



(12)发明专利

(10)授权公告号 CN 104395467 B

(45)授权公告日 2017.05.03

(21)申请号 201280074082.3

(22)申请日 2012.11.06

(65)同一申请的已公布的文献号
申请公布号 CN 104395467 A

(43)申请公布日 2015.03.04

(30)优先权数据
2012-140231 2012.06.21 JP

(85)PCT国际申请进入国家阶段日
2014.12.18

(86)PCT国际申请的申请数据
PCT/JP2012/078716 2012.11.06

(87)PCT国际申请的公布数据
W02013/190720 JA 2013.12.27

(73)专利权人 丰田自动车株式会社
地址 日本爱知县

(72)发明人 西村哲 铃木一代 都筑祥子

(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 金世煜 苗堃

(51)Int.Cl.
C12N 15/09(2006.01)
A01H 1/00(2006.01)
A01H 5/00(2006.01)

(56)对比文件
CN 1487997 A,2004.04.07,
CN 103298829 A,2013.09.11,
BESNARD,G.,et al.Characterisation of
the phosphoenolpyruvate carboxylase gene
family in sugarcane(Saccharum spp.).
《Theoretical and Applied Genetics》.2003,
第107卷(第3期),470-478.

审查员 靳春鹏

权利要求书1页 说明书12页
序列表6页 附图7页

(54)发明名称

甘蔗成花控制技术

(57)摘要

本发明提供能够使植物特别是甘蔗及其近缘属种进行有效的交配育种的技术。

1. 一种DNA,由下述DNA构成:序列号1所示的碱基序列构成的DNA和编码由序列号7所示的氨基酸序列构成的多肽的DNA;

具有促进植物的成花或出穗的功能和/或促进植物的分蘖的功能。

2. 一种重组载体,包含权利要求1所述的DNA。

3. 一种属于甘蔗属的转化植物的制造方法,导入权利要求1所述的DNA或权利要求2所述的重组载体。

甘蔗成花控制技术

技术领域

[0001] 本发明涉及包含具有对成熟叶特异性促进基因表达的活性的基因表达控制DNA和成花诱导基因的、具有促进植物的成花或出穗的功能和/或促进植物的分蘖的功能的DNA及其应用。

背景技术

[0002] 甘蔗是在巴西、印度等以约2000万ha生产、被用于制糖用·乙醇生产用途的资源作物,今后由于生物燃料的普及而有持续需求增长的前景,期待其增产。

[0003] 出于作为原料作物的甘蔗的廉价并且稳定的供给的目的,甘蔗的品种改良和育种盛行。一直以来,甘蔗的品种改良和育种通过交配来进行(非专利文献1)。

[0004] 然而,由于甘蔗是起源于热带地方的植物,所以在国土几乎都属于温带的日本国内不能出穗·开花,而且由于基因组构成复杂,因此有效的品种改良和育种是非常困难的。

[0005] 另外,以往的杂交育种,以依靠经验、直觉的组合杂交进行交配,对大量后代作出全面评价来进行选拔。一般而言,为了进行交配,必须经过成花诱导、开花·受粉、结实促进·采种的过程,对于甘蔗的情况,甚至在适合生长发育的地方,1年也只能组织1次该过程。因此,开发出一个品种必须花费非常长的时间。

[0006] 并且,在交配的品种是不易开花的品种、要交配的品种彼此的开花时期不吻合的情况下,进行所希望的交配是非常困难的。

[0007] 因此,在该领域中,迫切需要甘蔗的有效的交配育种方法。

[0008] 迄今为止,报告了通过在拟南芥、稻子中过量表达了FT基因或者OsHd3a基因之类的成花诱导基因,能够诱导成花(出穗)(专利文献1-3、非专利文献2-3)。另外,现在,FT基因或者OsHd3a基因被确认在许多植物种中发挥相同的效果。

[0009] 然而,没有在甘蔗中使这些基因过量表达的重组体的报告。

[0010] 现有技术文献

[0011] 专利文献

[0012] 专利文献1:日本特开2000-139250号公报

[0013] 专利文献2:日本特表2002-511270号公报

[0014] 专利文献3:日本特开2002-153283号公报

[0015] 非专利文献

[0016] 非专利文献1:宫里清松,甘蔗及其栽培,1986年,日本分蜜糖工业会发行

[0017] 非专利文献2:Kardailsky I.et al.,Science.1999Dec 3;286(5446):1962-5.

[0018] 非专利文献3:Kojima S.et al.,Plant Cell Physiol 2002Oct;43(10):1096-105.

发明内容

[0019] 本发明的目的在于提供能够使植物特别是甘蔗及其近缘属种进行有效的交配育

种的技术。

[0020] 本发明人等为了解决上述课题进行了深入研究,结果发现在具有对成熟叶特异性促进基因表达的活性的基因表达控制DNA的控制下,使成花诱导基因表达,从而能促进植物体的成花(出穗),能够有效地得到可交配的植物体,进而完成本发明。

[0021] 即,本发明包含以下[1]~[5]的特征。

[0022] [1]一种DNA,包含选自以下的(a)~(d)中的任一种DNA和编码选自以下的(e)~(g)中的任一种多肽的DNA,具有促进植物的成花或出穗的功能和/或促进植物的分蘖的功能,

[0023] (a)由序列号1所示的碱基序列构成的DNA,

[0024] (b)由在序列号1所示的碱基序列中缺失、取代、添加或插入了1个~多个碱基的碱基序列构成,并且具有对成熟叶特异性促进基因表达的活性的DNA,

[0025] (c)由与序列号1所示的碱基序列具有90%以上的序列同源性的碱基序列构成,并且具有对成熟叶特异性促进基因表达的活性的DNA,

[0026] (d)在严格的条件下与由跟序列号1所示的碱基序列互补的序列构成的DNA杂交,并且具有对成熟叶特异性促进基因表达的活性的DNA;

[0027] (e)由序列号7所示的氨基酸序列构成的多肽,

[0028] (f)由在序列号7所示的氨基酸序列中缺失、取代、添加或插入了1个~多个氨基酸的氨基酸序列构成,并且具有促进成花或出穗的活性的多肽,

[0029] (g)由与序列号7所示的氨基酸序列具有90%以上的序列同源性的氨基酸序列构成,并且具有促进成花或出穗的活性的多肽。

[0030] [2]一种重组载体,包含[1]的DNA。

[0031] [3]一种转化植物,导入了[1]的DNA或[2]的重组载体。

[0032] [4]根据[3]的转化植物,属于禾本科。

[0033] [5]根据[4]的转化植物,属于甘蔗属、高粱属或者芒属。

[0034] 本说明书包含作为本申请的优先权的基础的日本国专利申请2012-140231号的说明书和/或附图所记载的内容。

[0035] 根据本发明,能够提供可使植物特别是甘蔗及其近缘属种进行有效的交配育种的技术。

附图说明

[0036] 图1表示来源于甘蔗(*Saccharum officinarum*)的ecc0002EST的碱基序列(序列号4)(用DFCI Sugarcane Gene Index检索)。

[0037] 图2表示*Saccharum spp.cv.NiF8*的各组织中的ecc0002EST的表达量的分析结果。将表达最强的成熟叶的表达量表示为100%。

[0038] 图3表示在5'末端导入HindIII限制酶识别序列和在3'末端导入BlnI限制酶识别序列的ecc0002基因的基因表达控制区域的碱基序列(序列号5)。

[0039] 图4表示(A)来源于粳稻组(*Oryza sativa Japonica Group*)的Hd3a基因的碱基序列(CDS:153-692)(序列号6)和(B)该基因所编码的多肽的氨基酸序列(序列号7)。

[0040] 图5表示包含彼此连接的ecc0002基因的基因表达控制区域和 β -葡萄糖醛酸酶基

因的基因表达载体的示意图。

[0041] 图6表示利用ecc0002基因的表达控制DNA来控制 β -葡萄糖醛酸酶基因的表达的基因重组甘蔗的各组织中的GUS基因的表达量的分析结果。将观察到的最强的表达量的该基因重组甘蔗的成熟叶的表达量表示为100%。

[0042] 图7表示向基因表达载体插入的包含ecc0002基因的表达控制DNA和稻子Hd3a基因的DNA的碱基序列(序列号3)。大文字: ecc0002基因的表达控制DNA; 小文字: 稻子Hd3a基因。

[0043] 图8表示包含彼此连接的ecc0002基因的基因表达控制区域和稻子Hd3a基因的基因表达载体的示意图。

[0044] 图9表示利用ecc0002基因的表达控制DNA来控制稻子Hd3a基因的基因重组甘蔗、利用CaMV 35S启动子来控制稻子Hd3a基因的表达的基因重组甘蔗以及野生型甘蔗的各组织中的稻子Hd3a基因的表达量的分析结果。表中, “基因重组体ecc0002pro.”表示利用ecc0002基因的基因表达控制区域来控制稻子Hd3a基因的表达的基因重组甘蔗。另外, “基因重组体35SCaMV pro.”表示通过CaMV 35S启动子来控制稻子Hd3a基因的表达的基因重组甘蔗。

[0045] 图10表示利用ecc0002基因的表达控制DNA来控制稻子Hd3a基因的基因重组甘蔗的分蘖诱导能力的分析结果。分蘖数为基因重组甘蔗(n=5)和野生型(n=3)的平均值。

[0046] 图11表示利用ecc0002基因的表达控制DNA来控制稻子Hd3a基因的基因重组甘蔗的出穗诱导能力的分析结果。出穗数为基因重组甘蔗(n=5)和野生型(n=3)的平均值。

具体实施方式

[0047] 以下,对本发明进行详细说明。

[0048] 本发明的具有促进植物的成花或出穗的功能和/或促进植物的分蘖的功能的DNA,包含具有对成熟叶特异性促进基因表达的活性的基因表达控制DNA和成花诱导基因。

[0049] 本发明中,“促进植物的成花或出穗”是指在导入了具有该功能的DNA的植物体中,促进花芽、花器官的分化·形成和/或促进出穗和/或开花(缩短到出穗和/或开花为止的时期)。

[0050] 另外,本发明中,“促进植物的分蘖”是指在导入了具有该功能的DNA的植物体中,促进根附近的侧枝的形成,缩短该侧枝的形成开始时期和/或增大该侧枝的数目。

[0051] 首先,说明具有对成熟叶特异性促进基因表达的活性的基因表达控制DNA。应予说明,以下将“具有对成熟叶特异性促进基因表达的活性的基因表达控制DNA”记载为“基因表达控制DNA”,本说明书中这些术语可彼此互换使用。

[0052] 本发明的基因表达控制DNA包含以下的(a)~(d)中任一种DNA。

[0053] (a) 由序列号1所示的碱基序列构成的DNA,

[0054] (b) 由在序列号1所示的碱基序列中缺失、取代、添加或插入了1个~多个碱基的碱基序列构成,并且具有对成熟叶特异性促进基因表达的活性的DNA,

[0055] (c) 由与序列号1所示的碱基序列具有90%以上的序列同源性的碱基序列构成,并且具有对成熟叶特异性促进基因表达的活性的DNA,

[0056] (d) 在严格的条件下与由跟序列号1所示的碱基序列互补的序列构成的DNA杂交,并且具有对成熟叶特异性促进基因表达的活性的DNA。

[0057] 本发明的基因表达控制DNA可以通过如下方式取得：通过利用来源于甘蔗的各组织(例如,茎、成熟叶、幼叶)的总RNA或者来源于其的cDNA的基因表达分析,取得成熟叶特异性表达的候补基因,对该候补基因评价表达特性,从而确定被评价为成熟叶特异性表达的基因,基于确定的基因的cDNA或者基因组DNA鉴定该候补基因的5'上游区域的碱基序列。这里基因表达分析可以使用DNA芯片、差异显示法等本领域技术人员所公知的基因表达的各种分析方法。

[0058] 具体而言,序列号1的碱基序列存在于在甘蔗的成熟叶中特异性表达的基因(以下,记载为“ecc0002”)的5'上游区域。本说明书中“甘蔗”包含属于甘蔗属的植物,例如甘蔗(*Saccharum officinarum*)、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*)、甘蔗属杂交栽培种(*Saccharum spp. Hybrids cv. NiF8*)等(不特别限定于这些)及其近缘属种,例如高粱等。优选为甘蔗属杂交栽培种(*Saccharum spp. Hybrids cv. NiF8*)。

[0059] 分离5'上游区域的DNA的方法没有特别限定,可以使用本领域技术人员所公知的一般方法进行。例如,可以使用基于上述ecc0002基因的碱基序列(序列号4)克隆未知的区域(这时为5'上游区域)的公知方法进行分离。例如,作为其中之一,对包含ecc0002基因的5'上游区域的基因组DNA进行限制酶处理而使由规定的碱基序列构成的接头(adapter)结合,对ecc0002基因的碱基序列和接头设定引物而进行PCR。由此,能够将与ecc0002基因的碱基序列在5'上游邻接的未知的碱基序列扩增。而且测定扩增了的碱基序列的序列后,基于测定的碱基序列,再设计一对引物,将与测定的碱基序列邻接的另外的未知的碱基序列同样地进行扩增。该方法中,可以使用市售的克隆试剂盒,例如,RightWalk(注册商标)试剂盒(BEX)进行。另外,作为上述以外的方法,可举出基于反向PCR的方法。这时,基于ecc0002基因的碱基序列信息设计一对引物,使用这一对引物和用规定的限制酶处理后而自身连接的基因组DNA片段进行PCR,从而能够扩增ecc0002基因的上游区域。另外,作为其它方法,可举出将ecc0002基因的上游区域从基因组DNA文库分离的方法。此时,将包含ecc0002基因的cDNA作为探针,从根据常规方法制备的基因组DNA文库中筛选包含ecc0002基因的基因组DNA。其后,通过测定筛选出的基因组DNA的碱基序列来确定存在于ecc0002基因的上游区域的5'上游区域,并且,能够通过PCR等仅扩增5'上游区域。

[0060] 这样,依次扩增或者筛选位于ecc0002基因的上游的未知的碱基序列,根据常规方法测定碱基序列,从而能够得到由序列号1表示的碱基序列。这样,如果测定由序列号1表示的碱基序列,则其后使用基于序列号1所示的碱基序列设计的引物,利用以从甘蔗提取出的基因组DNA为模板的PCR,能够得到序列号1所示的碱基序列。

[0061] 序列号1所示的碱基序列,作为在成熟叶中特异性诱导基因表达的基因表达控制区域发挥功能。基因表达控制区域包含启动子区域、增强子区域、TATA盒和/或CAT盒等(不特别限定于这些)、涉及基因的转录调节的碱基序列。

[0062] 本说明书中“特异性”是指在构成植物体的各种组织中仅在成熟叶中具有基因表达诱导功能,除此之外是指成熟叶中的基因表达诱导功能与成熟叶以外的组织(例如,幼叶、茎髓、茎皮、根、分生组织等)中的基因表达诱导功能相比明显高,或者统计学有意义地高(例如,大约2倍、3倍、4倍、5倍、6倍、7倍、8倍、9倍、10倍或者其以上)。

[0063] 本说明书中“成熟叶”是指为了进行光合作用而将叶绿体蓄积在细胞内、带绿色的叶子,是除作为不具有用于光合作用的叶绿体的叶子的幼叶以外的叶组织。

[0064] 基因表达诱导功能可以通过本领域技术人员所公知的报告基因试验等来确认。报告基因试验可以通过如下方式确认:制备在研究基因表达诱导功能的碱基序列的控制下(下游区域)连接各种报告基因(例如, β -葡萄糖醛酸酶基因(GUS)、荧光素酶基因(LUC),绿色荧光蛋白基因(GFP)等)而成的载体,使用该载体导入或者短暂导入宿主的基因组后,测定该报告基因的表达水平。作为该报告基因,只要其表达是可检测的,就没有特别限制,例如,可以举出本领域技术人员一般使用的CAT基因、lacZ基因、荧光素酶基因(以下为LUC)、 β -葡萄糖醛酸酶(以下为GUS)基因以及绿色荧光蛋白(以下为GFP)基因等。

[0065] 报告基因的表达水平,可以根据该报告基因的种类,通过本领域技术人员所公知的方法进行测定。例如,报告基因为CAT基因时,通过检测该基因产物所致的氯霉素的乙酰化,从而能够测定报告基因的表达水平。各报告基因的表达水平可以用以下方法测定。报告基因为lacZ基因时,检测该基因表达产物的催化剂作用所致的色素化合物的发色。为LUC基因时,检测该基因表达产物的催化剂作用所致的荧光化合物的荧光。为GFP基因时,检测GFP蛋白质所致的荧光。例如,报告基因为GUS时,在宿主细胞内的启动子活性可以分别通过如下方式确认:通过(i)利用组织化学的GUS染色的方法(EMBO J.6,3901-3907(1987))和/或根据(ii)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.Schilperoort,Kluwer Academic Publishers)测定GUS活性,进一步根据Bradford的方法(Anal.Biochem.72,248-254(1976))测定蛋白质量,按蛋白质量换算GUS活性(计算为nmole 4-MU/min/mg protein)。

[0066] 另外,使用上述以外的基因作为报告基因时可以通过如下方式进行:使用Northern杂交法、RT-PCR法、DNA阵列技术等测定该基因的转录水平,或者使用SDS-PAGE等电泳法、Western印迹法等测定该基因所编码的蛋白质的表达量。

[0067] 本发明的基因表达控制DNA不限于由序列号1所示的碱基序列构成的DNA,如上述(b)所记载,只要具有对成熟叶特异性促进基因表达的活性,也可以是序列号1所示的碱基序列中缺失、取代、添加或插入了1个~多个碱基的碱基序列。

[0068] 例如,即使是序列号1所示的碱基序列中缺失、取代、添加或插入了1~100个的碱基、优选1~50个的碱基、更优选1~10个的碱基的碱基序列,只要显示对成熟叶特异性促进基因表达的活性,就包括在本发明的基因表达控制DNA中。

[0069] 另外,本发明的基因表达控制DNA不限于由序列号1所示的碱基序列构成的DNA,如上述(c)所记载,只要显示对成熟叶特异性促进基因表达的活性,也可以是与序列号1所示的碱基序列具有80%以上、更优选90%以上、进一步优选95%以上、最优选99%以上的序列同源性的碱基序列。碱基序列的比较可以通过公知的方法进行,例如,可以按例如默认设置使用BLAST(Basic Local Alignment Search Tool at the National Center for Biological Information(美国国立生物学信息中心的基本局部比对搜索工具))等实施。

[0070] 另外,本发明的基因表达控制DNA不限于由序列号1所示的碱基序列构成的DNA,如上述(d)所记载,只要显示对成熟叶特异性促进基因表达的活性,也可以是在严格的条件下与由跟序列号1所示的碱基序列互补的序列构成的DNA杂交的碱基序列。

[0071] 本发明中“严格的条件”是指形成所谓的特异性杂交体、不形成非特异性杂交体的

条件,例如指在含有 $2\sim 6\times\text{SSC}$ ($1\times\text{SSC}$ 的组成: 0.15MNaCl , 0.015M 柠檬酸钠, $\text{pH}7.0$)和 $0.1\sim 0.5\%$ SDS的溶液中进行杂交并在含有 $0.1\sim 0.2\times\text{SSC}$ 和 $0.1\sim 0.5\%$ SDS的溶液中进行清洗的条件。

[0072] 另外,本发明的基因表达控制DNA,只要显示对成熟叶特异性促进基因表达的活性,也可以是在序列号1所示的碱基序列中从5'末端和/或3'末端缺失了连续的100个、200个、300个、400个、500个、600个、700个、800个、900个、1000个、1100个、1200个、1300个、1400个、1500个或者其以上的碱基的DNA片段。碱基的缺失可以通过本领域技术人员所公知的一般方法(例如,PCR法、限制酶处理等)进行。该DNA片段可以是在本发明的基因表达控制DNA中的启动子区域。规定的基因表达控制DNA中的启动子区域可以使用本领域技术人员所公知的启动子分析工具(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)进行检索。作为这样的序列号1所示的碱基序列的片段,可举出由序列号1的第869~1119位的碱基构成的DNA。得到的该片段可以通过上述报告基因试验等来确认其基因表达诱导功能。

[0073] 这样,如果本发明的基因表达控制DNA的碱基序列被确定,则其后通过化学合成,或者通过以基因组DNA为模板的PCR,或者通过将具有该碱基序列的DNA片段作为探针进行杂交,能够得到本发明的基因表达控制DNA。并且,可以通过定点突变等合成具有突变的序列号1所示的碱基序列。应予说明,在向序列号1所示的碱基序列导入突变中,可以采用Kunkel法、Gapped duplex法等公知的方法或者基于它们的方法。例如使用利用了定点突变的突变导入试剂盒(例如Mutant-K (TAKARA公司制)、Mutant-G (TAKARA公司制))等或者使用TAKARA公司的LA PCR in vitro Mutagenesis系列试剂盒进行突变的导入。

[0074] 接下来,对成花诱导基因进行说明。

[0075] 本说明书中“成花诱导基因”是指对具有促进植物的成花或者出穗的活性的蛋白质进行编码、参与植物中的成花或者出穗的基因。

[0076] 成花诱导基因是公知的,在拟南芥、稻子中作为FT基因、OsHd3a基因等已经被确定(日本特开2000-139250号公报、日本特表2002-511270号公报、日本特开2002-153283号公报、Kardailsky I.et al.,上载,Kojima S.et al.,上载),本发明中可以利用这些。成花诱导基因可以通过公开的数据库(GenBank等)进行检索,例如,来源于*Oryza sativa Japonica Group*的Hd3a基因的碱基序列(图4,序列号6)和氨基酸序列(图4,序列号7)作为AB052944.1登录在GenBank上。

[0077] 成花诱导基因可以通过分子生物学的领域中公知的克隆方法取得。例如,可以通过使用基于公知的基因序列(例如,作为AB052944.1登录在GenBank上)设计的探针或者引物,筛选来源于植物的基因组文库或者cDNA文库而得到。来源于序列信息未知的植物的基因可以通过使用利用序列已知的植物的基因信息而设计的探针或者引物,筛选来源于该植物的基因组文库或者cDNA文库而得到。如果序列被分离,则使用聚合酶链反应(PCR)这样的标准扩增法来扩增DNA,能够得到适合于转化(基因导入)的量的基因(DNA)。

[0078] 优选成花诱导基因包含编码选自以下的(e)~(g)中的任一种多肽的DNA:

[0079] (e)由序列号7所示的氨基酸序列构成的多肽,

[0080] (f) 由在序列号7所示的氨基酸序列中缺失、取代、添加或插入了1个~多个氨基酸的氨基酸序列构成,并且具有促进成花或出穗的活性的多肽,

[0081] (g) 由与序列号7所示的氨基酸序列具有90%以上的序列同源性的氨基酸序列构成,并且具有促进成花或出穗的活性的多肽。

[0082] 序列号7所示的氨基酸序列表示来源于稻子的Hd3a蛋白质的氨基酸序列。

[0083] 本发明中成花诱导基因所编码的多肽不限于由序列号7所示的氨基酸序列构成的多肽,如上述(f)所记载,也可以是由在序列号7所示的氨基酸序列中缺失、取代、添加或插入了1个~多个氨基酸的氨基酸序列构成并且具有促进成花或出穗的活性的多肽。

[0084] 例如,即使是序列号7所示的氨基酸序列中缺失、取代、添加或插入了1~20个的碱基、优选1~10个、更优选1~5个的氨基酸的氨基酸序列,只要具有促进成花的活性,就能够用于本发明。

[0085] 另外,本发明中成花诱导基因所编码的多肽不限于由序列号7所示的氨基酸序列构成的多肽,如上述(g)所记载,只要具有促进成花的活性,也可以是与序列号7所示的氨基酸序列具有80%以上、更优选90%以上、进一步优选95%以上、最优选99%以上的序列同源性的氨基酸序列。碱基序列的比较,如上所述,例如可以按例如默认设置使用BLAST等实施。

[0086] 作为这样的成花诱导基因,可举出由序列号2所示的碱基序列构成的DNA。由该序列号2所示的碱基序列构成的DNA编码来源于稻子的Hd3a蛋白质。

[0087] 另外,作为本发明的成花诱导基因,不限于由序列号2所示的碱基序列构成的基因,只要编码具有促进成花或出穗的活性的蛋白质,就能够利用在序列号2所示的碱基序列中缺失、取代、添加或插入了1个~多个碱基的碱基序列。

[0088] 例如,即使是序列号2所示的碱基序列中缺失、取代、添加或插入了1~50个的碱基、更优选1~10个的碱基的碱基序列,只要编码具有促进成花的活性的蛋白质,就可以用于本发明。

[0089] 并且,作为本发明的成花诱导基因,不限于由序列号2所示的碱基序列构成的基因,只要编码具有促进成花或出穗的活性的蛋白质,也可以是与序列号2所示的碱基序列具有80%以上、更优选90%以上、进一步优选95%以上、最优选99%以上的序列同源性的碱基序列。碱基序列的比较,如上所述,例如可以按例如默认设置使用BLAST等实施。

[0090] 并且,作为本发明的成花诱导基因,不限于由序列号2所示的碱基序列构成的基因,只要编码具有促进成花或出穗的活性的蛋白质,也可以是在严格的条件下与由跟序列号2所示的碱基序列互补的序列构成的DNA杂交的碱基序列。

[0091] “严格的条件”如上所述。

[0092] 接着,对使用了上述的基因表达控制DNA和成花诱导基因的重组载体进行说明。

[0093] 本发明的重组载体,可以通过将以能发挥功能的形式对上述的基因表达控制DNA连接成花诱导基因而成的DNA导入适当的载体来构建。本说明书中,“以能发挥功能的形式连接”、“以能发挥功能的形式被连接”是指在导入了该载体的宿主细胞内,在该基因表达控制DNA的控制之下成花诱导基因被正确地表达的方式,含有彼此连接的基因表达控制DNA和成花诱导基因。这里“连接”可以是直接连接,也可以是隔开适当的长度和序列的间隔而间接连接。优选以利用存在于序列号1所示的碱基序列的3'末端的“ATG”作为成花诱导基因

中编码第一蛋氨酸的“ATG”的形式进行连接。作为这样的DNA,可举出以利用存在于序列号1所示的碱基序列的3'末端的“ATG”作为存在于序列号2的5'末端的“ATG”的形式被连接的、由序列号3所示的碱基序列构成的DNA。

[0094] 作为本发明中使用的载体,优选使用能够借助农杆菌属向植物导入功能性基因的pBI系、pBII系、pPZP系(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系(http://www.cambia.org/main/r__et__camvec.htm)、pSMA系的载体等。特别优选使用pBII系和pBI系的双元载体或者中间载体系,例如,可举出pBII221、pBI121、pBI101、pBI101.2、pBI101.3、pIG121等。双元载体是在大肠杆菌(*Escherichia coli*)和农杆菌属中可复制的穿梭载体,如果使植物感染保持双元载体的农杆菌属,则能够将由位于载体上的LB序列和RB序列构成的边界序列所围起的部分的DNA整合到植物核DNA中(*EMBO Journal*, 10(3), 697-704(1991))。另一方面,pUC系的载体能够向植物直接导入基因,例如,可举出pUC18、pUC19、pUC9等。另外,也可以使用花椰菜花叶病毒(CaMV)、菜豆花叶病毒(BGMV)、烟草花叶病毒(TMV)等植物病毒载体。

[0095] 为了使连接和/或向载体的插入变得容易,上述的基因表达控制DNA和/或成花诱导基因中可以适当地取代、插入或者添加限制酶识别序列。向载体插入时,可以使用如下方法:首先,将上述的基因表达控制DNA和包含成花诱导基因的纯化的DNA用适当的限制酶切断,插入到适当的载体DNA的限制酶识别部位或者多克隆位点与载体连接的方法等。

[0096] 还可以根据需要在载体中在基因表达控制DNA和/或成花诱导基因的上游、内部、或下游连接增强子、内含子、poly(A)加尾信号、5'-UTR序列、选择标记基因等。

[0097] 作为增强子,例如,用于提高成花诱导基因的表达效率,例如,可举出包含CaMV35S启动子内的上游侧的序列的增强子区域等。

[0098] 作为终止子,只要是能够将利用上述启动子转录的基因转录终结的序列即可,例如,可举出胭脂碱合成酶基因的终止子、章鱼碱合成酶基因的终止子、CaMV 35S RNA基因的终止子等。

[0099] 作为选择标记基因,例如,可举出潮霉素抗性基因、卡那霉素抗性基因、双丙氨磷抗性基因、灭瘟素S抗性基因、乙酰乳酸合成酶(Acetolactate synthase)基因等。选择标记基因,可以如上述那样与成花诱导基因一起与相同的质粒连接来制备重组载体,或者,分别制备将选择标记基因与质粒连接而得到的重组载体和将成花诱导基因与质粒连接而得到的重组载体。分别制备时,将各载体共转染(共导入)到宿主。

[0100] 可以使用这样制成的重组载体制备转化体。

[0101] 制备转化植物时,可以适当地利用已经报告并确立的各种的方法,作为其优选的例子,可举出农杆菌属法、PEG-磷酸钙法、电穿孔法、脂质体法、基因枪法、显微注射法等。使用农杆菌属法的情况有使用原生质体的情况、使用组织片的情况、以及使用植物体其本身的情况(*in planta*法)。使用原生质体时,可以通过与具有Ti质粒的农杆菌属共培养的方法,与原生质化的农杆菌属融合的方法(原生质法)进行,使用组织片时,可以通过使对象植物的无菌培养叶片(叶盘)感染的方法、使愈合组织感染等进行。另外,使用利用种子或植物体的*in planta*法时,即在不介由添加植物激素的组织培养的体系中,可以通过对吸水种子、幼植物(幼苗)、盆栽植物等以农杆菌属的直接处理等来实施。

[0102] 包含上述基因表达控制DNA和成花诱导基因的DNA是否被整合到植物体中的确认可以通过PCR法、Southern杂交法、Northern杂交法、Western印迹法等进行。例如,从转化植物制备DNA,设计DNA特异性引物而进行PCR。进行PCR后,对扩增产物进行琼脂糖凝胶电泳、聚丙烯酰胺凝胶电泳或者毛细管电泳等,利用溴化乙锭、SYBR Green液等染色,然后作为1根的条带的形式检测出扩增产物,从而能够确认已转化。另外,使用预先以荧光色素等标记的引物进行PCR,也能够检测扩增产物。此外,也可以是使扩增产物与微孔板等固相结合,通过荧光或者酶反应等确认扩增产物的方法。

[0103] 作为本发明中转化所使用的植物,例如,可举出属于禾本科、茄科、十字花科、豆科、蔷薇科、菊科、百合科、伞形科、石竹科、葫芦科、旋花科、藜科等的植物,不特别限定于这些植物。优选禾本科的植物,例如,可举出甘蔗、稻子、大麦、小麦、玉米、结缕草、高粱、小米、稗草、象草以及柳枝稷等植物。

[0104] 本发明中,作为成为转化的对象的植物材料,例如,可举出根、茎、叶、种子、胚、胚珠、子房、茎顶(植物的芽的前端的生长点)、花药、花粉等植物组织或其切片、未分化的愈合组织、将其进行酶处置而去除细胞壁的原生质体等植物培养细胞。另外使用in planta法时,可以使用吸水种子、植物体整体。

[0105] 另外,本发明中转化植物是指植物体整体、植物器官(例如根、茎、叶、花瓣、种子、种子、果实等)植物组织(例如表皮、韧皮部、薄壁组织、木质部、维管束等)、植物培养细胞中的任一种。

[0106] 将植物培养细胞作为对象时,为了从所得到的转化细胞再生出转化植物,可以通过已知的组织培养法使器官或者个体再生。这样的操作,是利用作为从植物细胞到植物体的再生方法而通常为人所知的方法,作为本领域技术人员能够容易进行的。从植物细胞到植物体的再生,例如,可以如下进行。

[0107] 首先,使用植物组织或者原生质体作为成为转化的对象的植物材料时,将它们在加入无机要素、维生素、碳源、作为能量源的糖类、植物生长调节物质(生长素、细胞分裂素等植物激素)等经灭菌的愈合组织形成用培养基中培养,形成不规则增殖的脱分化的愈合组织(以下称为“愈合组织诱导”)。将这样形成的愈合组织转移到包含生长素等植物生长调节物质的新的培养基进行进一步增殖(继代培养)。

[0108] 愈合组织诱导如果用琼脂等固体培养基进行,继代培养例如以液体培养进行,则能够有效且大量进行各自的培养。接下来,将通过上述的继代培养而增殖的愈合组织在适当的条件下培养,从而诱导器官的再分化(以下,称为“再分化诱导”),最终再生出完整的植物体。再分化诱导可以通过适当设定培养基中的生长素、细胞分裂素等植物生长调节物质、碳源等各种成分的种类、量、光、温度等而进行。通过上述再分化诱导,形成不定胚、不定根、不定芽、不定茎叶等,进而发育成完整的植物体。或者,以成为完整的植物体前的状态(例如胶囊化的人工种子、干燥胚、冻干细胞以及组织等)进行储藏等。

[0109] 本发明中“转化植物”中,除了实施了转化的作为再分化当代的“T1代”之外,还包括作为从该植物的种子得到的后代的“T2代”,将通过药剂选择或Southern法等进行分析而判断为基因重组的“T2代”植物的花进行自体受粉而得到的下一代(T3代)等后代植物。

[0110] 这样产生的转化植物以成熟叶特异性的方式表达导入的成花诱导基因。

[0111] 本发明的转化植物中,促进成花或者出穗,与野生型相比在早期就发生花芽、花器

官的分化·形成,和/或,发生出穗和/或开花。即,本发明的转化植物与野生型相比,从播种到出穗和/或开花的时间短。

[0112] 另外,本发明的转化植物中,促进分蘖,与野生型相比,在早期产生分蘖和/或发生分蘖数的增大。

[0113] 本发明的包含具有对成熟叶特异性促进基因表达的活性的基因表达控制DNA和成花诱导基因的DNA,具有促进植物特别是甘蔗和甘蔗近缘属种的出穗的功能和/或促进分蘖的功能。本发明的该DNA,即便是不出穗的品种也能够诱导出穗,能够提早出穗时期,能够提早分蘖时期和/或能够增大分蘖数。根据这些特征,导入了本发明的该DNA的转化植物,能够增加可出穗的分蘖数,即有效分蘖数,在不受系统所致的出穗时期和出穗能力的影响下,有效地进行系统间交配。另外,导入本发明的该DNA的转化植物的代与代的时间,与野生型相比能够缩短为1/2~1/5,优选1/3,能够提高交配育种的效率。

[0114] 实施例

[0115] 以下,使用实施例对本发明进一步详细进行说明。然而,本发明的技术的范围不限于这些实施例。

[0116] (实施例1)成熟叶特异性基因的克隆

[0117] 使用RNeasy Plant Mini Kit (QIAGEN) 从甘蔗 (*Saccharum spp.cv.NiF8*) 的成熟叶、幼叶以及茎的各组织中提取纯化总RNA,根据常规方法构建cDNA文库并用于基因表达分析。基因表达分析利用Sugarcane Genome Array (Affimetrix),根据制造商的说明书进行。

[0118] 基因表达分析的结果是在*Saccharum spp.cv.NiF8*的成熟叶中发现表达量特别高的基因,将该基因作为“ecc0002”。应予说明,来源于*Saccharum spp.cv.NiF8*的ecc0002基因的碱基序列与图1所示的来源于*Saccharum officinarum*的ecc0002EST的碱基序列一致。从*Saccharum spp.cv.NiF8*的成熟叶、幼叶、茎髓、茎皮、根以及分生组织中提取纯化总RNA,根据常规方法制备cDNA,通过使用ABI7500实时PCR装置 (Applied Biosystems) 的SYBRGreen法对各组织中的ecc0002基因的表达量进行分析。

[0119] 将结果示于图2 (将表达最强的成熟叶的表达量表示为100%)。根据该结果,可知*Saccharum spp.cv.NiF8*中ecc0002基因在成熟叶中被强烈诱导基因表达。

[0120] (实施例2)成熟叶特异性启动子的取得

[0121] 使用DNeasy Plant Mini Kit (QIAGEN) 从甘蔗 (*Saccharum spp.cv.NiF8*) 的成熟叶组织 (0.5g) 中提取纯化约300ng基因组DNA。以实施例1中得到的ecc0002基因的碱基序列为基础,使用RightWallk (注册商标) Kit (BEX),由上述基因组DNA得到存在于ecc0002基因的5'上游的基因表达控制区域。在取得的基因表达控制区域的5'末端导入HindIII限制酶识别序列 (AAGCTT) 作为接头序列,在存在于3'末端的ecc0002基因的翻译起始位点 (ATG) 的3'侧导入BlnI限制酶识别序列 (CCTAGG) 作为接头序列,制备编码ecc0002基因的表达控制区域的DNA (图3) (序列号5)。

[0122] 对上述DNA的序列,利用公知的启动子分析工具 (BioInformatics and Molecular Analysis Section (http://www-bimas.cit.nih.gov/mol_bio/proscan/)) 进行分析,结果在序列号5中的第875~1125位的碱基区域被推断为作为启动子发挥功能的区域。

[0123] (实施例3) β -葡萄糖醛酸酶基因表达载体的构建

[0124] 构建基因表达载体,该基因表达载体相连地包含实施例2中得到的编码基因表达

控制区域的DNA与编码 β -葡萄糖醛酸酶(GUS)基因的UidAcDNA。对于该基因表达载体,使用植物转化用载体(pBII221)。在编码基因表达控制区域的DNA与UidAcDNA连接时,使存在于UidAcDNA的5'末端侧的编码第一蛋氨酸的ATG序列与存在于编码基因表达控制区域的DNA的3'末端侧的ecc0002基因的编码第一蛋氨酸的ATG序列一致地连接(翻译融合型)。将该基因表达载体的示意图示于图5。

[0125] (实施例4)GUS表达基因重组植物的制作

[0126] 使用农杆菌属法将实施例3中制备的基因表达载体导入宿主植物体(Saccharum spp. cv. NiF8)中,制作利用ecc0002基因的表达控制DNA来控制GUS基因的表达的基因重组甘蔗。

[0127] 从该基因重组甘蔗的成熟叶、幼叶、茎髓、茎皮、根以及分生组织提取纯化总RNA,根据常规方法制备cDNA,通过使用了ABI7500实时PCR装置(Applied Biosystems)的SYBRGreen法对各组织中的GUS基因的表达量进行分析。

[0128] 作为比较,与上述同样地对非基因重组甘蔗的分生组织(对照),以及来源于利用花椰菜花叶病毒(CaMV)35S启动子来控制GUS基因的表达的基因重组甘蔗的成熟叶和幼叶中的GUS基因的表达量进行分析。

[0129] 将结果示于图6。各组织的表达量表示分析的组织中最高表达,以将利用ecc0002基因的表达控制DNA来控制GUS基因的表达的基因重组甘蔗的成熟叶中的GUS基因的表达量设为100%时的相对值表示。

[0130] 根据该结果,可知ecc0002基因的表达控制DNA在成熟叶中与利用CaMV35S启动子控制的GUS基因的表达量相比为约17倍,能够特异性诱导基因表达。

[0131] (实施例5)稻子Hd3a基因表达载体的构建

[0132] 构建基因表达载体,该基因表达载体包含彼此连接的实施例2中得到的编码基因表达控制区域的DNA与编码稻子Hd3a的cDNA(记载为“Hd3acDNA”)。该基因表达载体使用植物转化用载体(pIG121-Hm)。编码基因表达控制区域的DNA与Hd3acDNA连接时,使存在于Hd3acDNA的5'末端侧的编码第一蛋氨酸的ATG序列与存在于编码基因表达控制区域的DNA的3'末端侧的ecc0002基因的编码第一蛋氨酸的ATG序列一致地连接(翻译融合型)。将连接的编码基因表达控制区域的DNA和Hd3acDNA的序列(序列号3)示于图7。另外,将该基因表达载体的示意图示于图8。

[0133] (实施例6)稻子Hd3a表达基因重组植物的制作

[0134] 使用农杆菌属法将实施例5中制作的基因表达载体导入宿主植物体(Saccharum spp. cv. NiF8)中,制作利用ecc0002基因的表达控制DNA来控制稻子Hd3a基因的表达的基因重组甘蔗。

[0135] 作为比较,使用野生型以及制作并使用利用CaMV 35S启动子来控制稻子Hd3a基因的表达的基因重组甘蔗。

[0136] 从各基因重组甘蔗的愈合组织、再分化组织以及成熟叶中提取纯化总RNA,根据常规方法制备cDNA,通过使用了ABI7500实时PCR装置(Applied Biosystems)的SYBRGreen法对各组织中的稻子Hd3a基因的表达量进行分析。应予说明,“再分化组织”是指对愈合组织诱导再分化而得到的组织,不表示完全再生的植物的组织或者器官。同样地分析作为内标的肌动蛋白基因的表达量。

[0137] 将结果示于图9。

[0138] 利用ecc0002基因的表达控制DNA来控制稻子Hd3a基因的表达的基因重组甘蔗的各系统中,被导入的稻子Hd3a基因在再分化后的成熟叶组织中特异性表达。

[0139] 应予说明,对于利用CaMV 35S启动子来控制稻子Hd3a基因的表达的基因重组甘蔗,在用于使重组愈合组织再分化的再分化培养基置床后,2周左右愈合组织带有褐变色,其后全部的重组愈合组织枯死。这表明与稻子、拟南芥等其它植物不同,Hd3a基因在甘蔗的再分化阶段给予致命的影响。

[0140] (实施例7)稻子Hd3a表达基因重组植物中的出穗和分蘖的诱导能力

[0141] 对上述实施例6中制备的稻子Hd3a表达基因重组体的出穗和分蘖的诱导能力进行研究。

[0142] 将结果示于图10和11。

[0143] 在稻子Hd3a表达基因重组体中,确认了播种腋芽苗后从约20天左右的早期生长阶段分蘖(图10),而且在播种腋芽苗后约2个月后确认出穗(图11)。

[0144] 另一方面,在野生株中,播种腋芽苗后到确认分蘖的时间比稻子Hd3a表达基因重组体长(图10),另外播种腋芽苗后即使经过3个月也看不到出穗(图11)。

[0145] 根据这些结果,确认了利用ecc0002基因的表达控制DNA来控制稻子Hd3a基因的表达的基因重组甘蔗,具有分蘖和出穗的高诱导能力,能够从分蘖茎进一步出穗,这样能够在数月的长时间维持出穗茎。

[0146] 产业上的可利用性

[0147] 根据本发明,对植物特别是甘蔗及甘蔗近缘属种能够促进出穗和/或促进分蘖。由此,能够在不被系统的出穗时期和出穗的能力影响的情况下,有效地进行系统间交配,另外,由于能够缩短植物的代与代的时间,所以能够显著提高交配育种的效率。本发明可期待对植物特别是甘蔗及甘蔗近缘属种的交配育种作出巨大贡献。

[0148] 将本说明书中引用的全部出版物、专利以及专利申请直接作为参考引入本说明书。

序列表

<110> 丰田自动车株式会社

<120> 甘蔗成花控制技术

<130> PH-5254-PCT

<150> JP 2012-140231

<151> 2012-06-21

<160> 7

<170> PatentIn version 3.5

<210> 1

<211> 1536

<212> DNA

<213> Saccharum spp. cv. NiF8

<400> 1

taggagatgc ggtgtgttac taaatgcaag gtccaaattc aatgcitttt ccatgtttct	60
ttgaaacgca atgccacatc ttcttttaa gtaagaactg aggggtecca tgtttctttt	120
tgcacttttc acaagaatgi acaactgaaa atatcatgaa acatcattac cctctttata	180
tgcgtctgca tetattcacc taaactcaet gataggattg atgcaettca gtaactcat	240
acgtgacaac tactgttttt gaaagtgaac atttgtagtg ctactatttg catgtatggg	300
aaattgggaa ttctttcttg ccatggctga tccagatctc gacctgttg atetaatgca	360
aacatgeatg ttgatagcaa gcigaggatc tagagatata aggtgttagg agatgcggtg	420
tggtaactaa tgcgaagtca aaattccatg ctitttccat gctcaattac ctagcaittc	480
[0001] ctaattttta atttgtata ctaatgcate gagecatata taattcagta aataatgata	540
tttaagcata tatatatata tacagttaac atttctaatt cttctttttt tgtgtggagg	600
tccgcgacga tgcgaagtgc tccaaccca aattaalcca cctctcttaa atccgcagat	660
cttaccacc agcagetaca catcgtattg tgtcggcttg accgcattgg cgcgtgggt	720
tttggcagcg cctgaatgca gtacagccac ctgtatggtg cccttggtag agtaacaccc	780
ttatccctac ggcagccatg tataaccctt atccctacgg cagccatgta ttgtagccca	840
tcttcttaac cacaagtca ttttaaattt ccggccggtc tcttgaggaa atcaaatttt	900
attttcacia tttatatgga talaggataa tetatgttcc taacagtggc taacaggctc	960
ctctctctcc acatacatcg cgigcaagca ttctctcaaa ctcttccgat cccccgaatc	1020
cagccittgac tgcnaacaga cgcccctctc cacatcctgc acacceatca gccaacggaa	1080
taacagaaga aggcgagtga gcagtgacaa agcagctcaa cagcagcgag ccaagccaaa	1140
atggagccag gagcaagccg cggccgcagc tctcccggtc ccccttgcgg ttaccgctag	1200
caaacgcccc tctccacalc ctgcancaca aggaggcaag tgcgcagtga caaaglacgi	1260
ccacagcagc gagccaagcc aaaaggagct cagccacagc cgcagctctc ggetaccgtt	1320
accgccgata acatgatgc ctttccaaac gccaaaggcc gcgcaatccc gfgcacaccg	1380
accacaact cgccaactcc ccatccctat ttgaagccac cggcccgcgc actgcattga	1440
tcaactcgca gcagttagagc agcacgagca acacgcgcgc ccgctccaac catctcagct	1500
tcgcgttccc cggccccegc cgcccgcgcc gccatg	1536

<210> 2

<211>	540	
<212>	DNA	
<213>	Oryza sativa Japonica Group	
<400>	2	
	atggccggaa gttgcaggga cagggaacct ctgttggttg gtagggttgt ggggatgtg	60
	ctggacgggt tcgtccggag caccaacctc aaggcacct atggtccaa gaccgtgtcc	120
	aatgcttgcg agctcaagcc gtccatggtc acccaccagc ctagggtcga ggtcggcggc	180
	aatgacatga ggacattcta cacccttgtg atggtagacc cagatgcacc aagcccaagt	240
	gaaccetaacc ttagggagta tetacattgg ttggtcactg atattcctgg tactactgca	300
	gcgtcatttg ggcaagaggt gatgtctac gagagcccaa ggccaacct ggggatccac	360
	cggtgtgtgt tcgtctgttt ccagcagctg gggcgtcaga cagtgtacc gcccggttg	420
	cgtcagaact tcaacaccaa ggacttcgcc gagctctaca acctcggctc gccggtcggc	480
	gccgtctact tcaactgcc a gcgcaggca ggctccggcg gcaggagggt ctacccttag	540
<210>	3	
<211>	2079	
<212>	DNA	
<213>	人工的	
<220>		
<223>	合成的	
<400>	3	
	aagettttagg agatgcggtg iggtactaaa tgenaggtec anattcaatg ctttttccat	60
	gtttctttga aacgcaatgc cacatctttc tttaaagtaa gaactgaggg gtcccatgtt	120
[0002]	tcittttgca cttttcaca gaatgtacaa ctgaaaatat catgaaacat cattaccctc	180
	tttatatgcg tcgtcatcta ttcacctaaa ctcaactgata ggattgatgc aettcagtac	240
	actcatacgt gacaactact gtttttga aa gtgaacattt gtagtctac tattttcatg	300
	tatgggaaat tgggaattct ttcttgccat ggctgatcca gatctcgacc tgcctgatct	360
	aatgcaaaca tgcatgttga tagcaagctg aggaictaga gatataaggt gttaggagat	420
	gcggtgtggt actaaatgca aggtcaaaat tccatgcttt ttccatgctc aattacctag	480
	catttccetaa tttttaattg igataactaa tgcategagc catatataat tcagtaata	540
	tgtatattta agcatatata tatatataca gtttacattt ctaattcttc tttttttgtg	600
	lggaggiccg egacgatgca agttgctccc aaccctaaat aatccacctc tttaaatec	660
	gcagatcttc accaccagca gctacacatc gtattgtgtc ggcttgaccg catgtcgcg	720
	ctgggttttg geagegcttg aatgcagtac agccacctgt atggtgceet tggtagagta	780
	acacccttat ccttacggca gccatgtata acccttatec ctacggcagc catgtattgt	840
	ageccatctt cttaaccaca agttcatttt aaatttcgg ccggtctctt gaggaaatca	900
	aattttattt tcacaatita tatggatata ggataateta tgttcctaac agtggetaac	960
	aggetcectc tcttccacat acatcgcgtg caagcattcc tccaaactct tccgatcccc	1020
	cgaatccagc cttagactgca aacagacgcc cctctccaca tcttgcacac ccatcagcca	1080
	acggaataac agaagaaggc gagtgagcag tgacaaagca cgtcaacagc agcgagccaa	1140
	gccaaaatgg ageccaggagc aagcccgggc cgcagctctc ccggtcctcc ttgctgttac	1200
	cgctagcaaa cgcctctctc cacatctctc aacacaagga ggcaagtgcg cagtgacaaa	1260

gtacgtccac agcagcgagc caagccaaaa ggagctcagc cacagccgca gctctcggct	1320
accgttaccg ccgatcacat gcatgccttt ccaaacgcca agggccgcgc aatcctcgtc	1380
acaaccgacca cacactcgcg aactccccat ccctatttga agccaccgge ccgcgcactg	1440
cattgatcaa ctccgagcag tagagcagca cgagcaaac gcccgcgccg tccaaccatc	1500
tcagcttcgc gcttccccgc ccccgcgcc gcgccgcca tggccggaag tggcagggae	1560
agggaccctc ttggtgttgg tagggttgtg ggtgatgtgc tggacgcgtt cgtccggagc	1620
accaaacctca aggtcaccta tggctccaag accgtgtcca atggctgcca gctcaagccg	1680
tccatggta cccaaccgcc tagggctcag gtcggcgcca atgacatgag gacattctac	1740
acccttgtga tggtagaccg agatgcacca agcccaagtg accctaacct tagggagtat	1800
ctacattggt tggtaactga tattcctggt actactgcag cgtcatttgg gcaagaggtg	1860
atgtgctacg agagcccaag gccaacatg gggatccacc gcttggtgtt cgtgctgttc	1920
cagcagctgg ggcgtcagac agtgtacgcg cccgggtgce gtcagaactt caacaccaag	1980
gacttcgcgc agctcfacaa cctcggctcg cggctgcgcg cegtctactt caactgccag	2040
cgcgaggcag gctccggcgg caggagggtc tacccttag	2079
<210> 4	
<211> 3389	
<212> DNA	
<213> Saccharum officinarum	
<400> 4	
atcaaacctg cagcagtaga gcagcagcag caaacgcgcg cgcctctcca accatctcag	60
ettcgcgctt cccgcgcccc gccgccgcgc ccgccaatgc gtcagagcgg caccactcca	120
tcgacgcgca gctccgtgcc ctggcccccg gcaaggcttc cgaggagctc atccagtaag	180
acgcctctgt cgcgcaccgt ttctctgaca tcttccagga cctccatgce cctagccttc	240
gcgaatttgt ccaggagtcg tacgagggtg cggccgatta cgagggcaag aaggacacgt	300
cgaaactggg cgagctgggc accaagctca cggggctgce gccccgcgc gccatcctgg	360
tggcgagctc catcctgcac atgctcaacc tggccaacct ggccgaggaa gtggagctgg	420
cgcaccgcgc ccggaacagc aagctcaagc acggggactt ctccagcagc ggtcccgcca	480
ccaccgagtc ggacatcgag gagaogctca agcgcctcgt gtcgctgggc aagacccccg	540
aggaggtggt cgaggcgctc aagaaccaga gcgtcgacct cgtcttcace gcgcaéccca	600
égcagtcgc caggaggctg ctctcgcaga aaaacgccag gatccggaat tctctgagc	660
agetgagtc caaggacgtc acggacgacg acaagaagga gctcagcagc gctctgcaga	720
gagagatcca ageagcttcc agaactgatg agatccggag agcacaaccé accccacagg	780
atgaaatgcg ctatgggatg agctacatcc atgaaactgt atggaagggt gtccctaagt	840
ttttgcgcgc tgtggataca gccctgaaga atatcggcat caatcagcgc ctccctaca	900
atgttctctt cattaagttc tggcttggc tgggtgtgta ccgtgatgga aatccaagag	960
taactccgga ggtgacaaga gaagtatgct tgctgtccag aatgacggct gcaaaactgt	1020
acatcgatea ggtcgaagac ctgatgtttg agctctctat gtggcctgc aatgatgaac	1080
ttcgtgctcg agccgaagaa gtccagagta ctccagcttc aaagaaagtt accaagtatt	1140
acatagaatt ctggaagcaa attctctcaa acgagcccta ccgggtgata ctgggtgctg	1200

[0003]

faaggacaa gttatacaac acacgegagc gtgeacgcca tctgetgca actggatitt	1260
ctgaaatttc tgtggaactcg gtatttacca atatcgaaga gttccttgag ccccttgagc	1320
tatgctacaa atccctgtgt gactgcggcg acaaggceat cgeggaacggg agcctcctgg	1380
acctoctcgc ccagggtgtc acgttcgggc tctccctggg gaagttggac atccgtcagg	1440
agtcggagcg gcacaccgac gtagtcgacg ccatcaaccac gtaccttggc atcgggtcgt	1500
accgctcgtg gcccaggac aagcggatgg agtggctggg gtcggagctg aaaggcaagc	1560
ggccgctgct gccccggac ctcccatga ccgaggagat cggcagcgc atcggggcga	1620
tgcacgtctc cgcggagctc ccgtcggaca gcttcggccc ctacatcacc tccatgtgca	1680
cagcccccct cgacgtgctc gccgtggagc tcttcagcgc cgagtgtgga attcgcaga	1740
cgctgcccgt ggtgcgcctg ctgagagcc tggccgacct gcaggcggcg cccgctcgc	1800
tggagcgctt cttctccact gacttgtact tcgaccacat caaggccaag cagcaggtga	1860
tggctgggta ctccgactcc ggcaaggacg ccggccgctt gtcgcggcg tggcagctgt	1920
acgtggcgca ggaggagatg gccaaagggt ccaagaaata cggcgtgaag ctgaccttgt	1980
tccacggcgc cggcggcacc gtgggcaggg gtggcggccc gacgcaccig gccatcctgt	2040
cccagcgcgc ggacaccatc aacgggtcaa tccgcgtgac ggtgcagggc gaggctcag	2100
agttcatgtt cggggaggat cacctgtgct tccagtctct gcagccttc acggccgcca	2160
cgctggagca cggcatgcac ccgccggtgt ctcccaagcc cgagtggcgc aagctcatgg	2220
aggagatggc agtcgtggcc acggaggagt accgctcctt cgtcgtcaag gagccgagat	2280
[0004] tctcagagta cttcagatcg gctacccttg agactgagta cgggaagatg aacatcggca	2340
gccggccagc caagaggaag ccggcggcg gcacaccac cctgcgcgc atcccctgga	2400
tcttctcgtg gacccagacg aggttcacc tcccctgtg gctgggagtc ggcgcgcct	2460
tcaagtgggc cctcagacaag gacatcaaga acttccagaa gctcaagag atgtacaacg	2520
agtgccattt cttcagggtc accctggacc tgetggagat ggttttcgcc aaggagatc	2580
ctggcattgc cggcttgtat gacttctgc ttgtcgcga cgtctcaag cctttggga	2640
agcagctcag ggacaaatc gtggagacag agaagcttct cctacagatc gctgggcaca	2700
aggatattct tgaaggcgat ccttacctga agcaggggct gggctacgc aatccctaca	2760
tcaccacctt gaacgtgttg caggctaca cgtgaagcg gataaggat ccgagcttca	2820
aggtagcgc gcagccgccc ctgtccaagg agttcgcga cgagaacaag cccgccggac	2880
tggtagagct gaaccggcg agcgagtacc cggccggctt ggaagacacg ctcatcctca	2940
ccaigaaaag tctgcgcgc gccatgcaga acaccgcta ggcgccttc cttactcac	3000
ctgcagagta ctgcacggca ataataatca gcttccgat ggtgtcgttt tgcagtttt	3060
ggatgaaaat gctgaaaact gacaccttct gtttccacta tgtttatgtt tatgtaattt	3120
cctcggcttt ggccctttta tatttictt ctgtttgga agtccaagtg gaaaaatctt	3180
ggcatcttaa acatattgta ataataatca tcatacaatc tacaaattta ctattttgta	3240
ttaatctatc tggeaggga aatgtcactt tatatccag cccattggat ggactttttt	3300
accatgatgc tagtcaacc atcctctttt gattgtgcta aacaaattct gaaatcaaat	3360
gcciggcaat atatgttaec ggttgaatc	3389

<210>	5	
<211>	1548	
<212>	DNA	
<213>	人工的	
<220>		
<223>	合成的	
<400>	5	
	aagcttttagg agafgcggtg tggactactaaa tgaaggtcc aaattcaatg ctttttccat	60
	gtttctttga aacgcaatgc cacatctttc tttaaagtaa gaactgaggg gteccatgtt	120
	tcttttttga cttttcacia gaatgtacaa ctgaaaatat catgaaacat cattaccctc	180
	tttataatgc tgcgcateta ttcacctaaa ctactgata ggattgatgc acffcagtac	240
	acteatacgt gacaactact gtttttgaaa gtgaacattt gtagtgctac tatttgcatg	300
	tatgggaaat tgggaattct ttcttgccat ggcgatcca gatctegacc tgettgatct	360
	aatgcaaaac tgcattgta tagcaagctg aggatctaga gatataaggt gttaggagat	420
	gcggtgtggt actaaatgca aggtcaaaa tccatgcttt ttccatgctc aattaccctag	480
	catttccctaa ttttlaattg tgataaactaa tgcactgagc catatataat tcagtaaaata	540
	tgtatattta agcatatata tataatata gtttacattt ctaattcttc tttttttgtg	600
	tggaggtccg cgacgatgca agttgctccc aaccctaaatt aatccacctc tcttaaatec	660
	gcagatcttc accaccagca gctacacatc gtatttggtc ggcttgaccg catgtgcgcg	720
	ctgggttttg cgacgcctg aatgcagtae agccacctgt atggtgacct tggtagagta	780
	acacccttat ccttacggca gcatgtata acccttatcc ctacggcagc catgtatgt	840
[0005]	agccatcttt ctttaaceaca agttcatttt aaatttcogg ccggtctctt gaggaaaaca	900
	aattttatit tcaeaattta taigatata ggataafcta tgitctaac agtggtaac	960
	aggetccctc tctccacat acatcgctg caagcattcc tccaaactct tccgatcccc	1020
	cgaatccagc cttgactgca aacagacgcc cctctceaca tccatcacac ccatcagcca	1080
	acggaataac agaagaagge gactgagcag tgacaaaagca cgtcaacagc agcgagccaa	1140
	gccaaaatgg agccaggagc aagccgcggc cgcagctctc ccggtccccc ttgcggttac	1200
	cgctagcaaa cgcccctctc cacatctctc aacacaagga ggcaagtgcg cagtgcacaa	1260
	gtactctcac agcagcagc caagccaaaa ggagctcagc cacagccgca gctctcggt	1320
	accgttaccg ccgatcacat gcatgcttt ccaaacgcca agggccgcgc aatcccgtgc	1380
	acaaccgacca caeactcgcc aactcccctt cctatttga agccaccggc ccgcgactg	1440
	cattgatcaa ctgcagcag tagagcagca cgagcaaac gccgcgcgc tccaaccatc	1500
	tcagcttgcg gcttcccgcg ccccgccgcc gcgcccgcca tgcctagg	1548
<210>	6	
<211>	847	
<212>	DNA	
<213>	Oryza sativa Japonica Group	
<400>	6	
	tgaaccacac acagttcagc tagcagatca cctagctaga tagctgcctc tatecacgta	60
	tatttgcctc ctgcaacttg ctgctgctgc aatagctagc agctgcagct agtaagcaaa	120
	actataaacc ttcaggggttt ttgcaagat cgaaggccgg aagtgccagg gacagggacc	180
	ctcttgtggt tggtagggtt gtgggtgatg tgctggagcc gttctgctgg agcaccnacc	240

tcaaggtcac ctatggctcc aagaccgtgt ccaatggctg cgagctcaag cegtecatgg 300
 tcaceccacca gctatagggtc gaggtcggcg gcaatgacat gaggacattc tacacccttg 360
 tgatggtaga cccagatgea ccaagcccaa gtgacctaa ccttagggag tatctacatt 420
 ggttggtcac tgatattect ggtactactg cagcgtcatt tgggcaagag gtgatgtgct 480
 acgagagccc aagccaacc atggggatcc accggctggt gttcgtgctg ttccagcagc 540
 tggggcgtca gacagtgtac gcgcccgggt ggcgtcagaa cttaacacc aaggacttcg 600
 ccgagcteta caacctcggc tcgccgtcgc ccgccgteta cttaactgc cagcgcgagg 660
 cagctccgg cgccaggagg gtctaccct agctaacgat gatcccgatc gatctgtgc 720
 atgctacta tcatcctca gcatgetata cattgcaggt tcagacaatt gaaatgattc 780
 tcgacacaca acatatatat gatggtgtaa ttaattatgc aattaantag ctgagcaagg 840
 etaaggt 847

<210> 7
 <211> 179
 <212> PRT
 <213> Oryza sativa Japonica Group

<400> 7

Met Ala Gly Ser Gly Arg Asp Arg Asp Pro Leu Val Val Gly Arg Val
 1 5 10 15

Val Gly Asp Val Leu Asp Ala Phe Val Arg Ser Thr Asn Leu Lys Val
 20 25 30

Thr Tyr Gly Ser Lys Thr Val Ser Asn Gly Cys Glu Leu Lys Pro Ser
 35 40 45

Met Val Thr His Gln Pro Arg Val Glu Val Gly Gly Asn Asp Met Arg
 50 55 60

Thr Phe Tyr Thr Leu Val Met Val Asp Pro Asp Ala Pro Ser Pro Ser
 65 70 75 80

Asp Pro Asn Leu Arg Glu Tyr Leu His Trp Leu Val Thr Asp Ile Pro
 85 90 95

Gly Thr Thr Ala Ala Ser Phe Gly Gln Glu Val Met Cys Tyr Glu Ser
 100 105 110

Pro Arg Pro Thr Met Gly Ile His Arg Leu Val Phe Val Leu Phe Gln
 115 120 125

Gln Leu Gly Arg Gln Thr Val Tyr Ala Pro Gly Trp Arg Gln Asn Phe
 130 135 140

Asn Thr Lys Asp Phe Ala Glu Leu Tyr Asn Leu Gly Ser Pro Val Ala
 145 150 155 160

Ala Val Tyr Phe Asn Cys Gln Arg Glu Ala Gly Ser Gly Gly Arg Arg
 165 170 175

[0006]

[0007]

Val Tyr Pro

LOCUS	BASE COUNT	740 a	1007 c	968 g	574 t	
ORIGIN						
1	ATCAA	ACTCG	CAGCAG	TAGA	GCAGC	ACGAG
61	CTTCG	GGGTT	CCGGC	GGGG	GGGG	GGGG
121	TGACG	GGGCA	GCTCC	GTGGC	GGGG	GGGG
181	ACGCC	TGCT	CCCGA	CCCGT	TTCCT	GGACA
241	GCGAA	TTTGT	CCAGG	AGTGC	TACGAG	TGT
301	CGAAG	CTGGG	CGAGC	TGGGC	ACCAA	AGCTCA
361	TGGCG	AGCTC	CATCC	TGCAC	ATGCT	CAACC
421	CGCAC	GGCG	CCGGA	ACAGC	AAGCT	CAAGC
481	CCACG	AGTC	GGACA	TGGAG	GAGAC	GCTCA
541	AGGAG	GTGT	CGAGG	CGCTC	AAGA	ACCAGA
601	CGCAG	CTCCG	CAGGAG	TGCTC	GTCCG	AGACA
661	AGCTG	AGTGC	CAAGG	ACGAG	ACAAG	AAGGA
721	GAGAG	ATCCA	AGCAG	CTTC	AGA	ACTGATG
781	ATGAA	ATGG	CTA	TGGATG	AGCTA	CAATC
841	TTTTG	GGCG	IGTGG	ATACA	GGCCT	GAA
901	ATGTT	CTCT	CATTAA	AGTTC	TGGT	CTTGA
961	TAATC	CGGA	GGTGA	CAAGA	GAAGT	ATGCT
1021	ACATG	GATCA	GGTCG	AAGAC	CTGAT	GTGT
1081	TTCTG	TGCTG	AGCCG	AAGAA	GTCC	AGAGT
1141	ACATG	AATT	CTGGA	AGCAA	ATTCC	TCCAA
1201	TAAGG	GAACA	GTTAT	ACAAC	ACAGC	GGAGC
1261	CTGAA	TTTC	TGTTG	ACTCG	GTATT	TACCA
1321	TATGT	TACAA	ATCCG	TGTGT	GACTG	GGCG
1381	ACCTC	TGGG	CCAGG	TGTT	ACGTT	GGGG
1441	AGTCG	GAGCG	GCAC	ACCGAC	GTGAT	CGACG
1501	ACCGC	TGGT	GGCCG	AGGAC	AAGCG	ATGG
1561	GGCCG	TGCT	GGCCG	CGGAC	CTTCC	ATGA
1621	TGCAC	GTCT	CGCGG	AGCTC	CGTCC	GACA
1681	CAGCC	CCCT	CGAGT	GTCT	GGC	TGGAG
1741	CGCTG	CCCT	GGTCC	GGTG	CTCG	AGAGG
1801	TGGAG	CGGCT	CTTCC	ACT	GACT	TGACT
1861	IGTGG	GTA	CTCCG	ACTC	GGCA	AGGACG
1921	ACGTG	GGCA	GGAGG	AGATG	GCCA	AGGTG
1981	TCCAC	GGGG	CGGGC	CAAC	GTGGG	CAAGG
2041	CCAGC	CGCC	GCAC	ACCATC	AACGG	TCAA
2101	AGTTC	ATGTT	CGGGG	AGGAT	CACC	TGTCT
2161	CGCTG	GAGCA	CGGC	ATGCAC	CGCCG	TGT
2221	AGGAG	ATGG	AGTCC	TGGCC	ACGG	AGGAT
2281	TCTCG	AGTA	CTTCA	GATCG	GCTA	CCCTG
2341	GGCGG	CCAG	CAAG	AGGAAG	CGGG	CGGG
2401	TCTTC	TGTT	GACCC	AGACG	AGT	TCC
2461	TCAAG	TGGG	CATCG	ACAAG	GACAT	CAAG
2521	AGTGG	CCATT	CTTCA	AGGTC	ACCC	TGACC
2581	CTGGC	ATTGC	CGGCT	TGTAT	GACT	TGCTG
2641	AGCAG	CTCAG	GGACA	AAATAC	GTGG	AGACG
2701	AGGAT	ATTCT	TGAAG	CGGAT	CCTT	ACCTG
2761	TCACC	ACCT	GAA	CGTGTG	CAGGC	CTACA
2821	AGGTG	AGCC	GCAGC	GGCG	CTG	CCAAG
2881	TGGTG	AAGCT	GAAC	CCGGC	AGCG	AGTACC
2941	CCATG	AAAG	TATCG	CGGC	GGC	ATGCAG
3001	CTGC	AGAGT	CTGC	ACGGCA	ATAA	TATCA
3061	GGATG	GAAAT	GCTGA	AAACT	GAC	ACCTCT
3121	CCTCG	GCTTT	GGCC	CTTTA	TATTT	CACT
3181	GGCAT	CTTAA	ACAT	ATTGTA	ATAA	TGAACA
3241	TTAAT	CTATC	TGGC	AGGGAA	AATG	CACTT
3301	ACCAT	GATGC	TAGT	CAACC	ATCC	CTTTT
3361	GCCTG	GCAAT	ATA	TGTTACC	GGT	GAAAT

图1

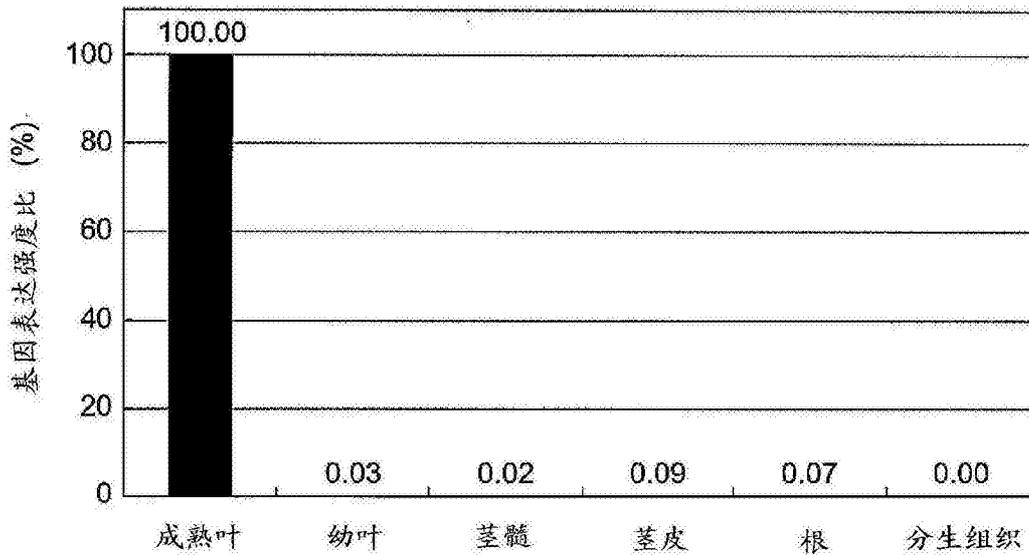


图2

HindIII

```

1 AAGCTTTAGG AGATCGGGTG TGGTACTAAA TGCAAGGTCC AAATTC AATG CTTTTCCAT
61 GTTTCTTTGA AACGCAATGC CACATCTTTC TTTAAAGTAA GAACTGAGGG GTCCCATGTT
121 TCTTTTTGCA CTTTTCACAA GAATGTACAA CTGAAAATAT CATGAAACAT CATTACCCCTC
181 TTTATATGCG TCGTCATCTA TTCACCTAAA CTCACTGATA GGATTGATGC ACTTCAGTAC
241 ACTCATACGT GACAACACT GTTTTTGAAA GTGAACATTT GTAGTGCTAC TATTTGCATG
301 TATGGGAAAT TGGGAATTCT TTCTTGCCAT GGCTGATCCA GATCTCGACC TGCTTGATCT
361 AATGCAAACA TGCATGTTGA TAGCAAGCTG AGGATCTAGA GATATAAGGT GTTAGGAGAT
421 GCGGTGTGGT ACTAAATGCA AGGTCAAAT TCCATGCTTT TTCCATGCTC AATTACCTAG
481 CATTTCTCTA TTTTAAATTG TGATAACTAA TGCATCGAGC CATATATAAT TCAGTAAATA
541 TGTATATTTA AGCATATATA TATATATACA GTTTACATTT CTAATTCCTC TTTTTTTGTG
601 TGGAGGTCCG CGACGATGCA AGTTGCTCCC AACCCAAATT AATCCACCTC TCTTAAATCC
661 GCAGATCTTC ACCACCAGCA GCTACACATC GTATTGTGTC GGCTTGACCG CATCTGCGCG
721 CTGGGTTTTG GCAGCGCCTG AATGCAGTAC AGCCACCTGT ATGGTGCCCT TGGTAGAGTA
781 ACACCCTTAT CCCTACGGCA GCCATGTATA ACCCTTATCC CTACGGCAGC CATGTATTGT
841 AGCCCATCTT CTTAACCACA AGTTCATTTT AAATTTCCGG CCGGTCTCTT GAGGAAATCA
901 AATTTTATTT TCACAATTTA TATGGATATA GGATAATCTA TGTTCTAAC AGTGGCTAAC
961 AGGCTCCCTC TCCTCCACAT ACATCGCGTG CAAGCATTCC TCCAAACTCT TCCGATCCCC
1021 CGAATCCAGC CTTGACTGCA AACAGACGCC CCTCTCCACA TCCTGCACAC CCATCAGCCA
1081 ACCGAATAAC AGAAGAAGGC GAGTGAGCAG TGACAAAGCA CGTCAACAGC AGCGAGCCAA
1141 GCCAAAATGG AGCCAGGAGC AAGCCGCGGC CGCAGCTCTC CCGGTCCCCC TTGGCGTTAC
1201 CGCTAGCAAA CGCCCTCTC CACATCCTGC AACACAAGGA GGCAAGTGGC CAGTGACAAA
1261 GTACGTCCAC AGCAGCGAGC CAAGCCAAAA GGAGCTCAGC CACAGCCGCA GCTCTGGGCT
1321 ACCGTTACCG CCGATCACAT GCATGCCTTT CCAAACGCCA AGGGCCGGCG AATCCCGTGC
1381 ACACCGACCA CACACTCGCC AACTCCCCAT CCCTATTTGA AGCCACGGGC CCGCGGCACTG
1441 CATTGATCAA CTCCGAGCAG TAGAGCAGCA CGAGCAACAC GCCCGCGCGC TCCAACCATC
1501 TCAGCTTCGC GCTTCCGGCG CCCC GCCGCC GCGCCGCCA TGCTTAGG
    
```

起始密码子 Bln I

图3

(A)

```

1  igcaccacac acagttcagc tagcagatca cctagctaga tagctgcctc taticacagta
61  tatttgctcc ctgcaacttg ctgctgctgc aatagctagc agctgcagct agtaagcaaa
121 actataaacc ttcagggttt ttgcaagat cgatggccgg aagtgccagg gacagggacc
181 ctcttggtgt tggtagggtt ggggtgatg tgctggacgc gttcgtccgg agcaccaacc
241 tcaaggtcac ctatggctcc aagaccgtg ccaatggctg cgagctcaag ccgtccaagg
301 tcaccacca gcctagggtc gaggtcggcg gcaatgacat gaggacattc tacacccttg
361 tgatggtaga cccagatgca ccaagcccaa gtgaccctaa cttagggag tatctacatt
421 ggttggtcac tgatattcct ggtactactg cagcgtcatt tgggcaagag gtgatgtgct
481 acgagagccc aaggccaacc atggggatcc accggctggt gttcgtgctg ttccagcagc
541 tggggcgtca gacagtgtac gcgcccgggt ggcgtcagaa cttcaacacc aaggacttgc
601 ccgagctcta caacctcggc tcgccggtcg ccgccgtcia cttcaactgc cagcgcgagg
661 caggctccgg cggcaggagg gtctaccctt agctaacgat gatcccgatc gatctgctgc
721 atgctacta tcatcatcca gcatgctata cattgcaggt tcagacaatt gaaatgattc
781 tcgacacaca acatatatat gatggtgtaa ttaattatgc aattaaatag ctgagcaagg
841 ctaaggt
    
```

(B)

```

1  Met Ala Gly Ser Gly Arg Asp Arg Asp Pro Leu Val Val Gly Arg Val
17  Val Gly Asp Val Leu Asp Ala Phe Val Arg Ser Thr Asn Leu Lys Val
33  Thr Tyr Gly Ser Lys Thr Val Ser Asn Gly Cys Glu Leu Lys Pro Ser
49  Met Val Thr His Gln Pro Arg Val Glu Val Gly Gly Asn Asp Met Arg
65  Thr Phe Tyr Thr Leu Val Met Val Asp Pro Asp Ala Pro Ser Pro Ser
81  Asp Pro Asn Leu Arg Glu Tyr Leu His Trp Leu Val Thr Asp Ile Pro
97  Gly Thr Thr Ala Ala Ser Phe Gly Gln Glu Val Met Cys Tyr Glu Ser
113 Pro Arg Pro Thr Met Gly Ile His Arg Leu Val Phe Val Leu Phe Gln
129 Gln Leu Gly Arg Gln Thr Val Tyr Ala Pro Gly Trp Arg Gln Asn Phe
145 Asn Thr Lys Asp Phe Ala Glu Leu Tyr Asn Leu Gly Ser Pro Val Ala
161 Ala Val Tyr Phe Asn Cys Gln Arg Glu Ala Gly Ser Gly Gly Arg Arg
177 Val Tyr Pro
    
```

图4

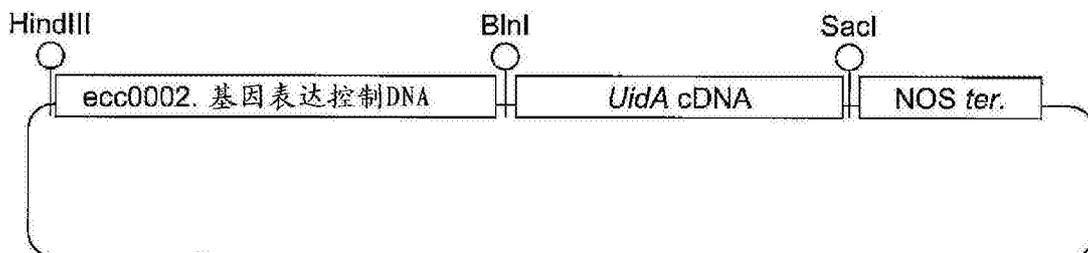


图5

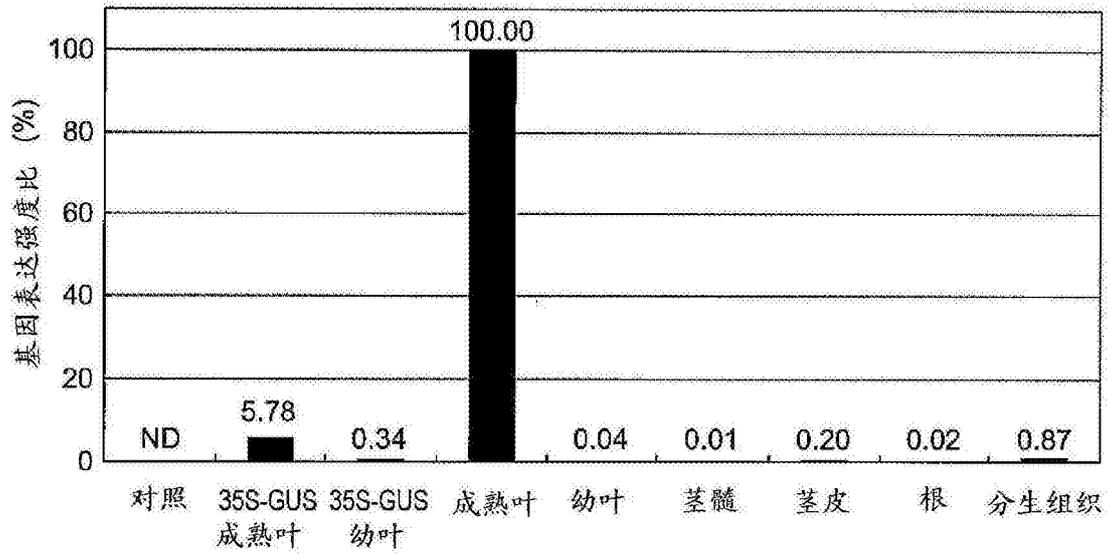


图6

```

1 AAGCTTTAGG AGATGCGGTG TGGTACTAAA TGCAAGGTCC AAATTCAATG CTTTTTCCAT
61 GTTTCTTTGA AACGCAATGC CACATCTTTC TTAAAGTAA GAAGTGAAGG GTCCCATGTT
121 TCTTTTGGCA CTTTTCACAA GAATGTAGAA CTGAAAATAT CATGAAACAT CATTACCGTC
181 TTTATATGCG TCGTCATCTA TTCACCTAAA CTCACTGATA GGATTGATGC ACTTCAGTAC
241 ACTCATACGT GACAAGTACT GTTTTTGAAA GTGAACATTT GTAGTGCTAC TATTTGCATG
301 TATGGGAAAT TGGGAATTCT TTCTTGCCAT GGCTGATCCA GATCTCGACC TGCTTGATCT
361 AATGCAAACA TGCATGTTGA TAGCAAGCTG AGGATCTAGA GATATAAGGT GTTAGGAGAT
421 GCGGTGTGGT ACTAAATGCA AGGTCAAAAT TCCATGCTTT TTCCATGCTC AATTACCTAG
481 CATTTCCTAA TTTTAAATTG TCATAACTAA TGCATCGAGC CATATATAAT TCAGTAAATA
541 TGTATATTTA AGCATATATA TATATATACA GTTTACATTT CTAATTCTTC TTTTTTGTG
601 TGGAGGTCCG CGACGATGCA AGTTGCTCCC AACCCAAAT AATCCACCTC TCTTAAATCC
661 GCAGATCTTC ACCACCAGCA GCTACACATC GTATTGTGTC GGCTTGACCG CATGTGCGCG
721 CTGGGTTTTG GCAGCGCTG AATGCAGTAC AGCCACCTGT ATGGTGCCCT TGGTAGAGTA
781 ACACCCTTAT CCCTACGGCA GCCATGTATA ACCCTTATCC CTACGGCAGC CATGTATTGT
841 AGCCCATCTT CTTAACCACA AGTTCATTTT AAATTTCCGG CCGGTCTCTT GAGGAAATCA
901 AATTTTATTT TCACAATTTA TATGGATATA GGATAATCTA TGTTCCCTAAC AGTGGCTAAC
961 AGGCTCCGTC TCCTCCACAT ACATCCCGTG CAAGCATTC CCAAACCTCT TCCGATCCCC
1021 CGAATCCAGC CTTGACTGCA AACAGACGCC CCTCTCCACA TCCTGCACAC CCATCAGCCA
1081 ACGGAATAAC AGAAGAAGGC GAGTGAGCAG TGACAAAGCA CGTCAACAGC AGCGAGCCAA
1141 GCCAAAATGG AGCCAGGAGC AAGCCCGGGC CGCAGCTCTC CCGGTCCGCC TTGCCGTTAC
1201 CGCTAGCAAA CGCCGCTCTC CACATCCTGC AACACAAGGA GGCAAGTGCG CAGTGACAAA
1261 GTACGTCCAC AGCAGCGAGC CAAGCCAAA GGAGCTCAGC CACAGCCGCA GCTCTCGGCT
1321 ACCGTTACCG CCGATCACAT GCATGCCTTT CCAAACGCCA AGGGCCGGCC AATCCCGTGC
1381 ACACCGACCA CACACTGGCC AACTCCCAT CCCTATTTGA AGCCACCGGC CCGGCCACTG
1441 CATTGATCAA CTCGCAGCAG TAGAGCAGCA CGAGCAACAC GCGCGCGCGC TCCAACCATC
1501 TCAGCTTCGC GCTTCCCGCG CCCC GCCCGCCGCa tggccggaag tggcagggac
1561 agggaccctc ttgtggttgg taggtttgtg ggtgatgtgc tggaccgctt cgtccggagc
1621 accaacctca aggtcaccta tggctccaag accggtgcca atggctgcca gctcaagccg
1681 tccatggtca ccaccagcc taggttcgag gtggcgggca atgacatgag gacattctac
1741 acccttgtga tggtagacc agatgcacca agcccaagt accctaacct tagggagtat
1801 ctacattggt tggtagactga tattcctggt actactgcag cgtcatttgg gcaagaggtg
1861 atgtgctacg agagcccaag gcccaacctg ggaatccacc ggctggtgtt cgtgctgttc
1921 cagcagctgg ggcgtcagac aggttacgca cccgggtggc gtcagaactt caacaccaag
1981 gacttcgccc agctctacaa cctcggtctg cgggtcggcg cgtctactt caactgccag
2041 cgcgaggcag gctccggcgg caggagggtc taccctag
    
```

图7

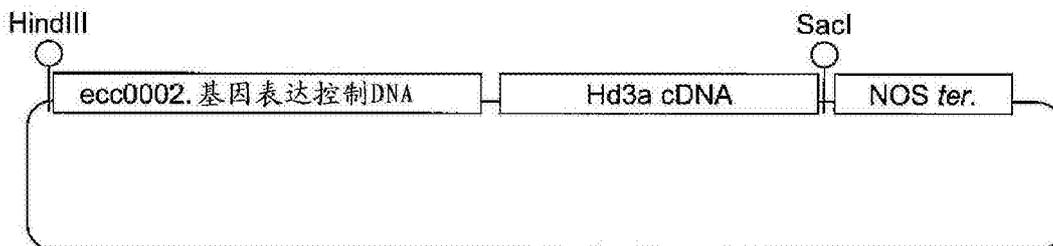


图8

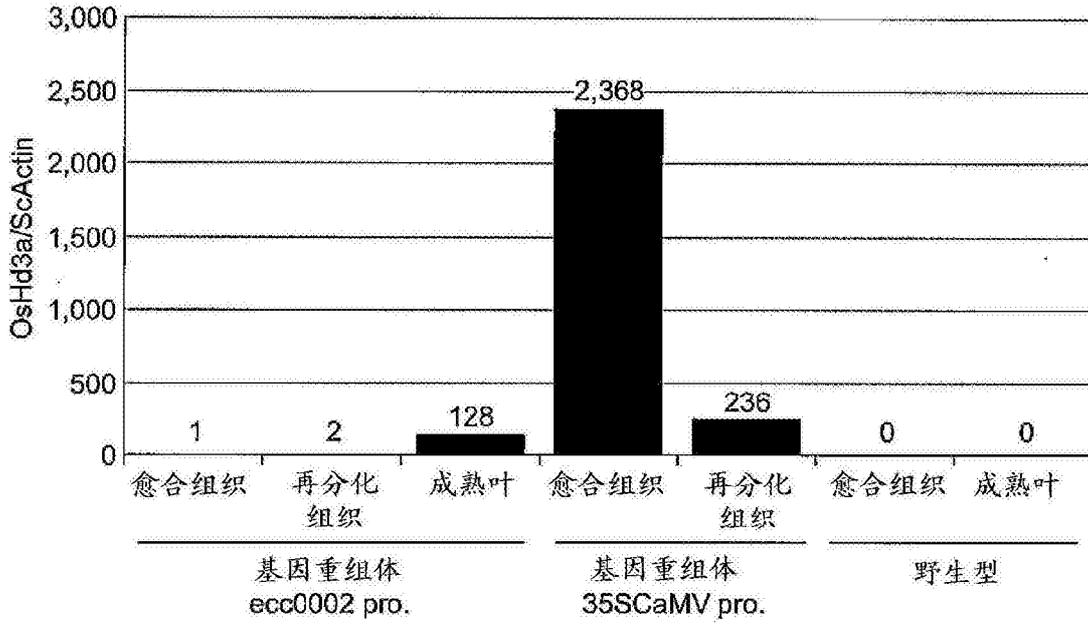


图9

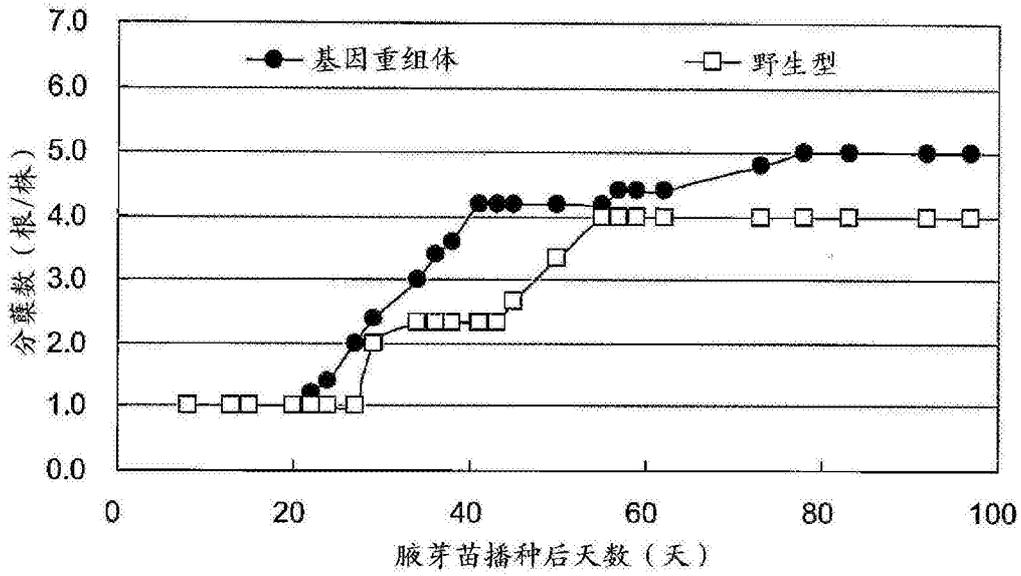


图10

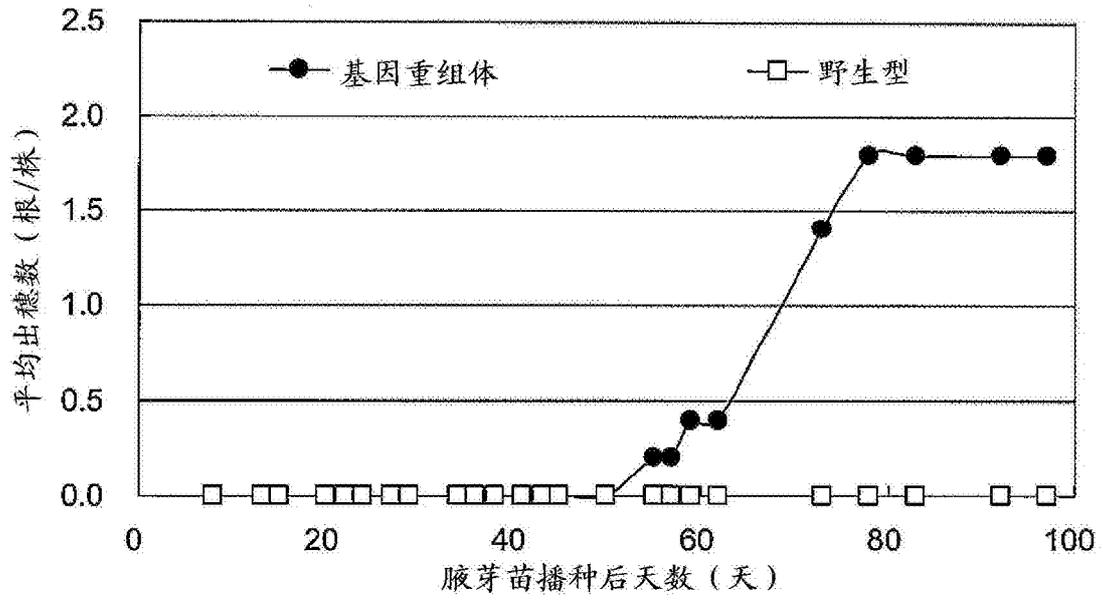


图11