟叶特异性方式启动基因表达的活性。

[0059] 本文中,术语“严格条件”指这样的条件:即在所述条件下形成特定特异性杂交体,但从不形成非特异性杂交体。例如,这样的条件包括在包含 2 至 $6\times\text{SSC}$ ($1\times\text{SSC}$ 组成: 0.15M NaCl、0.015M 柠檬酸钠, pH7.0) 和 0.1 至 0.5% SDS 的溶液中在 42°C 至 55°C 杂交,然后在包含 0.1 至 $0.2\times\text{SSC}$ 和 0.1 至 0.5% SDS 的溶液中在 55°C 至 65°C 洗涤。

[0060] 而且,具有以成熟叶特异性方式启动基因表达之活性的本发明基因表达调控性 DNA 可以在 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列 5' 端和 / 或 3' 端具有 100、200、300、400、500、600、700、800、900、1000、1100、1200、1300、1400 或 1500 个或多个连续核苷酸缺失的 DNA 片段,只要它表现出以成熟叶特异性方式启动基因表达的活性。可通过本领域技术人员众所周知的方法 (例如 PCR 或限制酶处理) 使核苷酸缺失。所述 DNA 片段可以是本发明基因表达调控性 DNA 的启动子区。可使用本领域技术人员众所周知的启动子分析技术 (例如 BioInformatics and Molecular Analysis Section (<http://www-bimas.cit.nih.gov/molbio/proscan/>); Prestridge, D. S. (1995), Predicting Pol II Promoter Sequences Using Transcription Factor Binding Sites, J. Mol. Biol. 249 :923-32) 检索预定基因表达调控性 DNA 的启动子区。SEQ ID NO :1 和 SEQ ID NO :4 所示核苷酸序列的此类片段的实例分别是由 SEQ ID NO :1 所示核苷酸序列的第 1412 位至第 1662 位核苷酸组成

的 DNA 以及由 SEQ ID NO :4 所示核苷酸序列的第 1413 位至第 1663 位核苷酸组成的 DNA。可通过以上的报告基因测定或类似测定来确认所得片段是否具有基因表达诱导型功能。

[0061] 确定了本发明基因表达调控性 DNA 的核苷酸序列后,通过化学合成、使用基因组 DNA 作为模板的 PCR、或使用具有所述核苷酸序列的 DNA 片段作为探针的杂交来得到本发明的基因表达调控性 DNA 就变得可能。另外,在 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列中具有突变的核苷酸序列可通过位点特异性诱变等方式来合成。可通过已知技术例如 Kunkel 法或缺口双链体 (Gapped duplex) 法或基于其的方法将突变引入 SEQ ID NO :1 或 SEQ ID NO :4 所示核苷酸序列中。例如,通过使用定点诱变 (例如 Mutant-K 或 Mutant-G (二者都是 TAKARA Bio 的商品名)) 等的诱变试剂盒或 LA PCR 体外诱变系列试剂盒 (in vitro Mutagenesis series kit) (商品名, TAKARA Bio) 来引入突变。

[0062] 接着,描述了重组载体,其包含以上具有以成熟叶特异性方式启动基因表达之活性的基因表达调控性 DNA。

[0063] 可通过将包含与以上基因表达调控性 DNA 有效连接的期望功能性基因的 DNA 引入适当载体中来构建本发明的重组载体。本文使用的术语“有效连接”指这样的情况:上述载体包含彼此连接的基因表达调控性 DNA 和功能性基因使得在经以上载体转化的宿主细胞中功能性基因在基因表达调控性 DNA 调控下正确表达。本文中,“连接”可以是直接连接或通过适当长度和适当序列的间隔子 (spacer) 间接连接。用于本发明的载体的优选实例包括 pBI 载体、pBII 载体、pPZP 载体 (Hajdukiewicz P, Svab Z, Maliga P. :The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation., Plant Mol Biol., 25 :989-94, 1994)、pCAMBIA 载体 (http://www.cambia.org/main/r_et_camvec.htm) 和 pSMA 载体,通过所述载体可使用农杆菌 (Agrobacterium) 将功能性基因引入植物中。特别优选地,使用 pBI 和 pBII 双元载体 (binary vector) 或中间载体。这样的载体的实例包括 pBI121、pBI101、pBI101. 2、pBI101. 3、pBII221 和 pIG121。双元载体是可在大肠杆菌和农杆菌中复制的穿梭载体。当植物感染包含双元载体的农杆菌时,载体上存在的对应于边界序列 (border sequence) (LB 序列和 RB 序列) 之间区域的 DNA 可合并到植物的核 DNA 中 (EMBO Journal, 10 (3), 697-704 (1991))。同时,使用 pUC 载体可将基因直接引入植物中。PUC 载体的实例包括 pUC18、pUC19 和 pUC9。此外,可使用植物病毒载体,例如花椰菜花叶病毒 (cauliflower mosaic virus, CaMV)、菜豆金黄花叶病毒 (bean golden mosaic virus, BGMV) 和烟草花叶病毒 (tobacco mosaic virus, TMV)。

[0064] 为了有助于连接和 / 或插入载体中,可通过限制酶识别序列的替换、插入或添加来适当修饰包含基因表达调控性 DNA 的 DNA 和 / 或包含功能性基因的 DNA。对于插入载体中,可使用例如包括以下的方法:用适当的限制酶切割包含基因表达调控性 DNA 和 / 或功能性基因的纯化 DNA,然后将每个所得片段插入适当载体 DNA 的限制酶识别位点或多克隆位点以与所述载体连接。

[0065] 术语“功能性基因”指目标植物的任意内源基因或期望引起基因产物在成熟叶中表达的任意外源基因。这样的基因的实例包括但不限于光合作用相关基因、易位 (translocation) 相关基因、能够产生有用物质 (例如药物、染料或芳香成分) 的基因、糖代谢相关基因、具有疾病 / 昆虫抗性 (例如昆虫损伤抗性、霉菌 (真菌) 和细菌疾病抗性、或病毒 (疾病) 抗性) 的基因、环境压力 (例如低温、高温、干燥、光损伤 (photolesion) 或紫

外线)抗性相关基因和植物生长调控(启动/抑制)基因。

[0066] 如果需要,则可将增强子、内含子、聚A加尾信号(poly-A addition signal)、5'-UTR序列、选择标记基因等连接到载体中基因表达调控性DNA和/或功能性基因的上游/下游或其之间。

[0067] 增强子用于例如提高功能性基因表达的效率。其实例是包含位于CaMV35S启动子上游之序列的增强子区。

[0068] 终止子可以是可终止由以上启动子引起的基因转录的序列。其实例包括胭脂碱(nopalins)合成酶基因终止子、章鱼碱合成酶基因终止子和CaMV35S RNA基因终止子。

[0069] 选择标记基因的实例包括潮霉素抗性基因、卡那霉素抗性基因、双丙氨磷抗性基因、灭瘟素S抗性基因和乙酰乳酸合酶基因。选择标记基因可以和功能性基因一起连接至如上所述的相同质粒以制备重组载体。或者,可分开制备通过连接选择标记基因与质粒得到的重组载体和通过连接功能性基因与质粒得到的重组载体。当它们分开制备时,用两种载体共转染宿主。

[0070] 可使用由此制备的重组载体产生转化体。

[0071] 当制备转化植物时,可适当地使用已报道和建立的多种方法。这样的方法的优选实例包括农杆菌法、PEG-磷酸钙法、电穿孔法、脂质体法、粒子枪(particle gun)法和显微注射法。对于农杆菌法,使用原生质体或组织切片或使用植物自身(在植物中(in planta)方法中)。当使用原生质体时,使用包括共培养具有Ti质粒的农杆菌与原生质体的方法或包括使农杆菌原生质球(spheroplast)与原生质体融合的方法(即原生质球法)。当使用组织切片时,可使用包括用农杆菌感染目标植物的无菌培养叶切片(叶盘)的方法、包括用农杆菌感染愈伤组织等的方法。此外,如果采用使用种子或植物(不需要进行添加植物激素的组织培养的系统)的植物中方法,则用农杆菌直接处理吸水种子、幼小植物(幼苗)、盆栽植物等。

[0072] 通过PCR法、Southern杂交法、Northern杂交法、Western印迹法等可确认基因并入植物是否发生。例如,由转化植物制备DNA。设计DNA特异性引物,然后进行PCR。PCR之后,使扩增产物经历琼脂糖凝胶电泳、聚丙烯酰胺电泳、毛细管电泳等,然后用溴化乙锭、SYBR Green液体等染色。然后,检测来源于扩增产物的单一条带。由此可确认转化的发生。此外,可使用预先标记有荧光染料等的引物通过PCR检测扩增产物。此外,还可进行包括使扩增产物与固相(例如微板)结合并通过例如荧光或酶反应确认扩增产物的方法。

[0073] 本发明用于转化的植物的实例包括但不限于属于禾本科(Gramineae)、茄科(Solanaceae)、十字花科(Brassicaceae)、豆科(Leguminosae)、蔷薇科(Rosaceae)、菊科(Asteraceae)、百合科(Liliaceae)、伞形科(Apiaceae)、石竹科(Caryophyllaceae)、葫芦科(Cucurbitaceae)、旋花科(Convolvulaceae)和藜科(Chenopodiaceae)的植物。其优选实例包括已从其中分离了上述基因表达调控性DNA的属于禾本科的植物,包括甘蔗、稻米、大麦、小麦、玉米、结缕草、高粱、粟、日本粟、象草(napier grass)和柳枝稷(switchgrass)。

[0074] 本发明的经历转化的植物材料的实例包括根、茎、叶、种子、胚(embryo)、胚珠、子房、枝条顶端(植物芽尖端的生长点)、花粉囊、花粉等的植物组织,此类植物组织的切片,未分化的愈伤组织,以及培养的植物细胞(例如通过使以上实例经历酶处理以除去细胞壁

从而得到的原生质体)。此外,当采用植物方法时,可使用吸水种子和整个植物。

[0075] 此外,本发明的经转化植物可以是整个植物、植物器官(例如根、茎、叶、花瓣、种子或果实)、植物组织(例如表皮、韧皮部、薄壁组织、木质部或维管束的组织)或植物培养细胞。

[0076] 当使用植物培养细胞时,可通过用于由所得经转化细胞再生经转化植物的已知组织培养方法再生植物器官或植物自身。本领域技术人员可容易地进行这样的再生操作,只要常规已知的方法作用于由植物细胞再生植物的方法。可如下所述由植物细胞再生植物。

[0077] 首先,当将植物组织或原生质体用作转化的目标植物材料时,在补充有无机元素、维生素、碳源、用作能量来源的糖、植物生长调节剂(例如植物激素,如植物生长素或细胞分裂素)等的无菌愈伤组织形成培养基中培养以形成能够不定(adventitiously)生长的去分化愈伤组织(下文中称为“愈伤组织诱导”)。将由此形成的愈伤组织转移至包含植物生长调节剂(例如植物生长素)的新鲜培养基中以进行进一步生长(继代培养)。

[0078] 在固体培养基(例如琼脂)上进行愈伤组织诱导。通过例如液体培养进行继代培养。在这种情况下,可有效地并且大规模地进行每一培养。接着,在用于诱导器官再分化的适当条件下培养通过上述继代培养生长的愈伤组织(下文中称为“再分化诱导”)。这最终导致完整植物的再生。可通过以下方法进行再分化诱导:确定待添加至培养基中的植物生长调节剂(例如植物生长素或细胞分裂素)和不同组分(例如碳源)的类型和量,然后以适当方式设定光、温度和其他条件。这种再分化诱导导致形成不定胚、根、芽、茎叶等,然后进一步培养以获得完整植物。或者,例如,可将再分化的产物维持在允许其变为完整植物的状态(例如胶囊化(encapsulated)人工种子、干燥胚、或冻干细胞或组织)下。

[0079] 根据本发明,术语“经转化植物”不仅表示转化后通过再分化得到的作为第一代植物的“T1代植物”,还表示作为由T1代植物种子后续产生的植物的“T2代植物”以及由通过药物选择、Southern分析等发现是转基因植物的“T2代”植物的花自花授粉得到的后代植物,例如第三代(T3代)植物。

[0080] 由此产生的经转化植物以成熟叶特异性方式表达引入的功能性基因。

[0081] 本发明的基因表达调控性DNA在植物(特别是甘蔗)中具有以成熟叶特异性方式启动基因表达的活性。将期望的功能性基因与本发明的基因表达调控性DNA连接并将连接产物引入植物中,从而使功能性基因在植物中以成熟叶特异性方式表达。由此赋予植物以期望特性。本发明的基因表达调控性DNA可适当地用于尚未针对其开发许多基因表达调控系统的甘蔗。使用本发明的基因表达调控性DNA能够产生具有适于生物质资源和其他应用之特性的转基因甘蔗。

实施例

[0082] 下文参照以下实施例对本发明进行更详细的描述,但是本发明的技术范围不限于此。

[0083] 实施例1

[0084] 成熟叶表达基因的克隆

[0085] 使用RNeasy Plant Mini Kit(QIAGEN)从甘蔗(甘蔗属栽培变种NiF8)的成熟叶、

幼叶和茎组织中提取并纯化总 RNA。根据标准方法构建 cDNA 文库并用于基因表达分析。使用甘蔗基因组阵列 (Sugarcane Genome Array) (Affimetrix) 根据制造商说明书进行基因表达分析。

[0086] 基因表达分析的结果是,发现了甘蔗属栽培变种 NiF8 成熟叶、幼叶、茎髓、茎表皮、根和分生组织中表达的基因。将所述基因命名为“ecc0001”。来源于甘蔗属栽培变种 NiF8 的 ecc0001 基因的核苷酸序列与图 1 所示来源于白甘蔗的 ecc0001EST 的核苷酸序列相同。从甘蔗属栽培变种 NiF8 的成熟叶、幼叶、茎髓、茎叶、根和分生组织中提取并纯化总 RNA。根据标准方法制备 cDNA。使用 ABI7500 实时 PCR 系统 (Applied Biosystems) 通过 SYBRGreen 方法分析每个组织中的 ecc0001 基因表达水平。

[0087] 图 2 示出结果。图 2 所示的结果表示为参比茎表皮中最强基因表达水平的相对值,所述茎表皮中最强基因表达水平指定为 100%。成熟叶、幼叶、茎髓、茎表皮、根和分生组织的表达水平分别为 12、13、90、100、17 和 16。结果表明甘蔗属栽培变种 NiF8 的成熟叶、幼叶、茎髓、茎表皮、根和分生组织中表达 ecc0001 基因。

[0088] 实施例 2

[0089] 成熟叶特异性启动子的获得

[0090] 使用 DNeasy Plant Mini Kit (QIAGEN) 从甘蔗 (甘蔗属栽培变种 NiF8) 的成熟叶组织 (0.5g) 中提取并纯化基因组 DNA (约 300ng)。使用 RightWalk (注册商标) Kit (BEX) 基于实施例 1 中所得 ecc0001 的核苷酸序列由以上基因组 DNA 得到位于 ecc0001 基因 5' 上游的基因表达调控区。HindIII 限制酶识别序列 (AAGCTT) 和 BlnI 限制酶识别序列 (CCTAGG) 用作接头序列。前者在所得基因表达调控区的 5' 端引入,而后者在基因表达调控区 3' 端存在的 ecc0001 基因翻译起始位点 (ATG) 的 3' 侧引入。由此制备了包含 ecc0001 基因表达调控区的 DNA (图 3) (SEQ ID NO :3)。

[0091] 使用众所周知的启动子分析工具 (BioInformatics and Molecular Analysis Section) (<http://www-bimas.cit.nih.gov/molbio/proscan/>) 分析包含 ecc0001 基因表达调控区的上述 DNA。结果,推测能够作为启动子起作用的区域存在于包含第 1419 至第 1669 位核苷酸的区域中 (图 4-1 和 4-2, SEQ ID NO :5)。所述区域对应于包含 SEQ ID NO :3 所示核苷酸序列的第 1418 至第 1668 位核苷酸的区域。

[0092] 实施例 3

[0093] 基因表达载体的构建

[0094] (1) β -葡糖醛酸糖苷酶基因表达载体

[0095] 构建了这样的基因表达载体:其包含与包含 β -葡糖醛酸糖苷酶 (GUS) 基因的 UidAcDNA 连接的包含实施例 2 所得基因表达调控区的 DNA (SEQ ID NO :5)。将植物转化载体 (pBII221) 用于基因表达载体。为了将包含基因表达调控区的 DNA 与 UidAcDNA 相连接,使编码 UidAcDNA 5' 端侧第一甲硫氨酸的 ATG 序列与编码包含基因表达调控区之 DNA 3' 端侧存在的 ecc0001 基因的第一甲硫氨酸的 ATG 序列匹配以连接 (翻译融合型)。图 5 示意性地示出所述基因表达载体。

[0096] 实施例 4

[0097] 转基因植物的产生

[0098] 使用农杆菌法将实施例 3 中产生的基因表达载体引入宿主植物 (甘蔗属栽培变种

NiF8) 中。由此产生了其中 GUS 基因表达受 ecc0001 基因的表达调控性 DNA 调控的转基因甘蔗。

[0099] 从转基因甘蔗的成熟叶、幼叶、茎髓、茎表皮、根和分生组织中提取并纯化总 RNA。根据标准方法制备 cDNA。使用 ABI7500 实时 PCR 系统 (Applied Biosystems) 通过 SYBRGreen 方法分析每个组织中的 GUS 基因表达水平。

[0100] 作为对比, 以上述方式分析其中 GUS 基因表达受花椰菜花叶病毒 (CaMV) 35S 启动子调控的转基因甘蔗成熟叶和幼叶中的 GUS 表达水平以及转基因甘蔗分生组织中的 GUS 表达水平 (对照)。

[0101] 图 6 示出结果。各个组织的表达水平表示为相比于在所分析的转基因甘蔗组织中确认的成熟叶中最强 GUS 基因表达水平 (其中 GUS 基因表达受 ecc0001 基因表达调控性 DNA 调控) (指定为 100%) 的相对值。

[0102] 以上结果表明 ecc0001 基因表达调控性 DNA 可以成熟叶特异性方式诱导基因表达, 其程度比受 CaMV35S 启动子调控的 GUS 基因表达水平大约高 36 倍。

[0103] 工业实用性

[0104] 本发明的基因表达调控性 DNA 具有以成熟叶特异性方式启动基因表达的活性。因为所述 DNA 具有这样的特征, 所以其可用于产生具有期望特性的植物, 所述期望特性通过使期望基因以成熟叶特异性方式表达来实现。

[0105] 本文引用的所有公开、专利和专利申请通过引用以整体并入本文。

[0001]

IP1346628P. 序列表.txt

序列表

<110> 丰田自动车株式会社
 <120> 参与成熟叶中基因表达调控的DNA
 <130> PH-4976-PCT
 <150> JP 2010-293783
 <151> 2010-12-28
 <160> 5
 <170> PatentIn version 3.4
 <210> 1
 <211> 2019
 <212> DNA
 <213> 甘蔗属栽培变种NIF8
 <400> 1
 gatctgtggt attttggact tetattacaa gtttacagac acagggctca tctgcagtcg 60
 gtgggggaga taaggagega atggaggggc acaaatgtga cgccacgceg gggaaggttc 120
 tctgatcgga gtgtattacg gacgggcegt ctgctccgtc ggcgttttgg gcttttgggg 180
 tagggcggat ggactcgacg tcggaactac caacatgatg attgatgaac gcgcgccgca 240
 cgcattggcg cgtgcgtccg tgcctccgtg ctgctacgc cccaceaacc ttctccacta 300
 cgcccggaca aatgccgtgg caaggagga caacacccac gctaaactaac gctcaataaa 360
 tcaactttgt actatttgta tagttattaa actggcagca ttccaagaaa tgaccgcacg 420
 ttaattccct ctgactgcct attgccaacc gttecttttc tagaceccat ccctccagtt 480
 ccaaaacaaga tgtattagaa atgettacat gtgtaaatat atgtacataa ataactgctg 540
 cctcgttttt callcattat tlaaaactat gactgagttt gatatglagt agttttgcat 600
 tgattggcac acaatatgta taatgagagt gtgtcaaaact ttctectctca ttttcaactca 660
 tccaaattcg atgcaegaat cctacggcc acaaccgact attaagaata gtccaatttg 720
 gcaactactt aaatagttgc tggagtttga tgttgtctct ccaactccaa caccgatate 780
 ttgtgcgtc cagctgcaaa accatttate actacacatt agcactctcc tcatctttt 840
 ettccaacca tegectctca tggtttgcga taaggatggg gtggagcgat ccattgtatct 900
 cctctagttc aacttaatga aatatatgtg cgaaaacaag gcaaggaatt gtctaattag 960
 ctcttctcagt tgaacecaga ggcgaaccac atatactcat cggaaatca tccccatct 1020
 tagtttggga gacattgate tcatgatctg tcaaggcgtt agagctagcg ttgecttgtt 1080
 gaacattggc ttgatgttg caaagactac atcgttcat taaggggtcg aagacgtgga 1140
 aattgtgcag gacgaagcaa aagccacaca ctgcgttgtt tgaatacttc acegtaattc 1200
 ttggttcaag ctttgagcat acaaccagaa tcaatcttgt ottgaagcct tagtcttgc 1260
 atccaccaac ctcagcgacc ttgacgctt cttctcggag gcaaaagttt tgcagtgat 1320
 caatgacatt tcatcagaca aggcgacatt accggacggg ttcactgggc tcttetacca 1380
 acgcgcgcat naacaagecc cagagatgta aaacaacag tttatataca taattaaaac 1440
 tagtaaatte aagcgtgccc gagtgaatct caagctgacg tgtgtagtte cgcgcgtcgc 1500
 actcaagaaa ggtcaaaagt ggacacgaat ctcaagcaa tgcataatgca tatgcacatg 1560
 caaagcaca acccgacgtc cccgtgccag acccggtggc ccaccgcatg aagtacgtac 1620

[0002]

IP1346628P. 序列表.txt

```

cgaattagca gatggaggaa tcttcttcgg ttgagacggt gctttccgct acaccttctc 1680
tgccagtcig ctgtctgctc tgcctcaggg tgctctgtag ctgagccctg caggtgatgt 1740
gctagctttt gactggeaca tgcctgtgtt ttgtgtagtc tagtgtagec ggcgtagcaa 1800
cgcttgggtg gccactaccg atttttaaag ccattacgct ttctacace gcaaggctca 1860
gtcaccacc caccacactc ctgagaacga aagcccccag aagctcagct cagctgctca 1920
caccctcac ccacaggcca cactcacage ggagcagagt agtgageact gagcaggcca 1980
gcagagcgcg cggctcagttc aggtctgctg tcagccatg 2019

```

<210> 2
 <211> 1321
 <212> DNA
 <213> 白甘蔗 (*Saccharum officinarum*)

```

<400> 2
acgcaccac ccactcactc agaacgaaa ccccagcaag ctacgctcag ctgctcacac 60
ctcacacca caggccacac tcacagcgga gcagagtagt gageactgag cagggeagca 120
gagcgcgcg tcagttcagg tcatctgta gccatgaca aggatgaca gcaccaaccg 180
cagcagcacc acctgctgct ctgctcgggg tgcattgctt cgtgctctgt ggcctggtc 240
ctcagctgct ccttcactgc gctgctcact tacctctccc tccgcccctc gaagcctgct 300
ttctacctgc aggacctgca gctgctgggg cccatctctc tggccgacc gtcctgacg 360
gctgctggcg aggtgactct ggctctccc aaccccaac agcactgagg catctctcac 420
aagcctctgg acgtgttctt gacctaccag aacgagcgg tcacggtgct cgtgctgctg 480
ccgcgcagt accagggeca ccgcagctc accgtctggt cccctgctgt gtcgcccag 540
tcctgctccg tggcgggta ctgctccag gacctcaagc gtcagctgct ggcggcttc 600
gtcctctctc aggtcaaggt ggacggccc gtcaggtgga aggtcggcag ctgggtctcc 660
gggagctacc acctcttctt cagctgccc cccgtctctt cccgagggta ccccgctctc 720
gtggcgccg cgcgcaacaa caccgtgctt tctctcaagt tctgctgctc caccggatgc 780
agcctgaggg tctgctgctg tccactgatg acgcactcaa gaacagttgc caagtactgc 840
tactcaaat tctttgggta caccgttaat tacctgttta atttgttca tctctcattc 900
tcaattctca tctctctgtg gggaataagt tctctgctca ggccatag ctgttcaaaa 960
acttgaagtc aattgtacca acatgtgttt attcaagaa gaaaagagg ccaactgcaaa 1020
aacatacata gatgcctaa ccccagcat gctcaaaact agaagaacte ccaccgagc 1080
ctatagtggc gactgaatac tgccacaac tgggggggg cgcctaaaag attcactct 1140
aatttctctg tctggaaaaga attaatctct tggaggggc ccaaaaataa ggcaccggc 1200
cccagtttaa aagaggatga ggggagaaac accggcgttc acaaaatata atgcccctga 1260
cagtaaaccc tccctctac catggggtgt ataaacaata aaaggccaca ctataaggc 1320
c 1321

```

<210> 3
 <211> 2031
 <212> DNA
 <213> 人工

<220>
 <223> 合成的

[0003]

IP1346628P. 序列表.txt

<400> 3
 aagcttggatc tgggtatatt tggactteta ttacaagitt acagacacag ggctcatctg 60
 cagtoggtgg gggagataag gagegaatgg aggggcacaa atgtgacgce acgccgggga 120
 aggttctctg atcggagtgt attacggacg ggcgctctgc tccgtggcg ttttgggett 180
 ttggggtagg gggatggac tegacgtcgg aactaccaac atgatgattg atgaacgege 240
 gcccgacgca ttggcgcgtg cgtccgtgce tccgtgcttg ctacgcccca ceaacctet 300
 ccaactacgcc cggacaaaatg ccgtggcaag gagggacaac acccacgcta actaacgctc 360
 atataatcaa ctttgtacta tttgtatagt tattaaactg gcagcattcc aagaaatgac 420
 cgcacgttaa ttccctctga ctgcctattg ccaaccgttc cttttctaga ccccatecat 480
 ccagttccaa acaagatgta ttagaatgce ttacatgtgt aaafatatgt acataaatac 540
 tgcttgccct gtttttcatt cattatttaa aactatgact gagtttgata ttagtagtatt 600
 ttgcattgat tggcacacaa tatgtataat gagagtgtgt caaactttcc tctcattttt 660
 cactcatcca aattcgatgc acgaatecct acggccacaa ccgactatta agaatagtcc 720
 aatttggcac tactetaaat agttgctgga gtttgatggt gtctctccaa ctccaacacg 780
 gatattcttg gcgctccagc tgcaaaacca tttatcacta cacattagca tctctccat 840
 ccttttcttc caaccatgct ctctcatggt ttgcgataag gatggggtgg agcgatccat 900
 gtatctctct tagttcaact taatgaaata tatgtgcgaa aacaaggcaa ggaattgtct 960
 aattagctct ttcagttgaa accagagcgc aaccacatat actcatcgga acatcatccc 1020
 catccttagt ttgggagaca ttgatctcat gatctgcaa ggcgttagag ctagegttgc 1080
 cttgttgaac attgctttg atgttgcaaa gactacatcg ctteattaa gggtcgaaga 1140
 cgtggaaatt gtgcaggacg aageaaaagc cacacactgc gttgtttgaa tacttcaccg 1200
 taattcttgg tcaagcttt gagcacaaca ccagaatcaa tcttctcttg aagccttagt 1260
 ctgcaatcc accaacctca ggcacctga cgtttgcttc tggaggcaa aagttttgct 1320
 agtgatcaat gacatttcat cagacaagge gacattaccg gacgggttca ctgggetett 1380
 ctaccaacgc cgcataaac aagccccaga gatgtaaaac aaacagtta tatacataat 1440
 taaaactagt aaattcaagc gtggccgagt gaatctcaag ctgacgtgtg tagttccgc 1500
 cgtcgcactc aagaaagtc aaaggtggac acgaatctca aagcaatgca tatgcatatg 1560
 cactatgaca agcacaaccc gacgtccccg tgcagacccc ggtgcccac cggatgaagt 1620
 acgtaccgaa tttagcagatg gaggaatctt cttegggtga gacgttgctt tccgtacac 1680
 ctctctgccc agtctgctgt ctgctctgct gcaggetgct tgetagetga gccctgcagg 1740
 tgatgtgcta gcttttgagt ggcacatgct ctgtgtttgt gtagcttagt gtagccggcg 1800
 tagcaacgct tgggtgtgcca ctaccgattt taaaagecat tacgctttcc tacaccgca 1860
 ggctcagtea cccacceacc cactcactca gaacgaaagc ccagcaage tcagctcage 1920
 tgctcacacc tcacaccacc aggccacact cacagcgag cagagtagtg agcactgagc 1980
 agggcagcag agcgcgcggt cagttcaggt cgatcgtcag ccatgcctag g 2031

<210> 4
 <211> 2044
 <212> DNA
 <213> 甘蔗属栽培变种NiF8

[0004]

IP1346628P. 序列表.txt

<400> 4
gatctgtggt attttggact tetattacaa gtttacagac acagggetca tctgcagtcg 60
gtgggggaga taaggagega atggagggge acaaatgtga cgccacgcgc gggaaggite 120
tetgatogga gtgtattacg gacgggecgt ctgctcogtc ggcgttttgg gottttgggg 180
tagggcggat ggactcgaec tcggaactac caacatgatg attgatgaac gcgcgcccga 240
cgcatlggcg cgtgcgtccg tgcctcogtg cttgctacgc cccaccaacc ttctccaacta 300
cgcccggaca aatgcgctgg caaggaggga caacaccac gctaactaac gctcatataa 360
tcaactttgt actatttga tagttattaa actggcagca ttccaagaaa tgaccgcacg 420
ttaattcctt ctgactgect atggccaacc gttccttttc tagaccceat ccatgccagt 480
tccaaacaag atgtattaga aatgettaca tgtgtaaata tatgtacata aatactgett 540
gcctegtttt tcattcatta tttaaaacta tgaetgagtt tgatatgtag tagttttgca 600
ttgatltgca cacaatatgi ataagagag tgtgtcaaac ittcctcctc atttcaecte 660
atccaaatc gatgcacgaa tccctacgcg cacaaccgac tattaagaat agtccaattt 720
ggcaactcct taaatagttg ctggagtttg atgttgtctc tccaactcca acacggatct 780
ctttgtcctt ccagctgcaa aaccatttat cactacacat tagcctcctc ctcatccttt 840
tcttccaacc atgcctctc atggtttgag ataaggatgg ggtggagega tccatgtatc 900
tcctctagtt caacttaatg aaatatatgt gcgaaaacaa ggcaaggaat tgtctaatta 960
gctctttcag ttgaacceag aggegaacca catatactca tcggaacatc atcccctcc 1020
ttagtttggg agacattgat ctcatgatct gccaagcgt tagagctagc gttgccttgt 1080
tgaacattgg ctttgatgtt gcaaagacta catccttca ttaaggggtc gaagacgtgg 1140
aaattgtgca ggacgaagea aaagccacac actgcgttgt ttgaatactt caccgtaatt 1200
ctttgttcaa gclttgagca tacaaccaga atcaactttg tcttgaagcc ttagtcttgc 1260
aatccacca cctcagcagc cttagcctt gcttctcggg ggcaaaagtt ttgtcagtga 1320
tcaatgacat ttcatcagac aagcgacat taccggagg gttcaactgg ctctctacc 1380
aacgcgcct caaacaagcc ccagagatgt aaaacaaca gtttatatac ataattaaaa 1440
ctagtaaat caagcgtgga cgagtgaatc tcaagctgac gtgtgtagtt cccgcctgag 1500
cactcaagaa aggtcaaagg tggacacgaa tctcaagca atgcatatgc atatgcacat 1560
gcacaagcgc aaccgcagct cccctgcca gaccgggtgg cccaccgat gaagtaagta 1620
ccgaattage agatggagga atctctctg gttgagacgt tcttctcgc tacacctct 1680
ctgccagctt gctgtctgct ctgctgcagg ctgtctgcta gctgagcct gcagcgagg 1740
ttaaccaate gtccattctg gatgtgctag cttttgagtg gcacatgcgc tgtgtttgtg 1800
tagtctagtg tagccggcgt agcaacgctt ggtgtgccac taccgatttt aaaagccatt 1860
acgtttctt acaccgcag cctcagtcac ccaaccacc actcactcag aacgaaagcc 1920
ccageaaget cagctcagct gctcacact caccaccaca ggccacactc acagcggage 1980
agagttagtg cactgagca gggcagcaga gcgcgcgttc agttcaggtc gatcgtcagc 2040
catg 2044

<210> 5
<211> 2056

[0005]

IP1346628P. 序列表.txt

| | |
|--|------|
| <212> DNA | |
| <213> 人工 | |
| <220> | |
| <223> 合成的 | |
| <400> 5 | |
| aagcttgatc tgtggtatit tggactteta ttacaagttt acagacacag ggctcatctg | 60 |
| cagtcgggtgg gggagataag gagcgaatgg aggggcacaa atgtgacgcc acgccgggga | 120 |
| aggtttctctg atcggagtgt attacggacg ggccgtctgc tccgtcggcg ttttgggctt | 180 |
| ttgggtagg gggatggac togaegtgg aactaccaac atgatgattg atgaacgcgc | 240 |
| gcccagcga ttggcgcgig cgtccgtgcc tccgtgcttg ctacgcccca ccaaccttct | 300 |
| ccactacgcc cggacaaaig ccgtggcaag gagggacaac acccacgcta actaacgctc | 360 |
| atataatcaa ctttgtacta ttgtatagt tattaacctg gcagcattcc aagaatgac | 420 |
| cgcacgttaa ttcttctga ctgaccttg ccaaccgttc cttttctaga ccccatccat | 480 |
| gcagttcca aacaagatgt attagaaatg cttacatgtg taaatataig tacataaata | 540 |
| ctgcttgcct cgtttttcat tcattattta aaactatgac tgagtttgat atgtagtatg | 600 |
| ttlgcaatga ttggcacaca ataigtataa tgagagtgig taaaacttgc ctctcattt | 660 |
| tcactcatcc aaattgatg cagcaatecc tacggccaca accgactatt aagaatagtc | 720 |
| caatttggca ctactctaaa tagttgctgg agtttgatgt tgfctctcca actccaacac | 780 |
| ggatatcttg tgcgcaccag ctgcaaaaacc atttatcact acacattagc atccctctca | 840 |
| tccttttctt ccaaccatcg cctctcatgg ttgagataa ggatggggig gagegateca | 900 |
| tgtatctcct ctagtccaac ttaatgaaat atatgtgoga aaacaaggca aggaattgtc | 960 |
| taattlagic ttlcagtga aaccagagge gaaccacata tactcatcgg aacatcacc | 1020 |
| ccatccttag ttggggagac attgatctca tgatctgta aggegttaga gctagegttg | 1080 |
| ccttgttgaa cattgcttt gatgttgea agactacatc gcttcattaa ggggtcgaag | 1140 |
| acgtgaaaat tglgcaggac gaagcaaaaag ccacacactg cgttglllga atactcacc | 1200 |
| gtaattcttg gtccaagctt tgagcataca accagaatea atcttctctt gaageccttag | 1260 |
| tcttgcaate caccaacctc agcagacctg acgcttgctt ctggaggca aaagttttgt | 1320 |
| cagtgatcaa tgacatttca tcagacaagg cgacattacc ggacgggttc actgggctct | 1380 |
| ctfaccaaag ccgatcaaaa caagccccag agatgtaaaa caaacagttt atatacataa | 1440 |
| ttaaaactag taaattcaag cgtggccgag tgaatctcaa gctgacgtgt gtagttcccg | 1500 |
| ccgtcgcact caagaaaggt caaagggtga cacgaatctc aaagcaatgc atatgcatal | 1560 |
| geacatgeac aagcacaacc cgacgtcccc gtgccagacc cgggtgcccc ccggatgaag | 1620 |
| tacgtaccga attagcagat ggaggaatct tcttcggttg agacgttget ttccgetaca | 1680 |
| ccittctctg cagtelgelg tctgctctg tgcaggctgt ctgclagctg agccctgcag | 1740 |
| cgagggttaa ccaatgctcc attcgtgatg tgctagcttt tgagtggcac atgcctgtg | 1800 |
| tttgttagt ctagttagc egcgtagca acgcttggtg tgccactacc gattttaaaa | 1860 |
| gccattacgc ttctctaac cgcacggctc agtcccccac ccaccacctc actcagaacg | 1920 |
| aaagccccag caagctcagc tcagctgctc acacctcaca cccacagccc aaactcacag | 1980 |
| cggagcagag tagtgagcac tgagcagggc agcagagcgc ggggtcagtt caggctcagc | 2040 |

[0006]

| | | |
|-------------------|---------------------|------|
| gtcagccatg cctagg | IP1346628P. 序列表.txt | 2056 |
|-------------------|---------------------|------|

1 AGGCACCCAG CCACTCACTC AGAAGGAAG CCCCAGCAAG CTCAGCTCAG CTGCTCACAC
61 CTCACACCCA CAGGCCACAC TCACAGCGGA GCAGAGTAGT GAGCACTGAG CAGGGCAGCA
121 GAGCGCGGG TCAGTTCAGG TCGATCGTCA GCCATGAGCA AGGATGACAA GCACCACCGC
181 CACGAGCAC ACCTGCCTCG CTTTCATCGC GCTGGTCATC TACCTGGCC TCCGCCCTC GAAGCCGTCC
241 CTCATCGTC CTTTCATCGC GCTGGTCATC TACCTGGCC TCCGCCCTC GAAGCCGTCC
301 TTCTACCTGC AGGACCTGCA GCTGGCGGG CCCATCTGG TGGCCGACCC GTCGCTGACG
361 GCGTCGGGC AGGTGACGCT GCGTCCCGC AACCCCAAG AGCACGTGG CATCTTCTAC
421 AAGCGCCTGG ACGTGTTCGT GACCTACCAG AACGAGCGG TCACGGTGC CGTGTCTGCTG
481 CCGCCGCAGT ACCAGGCCA CCGGACGTC ACCGTCTGGT CCCCCTGCT GTCGGCCGAG
541 TCCGTGCCCC TGGCCGGCTA CGTCGCCGAC GACCTCAAG GTGACGTGC GCGCCGCTTC
601 GTCGCTCTGC AGGTCAAGGT GGACGGCCG GTCAAAGTGA AGTTCGGCAG CTGGTCTCC
661 GGGAGCTACC ACCTCTTCGT CAGCTGCCC GCGTGTCTT CCGCGGGTA CCCCAGCGTC
721 GTGGCGGCG GCGGCAACA CACCGTGTG TCGCTCAGT TCGCGCAGC CACGGGATGC
781 AGCGTCGAGG TGTGATGTA TCCACTGATG ACGCACTCAA GAACAGTTGC CAAGTACTGC
841 TACTCAAAAT TCTTTGGGT TACTGTTAAT TACGCTGTTA ATTTTGTTC TGTCTCATTC
901 TCAATTCTCA TTCTCTTGTG GGAATAAGT TTCTGTGCA GGCCATATAG CTGTTCAAAA
961 ACTTGAAGTC AATTGTACCA ACATGTGTTT ATTCAAAGA GAAAAGAGG CCACTGCAAA
1021 AACATACATA GATGGGCTAA CCGGAGCAT GCTCAAACT AGAAGAATC CCACCGACGC
1081 CTATAGTGG GACTGAATAC TGCCACAAC TCGGGGGGG CGCCTAAAAG ATTCACATCT
1141 AATTTGCTG TGTGGARAGA ATTAATCCTC TCGGAGGGG CCAAAAATAA GGCACCGGGC
1201 CCCAGTTAA AAGAGGATGA GGGGAGAAC ACCGGGCTC ACAAATAATA ATGCCCTTGA
1261 CAGTAAACCC TCCCTTCTAC CATGGGTTGT ATAAACAATA AAAGGCCACA CCTATAAGGC
1321 C

图 1

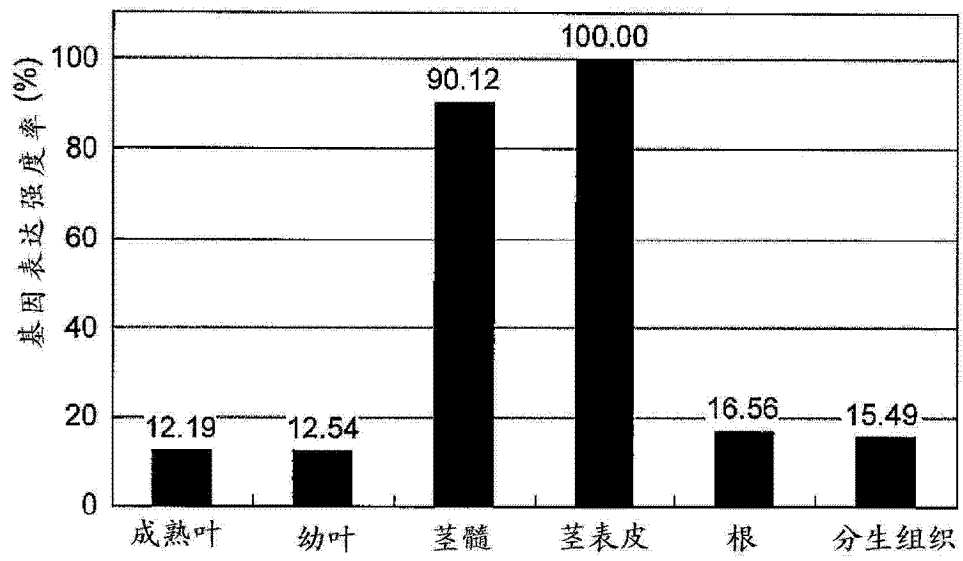


图 2

HindIII

```

AAGCTT GATC TGTGGTATTT TGGACTTCTA TTACAAGTTT ACAGACACAG GGCTCATCTG
CAGTCGGTGG GGGAGATAAG GAGCGAATGG AGGGGCACAA ATGTGACGCC ACGCCGGGGA
AGGTTCTCTG ATCGGAGTGT ATTACGGACG GGCCGTCTGC TCCGTCGGCG TTTTGGGCTT
TTGGGGTAGG GCGGATGGAC TCGACGTCGG AACTACCAAC ATGATGATTG ATGAACGCGC
GCCGACGCA TTGGGCGGTG CGTCCGTGCC TCCGTGCTTG CTACGCCCA CCAACCTTCT
CCACTACGCC CGGACAAATG CCGTGGCAAG GAGGGACAAC ACCCAGGCTA ACTAACGCTC
ATATAATCAA CTTTGTACTA TTTGTATAGT TATTAACTG GCAGCATTC AAGAAATGAC
CGCACGTAA TTCCTTCTGA CTGCCTATTG CCAACCGTTC CTTTTCTAGA CCCCATCCAT
CCAGTTCCAA ACAAGATGTA TTAGAAATGC TTACATGTGT AAATATATGT ACATAAATAC
TGCTTGCCCTC GTTTTTTCATT CATTATTTAA AACTATGACT GAGTTTGATA TGTAGTAGTT
TTGCATTGAT TGGCACACAA TATGTATAAT GAGAGTGTGT CAAACTTTCC TCCTCATTTT
CACTCATCCA AATTGATGCG ACGAATCCCT ACGGCCACAA CCGACTATTA AGAATAGTCC
AATTTGGCAC TACTCTAAAT AGTTGCTGGA GTTTGATGTT GTCTCTCCAA CTCCAACACG
GATATCTTGT GCGCTCCAGC TGCAAAACCA TTTATCACTA CACATTAGCA TCCTCCTCAT
CCTTTTCTTC CAACCATCGC CTCTCATGGT TTGCGATAAG GATGGGGTGG AGCGATCCAT
GTATCTCCTC TAGTTCAACT TAATGAAATA TATGTGCGAA AACAAGGCAA GGAATTGTCT
AATTAGCTCT TTCAGTTGAA ACCAGAGGCG AACCACATAT ACTCATCGGA ACATCATCCC
CATCCTTAGT TTGGGAGACA TTGATCTCAT GATCTGTCAA GGCCTTAGAG CTAGCGTTGC
CTTGTTGAAC ATTGGCTTTG ATGTTGCAAA GACTACATCG CTTCAATTAAG GGGTCGAAGA
CGTGGAAATT GTGCAGGACG AAGCAAAAGC CACACACTGC GTTGTTTGAA TACTTCACCG
TAATTCCTGG TTCAAGCTTT GAGCATACAA CCAGAATCAA TCTTGTCTTG AAGCCTTAGT
CTTGCAATCC ACCAACCTCA GCGACCTTGA CGCTTGCTTC TCGGAGGCAA AAGTTTGTG
AGTGATCAAT GACATTTCAAT CAGACAAGGC GACATTACCG GACGGGTTCA CTGGGCTCTT
CTACCAACGC CGCATCAAAC AAGCCCCAGA GATGTAAAAC AAACAGTTTA TATACATAAT
TAAAAC TAGT AAATTC AAGC GTGGCCGAGT GAATCTCAAG CTGACGTGTG TAGTTC CCGC
CGTCGCACTC AAGAAAGGTC AAAGGTGGAC ACGAATCTCA AAGCAATGCA TATGCATATG
CACATGCACA AGCACAACCC GACGTCCCCG TGCCAGACCC GGTGGCCAC CCGATGAAGT
ACGTACCGAA TTAGCAGATG GAGGAATCTT CTTGCGTTGA GACGTTGCTT FCCGCTACAC
CTTCTCTGCC AGTCTGCTGT CPGCTCTGCT GCAGGCTGTC TGCTAGCTGA GCCCTGCAG
GTGATGTGCT AGCTTTTGAG TGGCACATGC GCTGTGTTG TGTAGTCTAG TGTAGCCGGC
GTAGCAACGC TTGGTGTGCC ACTACCGATT TAAAAGCCA TTACGCTTTC CTACACCGCA
CGGCTCAGTC ACCCACCCAC CCACTCACTC AGAACGAAAG CCCAGCAAG CTCAGCTCAG
CTGCTCACAC CTCACACCCA CAGGCCACAC TCACAGCGGA GCAGAGTAGT GAGCACTGAG
CAGGGCAGCA GAGCGCGCG TCAGTT CAGG TCGATCGTCA GCCATG CCTA GG

```

起始
 密码子 Bln I

图 3

处理的序列:

```

1      AAGCTTGATCTGTGGTATTTTGGACTTCTATTACAAGTTTACAGACACAG
51     GGCTCATCTGCAGTCGGTGGGGGAGATAAGGAGCGAATGGAGGGGCACAA
101    ATGTGACGCCACGCCGGGAAGGTTCTCTGATCGGAGTGTATTACGGACG
151    GGCCGTCTGCTCCGTGGCGTTTTGGGCTTTGGGTAGGGCGGATGGAC
201    TCGACGTGGAACTACCAACATGATGATTGATGAACGGCGGCCGACGCA
251    TTGGGGCGTGGTCCGTGCCTCCGTGCTTGTACGCCCCACCAACCTTCT
301    CCACTACGCCCGGACAAATGCCGTGGCAAGGAGGGACAACACCACGCTA
351    ACTAACGCTCATATAATCAACTTTGTACTATTTGTATAGTTATTAACCTG
401    GCAGCATTCCAAGAAATGACCGCACGTTAATTCCTTCTGACTGCCTATTG
451    CCAACGGTTCCTTTTCTAGACCCCATCCATGCCAGTTCAAACAAGATGT
501    ATTAGAAATGCTTACATGTGTAATATATGTACATAAACTGCTTGCCCT
551    CGTTTTTCATTCATTATTTAAACTATGACTGAGTTTGATATGTAGTAGT
601    TTTGCATTGATTGGCACACAATATGTATAATGAGAGTGTGTCAAACCTTC
651    CTCGTCATTTTCACTCATCCAAATTCGATGCACGAATCCCTACGGCCACA
701    ACCGACTATTAAGAATAGTCCAATTTGGCACTACTCTAAATAGTTGCTGG
751    AGTTTGATGTTGTCTCTCCAACCTCCAACACGGATATCTTGTGCGCTCCAG
801    CTGAAAACCATTATCACTACACATTAGCATCCTCCTCATCCTTTTCTT
851    CCAACCATCGCCTCTCATGGTTTGGGATAAGGATGGGGTGGAGCGATCCA
901    TGTATCTCCTCTAGTTCAACTTAATGAAATATATGTGGAAAACAAGGCA
951    AGGAATTGTCTAATTAGCTCTTTCAGTTGAAACCAGAGGCGAACCACATA
1001   TACTCATCGGAACATCATCCGCATCCTTAGTTTGGGAGACATTGATCTCA
1051   TGATCTGTCAAGGCGTTAGAGCTAGCGTTGCCTTGTTGAACATTGGCTTT
1101   GATGTTGCAAAGACTACATCGCTTCATTAAGGGTTCGAAGACGTGGAAT
1151   TGTGCAGGACGAAGCAAAAGCCACACACTGCGTTGTTTGAATACTTCACC
1201   GTAATTCTTGGTTCAAGCTTTGAGCATACAACCAGAATCAATCTTGTCTT
1251   GAAGCCTTAGTCTTGCAATCCACCAACCTCAGCGACCTTGACGCTTGCTT
1301   CTCGGAGGCAAAAGTTTTGTCAGTGATCAATGACATTTCATCAGACAAGG
1351   CGACATTACGGGACGGGTTCACTGGGCTTCTACCAACGCCGCATCAAA
1401   CAAGCCCCAGAGATGTAACAAACAGTTTATATACATAATTAACCTAG
1451   TAAATTCAAGCGTGGCCGAGTGAATCTCAAGCTGACGTGTGTAGTTCCCG
1501   CCGTCGCACTCAAGAAAGGTCAAAGGTGGACACGAATCTCAAAGCAATGC
1551   ATATGCATATGCACATGCACAAGCACAACCCGACGTCCCGTGCCAGACC
1601   CCGTGGCCACCGGATGAAGTACGTACCGAATTAGCAGATGGAGGAATCT
1651   TCTTCGGTTGAGACGTTGCTTCCGCTACACCTTCTCTGCCAGTCTGCTG

```

图 4-1

```

1701 TCTGCTCTGCTGCAGGCTGTCTGCTAGCTGAGCCCTGCAGCGAGGGTTAA
1751 CCAATCGTCCATTTCGTGATGTGCTAGCTTTTGAGTGGCAGATGCGCTGTG
1801 TTTGTGTAGTCTAGTGTAGCCGGGTAGCAACGCTTGGTGTGCCACTACC
1851 GATTTTAAAAGCCATTACGCTTTCCTACACCGCAGGGCTCAGTCACCCAC
1901 CCAGCCACTCACTCAGAACGAAAGCCCCAGCAAGCTCAGCTCAGCTGCTC
1951 ACACCTCACACCCACAGGCCACACTCACAGCGGAGCAGAGTAGTGAGCAC
2001 TGAGCAGGGCAGCAGAGCGCGGGTCAGTTCAGGTCGATCGTCAGCCATG
2051 CCTAGGN

```

□ : 预测的启动子区

(HindIII限制酶识别序列 (AAGCTT) 和BlnI限制酶识别序列 (CCTAGG) 分别存在于5'端和3'端。)

Proscan: 1.7版

处理的序列: 2057个碱基对

在反义链1669位至1419位预测的启动子区

启动子评分: 64.48 (启动子截断值 = 53.000000)

在1437位发现的TATA, Est.TSS = 1405

显著信号:

| 名称 | 链 | 位置 | 权重 |
|----------|---|------|----------|
| AP-2 | + | 1607 | 1.091000 |
| Sp1 | - | 1502 | 3.191000 |
| Sp1 | + | 1497 | 3.119000 |
| ATF/CREB | - | 1488 | 1.564000 |
| E4F1 | - | 1488 | 1.201000 |
| NF-S | - | 1487 | 1.019000 |
| CREB | + | 1483 | 5.737000 |
| ATF | + | 1483 | 1.591000 |
| UCE.2 | + | 1464 | 1.216000 |
| TFIID | - | 1435 | 1.971000 |
| TFIID | - | 1433 | 2.618000 |

图 4-2

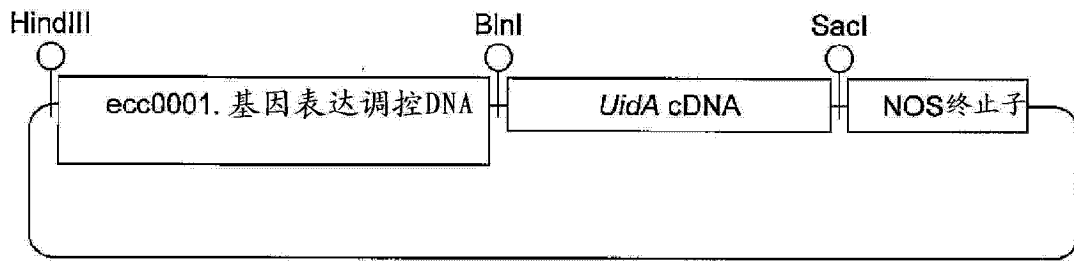


图 5

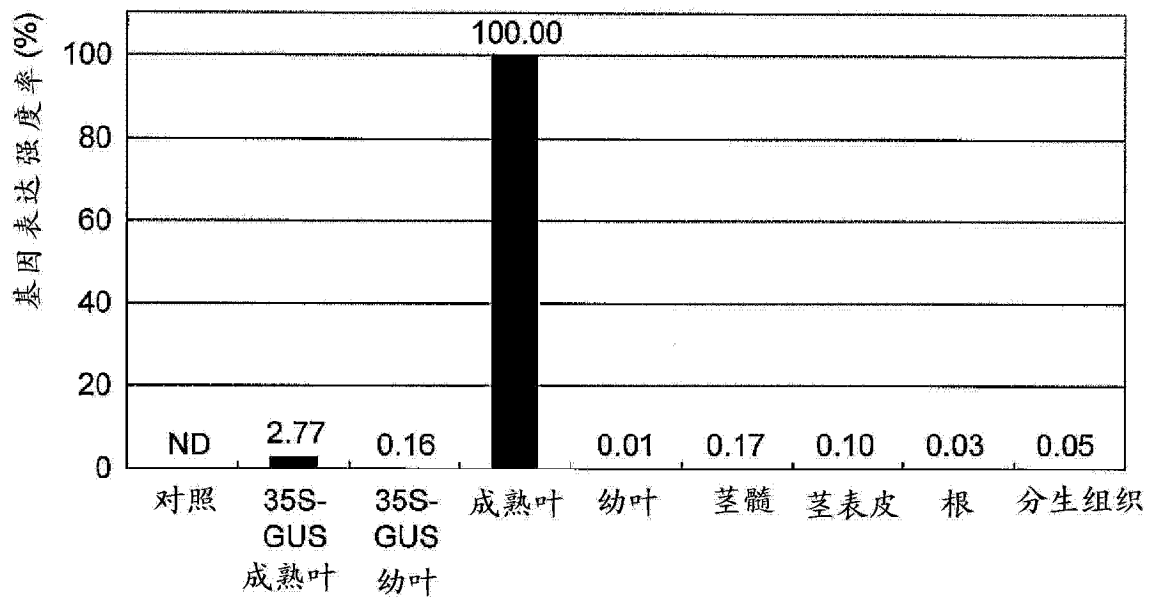


图 6