



(12)发明专利申请

(10)申请公布号 CN 106414746 A

(43)申请公布日 2017.02.15

(21)申请号 201480058842.0

(22)申请日 2014.08.29

(30)优先权数据

61/872,134 2013.08.30 US

(85)PCT国际申请进入国家阶段日

2016.04.26

(86)PCT国际申请的申请数据

PCT/US2014/053328 2014.08.29

(87)PCT国际申请的公布数据

W02015/031706 EN 2015.03.05

(71)申请人 美国陶氏益农公司

地址 美国印第安纳州

(72)发明人 S·库马尔 J·贝林格

(74)专利代理机构 北京坤瑞律师事务所 11494

代理人 罗天乐

(51)Int.Cl.

C12N 15/82(2006.01)

权利要求书2页 说明书49页

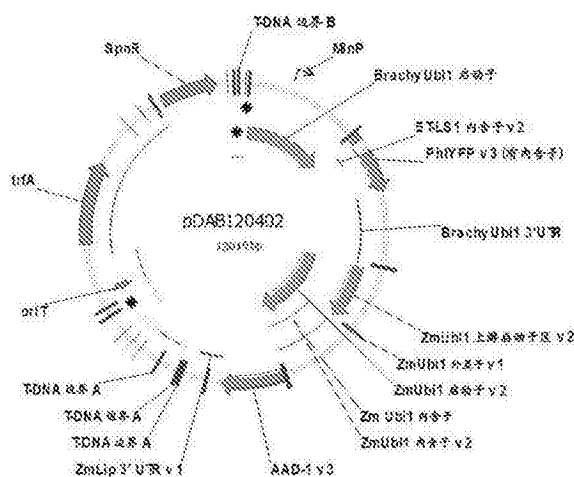
序列表28页 附图41页

(54)发明名称

利用来自短柄草属泛素基因的调节元件的转基因表达用构建体

(57)摘要

提供了利用调节元件在植物细胞和/或植物组织中表达转基因的构建体和方法,所述调节元件包括分离自二穗短柄草泛素基因的启动子和/或3'-UTR。在一个实施方案中,提供表达载体,其中泛素基因的调节元件与多接头序列可操作连接。依照一个实施方案,提供了植物、植物组织或植物细胞,其包含与非泛素转基因可操作连接的启动子。



1. 一种包含启动子的核酸载体,所述启动子可操作地连接于
 - i)多接头序列;
 - ii)非泛素转基因;或
 - iii)i)和ii)的组合,其中所述启动子包含SEQ ID NO:2或与SEQ ID NO:2具有90%序列同一性的序列。
2. 权利要求1的核酸载体,其中所述启动子长度小于3kb。
3. 权利要求1的核酸载体,其中所述启动子由SEQ ID NO:2或与SEQ ID NO:2具有90%序列同一性的序列组成。
4. 权利要求1-3中任一项的核酸载体,其进一步包含编码选择标志物的序列。
5. 权利要求4的核酸载体,其中所述启动子与转基因可操作连接。
6. 权利要求5的核酸载体,其中所述转基因编码选择标志物或赋予杀虫剂抗性、除草剂耐受性、氮利用效率、水分利用效率、或营养品质的基因产物。
7. 权利要求1-3或5中任一项的核酸载体,其进一步包含3'非翻译序列,所述3'非翻译序列包含SEQ ID NO:5或与SEQ ID NO:5具有90%序列同一性的序列,其中所述3'非翻译序列与所述多接头或所述转基因可操作连接。
8. 权利要求1-3或5中任一项的核酸载体,进一步包含5'非翻译序列,所述5'非翻译序列包含SEQ ID NO:12或与SEQ ID NO:12具有90%序列同一性的序列,其中所述5'非翻译序列插入在所述启动子序列与所述多接头或转基因之间,并且与它们可操作连接。
9. 权利要求8的核酸载体,进一步包含插入在所述5'非翻译序列之后的内含子序列。
10. 权利要求9的核酸载体,其中所述内含子序列包含SEQ ID NO:8。
11. 权利要求1的核酸载体,其中所述启动子由SEQ ID NO:40或与SEQ ID NO:40具有90%序列同一性的序列组成,并且所述启动子与转基因可操作连接。
12. 权利要求1的核酸载体,其中所述启动子由SEQ ID NO:16或与SEQ ID NO:16具有90%序列同一性的序列组成,并且所述启动子与转基因可操作连接。
13. 权利要求11或12的核酸载体,其进一步包含这样的3'非翻译序列,该3'非翻译序列包含SEQ ID NO:5或与SEQ ID NO:5具有90%序列同一性的序列,其中所述3'非翻译序列与所述转基因可操作连接。
14. 一种非短柄草属植物,其包含与转基因可操作连接的SEQ ID NO:2或与SEQ ID NO:2具有90%序列同一性的序列。
15. 权利要求14的植物,其中所述植物选自玉米、小麦、水稻、高粱、燕麦、黑麦、香蕉、甘蔗、大豆、棉花、向日葵、和芸苔。
16. 权利要求14的植物,其中所述植物是玉米。
17. 权利要求14-16中任一项的植物,其中所述转基因被插入所述植物的基因组中。
18. 权利要求14的植物,进一步包含5'非翻译序列,所述5'非翻译序列包含SEQ ID NO:12或与SEQ ID NO:12具有90%序列同一性的序列,其中所述5'非翻译序列插入在所述启动子与所述转基因之间,并且与它们可操作连接。
19. 权利要求18的植物,进一步包含插入在所述5'非翻译序列之后的内含子序列。
20. 权利要求19的植物,其中所述内含子序列包含SEQ ID NO:8。
21. 权利要求18的植物,进一步包含3'非翻译序列,所述3'非翻译序列包含SEQ ID NO:

5或与SEQ ID NO:5具有90%序列同一性的序列,其中所述3'非翻译序列与所述转基因可操作连接。

22.权利要求14的植物,其中所述启动子由SEQ ID NO:40或与SEQ ID NO:40具有90%序列同一性的序列组成,并且所述启动子与转基因可操作连接。

23.权利要求14的植物,其中所述启动子由SEQ ID NO:16或与SEQ ID NO:16具有90%序列同一性的序列组成,并且所述启动子与转基因可操作连接。

24.权利要求22或23的植物,其进一步包含3'非翻译序列,所述3'非翻译序列包含SEQ ID NO:5或与SEQ ID NO:5具有90%序列同一性的序列,其中所述3'非翻译序列与所述转基因可操作连接。

25.一种包含转录终止子的核酸载体,所述转录终止子可操作地连接于

- i)多接头序列;
- ii)非泛素转基因;或
- iii)i)和ii)的组合,其中所述转录终止子包含SEQ ID NO:5或与SEQ ID NO:5具有90%序列同一性的序列。

26.权利要求25的核酸载体,其中所述转录终止子长度小于1kb。

27.权利要求26的核酸载体,其中所述转录终止子由SEQ ID NO:5的3'UTR序列组成。

利用来自短柄草属泛素基因的调节元件的转基因表达用构建体

[0001] 相关申请的交叉引用

[0002] 本申请根据35U.S.C. §119(e)要求2013年8月30日提交的美国临时申请No. 61/872,134的优先权,通过提述将上述申请全部并入本申请。

[0003] 通过提述并入以电子方式提交的材料

[0004] 通过提述将其全部内容并入与此申请同时提交的计算机可读核苷酸/氨基酸序列,其如下:建立于2014年8月15日的名称为“Brach_UBI_SEQ_LIST_ST25”的一个63KB的ACII(文本)文件”。

[0005] 背景

[0006] 植物转化是一种有吸引力的用于将农艺学上期望性状或特性导入不同作物物种中的技术。开发具有特定的期望性状的植物物种,和/或将植物物种修饰为具有特定的期望性状。通常,期望性状的例子包括:提高营养价值质量,增加产量,赋予抗虫性或抗病性,增加干旱和胁迫耐受性,提高园艺质量(例如,着色和生长),赋予除草剂抗性,使得可能从植物生产在工业上有用的化合物和/或材料的能力,和/或使得可能生产药品。

[0007] 通过植物转化技术来生产包含堆叠在单个基因组基因座处的多个转基因的转基因植物。植物转化技术导致转基因导入植物细胞中、在植物基因组中含有转基因的稳定整合拷贝的能育转基因植物的恢复、并且通过转基因的转录和翻译的后续转基因表达导致具有期望性状和表型的转基因植物。堆叠中的每个转基因一般需要独立的启动子用于基因表达,因此在转基因堆叠中需使用多个启动子。

[0008] 经常由于需要共表达调节同一性状的多个转基因而重复使用同一启动子来驱动多个转基因的表达。然而,重复使用包含享有高水平序列同一性的序列的多个启动子可能导致基于同源性的基因沉默(HBGS)。观察到转基因植物(Peremarti et al., 2010)中当在转基因中使用重复的DNA序列时频繁出现HBGS。另外,在土壤杆菌中,已经证明由于质粒的重组和不稳定性,在转基因构建体中重复使用相似的DNA序列是困难的。

[0009] 本文描述了与玉米泛素1启动子享有低水平序列同一性或同源性的泛素调节元件(例如,启动子和3'-UTR)。进一步描述了利用泛素调节元件的构建体和方法。

[0010] 概述

[0011] 本文公开了用于在植物细胞和/或植物组织中表达转基因的构建体和方法。在一个实施方案中,从柳枝稷(*Panicum virgatum*)、二穗短柄草(*Brachypodium distachyon*)、或粟(*Setaria italica*)基因组中纯化泛素基因调节元件,并且将其与不 该调节元件天然连接的序列重组,从而创制用于在对于所述泛素调节序列而言非天然的植物细胞中表达转基因的表达载体。在一个实施方案中,提供了表达载体,其中泛素基因的调节元件与多接头序列可操作地连接。这样的表达载体使得基因或基因盒以与泛素基因调节序列的可操作连接状态插入到所述载体中变得容易。

[0012] 在一个实施方案中,提供了构建体,其包含柳枝稷、二穗短柄草、或粟泛素启动子。在一个实施方案中,提供了基因表达盒,其包含与转基因可操作连接的柳枝稷、二穗短柄草

或粟泛素启动子。在一个实施方案中,基因表达盒包含与转基因可操作连接的柳枝稷、二穗短柄草或粟泛素5'-UTR。在一个实施方案中,基因表达盒包含与启动子可操作连接的柳枝稷、二穗短柄草或粟泛素5'-UTR。在一个实施方案中,基因表达盒包含与转基因可操作连接的柳枝稷、二穗短柄草或粟泛素内含子。在一个实施方案中,基因表达盒包含与启动子可操作连接的柳枝稷、二穗短柄草或粟泛素内含子。在一个实施方案中,构建体包含含有柳枝稷、二穗短柄草、或粟泛素3'-UTR的基因表达盒。在一个实施方案中,基因表达盒包含与转基因可操作连接的柳枝稷、二穗短柄草或粟泛素3'-UTR。在一个实施方案中,基因表达盒包含至少一个、两个、三个、五个、六个、七个、八个、九个、十个、或更多个转基因。

[0013] 在一个实施方案中,基因表达盒独立地包含:a)柳枝稷、二穗短柄草、或粟泛素启动子,b)柳枝稷、二穗短柄草或粟泛素内含子,c)柳枝稷、二穗短柄草或粟泛素5'-UTR,以及d)柳枝稷、二穗短柄草或粟泛素3'-UTR。

[0014] 根据一个实施方案,提供了核酸载体,其包含与非泛素转基因可操作连接的启动子,其中所述启动子由SEQ ID NO:16或与SEQ ID NO:16具有90%序列同一性的序列组成。在另一个实施方案中,所述核酸载体包括基因盒,其中所述基因盒包含启动子、非泛素转基因和3'非翻译区,其中所述启动子由与转基因的第一端可操作连接的SEQ ID NO:16组成,其中所述转基因的第二端与由SEQ ID NO:5组成的3'非翻译序列可操作连接。

[0015] 本文公开了种植利用柳枝稷、二穗短柄草、或粟启动子、5'-UTR、内含子、和3'-UTR表达转基因的植物的方法。本文还公开了培养利用柳枝稷、二穗短柄草或粟启动子、5'-UTR、内含子、和3'-UTR表达转基因的植物组织和细胞的方法。

[0016] 根据一个实施方案,提供了包含与非泛素转基因可操作连接的启动子的植物、植物组织、或植物细胞,其中所述启动子包含SEQ ID NO:2。根据一个实施方案,提供了包含与转基因可操作连接的SEQ ID NO:2、或与SEQ ID NO:2具有90%序列同一性的序列的非短柄草属植物或植物细胞。在一个实施方案中,所述植物为玉米品种。在一个实施方案中,提供了包含与非泛素转基因可操作连接的启动子的植物、植物组织、或植物细胞,其中所述启动子由SEQ ID NO:16或SEQ ID NO:40组成。在一个实施方案中,提供了包含基因盒的非短柄草属(non-Brachypodium)植物或植物细胞,其中所述基因盒包含与转基因可操作连接的启动子,且其中所述启动子由SEQ ID NO:16组成。在另一个实施方案中,所述启动子与转基因的第一端可操作连接,其中所述转基因的第二端与由SEQ ID NO:5组成的3'非翻译序列可操作连接。

[0017] 附图简述

[0018] 图1显示了玉米泛素(ZM Ubi1)蛋白质序列与用于启动子鉴定的二穗短柄草和粟泛素序列的蛋白质比对。Zm Ubi1蛋白质序列在本文中公开为SEQ ID NO:22。粟Ubi2蛋白质序列在本文中公开为SEQ ID NO:23。二穗短柄草Ubi1启动子序列在本文中公开为SEQ ID NO:24。二穗短柄草Ubi1C蛋白质序列在本文中公开为SEQ ID NO:25。共有序列在本文中公开为SEQ ID NO:26。

[0019] 图2显示了玉米泛素(ZM Ubi1)启动子多核苷酸序列与本文中鉴定的二穗短柄草和粟泛素启动子多核苷酸的比对。玉米泛素1(Zm-Ubi)启动子序列在本文中公开为SEQ ID NO:27。二穗短柄草Ubi1启动子序列在本文中公开为SEQ ID NO:16。二穗短柄草Ubi1-C启动子序列在本文中公开为SEQ ID NO:15。粟Ubi2启动子序列在本文中公开为SEQ ID NO:17。

[0020] 图3为显示合成的粟泛素2启动子遗传元件的质粒图谱。

[0021] 图4为显示合成的二穗短柄草泛素1C启动子遗传元件和侧翼无缝克隆突出端位置的质粒图谱。

[0022] 图5为显示合成的二穗短柄草泛素1启动子遗传元件和侧翼无缝克隆突出端位置的质粒图谱。

[0023] 图6为显示含有融合于PhiYFP报告基因的粟泛素2(SI-Ubi2)启动子的表达载体的质粒图谱。

[0024] 图7为显示含有融合于PhiYFP报告基因的二穗短柄草泛素1C启动子的表达载体的质粒图谱。

[0025] 图8为显示含有融合于PhiYFP报告基因的二穗短柄草泛素1启动子的表达载体的质粒图谱。

[0026] 图9为显示含有融合于PhiYFP报告基因的OS Act1(水稻肌动蛋白1)启动子的表达载体的质粒图谱。

[0027] 图10为显示含有融合于PhiYFP报告基因的ZM Ubi1启动子的表达载体的质粒图谱。

[0028] 图11为显示利用Gateway技术用来建立二元表达载体的二元目标载体的质粒图谱。

[0029] 图12为显示二元表达载体的质粒图谱,该二元表达载体含有粟泛素2(SI-Ubi2)启动子,其融合于含有ST-LS1内含子的黄色荧光蛋白(Phi YFP)标志物基因编码区,之后为包含来自马铃薯的StPinII 3'UTR的片段。

[0030] 图13为显示二元表达载体的质粒图谱,该二元表达载体含有二穗短柄草泛素1C启动子,其融合于含有ST-LS1内含子的黄色荧光蛋白(Phi YFP)标志物基因编码区,之后为包含来自马铃薯的StPinII 3'UTR的片段。

[0031] 图14为显示二元表达载体的质粒图谱,该二元表达载体含有二穗短柄草泛素1启动子,其融合于含有ST-LS1内含子的黄色荧光蛋白(Phi YFP)标志物基因编码区,之后为包含来自马铃薯的StPinII 3'UTR的片段。

[0032] 图15为显示二元表达载体的质粒图谱,该二元表达载体含有OS Act1启动子,其融合于含有ST-LS1内含子的黄色荧光蛋白(Phi YFP)标志物基因编码区,之后为包含来自马铃薯的StPinII 3'UTR的片段。

[0033] 图16为显示二元表达载体的质粒图谱,该二元表达载体含有ZM Ubi1启动子,其融合于含有ST-LS1内含子的黄色荧光蛋白(Phi YFP)标志物基因编码区,之后为包含来自马铃薯的StPinII 3'UTR的片段。

[0034] 图17显示了T₀叶中的YFP表达,其中YFP由如图12、13、14、15、和16描绘的跨物种泛素启动子和Os Act 1启动子驱动。

[0035] 图18显示了T₀叶中的AAD1表达,其中AAD1由如图12、13、14、15、和16描绘的Zm Ubi 1启动子驱动。

[0036] 图19显示了由新颖的二穗短柄草和粟启动子驱动的瞬时YFP表达,与由ZM Ubi1和OS Act1启动子驱动的YFP表达比较。

[0037] 图20显示了由新颖的二穗短柄草和粟启动子驱动的愈伤组织中的YFP表达,与由

ZM Ubi1和OS Act1启动子驱动的YFP表达比较。

[0038] 图21显示了由新颖的二穗短柄草和粟启动子驱动的组织中的YFP表达,与由ZM Ubi1和OS Act1启动子驱动的YFP表达比较。

[0039] 图22为显示合成的柳枝稷泛素1启动子遗传元件和侧翼无缝克隆突出端位置的质粒图谱。

[0040] 图23为显示合成的柳枝稷泛素1 3'UTR遗传元件和侧翼无缝克隆突出端位置的质粒图谱。

[0041] 图24为显示合成的二穗短柄草泛素1C 3'UTR遗传元件和侧翼无缝克隆突出端位置的质粒图谱。

[0042] 图25为显示合成的二穗短柄草泛素1 3'UTR遗传元件和侧翼无缝克隆突出端位置的质粒图谱。

[0043] 图26为显示合成的粟泛素2(SI-Ubi2)3'UTR遗传元件和侧翼无缝克隆突出端位置的质粒图谱。

[0044] 图27为显示含有融合于PhiYFP报告基因的柳枝稷泛素1启动子和3'UTR的表达载体的质粒图谱。

[0045] 图28为显示含有融合于PhiYFP报告基因的二穗短柄草泛素1C启动子和3'UTR的表达载体的质粒图谱。

[0046] 图29为显示含有融合于PhiYFP报告基因的粟泛素2启动子和3'UTR的表达载体的质粒图谱。

[0047] 图30为显示含有融合于PhiYFP报告基因的二穗短柄草泛素1启动子和3'UTR的表达载体的质粒图谱。

[0048] 图31为显示二元表达载体的质粒图谱,该二元表达载体含有二穗短柄草泛素1C启动子,融合于含有ST-LS1内含子的黄色荧光蛋白(Phi YFP)标志物基因编码区,之后为包含二穗短柄草泛素1C 3'UTR的片段。

[0049] 图32为显示二元表达载体的质粒图谱,该二元表达载体含有柳枝稷泛素1启动子,融合于含有ST-LS1内含子的黄色荧光蛋白(Phi YFP)标志物基因编码区,之后为包含柳枝稷泛素1 3'UTR的片段。

[0050] 图33为显示二元表达载体的质粒图谱,该二元表达载体含有粟泛素2启动子,融合于含有ST-LS1内含子的黄色荧光蛋白(Phi YFP)标志物基因编码区,之后为包含粟泛素2 3'UTR的片段。

[0051] 图34为显示二元表达载体的质粒图谱,该二元表达载体含有二穗短柄草泛素1启动子,融合于含有ST-LS1内含子的黄色荧光蛋白(Phi YFP)标志物基因编码区,之后为包含二穗短柄草泛素1 3'UTR的片段。

[0052] 图35呈现了二穗短柄草泛素1C编码序列和推定的启动子(ATG的上游序列)。上游启动子序列加下划线,5'-UTR序列以大写字母表示,内含子加框,Ubi1CDS为斜体,3'-UTR(加有下划线)和转录终止序列在TAA(翻译终止密码子)的下游。

[0053] 图36呈现了二穗短柄草泛素1编码序列和推定的启动子。上游启动子加下划线,5'UTR序列以大写字母表示,内含子加框,CDS为斜体,3'-UTR(加有下划线)和转录终止序列在TAA(翻译终止密码子)的下游。

[0054] 图37呈现了粟泛素2编码序列和推定的启动子。上游启动子加下划线,5'UTR序列以大写字母表示,内含子加框,CDS为斜体,3'-UTR(加有下划线的)和转录终止序列在TAA(翻译终止密码子)的下游。

[0055] 图38呈现了柳枝稷(Switchgrass)泛素1编码序列和推定的启动子。上游启动子加下划线,5'UTR序列以大写字母表示,内含子加框,CDS为斜体,3'-UTR(加有下划线)和转录终止序列在TAA(翻译终止密码子)的下游。

[0056] 发明详述

[0057] 定义

[0058] 在描述和要求本发明时,下列术语将根据下文陈述的定义使用。

[0059] 如本文中使用的术语“大约”意指比言明的值或值的范围大或小10%,但不意图将任何值或值的范围仅仅指定这个更宽的定义。在每个值或值的范围之前的术语“大约”同样意图涵盖所陈述的绝对值或值的范围的实施方案。

[0060] 如本文中使用的,术语“回交”是指其中育种者使杂交后代与亲本之一杂交,例如,第一代杂种F1与F1杂种的亲本基因型之一杂交的过程。

[0061] “启动子”是能够结合细胞中的RNA聚合酶并启动编码序列的下游(3'方向)转录的DNA调节区。启动子可含有被转录因子识别的特异性序列。这些因子可与启动子DNA序列结合,导致RNA聚合酶的募集。为了定义本发明的目的,启动子序列在其3'端以转录起始位点为界,并向上游(5'方向)延伸,以包含为启动高于背景的可检测水平的转录所需的最小数目的碱基或元件。在启动子序列内可见转录起始位点(通过例如核酸酶S1作图来方便地界定)、以及负责结合RNA聚合酶的蛋白结合结构域(共有序列)。启动子可与其他表达控制序列(包括增强子和阻抑物序列)可操作地关联。

[0062] 为本公开的目的,“基因”包括编码基因产物(如上所述)的DNA区域,也包括所有调节基因产物之产生的DNA区域,无论这类调控序列是否位于编码和/或转录序列的附近。因此,基因包括,而不一定限于,启动子序列、终止子、翻译调节序列(如核糖体结合位点和内部核糖体进入位点)、增强子、沉默子、隔离子、边界元件、复制起点、核基质附着位点和基因座控制区。

[0063] 如本文中使用的术语“天然的”或“自然的”定义了自然界中发现的状况。“天然DNA序列”是通过自然手段或传统育种技术产生的而不是通过基因工程(例如,使用分子生物学/转化技术)产生的存在于自然界中的DNA序列。

[0064] 如本文中使用的“转基因”定义为编码基因产物(包括,例如但不限于mRNA)的核酸序列。在一个实施方案中,转基因为外源核酸,其中转基因序列已通过基因工程导入通常未发现所述转基因的宿主细胞(或其后代)中。在一个实例中,转基因编码在工业上或药理学上有用的化合物、或编码期望的农艺性状的基因(例如,除草剂抗性基因)。在又一个实例中,转基因为反义核酸序列,其中所述反义核酸序列的表达抑制靶核酸序列的表达。在一个实施方案中,转基因是内源核酸,其中所述内源核酸的额外基因组拷贝是期望的,或者是相对于宿主生物中的靶核酸序列处于反义取向的核酸。

[0065] 如本文中使用的术语“非泛素转基因”是与玉米泛素1编码序列(SEQ ID NO:27)具有小于80%序列同一性的任何转基因。

[0066] 如本文中定义的“基因表达”是包含在基因中的信息到基因产物的转化。

[0067] 如本文中定义的“基因产物”是由基因产生的任何产物。例如,基因产物可以是基因直接转录的产物(例如,mRNA、tRNA、rRNA、反义RNA、干扰RNA、核酶、结构RNA或任何其他类型的RNA)或mRNA翻译产生的蛋白质。基因产物还包括修饰过的RNA,例如通过加帽、多聚腺苷酸化、甲基化和编辑修饰过的RNA,以及修饰的蛋白质,例如通过甲基化、乙酰化、磷酸化、泛素化、ADP核糖基化、豆蔻酰化(myristylation)和糖基化修饰的蛋白质。基因表达可能受到外部信号的影响,例如细胞、组织、或生物被暴露于可增加或降低基因表达的物质。基因表达也可以在从DNA到RNA再到蛋白质的途径中的任何位置受到调节。基因表达的调节发生的方式有,例如,对转录、翻译、RNA运输和加工的控制;对于中间分子比如mRNA的降解;或在特定蛋白分子已经产生之后对它们的活化、失活化、区室化、或降解,或上述方式的组合。基因表达可以在RNA水平或蛋白质水平上通过本领域已知的任何方法来测量,这些方法包括但不限于,Northern印迹、RT-PCR、Western印迹,或体外、原位或体内蛋白活性测定。

[0068] 如本文中使用的,术语“内含子”定义为包含在基因(或表达的感兴趣核苷酸序列)中的任何被转录但不被翻译的核酸序列。内含子包括表达的DNA序列之内的非翻译核酸序列、以及从所述DNA序列转录的RNA分子中的相应序列。本文中描述的构建体还可以含有强化翻译和/或mRNA的稳定性的序列,如内含子。一种这样的内含子的实例是拟南芥组蛋白H3变体的基因II的第一内含子或任何其他公知的内含子序列。内含子可与启动子序列联合使用来强化翻译和/或mRNA的稳定性。

[0069] 如本文中使用的,术语“5'非翻译区”或“5'-UTR”定义为前mRNA或成熟mRNA的5'端中的非翻译区段。例如,在成熟mRNA上,5'-UTR一般在其5'端上包含7-甲基鸟苷帽,且涉及许多过程,如剪接、多腺苷酸化、mRNA朝向细胞质输出、翻译机构识别mRNA的5'端、以及保护mRNA免于降解。

[0070] 如本文中使用的,术语“转录终止子”定义为前mRNA或成熟mRNA的3'端中的转录区段。例如,超过“多腺苷酸化信号”位点的更长DNA段被转录为mRNA。这个DNA序列常常含有用于将前mRNA正确加工为成熟mRNA的一个或多个转录终止信号。

[0071] 如本文中使用的,术语“3'非翻译区”或“3'-UTR”定义为前mRNA或成熟mRNA的3'端中的非翻译区段。例如,在成熟的mRNA上,这个区域包含多聚(A)尾,并且已知其在mRNA稳定性、翻译起始、和mRNA输出中具有多种作用。

[0072] 如本文中使用的,术语“多腺苷酸化信号”指示存在于mRNA转录物中的一种核酸序列,其允许转录物在多聚(A)聚合酶的存在下在多腺苷酸化位点上(例如位于多聚(A)信号的下流的10到30个碱基的多腺苷酸化位点上)多腺苷酸化。许多多腺苷酸化信号在本领域是已知的,并可用于本发明。示例序列AAUAAA及其变体,正如描述于Loke J.等人的(2005) Plant Physiology 138(3);1457-1468。

[0073] 如本文中使用的术语“分离的”意指已经从其自然环境中移出,或者已经与该化合物最初形成时存在的其他化合物分开。术语“分离的”包括从天然来源分离的物质,以及在宿主细胞中重组表达之后回收的物质(例如,核酸和蛋白质),或化学合成的化合物,如核酸分子、蛋白质、和肽类。

[0074] 如本文中使用的术语“纯化的”涉及分子或化合物以这样的形式分离:基本上不含通常在天然或自然环境中与所述分子或化合物相关联的污染物,或相对于在所述化合物最初形成时存在的其他化合物而言在浓度上高度富集;该术语还意指作为与原始组成中其他

组分分离的结果,在纯度上已经增加。术语“纯化的核酸”在本文中用来描述已经从其他生物化合物(包括但不限于,多肽、脂质和碳水化合物)分离、分开产生、或纯化出来,同时影响组分的化学或功能变化(例如,通过移除蛋白质污染物并使连接核酸与染色体中的其余DNA的化学键断裂,可将核酸从染色体中纯化出来)的核酸序列。

[0075] 如本文中使用的,术语“基于同源性的基因沉默”或“HBGS”为通用术语,包括转录基因沉默和转录后基因沉默两者。非连锁的沉默基因座对靶基因座的沉默作用可以是转录抑制(转录基因沉默;TGS)或mRNA降解(转录后基因沉默;PTGS)的结果,二者分别是由对应于启动子的双链RNA(dsRNA)或对应于转录序列的双链RNA的产生所导致的。每一过程涉及独特的细胞组分,这表明dsRNA诱导的TGS和PTGS很可能是从某个古老的共同机制的多样化而来。然而,迄今为止TGS和PTGS的严格比较是难于实现的,因为这种比较通常依赖于对独特的沉默基因座的分析。由于产生对应于不同靶基因的启动子和转录序列的dsRNA,单转基因基因座可以说既可触发TGS又可触发PTGS。

[0076] 如本文中使用的,术语“核酸分子”、“核酸”、或“多核苷酸”(所有这三个术语彼此同义)是指核苷酸的聚合形式,其可包括RNA、cDNA、基因组DNA、和合成形式的有义和反义链、及其混合聚合物。“核苷酸”可以指核糖核苷酸、脱氧核糖核苷酸、或这两种类型核苷酸的任一者的修饰形式。除非另外指明,核酸分子通常长度为至少10个碱基。前述术语可以指具有不确定长度的RNA或DNA分子。所述术语包括DNA的单链和双链形式。核酸分子可以包括天然存在的核苷酸和/或修饰核苷酸,藉由天然存在和/或非天然存在的核苷酸联接而连接在一起。

[0077] 正如本领域技术人员易于理解的,核酸分子可被化学修饰或生物化学修饰,或者可以含有非天然或衍生的核苷碱基。这些修饰包括,例如,标记物、甲基化、用类似物取代一个或多个天然存在的核苷酸、核苷酸间修饰(例如,如不带电联接(uncharged linkage)的修饰:例如甲基膦酸盐、磷酸三酯、氨基磷酸酯、氨基甲酸酯等;带电连接(charged linkage)的修饰:例如,硫代磷酸酯、二硫代磷酸酯等;悬垂部分的修饰:例如肽类;嵌入剂修饰:例如吡啶、补骨脂素等;螯合剂修饰;烷化剂(alkylator)修饰;和修饰的连接:例如 α 异头核酸等)。术语“核酸分子”还包括任何拓扑结构,包括单链的、双链的、部分双链体的、三链体的、发夹形的(hairpinned)、圆形的、挂锁形的结构。

[0078] 转录以5'到3'的方式沿着DNA链行进。这意味着通过向生长中的链的3'端连续添加5'-核糖核苷三磷酸而产生RNA(消除焦磷酸盐是必不可少的)。在线性或环形核酸分子中,如果离散元件(例如,特定的核苷酸序列)结合到或会结合到同一核酸上另一个元件的5'方向,则所述离散元件可称为处于相对所述另一个元件的“上游”。类似地,如果离散元件结合到或将要结合到同一核酸上另一个元件的3'方向,则所述离散元件可处于相对于所述另一个元件的“下游”。

[0079] 如本文中使用的,术语“碱基位置”是指在指定核酸之内的给定碱基或核苷酸残基的位置。指定的核酸可通过与参考核酸的比对来定义。

[0080] 如本文中使用的,术语“杂交”是指其中寡核苷酸及其类似物通过互补碱基之间的氢键杂交的过程,所述氢键包括Watson-Crick、Hoogsteen或反Hoogsteen氢键。通常,核酸分子由含氮碱基组成,所述碱基为嘧啶(胞嘧啶(C)、尿嘧啶(U)和胸腺嘧啶(T))或嘌呤(腺嘌呤(A)和鸟嘌呤(G))。这些含氮碱基在嘧啶和嘌呤之间形成氢键,并且嘧啶与嘌呤的键合

被称为“碱基配对”。更具体地，A会与T或U键合，而G会与C键合。“互补”指在两条不同的核酸序列或同一核酸序列的两个不同区域之间发生的碱基配对。

[0081] 如本文中使用的，术语“可特异性杂交”和“特异性互补”是指足够的互补程度，使得在寡核苷酸与DNA或RNA靶标之间发生稳定的和特异性的结合。寡核苷酸不必与其特异性杂交的靶序列100%互补。当寡核苷酸与靶DNA或RNA分子的结合干扰DNA或RNA的正常功能，并且存在足够的互补程度以避免在期望特异性结合的条件（例如对于体内测定或系统的情形而言，生理条件）下的寡核苷酸与非靶序列的非特异性结合时，所述寡核苷酸是可特异性杂交的。这样的结合称为特异性杂交。导致具体的严格程度的杂交条件会根据选择的杂交方法的性质和杂交的核酸序列的组成和长度而变化。通常，杂交温度和杂交缓冲液的离子强度（尤其是 Na^+ 和/或 Mg^{2+} 浓度）有助于杂交的严格性，尽管洗涤时间也影响严格性。关于得到特定严格程度所需要的杂交条件的考虑论述于Sambrook等（编），*Molecular Cloning: A Laboratory Manual*，第2版，1-3卷，Cold Spring Harbor Laboratory Press，Cold Spring Harbor，New York，1989，第9章和第11章中。

[0082] 如本文中使用的，“严格条件”涵盖这样的条件：在该条件下，只有当杂交分子与DNA靶之间的错配小于50%时才发生杂交。“严格条件”包括另外的具体严格性水平。因此，如本文中使用的，“中等严格性”条件为具有50%以上序列错配分子不会杂交的条件；“高严格性”条件为具有20%以上错配分子不会杂交的条件；“极高严格性”条件为具有10%以上错配的序列不会杂交的条件。在特定的实施方案中，严格条件可包括：在65℃杂交，然后在65℃用0.1x SSC/0.1% SDS洗涤40分钟。以下是代表性的非限制性杂交条件：

[0083] • 极高严格性：在5x SSC缓冲液中65℃下杂交16小时；在2x SSC缓冲液中室温下洗涤两次，每次15分钟；再在0.5x SSC缓冲液中65℃下洗涤两次，每次20分钟。

[0084] • 高严格性：在5-6x SSC缓冲液中65-70℃杂交16-20小时；在2x SSC缓冲液中在室温下洗涤两次，每次5-20分钟；再在1x SSC缓冲液中55-70℃洗涤两次，每次30分钟。

[0085] • 中等严格性：在6x SSC缓冲液中室温到55℃杂交16-20小时；再在2x-3x SSC缓冲液中室温到55℃下洗涤至少两次，每次20-30分钟。

[0086] 在一个实施方案中，可特异性杂交的核酸分子可在极高严格性杂交条件下保持结合。在一个实施方案中，可特异性杂交的核酸分子可在高严格性杂交条件下保持结合。在一个实施方案中，可特异性杂交的核酸分子可在中等严格性杂交条件下保持结合。

[0087] 如本文中使用的，术语“寡核苷酸”是指短的核酸聚合物。寡核苷酸可通过切割较长核酸区段、或通过使单独的核苷酸前体聚合而形成。自动合成仪能够合成长度高达几百个碱基对的寡核苷酸。由于寡核苷酸可与互补的核苷酸序列结合，故它们可用作检测DNA或RNA的探针。由DNA组成的寡核苷酸（寡脱氧核糖核苷酸）可用于PCR——一种用于扩增小DNA序列的技术中。在PCR中，寡核苷酸一般被称为“引物”，其允许DNA聚合酶延伸寡核苷酸并复制互补链。

[0088] 如本文中使用的，术语“聚合酶链式反应”或“PCR”定义了如1987年7月28日公告的美国专利No. 4,683,195中所描述的扩增微量核酸、RNA和/或DNA的程序或技术。通常，需要知晓感兴趣区域的末端或在此之外的序列信息，以便能够设计寡核苷酸引物；这些引物与有待扩增的模板的相对链在序列上相同或相似。两个引物的5'端核苷酸可与扩增材料的末端一致。PCR可用来扩增特异性RNA序列、来自总的基因组DNA的特异性DNA序列、和从总的细

胞RNA、噬菌体或质粒序列转录的cDNA,等等。一般可参见Mullis等,Cold Spring Harbor Symp.Quant.Biol.,51:263(1987);Erlich编辑,PCR Technology,(Stockton Press,NY,1989)。

[0089] 如本文中使用的,术语“引物”是指当条件适合于引物延伸产物的合成时,能够充当沿着互补链合成的起始点的寡核苷酸。合成条件包括四种不同的脱氧核糖核苷三磷酸、以及至少一种聚合诱导剂(诸如逆转录酶或DNA聚合酶)的存在。这些物质处于适合的缓冲液中,所述缓冲液可包括作为辅因子、或在不同的适当温度下影响pH等条件的组分。引物优选地是单链序列,使得扩增效率最优化,但也可以利用双链序列。

[0090] 如本文中使用的,术语“探针”是指与靶序列杂交的寡核苷酸。在TaqMan[®]或TaqMan[®]型测定程序中,探针与位于两个引物的退火位点之间的靶部分杂交。探针包含约八个核苷酸、约十个核苷酸、约十五个核苷酸、约二十个核苷酸、约三十个核苷酸、约四十个核苷酸、或约五十个核苷酸。在一些实施方案中,探针包含从约八个核苷酸到约十五个核苷酸。探针可进一步包含可检测标记,例如,荧光团(Texas-Red[®]、异硫氰酸荧光素,等)。可检测标记可直接共价附接到探针寡核苷酸上,例如位于探针的5'端或探针的3'端。包含荧光团的探针还可进一步包含淬灭剂,例如,Black Hole Quencher[™]、Iowa Black[™],等。

[0091] 如本文中使用的,术语“序列同一性”或“同一性”可以互换地使用,指两个序列以指定比较窗口上的最大对应性对齐时,两个序列中相同的核酸残基数。

[0092] 如本文中使用的,术语“序列同一性百分比”是指通过在比较窗口上比较两个最优对齐的序列(例如核酸序列或氨基酸序列)而确定的值,其中为了实现两个序列的最优对齐,该比较窗口中的序列部分相比于参考序列(不包含添加或缺失)可以包含添加或缺失(即,空位)。通过确定在两个序列中出现相同核酸或氨基酸残基的位置的数目来产生匹配位置的数目,用匹配位置的数目除以比较窗口中的位置的总数,结果乘以100产生序列同一性的百分比,从而计算出该百分比。用于比较的序列比对方法是熟知的。各种程序和比对算法例如描述于Smith和Waterman(1981)Adv. Appl. Math. 2:482;Needleman和Wunsch(1970)J. Mol. Biol. 48:443;Pearson和Lipman(1988)Proc. Natl. Acad. Sci. U.S.A. 85:2444;Higgins和Sharp(1988)Gene 73:237-44;Higgins和Sharp(1989)CABIOS 5:151-3;Corpet等(1988)Nucleic Acids Res. 16:10881-90;Huang等(1992)Comp. Appl. Biosci. 8:155-65;Pearson等(1994)Methods Mol. Biol. 24:307-31;Tatiana等(1999)FEMS Microbiol. Lett. 174:247-50。

[0093] 美国国家生物技术信息中心(NCBI)基本局部比对搜索工具(BLAST[™];Altschul等(1990)J. Mol. Biol. 215:403-10)可从几个来源获得,包括美国国家生物技术信息中心(Bethesda,MD),并且在互联网上也可供与几种序列分析程序结合使用。如何使用该程序确定序列同一性的描述可在互联网的BLAST[™]的“帮助”部分获得。为了比较核酸序列,可以采用利用默认参数的BLAST[™](Blastn)程序的“Blast 2序列”函数。在用这种方法评估时,与参考序列具有较大相似性的核酸序列将显示出同一性百分比的增加。

[0094] 如本文中使用的,术语“可操作连接”指已被置于相互之间的功能关系中的两种组分。在有关调节序列和编码序列的情况下使用时,术语“可操作连接”意味着该调节序列影响该连接的编码序列的表达。“调节序列”、“调节元件”、或“控制元件”是指影响转录的周期

和水平/量、RNA加工或稳定性、或相关编码序列的翻译的核酸序列。调节序列可包括启动子；翻译前导序列；5'和3'非翻译区，内含子；增强子；茎环结构；抑制子结合序列；终止序列；多腺苷酸化识别序列等。具体的调节序列可位于与其可操作连接的编码序列的上游和/或下游。并且，与编码序列可操作连接的具体调节序列可位于双链核酸分子的相关互补链上。连接可通过合适的限制性位点处的连接作用(ligation)来完成。如果这样的位点不存在，则根据常规做法使用合成的寡核苷酸衔接子或接头。然而，可操作连接的元件不需要是毗连的。

[0095] 如本文中使用的，术语“转化”涵盖所有可将核酸分子导入这种细胞中的技术。实例包括但不限于：用病毒载体转染；用质粒载体转化；电穿孔；脂质体转染；显微注射(Mueller等(1978)Cell 15:579-85)；土壤杆菌介导的转移；直接DNA摄取；晶须介导的转化；和微粒轰击。

[0096] 如本文中使用的，术语“转导”是指病毒将核酸转移到细胞中的过程。

[0097] 如本文中使用的术语“多接头”或“多克隆位点”定义核酸序列上成簇的三个或更多个2型限制酶位点，彼此之间相距10个核苷酸以内。包含多接头的构建体用于核酸序列(诸如基因的编码区)的插入和/或切除。

[0098] 如本文中使用的，术语“限制性内切核酸酶”和“限制酶”是指多种细菌酶，它们各自切割双链DNA上的特异性核苷酸序列或附近。2型限制酶在同一位点识别和切割DNA，2型限制酶包括但不限于XbaI、BamHI、HindIII、EcoRI、XhoI、SalI、KpnI、AvaI、PstI和SmaI。

[0099] 术语“载体”与术语“构建体”、“克隆载体”和“表达载体”可互换使用，意指一种媒介物，可用其将DNA或RNA序列(例如，外来序列)导入宿主细胞中，从而转化宿主并促进导入序列的表达(例如转录和翻译)。“非病毒载体”意图表示任何不包含病毒或反转录病毒的载体。在一些实施方案中，“载体”为包含至少一个DNA复制起点和至少一个选择标志物基因的DNA序列。实例包括但不限于质粒、粘粒、噬菌体、细菌人工染色体(BAC)、或携带外源DNA进入细胞的病毒。载体还可包括一个或多个基因、反义分子、和/或选择标志物基因和本领域已知的其他遗传元件。载体可以转导、转化或感染细胞，由此导致细胞表达核酸分子和/或由该载体编码的蛋白质。

[0100] 术语“质粒”定义能够在原核或真核宿主细胞中进行常染色体复制的环链核酸。该术语包括可为DNA或RNA并且可为单链或双链的核酸。质粒的定义还可包括对应于细菌复制起点的序列。

[0101] 如本文中使用的术语“选择标志物基因”定义编码蛋白质的基因或其他表达盒，所述蛋白质可以帮助鉴定这样的细胞：该细胞中插入有选择标志物基因。例如，“选择标志物基因”涵盖报告基因以及在植物转化中使用的、例如用于保护植物细胞免受选择剂的影响或提供针对选择剂的抗性/耐受性的基因。在一个实施方案中，只有那些接收功能性选择标志物的细胞或植物才能在有选择剂的条件下分化或生长。选择剂的实例可以包括，例如包括大观霉素、新霉素、卡那霉素、巴龙霉素、庆大霉素、和潮霉素之类的抗生素。这些选择标志物包括表达赋予抗生素卡那霉素抗性的酶的新霉素磷酸转移酶基因(npt II)，和针对有关抗生素新霉素、巴龙霉素、庆大霉素和G418的基因，或表达赋予潮霉素抗性的酶的潮霉素磷酸转移酶基因(hpt)。其他选择标志物基因可以包括编码除草剂抗性的基因，包括bar或pat(针对草胺磷或膦丝菌素的抗性)、乙酰乳酸合酶(ALS，针对抑制剂诸如磺酰脲类(SU)、

咪唑啉酮类(IMI)、三唑并嘧啶类(TP)、嘧啶氧基苯甲酸类(POB)、和阻止支链氨基酸合成第一步的磺酰基氨基羰基三唑啉酮类的抗性)、草甘膦、2,4-D、和金属抗性或敏感性。可用作选择标志物基因的“报告基因”的实例包括表达的报告基因蛋白的视觉观察,所述报告基因蛋白例如编码的蛋白质:β-葡萄糖醛酸糖苷酶(GUS)、萤光素酶、绿色荧光蛋白(GFP)、黄色荧光蛋白(YFP)、DsRed、β-半乳糖苷酶、氯霉素乙酰转移酶(CAT)、碱性磷酸酶等。短语“标志物阳性”是指已被转化而包含选择标志物基因的植物。

[0102] 如本文中使用的,术语“可检测标志物”是指能够检测的标记,例如放射性同位素、荧光化合物、生物发光化合物、化学发光化合物、金属螯合剂、或酶。可检测标志物的实例包括但不限于下列各项:荧光标记(例如,FITC、罗丹明、镧系磷光体)、酶标记(例如,辣根过氧化物酶、β-半乳糖苷酶、萤光素酶、碱性磷酸酶)、化学发光标记、生物素基;由二级报道分子识别的预定多肽表位(例如,亮氨酸拉链对序列、第二抗体的结合位点、金属结合结构域、附加表位)。在一个实施方案中,可检测标志物可通过不同长度的间隔臂附接,以降低潜在的空间位阻。

[0103] 如本文中使用的,术语“检测”以最广义使用,包括特异性分子的定性和定量测量,例如特异性多肽的测量。

[0104] 如本文中使用的,术语“盒”、“表达盒”和“基因表达盒”是指可在特异的限制性位点插入核酸或多核苷酸中或通过同源重组插入核酸或多核苷酸中的DNA区段。如本文中使用的,DNA区段包括编码感兴趣多肽的多核苷酸,所述盒和限制性位点被设计为确保所述盒按照适于转录和翻译的正确阅读框插入。在一个实施方案中,表达盒可包含编码感兴趣多肽的多核苷酸,并且除了所述多核苷酸之外,还具有便于特定宿主细胞转化的元件。在一个实施方案中,基因表达盒还可包括允许编码感兴趣多肽的多核苷酸在宿主细胞中增强表达的元件。这些元件可包括但不限于:启动子、最小启动子、增强子、应答元件、终止子序列、多聚腺苷酸化序列,等。

[0105] 如本文中使用的“接头”或“间隔区”是使两个分开的实体彼此结合的键、分子或分子基团。接头和间隔区可为两个实体提供最佳间距,或者可进一步提供允许两个实体彼此分开的连接。不稳定连接包括可光裂解的基团、酸不稳定性部分、碱不稳定性部分和酶可裂解的基团。

[0106] 如本文中使用的,术语“对照”是指为比较目的而在分析程序中使用的样品。对照可为“阳性”的或“阴性”的。例如,在分析程序的目的是检测细胞或组织中的差异表达的转录物或多肽的情况下,通常优选包括阳性对照,例如来自表现出期望的表达的已知植物的样品,以及阴性对照,例如来自缺乏该期望的表达的已知植物的样品。

[0107] 如本文所使用的,术语“植物”包括整个植物及植物的任何后代、细胞、组织、或部分。可用于本发明中的植物类别通常与适合于诱变的高等和低等植物类别一样宽,包括被子植物(单子叶和双子叶植物)、裸子植物、蕨类和多细胞藻类。因此,“植物”包括双子叶和单子叶植物。术语“植物部分”包括植物的任何部分,包括,例如但不限于:种子(包括成熟种子和未成熟种子);植物插条;植物细胞;植物细胞培养物;植物器官(如花粉、胚、花、果实、芽(shoot)、叶、根、茎、和外植体)。植物组织或植物器官可以是种子、原生质体、愈伤组织、或者任何其他被组织成一定结构或功能单元的植物细胞群。植物细胞或组织培养物可能能够再生出具有该细胞或组织所来源的植物的生理学和形态学特征的植物,并且可能能够再

生出与该植物具有基本上相同的基因型的植物。与之相反,一些植物细胞不能再生而产生植物。植物细胞或组织培养物中的可再生细胞可以是胚、原生质体、分生组织细胞、愈伤组织、花粉、叶、花药、根、根尖、须、花、果仁、穗、穗轴、壳、或茎。

[0108] 植物部分包括可收获的部分和可用于繁殖后代植物的部分。可用于繁殖的植物部分包括,例如但不限于:种子;果实;插条;苗;块茎;和根砧木。植物的可收获部分可以是植物的任何有用部分,包括,例如但不限于:花;花粉;苗;块茎;叶;茎;果实;种子;和根。

[0109] 植物细胞是植物的结构和生理的单位,包括原生质体和细胞壁。植物细胞可以是分离的单细胞或细胞聚集体的形式(例如,松散型(friable)愈伤组织和培养细胞),也可以是更高级的有组织单元(例如,植物组织、植物器官、和植物)的一部分。因此,植物细胞可以是原生质体、配子产生细胞、或可以再生成完整植物的细胞或细胞集合。这样,种子,因其包含多个植物细胞并能够再生为完整植物,在本文的实施方案中被认为是“植物细胞”。

[0110] 如本文中使用的术语“原生质体”是指其细胞壁已经被完全或部分去除的植物细胞,其脂质双层膜裸露,因此包括细胞壁已经完全去除的原生质体、以及其细胞壁仅被部分去除的原生质球,但并非仅限于此。原生质体一般为没有细胞壁的分离的植物细胞,具有再生为细胞培养物或完整植物的潜能。

[0111] 除非另外特别地解释,本文中使用的全部技术术语和科学术语具有与属于本公开的领域之内的普通技术人员通常所理解的相同的含义。可以在以下出版物中找到分子生物学中常见术语的定义:例如,Lewin,Genes V,Oxford University Press,1994(ISBN 0-19-854287-9);Kendrew等(编),The Encyclopedia of Molecular Biology,Blackwell Science Ltd.,1994(ISBN 0-632-02182-9);和Meyers(编),Molecular Biology and Biotechnology:A Comprehensive Desk Reference,VCH Publishers,Inc.,1995(ISBN 1-56081-569-8)。

[0112] 实施方案

[0113] 如本文中公开的,提供了新颖的重组构建体,它们使用来自柳枝稷、二穗短柄草、或粟的泛素基因的调节序列来表达非泛素转基因。这些构建体可用来转化包括植物细胞在内的细胞,以便产生在其细胞中表达转基因的基因产物的完整生物。

[0114] 调节元件

[0115] 用于基础研究或生物技术应用植物启动子通常是单向的,其仅仅指导一个已融合在其3'端(下游)的基因。代谢工程和性状堆叠常常需要将多个基因导入植物中,因此,在转基因作物中一般需要多个启动子来驱动多个基因的表达。

[0116] 转基因产品的开发变得日益复杂,其需要在单基因座中堆叠多个转基因。传统上,每个转基因的表达通常都需要启动子,其中需要多个启动子来在一个基因堆叠中表达不同的转基因。这经常导致在一个转基因堆叠内重复使用相同的启动子,以使得用于表达单一多基因性状的不同转基因获得表达模式的相似水平。已知由相同启动子驱动的多基因构建体可导致基因沉默,结果降低转基因产物在田间的效用。由于启动子重复导致的转录因子(TF)结合位点的过量可引起内源TFs的消耗,从而导致转录失活。转基因的沉默将很可能不良地影响产生的转基因植物表达转基因的性能。转基因内的重复序列可引起基因座内(intra-locus)同源重组,导致多核苷酸重排。

[0117] 使用多样化的启动子表达基因堆叠中的不同转基因是有利的。在一个实施方案

中,从不同植物物种获得的多样化的组成型泛素能驱动多个转录单位,包括RNAi、人工miRNA、或发夹环RNA序列的转录。

[0118] 提供了利用组成型泛素(Ubi1)启动子在植物中表达非泛素转基因的方法和构建体。在一个实施方案中,启动子可以是二穗短柄草泛素1C(Ubi1C)启动子。

[0119]

```
CTGCTCGTTTCAGCCACAGTAACACGCCGTGCGACATGCAGATGCCCTCCACCACGCCGACCAACCCCAAGTCCGCC
GCGCTCGTCCACGGCGCCATCCGCATCCGCGCGTCAACGTCATCCGGAGGAGGCGAGCGCGATGTCGACGGCCACGG
CGGCGGCGGACACGACGGCGACGCCCCGACTCCGCGCGCGCGTCAAGGCTGCAGTGGCGTCGTGGTGGCCGTCCGCC
TGCACGAGATCCCCGCGTGGACGAGCGCCGCTCCACCCAGCCCCCTATATCGAGAAATCAACGGTGGGCTCGAGCTC
CTCAGCAACCTCCCCACCCCCCTTCCGACCACGCTCCCTTCCCCCGTGCCCCCTTCTCCGTAAACCCGAGCCGCC
GAGAACAACACCAACGAAAGGGCGAAGAGAATCGCCATAGAGAGGAGATGGGCGGAGGCGGATAGTTTCAGCCATTC
ACGGAGAAATGGGGAGGAGAGAACACGACATCATACGGACGCGACCCTCTAGCTGGCTGGCTGTCCTAAAGAATCGA
ACGGAATCGCTGCGCCAGGAGAAAAACGAACGGTCTCTGAAGCATGTGCGCCCGTTCTTCCAAAACACTTATCTTTAA
GATTGAAGTAGTATATATGACTGAAATTTTTACAAGGTTTTTCCCCATAAAACAGGTGAGCTTATCTCATCCTTTTG
TTTAGGATGTACGTATTATATATGACTGAATATTTTTTATTTCATTGAATGAAGATTTTCGACCCCCCAAAATAA
AAAACGGAGGGAGTACCTTTGTGCGGTGTATATGGACTAGAGCCATCGGGACGTTTCCGGAGACTGCGTGGTGGGGG
CGATGGACGCACAACGACCGCATTTTCGGTTGCCGACTCGCCGTTTCGCATCTGGTAGGCACGACTCGTCGGGTTCGG
CTCTTGCGTGAGCCGTGACGTAACAGACCCGTTCTCTTCCCCCGTCTGGCCATCCATAAATCCCCCCTCCATCGGCT
TCCCTTTCCTCAATCCAGCACCCCTGATT(SEQ ID NO:1)
```

[0120] 在一个实施方案中,启动子可以是二穗短柄草泛素1(Ubi1)启动子。

[0121]

```
GGCGTCAGGACTGGCGAAGTCTGGACTCTGCAGGGCCGAAGTCTGAAGACGAAGCAGAGGAAGAGAAAGGGAAGTG
TTCGACTTGTAATTTGTAGGGGTTTTTTTTAGAGGAACTTGTAATTTGTAGGTGGGCTGGCCTCGTTGGAAAAACGA
TGCTGGCTGGTTGGGCTGGGCCGATGTACGCTTGCAAACAAGTGTGGCGGCCGTTCTGGACGAGCAGGAGTTTCT
TTTTGTTCTCACTTTTCTGGTCTTCTTTAGTTACGGAGTACCTTTTGTTTTTTAAAGGAGTTACCTTTTTTTTAGG
AATTCTTTAGTTACCTTTTCGCTTGCTCTCAAAAAATATTTAACTTTCGCTTTTTTTTCATTTTAATTTTTGCAACTAT
TTACGAGTTTCATGAATGCTTATTTTCCAGCATATCATTATTTGCAAGTATTTTTATGCCGTATGTATTGGACGAGA
GCCATCGGGACTGTTCCAGAGACTGCGTGGTGGGGACGGCTCCCAACCGCCTTTTCTATCTCTGTTGCATCCGGTG
GCCGACTTGGCTCGCGCGTGAGCCGTGACGTAACAGACTTGGTCTCTTCCCCATCTGGCCATCTATAAATTCCCCCA
TCGATCGACCCTCCCTTTCC(SEQ ID NO:2)
```

[0122] 在一个实施方案中,启动子可以是粟泛素2(Ubi2)启动子。

[0123]

```
TGCGTCTGGACGCACAAGTCATAGCATTATCGGCTAAAATTTCTTAATTTCTAAATTAGTCATATCGGCTAAGAAAG
TGGGGAGCACTATCATTTTCGTAGAACAAGAACAAGGTATCATATATATATATATAATATTTAACTTTGTAA
GTGGAATCAAAGTGCTAGTATTAATGGAGTTTCATGTGCATTAAATTTTATGTCACATCAGCAATTTTGTGACTTG
GCAAGGTCATTTAGGGTGTGTTTGAAGACAGGGGCTATTAGGAGTATTAACATAGTCTAATTACAAAATAATTG
CACAACCGCTAAGCTGAATCGCGAGATGGATCTATTAAGCTTAATTAGTCCATGATTTGACAATGTGGTGCTACAAT
AACCATTTGCTAATGATGGATTACTTAGGTTTAATAGATTTCGTCTCGTGATTTAGCCTATGGGTTCTGCTATTAATT
TTGTAATTAGCTCATATTTAGTTCTTATAATTAGTATCCGAACATCCAATGTGACATGCTAAAGTTTAACCCTGGTA
```

TCCAAATGAAGTCTTATGAGAGTTTCATCACTCCGGTGGTATATGTACTTAGGCTCCGTTTTCTCCACCGACTTAT
TTTTAGCACCCGTCACATTGAATGTTTAGATACTAATTAGAAGTATTAACGTAGACTATTTACAAAATCCATTACA
TAAGACGAATCTAAACGGCGAGACGAATCTATTAAACCTAATTAGTCCATGATTTGACAATGTGTTGCTACAGTAAA
CATTGCTAATGATGGATTAATTAGGCTTAATAGATTCTGCTCGCCGTTTAGCCTCCACTTATGTAATGGGTTTTCT
AAACAATCTACGTTTAATACTCCTAATTAGTATCTAAATATTCAATGTGACACGTGCTAAAAATAAGTCAGTGAAG
GAAGAGAACGTCCCCTTAGTTTTCCATCTTATTAATTGTACGATGAAACTGTGCAGCCAGATGATTGACAATCGCAA
TACTTCAACTAGTGGGCCATGCACATCAGCGACGTGTAACGTCGTGAGTTGCTGTTCCCGTAG(SEQ ID NO:3)

[0124] 在一个实施方案中,启动子可以是柳枝稷(Switchgrass)泛素1启动子。

[0125]

TTGAATTTTAATTTCAAATTTTGCAGGGTAGTAGTGGACATCACAATACATATTTAGAAAAAGTTTTATAATTTTCC
TCCGTTAGTTTTTCATATAATTTTGAACCTCCAACGATTAATCTATTATTAATATCCCGATCTATCAAAATAATGATA
AAAATTTATGATTAATTTTTCTAACATGTGTTATGGTGTGTACTATCGTCTTATAAAATTTCAACTTAAACTCCAC
CTATACATGGAGAAATGAAAAAGACGAATTACAGTAGGGAGTAATTTGAACCAAATGGAATAGTTTGAGGGTAAAT
GAACTAAACAATAGTTTAGGAGGTTATTCAGATTTTAGTTATAGTTGAGAGGAGTAATTTAGACTTTTTCTATCTT
GAATTGTTGACGGCTCTCCTATCGGATATCGGATGGAGTCTTTCAGCCCAACATAACTTCATTGGGGCCAAACGTT
CGTCCATCCAGCCTAGGGAGAACATTTTGCCCATGATATCTGTTTTCTTTTTTCTATTTTCACTGGTATTATAGG
AGGGAAATATACAACGTGTTACCTTTGGTTTCATTCTTGTTCATCTGAATTTATCTAAACTGTGTTTGAACCTC
GTAAGAATTTTGTTTCGATCTGTCCGGTACATCGTGTGATAGGTGGCCTCCGAGATTCTTCTTTTAAACGGCAAAG
TAAATAATCTCAGCTCCAGCCTAACGTCAATTATCAGAGAGAGAAAAAATATTTTTTTATGATTGATCGGAAACC
AACCGCCTTACGTGTCGATCCTGGTTCCTGGCCGGCAGGCGGAGGAAAGCGACCGACCTCGCAACGCCGGCGCAG
GCGCCGCCGTGTTGGACTTGGTCTCCCGCGACTCCGTGGGCCTCGGCTTATCGCCGCCGCTCCATCTCAACCGTCCG
CTTGACACGTGGAAGTTGATCCGTCGCGCACCAGCCTCGGAGGTAACCTAACTGCCCCGTACTATAAATCCGGGATC
CGGCCTCTCCAATCCCCATCGCCA(SEQ ID NO:35)

[0126] 在一个实施方案中,提供了包含泛素启动子的核酸构建体。在一个实施方案中,所述泛素启动子是柳枝稷、二穗短柄草或粟泛素启动子。在一个实施方案中,提供了包含启动子的核酸构建体,其中所述启动子与SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、或SEQ ID NO:35至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同。在一个实施方案中,提供了包含与多接头可操作连接的泛素启动子的核酸构建体。在一个实施方案中,提供了包含与非泛素转基因可操作连接的泛素启动子的基因表达盒。在一个实施方案中,所述启动子由SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、或SEQ ID NO:35组成。在说明性实施方案中,基因表达盒包含与转基因的5'端可操作连接的泛素启动子,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。

[0127] 除了启动子之外,3'非翻译基因区(即,3'UTR)或终止子是转录终止和mRNA的多腺苷酸化所需要的。正确的转录终止和mRNA的多腺苷酸化对于转基因的稳定表达是重要的。转录终止对于多基因堆叠更为关键,以免转录通读到下一个转基因中。类似地,非多腺苷酸化的异常RNA(aRNA)是植物RNA依赖性RNA聚合酶(RdRPs)将aRNA转化为双链RNA(dsRNA)的底物,从而导致小RNA的产生和转基因的沉默。因此,强转录终止子对于单基因和多基因堆

叠都很有用。虽然启动子对于驱动转录是必要的,但3'-UTR基因区能终止转录并起始用于翻译和蛋白质合成的所得mRNA转录物的多腺苷酸化。3'-UTR基因区有助于转基因的稳定表达。

[0128] 根据一个实施方案,提供了包含泛素转录终止子的核酸构建体。在一个实施方案中,所述泛素转录终止子是柳枝稷、二穗短柄草或粟泛素转录终止子。在一个实施方案中,提供了包含转录终止子的核酸构建体,其中所述转录终止子与SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同。在一个实施方案中,提供了包含与多接头可操作连接的泛素转录终止子的核酸构建体。在一个实施方案中,提供了包含与非泛素转基因的3'端可操作连接的泛素转录终止子的基因表达盒。在一个实施方案中,所述转录终止子由SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36组成。在说明性实施方案中,基因表达盒包含与转基因可操作连接的泛素转录终止子,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。在一个实施方案中,提供了包含与多接头序列、非泛素转基因、或二者的组合可操作连接的转录终止子的核酸载体,其中所述转录终止子包含SEQ ID NO:5或与SEQ ID NO:5具有90%序列同一性的序列。在一个实施方案中,转录终止子长度小于1kb,在另一个实施方案中,转录终止子由SEQ ID NO:5的3' UTR序列组成。

[0129] 在一个实施方案中,提供了核酸构建体,其包含如本文中所述的泛素启动子和3'-UTR。在一个实施方案中,所述核酸构建体包含泛素3'-UTR。在一个实施方案中,所述泛素3'-UTR是柳枝稷、二穗短柄草或粟泛素3'-UTR。在一个实施方案中,3'-UTR可以是二穗短柄草泛素1C(Ubi1C)3'-UTR。

[0130]

```
GTGGTCAAAAAGTGGCCTACAGTCTGCTGCCCCCTGTTGGTCTGCCCCCTTGGAAGTAGTCGTGTCTATGGTTATGTG
AGAAGTCGTTGTGTTCTTTCTAATCCCGTACTGTTTGTGTGAACATCTGCTGCTGTCGTATTGCATCGTGAAGAATC
CTGTTATGAATAAGTGAACATGAACCTTGTCTGTGATTACGGCTTCGTGGTTATGCCAACGTTCTTACAAACGCAA
TTGCACCTGATGTAAATCGTTTTTGTCTAGCTGTATGGAACAAGTGCTCATGATGTTTCATGCAAGATGCAATTCAG
CTTTTGTGGTTTGTCTATCTTTGTACTGTGCTTACCGCACATAAAGATTGCATCTTGCTTATTGCTTTGTGCTTTG
GTGCTCGTCCGCTTCTCCTTGACCTTATCAAACCTTTGTTTAGATTCTCTTCTTATAGCACTTGGTAACTCTCAGC
TTTACAACGCCAGTACTGTTTCTGAAATTTTCATGACTGATAAAGCTGATAGATGGAGTACTAATATATGACATCTTT
CCATAAATGTTCCGGTGCAGAGATATGGAGGCCCCAGGATCCTATTTACAGGATGAACCTACCTGGGCCGCTGTACG
CATGACATCCGCGAGCAAGTCTGAGGTTCTCAATGTACACATGAAATTGATTTTTGCTGCGTTTGGCTTGGCTGATC
GTTGCATTTGTTCTGATTCATCAGAGTTAAATAACGGATATATCAGCAAATATCCGCAGCATCCACACCGACCACAC
GTCCGGTTAACAGAGTCCCCCTGCCTTGCTTTAATTATTACGGAGTACTCCGCTATTAATCCTTAGATATGTTTCGA
AGGAACTCAAACCTTCTCCATCTGCAAATCTCAGTGCTTCAAACTGGAATTAGATAATTGAAACCTTCATTCCGT
TGCAATTCAAACTGCAAATTGAACAGCACTGTCAATTTCAATTTCCGGTTCACGATTCACCGATAGGTTGACATG
ATCCATGATCCACCCATTGTACAAC(SEQ ID NO:4)
```

[0131] 在一个实施方案中,3'-UTR可以是二穗短柄草泛素1(Ubi1)3'-UTR。

[0132]

GCTTCTGCCGAAC TGGTTCACAGTCTGCTGCCCTTGGTGGTCTGCCCCCTAGTGGTCATGCCTTTTGTATGTGTCT
TGCGTCCCAATCCTGTATCGTTTGTGTGAACATCTCTGCTGCTGTATAGCAGCTTGAATCCTGTTATGAATTTGTGA
ACCTGAACCTTGTTCCTGGAATCATGTTATGAATAAGTGAACCTGAACCTTGTTCCTGATTATTGTTACAATCTGT
TGTGCCGTATGGTTGGTGTGTGATTATGTTGAACTGGAGAACCAAGTTCGTTCCAGGACATATTGCAACCTAA
GCTAAACCATGTAGAACTACTTGTCTGGGAGACATAAAACGTCATTTTTATGCATTTCGTAACATTTAAGCATACTA
CAATAATTGTATTGTCTTTTCTACTCATCCTTGAAACCATATGCCTCTTCTCAGCGCCTCTACATGCAGTGTGCT
CAGAACAAACAGGCCCTGCCAGCTGCTTTTCAATTTTCCAATTAATAACCACAATAGTCGGACTATGGCATCTGTGG
GTGACTATGCAAGATGTTGCTGTGAGGTCTCTGAAACTTTTCCCATGTATCTGTTGAAATTACCCAGTAAATTCATG
CCTCTATTTAATCTGGCATGGTTGATTTTCAAACAGAATGTGTTTTTTTTTGTCTGGAAGCTATATTGGTAAATAA
ATACAAAGCTGGAGTGTGATTATATTTCCAACAGATATTCAAGAAAATCTCAGTTGATTTATTTACTACTGTAGTAT
ATATATATATCTTACAGTTGACTTCTCATATTTCAAACGACATGTGAGCACATTGTTTCAGTTTCTTAGGATGTGTTG
TGTGCTCAAAGGTGTAATTTTGCATTCTGCCCTCCGAGTAAACACTACAGTATTTTTTTGAGTGGCAGTGCATTTG
ATTACAAGGCAACAACAACAAAACCTATGGCAAGATATCCTTCTTAGAGGCTGCCAGGATCATTTTGAAGTGAACCTA
TGTAAGGCTGAAGAAAAGG (SEQ ID NO:5)

[0133] 在一个实施方案中,3'-UTR可以是粟泛素2(Ubi2)3'-UTR。

[0134]

GCCCATCGGTCATGGATGCTTCTACTGTACCTGGGTCGTCTGGTCTCTGCCTGTGTACCTTTGAAGTACCTGTGTC
GGGATTGTGTTTGGTCATGAACTGCAGTTTGTCTTTGATGTTCTTTTGTCTGGTCTTATGAACTGGTTGTATCTGTA
TGTTTACTGTAACTGTTGTTGCGGTGCAGCAGTATGGCATCCGAATGAATAAATGATGTTTGGACTTAAATCTGTA
CTCTGTTTGTTCGTTATGCCAGTTCTATATTGCCTGAGATCAGAATGTTTAGCTTTTGAGTTCTGTTTGGCTTG
TGGTCGACTCCTGTTTCTTACTTGAGGCGTAACTCTGTTCTGGCAAACCTCAAATGTCTAACTGAATGTTTTAGGACT
TAATTGTTGGACAGATTAACGTGTTTGGTTTGTCTAGATTGTGATTCGGAAGGCTTGTTAGTTGTGGAATCAAGG
AGAGCAGCTAGGTCTGTGCAGAACGTTATTTTGGATTTAAGCCTTCTCAGATTATGCCATTACTCTAAACCTAATGA
TATCATATTTCACTCGGGGATGTTGGAGTAGTCTTTTCTTCTCCTGCAGACAAAATGATTTTGCTTTCTGTGTGT
ACATGATTTTGTGCAACTGTTGCAACAACCTGAAGTAGACAAGTTTGTGACCTCACCAGAAGAATGAAAAAGATTTTGG
AATTTGTTACATCGACAAACCATTTGTAACCTTGCCCATCAGAATGCACAGAAGAGCGGCTACAAATTGACATGCGTT
GCAAACCTTTGCAATAGTTGATGCACATGTTTGCCTATGCCTGCCAGTCTTAGGAAAAGTGTGTGGTTTCGAGAAATCT
AAGCATATGTGCTCTGCTCACATTGCGTGGAACCCACACAGCTTTGTCACACTCTGTCCACTCCAGAAGTCATTCC
TGGCGCTGTTTACCCCTGGTAAAAGGTAACCGAAAACTTCTCAAGGCTGTACCCAAAACCTGGAAGGAAATTTGGAGG
AAATCTTTGCTTTTGATCGGCTCACTCTTTC (SEQ ID NO:6)

[0135] 在一个实施方案中,3'-UTR可以是柳枝稷(Switchgrass)泛素1 3'-UTR。

[0136]

GCCTAGTGCTCCTGAGTTGCCTTTTGTGCTTATGGTCAACCTCTGGTTTAAGTCGTGTGAACTCTCTGCATTGCGTT
GCTAGTGTCTGGTTGTGGTTGTAATAAGAACATGAAGAACATGTTGCTGTGGATCACATGACTTTTTTTTTTTGAACC
GGAAGATCACATGACTTTTCATGGCTTTAAGTTCCTGAACTCTGAAATCTGGACCCCTTTTTAAGCTCTGAACTCATC
ATTCTTGATTTACATCTGGTGTGATCTTATTGATGTGATGCAGTCCTGCTGAAATAGTCAATGTAGATTTCATGAC
TGACTGATTGCGTTTATGGTGTGTATGTTGTTAACAAGCTGAAGGTCGTGTGGTGTCTTTCCAGTTAGACGAAGTGT
GCTTTATTGTAGCGTGTAGTGCTGCTGGATGATTGATGAACTGAAACATTCTGCATTTAGCAACTAGCGAGCCAAAG

GTGATGACTGAGTTTCTGTAGACCTGTTTTTTTATGCCCATGGTCGTTCTTCAATTGCACTTGATTTTCACATTAGC
TGGATCATAATCTGAGCAGACTACTCAAAAGTACAAAGTTCATCTTCGCTATGACGCTTTGCCACTAGGATTTTCTT
TGTATGATTTGTTTACAAATCCTGTAATCTAGTCAAAAGAAAAGCCAAATTTTTCTTTGTATGATTTGTTTACAAA
TCCTCTAATCTAGTCAAAGAAAAGCCAAATTTATCCCTCCTGGTCCCCTACATCACGTAGCTATGTGGCCCGCAAGC
AGATGAAAGCAGCCCCGTCAGCCGACGCCGACGCCGACGCCAACACATCCTGCTCCTCCCTCGCCGGCGCCGGCGCC
GGCGAGGCCGCACCGCCGCTGCCCCGTGGCCGACGGTGCCGCACTGCCGCCGCCCGTGCCGCGAGGCAC
ACGGTGCCGCACTGCCGCCGCTCCCCTTCCGGCATTGCCGGACGGCTGGGCTACTGTCCCCGCCGCTTCCCAAT
(SEQ ID NO:36)

[0137] 在一个实施方案中,提供了核酸构建体,其包含如本文中所述的泛素启动子和3'-UTR,其中所述3'-UTR与SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同。在一个实施方案中,提供了包含如本文中所述的泛素启动子和3'-UTR的核酸构建体,其中所述泛素启动子和3'-UTR可操作连接于多接头的相反末端。在一个实施方案中,提供了包含如本文中所述的泛素启动子和3'-UTR的基因表达盒,其中所述泛素启动子和3'-UTR可操作连接于非泛素转基因的相反末端。在一个实施方案中,3'-UTR由SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36组成。在一个实施方案中,提供了包含如本文中所述的泛素启动子和3'-UTR的基因表达盒,其中所述泛素启动子包含SEQ ID NO:2且3'-UTR包含SEQ ID NO:5,其中所述启动子和3'-UTR可操作连接于非泛素转基因的相反末端。在一个实施方案中,3'-UTR由SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36组成。在一个实施方案中,所述启动子由SEQ ID NO:2、16或40组成,并且3'-UTR由SEQ ID NO:5组成。在说明性实施方案中,基因表达盒包含与转基因可操作连接的泛素3'-UTR,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。在另一个实施方案中,所述转基因与泛素启动子和自柳枝稷、二穗短柄草、或粟分离的相同泛素基因的3'-UTR可操作连接。

[0138] 在一个实施方案中,提供了载体,其包含第一转基因和/或多接头以及第二转基因和/或多接头,其中所述第一转基因和/或多接头与包含选自SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、和SEQ ID NO:35的序列的启动子可操作连接、并且与包含选自SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36的序列的3'-UTR可操作连接,并且所述第二转基因和/或多接头与包含选自SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、和SEQ ID NO:35的序列的启动子可操作连接,并且与包含选自SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36的序列的3'-UTR可操作连接;进一步地,其中第一转基因和/或多接头与第二转基因和/或多接头的启动子来源于来自不同植物物种的Ubi基因。在另一个实施方案中,所述载体提供有第三转基因和/或多接头,其中所述第三转基因和/或多接头与包含选自SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、和SEQ ID NO:35的序列的启动子可操作连接、并且与包含选自SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36的序列的3'-UTR可操作连接,进一步地,其中第三转基因的启动子和/或多接头来源于来自与第一和第二转基因和/或多接头的启动子不同的植物物种的Ubi基因。

[0139] 转基因表达还可被位于启动子序列下游的内含子区所调节。启动子和内含子两者

都能调节转基因表达。启动子对于驱动转录是必要的,而内含子的存在能增加表达水平,从而产生用于翻译和蛋白质合成的mRNA转录物。内含子基因区有助于转基因的稳定表达。

[0140] 在一个实施方案中,提供了包含如本文中所述的泛素启动子和内含子的核酸构建体。在一个实施方案中,内含子与启动子的3'端可操作连接。在一个实施方案中,提供了包含泛素内含子的核酸构建体,所述泛素内含子与分离自柳枝稷、二穗短柄草或粟的泛素启动子的3'端,或这种启动子序列的衍生物的3'端可操作连接。在一个实施方案中,所述泛素内含子是柳枝稷、二穗短柄草或粟泛素内含子,或这种内含子序列的衍生物。

[0141] 在一个实施方案中,内含子可以是二穗短柄草泛素1C内含子。

[0142]

GTATGCAGCCTCGCTTCCTCCTCGCTACCGTTTCAATTCTGGAGTAGGTCGTAGAGGATACCATGTTGATTTGACAG
AGGGAGTAGATTAGATACTTGTAGATCGAAGTGCGGATGTTCCATGGTAGATGATACCATGTTGATTTGATTAGAT
CGGATTAAATCTTTGTAGATCGAAGTGCGCATGTTCCATGAATTGCCTGTTACCAGTAGATTCAAGTTTTTCTGTGT
TATAGAGGTGGGATCTACTCGTTGAGATGATTAGCTCCTAGAGGACACCATGCCGTTTTGGAAAATAGATCAGAACC
GTGTAGATCGATGTGAGCATGTGTTCTGTAGATCCAAGTTCTTTCGCATGTTACTAGTTGTGATCTATTGTTTGTG
TAATACGCTCTCGATCTATCCGTGTAGATTTCACTCGATTACTGTTACTGTGGCTTGATCGTTCATAGTTGTTGCTT
AGGTTTGATCGAACAGTGTCTGAACCTAATTGGATATGTATTCTTGATCTATCAACGTGTAGGTTTCAGTCATGTAT
TTATGTACTCCCTCCGTCCCAAATTAAGTACGTTGGATTTTGTATAAGAATCTATACAAATCCATGTCAGTTAATTC
GGGATGGAGTACCATATTCAATAATTTGTTTATTGCTGTCCACTTATGTACCATATGTTTGTGTTCCCTCATGTGGA
TTCTACTAATTATCATTGATTGGTGATCTTCTATTTTGCTAGTTTCCTAGCTCAATCTGGTTATTCATGTAGATGTG
TTGTTGAAATCGGAGACCATGCTTGTTATTAGATAGTTTATTGCTTATCAGTTTCATGTTCTGGTTGATGCAACACA
TATTCATGTTTCGCTATCTGGTTGCTGCTTGATATTCTCTGATTTACATTATTATAAGAATATATTCTGCTCTGGTT
GTTGCTTCTCATGACTTTACCTACTCGGTAGGTGACTTACCTTTTGGTTTACAATTGTCAACTATGCAG(SEQ ID
NO:7)

[0143] 在一个实施方案中,内含子可以是二穗短柄草泛素1(Ubi1)内含子。

[0144]

GTATGTAGCCTCTCGATTCCCTCCTCAGCCCTGCCCTCGATTTGGTGTACGCGTTGAGATGATGATCTCGTAGATGTC
TAGATGACACCATGTCGATTTGAAATAGATCAGATCCGTGTAGATCGATGAGCTCCTGTGTACCTGTGGATTCAAGT
TATTTTCGCATGCTATTGTTGTGATCTACTAGATCTAGTGTGTGTTATTCTATGCTATCGATTTCTCCGTGTAGATTT
CACTCGATTACTGTTACTGTGGCTTGATCGGCCATAGATGTTGGTTAAGGTTTGATCGGTTAGTGTGTTGAACCTGCG
TGGATATCTAGCATCCATCTATTATCGTGTAGGTTTCGAACAAACAAGCACTATTATTGTACTGATGGTTTCGTCTAT
GGTTGGTTTTGACCGTTTTAGTGTGTTGAACGAGCCTTCTGTATTTGTTTATTGCTGTCCAGTGATGTACCATGTTGCT
TGAGTGTGCGATTATACTAATTATTGTTGATTGATAATCTTGTAGTTTGCTTTTCCTAATTTATTTATCGTAGTCCT
GATTTGCCTCAGCTGTGCCTCACCCGTGCGATGGTCAATCAACTTGTTAGCCCAATCTGCTTAATCATGTACATTTG
TTGTTAGAATCAGAGATCAAGCCAATTAGCTATCTTATTGCTTATCTGTTCCATGTTCTGATCGATGTAACAGTCTA
CACTTTTGCTCTGTGCTACTTGATTAACAACTTCTGACTTAAATTCATGATTGGAAGTTTCAGATCTGATTGTTGCC
TTACTTGACTAATATCTATTATCATGTGACACCTCTCTGTCTTGGTAACTTACCGCTGTTTGTGTTGTAATTTCTGACTA
TGCAG(SEQ ID NO:8)

[0145] 在一个实施方案中,内含子可以是粟泛素2(Ubi2)内含子1。

[0146] GTCACGGGTTTCCTTCCCCACCTCTCCTCTTCCCCACCGCCATAAATAG(SEQ ID NO:9)

[0147] 在一个实施方案中,内含子可以是粟泛素2(Ubi2)内含子2。

[0148]

GTACGGCGATCGTCTTCCTCCTCTAGATCGGCGTGATCTGCAAGTAGTTGATTTGGTAGATGGTTAGGATCTGTGCA
CTGAAGAAATCATGTTAGATCCGCGATGTTTCTGTTCGTAGATGGCTGGGAGGTGGAATTTTTGTGTAGATCTGATA
TGTTCTCCTGTTTATCTTGTACGCTCCTGCGATTTGTGGGGATTTTAGGTCGTTGATCTGGGAATCGTGGGGTTGC
TTCTAGGCTGTTTCGTAGATGAGGTCGTTCTCACGGTTTACTGGATCATTGCCTAGTAGATCAGCTCGGGCTTTCGTC
TTTGTATATGGTGCCATACTTGCATCTATGATCTGGTTCCGTGGTGTACCTAGGTTTCTGCGCCTGATTTCGTCG
ATCGATTTTGTTAGCATGTGGTAAACGTTTGGTCATGGTCTGATTTAGATTAGAGTCGAATAGGATGATCTCGATCT
AGCTCTTGGGATTAATATGCATGTGTACCAATCTGTTCCGTGGTAAAGATGATGAATCTATGCTTAGTTAATGGGT
GTAGATATATATGCTGCTGTTCCCTCAATGATGCCGTAGCTTTTACCTGAGCAGCATGGATCCTCCTGTTACTTAGGT
AGATGCACATGCTTATAGATCAAGATATGTACTGCTACTGTTGGAATCTTTAGTATACCTGATGATCATCCATGCT
CTTGTTACTTGTTTTGGTATACTTGGATGATGGCATGCTGCTGCTTTTTGTTGATTTGAGCCCATCCATATCTGCAT
ATGTCACATGATTAAGATGATTACGCTGTTTCTGTATGATGCCATAGCTTTTATGTGAGCAACATGCATCCTCCTGG
TTATATGCATTAATAGATGGAAGATATCTATTGCTACAATTTGATGATTATTTGTACATACGATGATCAAGCATGC
TCTTCATACTTTGTTGATATACTTGGATAATGAAATGCTGCTGCACGTTTCTATAGCACTAATGATGTGATGAA
CACGCACGACCTGTTTGTGGCATCTGTTTGAATGTGTTGCTGTTCACTAGAGACTGTTTTATTAACCTACTGCT
AGATACTTACCCTTCTGTCTGTTTATTCTTTTGCAG(SEQ ID NO:10)

[0149] 在一个实施方案中,内含子可以是柳枝稷(Switchgrass)泛素内含子。

[0150]

GTACTCCTACCTAATCCTCCTTAACTGATCTCTCCTCTATCACGTTGGTAATCTTCGAATGATCTGCTGCCTGGCTC
GCTGTTCCCCCTCGTTATGCACTGTTTCCATCACGAGTTTTTTTTTTCATCATCTAATCTATGCGGTTGCGGAAGAA
TTGTGGCTAGTGGAGTAGTTTTCTGTGCTTGATCGGTAGATTCGATGTGTGGGTGTATGGATGTTTTCTGAAAAGTT
GCTGGATTAGTTTACGCTTTCAGGCCGAGGTCTGTTGAAAATTGATTATGAAGTCTATATGCTTTGGATCTATCGA
TTTCCAGTTTTATTAGATGTAGGCCAAAAAATTGTCGGCATTGTGTGGAATTAGTTGGCCTTTAGGTCTGCACAT
TCATGGTGACGGCACAGTTGCTGCTGGCTGTTGCGTGGGACGAGTTATTATAGTTGTTTTTGTTCCTGATTGA
TTCACATTTTCAATGATAACTAGCCTTTGTACCTAACCAAGTCCAGGTTGATCCTATCTGTGTTCTTCAGCTACCA
GTTTGCATAGATGATGGTGTATTCGATTGCTTTAGTAGGCCTTCTGATTTACATCTAATTCTGTGATGAATATAGA
TAACTTTACATGCTTTTGTATATACTTTATATTTGAACTGTTCACTGTCCAGCCTATTTTGGATAATTGAGTGCATTG
GCTTTTGTATGCCTGAATTATTCACATGTTCTGGATAATTGACCTGTGTACCTAGTTGACTGTTTTTTGAGGTGCC
ACCCGTCTGTTGAGCTGATTTGTGTATTGCTGCTAGTTAATCTTTGATTATGCAGCTAGTGCTTTGTGATAT
GTAGCTTTATAGGCTTCTGATGTCTTGGATATAGTTCAGTCTACTTGTCAAGTTGCTTTACAAGTAGTAGCTCTGA
TTCTATTTGGCTTCTGAGTCAGAGCTTTGCAAATTGCTTGTGTTACATTACATCATATTACTGAATTGCAGTTA
TTAATGGTTGGATTGTTGCTGTTTACTTCTACATTTTTTGTGCTGTTTTATATTATACTAAAATGTTTGTGTTGCTGC
TTTTAG(SEQ ID NO:37)

[0151] 在一个实施方案中,提供了包含如本文中所述的泛素启动子和内含子的核酸构建体,其中所述内含子与SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、或SEQ ID NO:37至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同。在一个实施方案中,提供了包含如本文中所述的泛素启动子、内含子序列和多接头的核酸构建体,其中所述启动子和内含子与多接头可操作连接。在一个实施

方案中,提供了包含如本文中所述的泛素启动子、内含子序列和非泛素转基因的基因表达盒,其中所述启动子和内含子与所述转基因的5'端可操作连接。任选地,所述构建体包含与非泛素转基因或多接头的3'端可操作连接的3'-UTR。在一个实施方案中,所述启动子和3'-UTR序列选自本文中所述的那些,并且所述内含子序列由SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、或SEQ ID NO:37组成。在一个实施方案中,基因表达盒包含与启动子可操作连接的泛素内含子,其中所述启动子为柳枝稷、二穗短柄草或粟泛素启动子,或为源自植物(例如,玉米泛素1启动子)、病毒(例如,木薯叶脉花叶病毒启动子)或细菌(例如,根癌土壤杆菌 δ_{mas})的启动子。在说明性实施方案中,基因表达盒包含与转基因可操作连接的泛素内含子,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。

[0152] 转基因表达还可受到位于启动子序列下游的5'-UTR区的调节。启动子和5'-UTR两者都能调节转基因表达。启动子对于驱动转录是必要的,而5'-UTR的存在能增加表达水平,结果产生用于翻译和蛋白质合成的mRNA转录物。5'-UTR基因区有助于转基因的稳定表达。

[0153] 在一个实施方案中,提供了包含如本文中所述的泛素启动子和5'-UTR的核酸构建体。在一个实施方案中,5'-UTR与启动子的3'端可操作连接。在一个实施方案中,提供了包含泛素5'-UTR的核酸构建体,所述泛素5'-UTR与分离自柳枝稷、二穗短柄草或粟的泛素启动子的3'端、或这种启动子序列的衍生物的3'端可操作连接。在另一个实施方案中,5'-UTR的3'端与来自柳枝稷、二穗短柄草或粟的泛素内含子的5'端可操作连接,如本文所述。

[0154] 在一个实施方案中,5'-UTR可以是二穗短柄草泛素1C(Ubi1C)5'-UTR。

[0155] CCGATCGAAAAGTCCCCGCAAGAGCAAGCGACCGATCTCGTGAATCTCCGTCAAG(SEQ ID NO:11)

[0156] 在一个实施方案中,5'-UTR可以是二穗短柄草泛素1(Ubi1)5'-UTR。

[0157]

CCAATCCAGCACCCCCGATCCCGATCGAAAATTCTCCGCAACAGCAAGCGATCGATCTAGCGAATCCCCGTCAAG
(SEQ ID NO:12)

[0158] 在一个实施方案中,5'-UTR可以是粟泛素2(Ubi2)5'-UTR1。

[0159]

AGAAATATCAACTGGTGGGCCACGCACATCAGCGTCGTGTAACGTGGACGGAGAGCCCCGTGACGGCGTCGACATC
GAACGGCCACCAACCACGGAACCACCCGTCCCCACCTCTCGGAAGCTCCGCTCCACGGCGTCGACATCTAACGGCTA
CCAGCAGGCGTACGGGTTGGAGTGGACTCCTTGCTCTTTGCGCTGGCGGCTTCCGGAATTGCGTGGCGGAGACGA
GGCGGGCTCGTCTCACACGGCACGGAAGAC(SEQ ID NO:13)

[0160] 在一个实施方案中,5'-UTR可以是粟泛素2(Ubi2)5'-UTR2。

[0161]

CCGACCCCTCGCCTTTCTCCCCAATCTCATCTCGTCTCGTGTGTTTCGGAGCACACCACCCGCCCAAATCGTTCT
TCCGCAAGCCTCGGCGATCCTTACCCGCTTCAAG(SEQ ID NO:14)

[0162] 在一个实施方案中,5'-UTR可以是柳枝稷(Switchgrass)泛素5'-UTR。

[0163]

CAAGTTCGCGATCTCTCGATTTACAAATCGCCGAGAAGACCCGAGCAGAGAAGTTCCTCCGATCGCCTTGCCAAG
(SEQ ID NO:38)。

[0164] 在一个实施方案中,提供了包含如本文中公开的泛素启动子和5'-UTR的核酸构建体,其中所述5'-UTR与SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、或SEQ ID NO:38至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同。在一个实施方案中,提供了包含泛素启动子和与多接头可操作连接的5'-UTR的核酸构建体,其中所述启动子与SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、或SEQ ID NO:35至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同。在一个实施方案中,提供了包含泛素启动子和与非泛素转基因可操作连接的5'-UTR序列的基因表达盒,其中所述启动子与SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、或SEQ ID NO:35至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同。任选地,所述构建体可进一步包含如本文中公开的泛素内含子,其与5'-UTR的3'端及非泛素转基因的5'端可操作连接;而且所述构建体任选地还可包含与非泛素转基因的3'端可操作连接的3'-UTR。在一个实施方案中,所述启动子、内含子和3'-UTR序列选自本文中所述的那些,且所述5'-UTR序列由SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、或SEQ ID NO:38组成。在一个实施方案中,3'-UTR由SEQ ID NO:12组成。

[0165] 在一个实施方案中,基因表达盒包含与启动子可操作连接的泛素5'-UTR,其中所述启动子为柳枝稷、二穗短柄草或粟泛素启动子,或为源自植物(例如,玉米泛素1启动子)、病毒(例如,木薯叶脉花叶病毒启动子)或细菌(例如,根癌土壤杆菌 δ mas)的启动子。在说明性实施方案中,基因表达盒包含与转基因可操作连接的泛素5'-UTR,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。

[0166] 在一个实施方案中,提供了核酸构建体,其包含启动子和多接头以及任选地下列元件中的一种或多种:

[0167] a)5'非翻译区;

[0168] b)内含子;和

[0169] c)3'非翻译区,

[0170] 其中

[0171] 所述启动子由SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、或SEQ ID NO:35或与SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、或SEQ ID NO:35具有98%序列同一性的序列组成;

[0172] 所述5'非翻译区由SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、或SEQ ID NO:38或与SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、或SEQ ID NO:38具有98%序列同一性的序列组成;

[0173] 所述内含子由SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、或SEQ ID NO:37或与SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、或SEQ ID NO:37具有98%序列同一性的序列组成;

[0174] 所述3'非翻译区由SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36或与SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36具有98%序列同一性的序列组成;进一步地,其中所述启动子与所述多接头可操作连接,并且每个任选元件当存在时也与启动子及多接头二者可操作连接。

[0175] 在一个实施方案中,提供了核酸构建体,其包含启动子和非泛素转基因以及任选地下列元件中的一种或多种:

[0176] a)5'非翻译区;

[0177] b)内含子;和

[0178] c)3'非翻译区,

[0179] 其中

[0180] 所述启动子由SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、或SEQ ID NO:35或与SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、或SEQ ID NO:35具有98%序列同一性的序列组成;

[0181] 所述5'非翻译区由SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、或SEQ ID NO:38或与SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、或SEQ ID NO:38具有98%序列同一性的序列组成

[0182] 所述内含子由SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、或SEQ ID NO:37或与SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、或SEQ ID NO:37具有98%序列同一性的序列组成

[0183] 所述3'非翻译区由SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36或与SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36具有98%序列同一性的序列组成;进一步地,其中所述启动子与所述转基因可操作连接,并且每个任选元件当存在时也与启动子和转基因二者可操作连接。在另一个实施方案中,提供了包含上文刚才公开的核酸构建体的转基因细胞。在一个实施方案中,所述转基因细胞为植物细胞,并且在另一个实施方案中提供了植物,其中所述植物包含所述转基因细胞。

[0184] 根据一个实施方案,转基因表达被与内含子和5'-UTR区可操作连接的启动子所调节,其中所述内含子和5'-UTR区位于所述启动子序列的下游。与内含子和5'-UTR区可操作连接的启动子可用来驱动转基因表达。启动子对于驱动转录是必要的,而内含子和5'-UTR的存在能增加表达水平,从而产生用于翻译和蛋白质合成的mRNA转录物。

[0185] 在一个实施方案中,基因表达盒包含与5'-UTR和内含子区可操作连接的启动子。在一个实施方案中,基因表达盒包含与泛素5'-UTR和泛素内含子可操作连接的泛素启动子。在一个实施方案中,与5'-UTR和内含子区可操作连接的泛素启动子是与内含子和5'-UTR可操作连接的柳枝稷、二穗短柄草或粟泛素启动子。

[0186] 在一个实施方案中,与5'-UTR和内含子可操作连接的启动子可以是与内含子和5'-UTR可操作连接的二穗短柄草泛素1C(Ubi1C)启动子。在一个实施方案中,所述启动子包含或为SEQ ID NO:15:

[0187]

```
CTGCTCGTTCAGCCACAGTAACACGCCGTGCGACATGCAGATGCCCTCCACCACGCCGACCAACCCCAAGTCCGCC
GCGCTCGTCCACGGCGCCATCCGCATCCGCGCGTCAACGTCATCCGGAGGAGGCGAGCGCGATGTCGACGGCCACGG
CGGCGGCGGACACGACGGCGACGCCCCGACTCCGCGCGCGTCAAGGCTGCAGTGGCGTCGTGGTGGCCGTCCGCC
TGCACGAGATCCCCGCGTGACGAGCGCCGCTCCACCCAGCCCCCTATATCGAGAAATCAACGGTGGGCTCGAGCTC
CTCAGCAACCTCCCCACCCCCCTTCCGACCACGCTCCCTTCCCCCGTGCCCCCTTCTCCGTAAACCCGAGCCGCC
GAGAACAACACCAACGAAAGGGCGAAGAGAATCGCCATAGAGAGGAGATGGGCGGAGGCGGATAGTTTCAGCCATTC
ACGGAGAAATGGGGAGGAGAGAACACGACATCATACGGACGCGACCCTCTAGCTGGCTGGCTGTCCTAAAGAATCGA
```

ACGGAATCGCTGCGCCAGGAGAAAACGAACGGTCCCTGAAGCATGTGCGCCCGTTCTTCCAAAACACTTATCTTTAA
GATTGAAGTAGTATATATGACTGAAATTTTTACAAGGTTTTTCCCCATAAAACAGGTGAGCTTATCTCATCCTTTTG
TTTAGGATGTACGTATTATATATGACTGAATATTTTTTATTTTCATTGAATGAAGATTTTCGACCCCCAAAAATAA
AAAACGGAGGGAGTACCTTTGTGCCGTGTATATGGACTAGAGCCATCGGGACGTTTCCGGAGACTGCGTGGTGGGGG
CGATGGACGCACAACGACCGCATTTTCGGTTGCCGACTCGCCGTTTCGCATCTGGTAGGCACGACTCGTCGGGTTCGG
CTCTTGCGTGAGCCGTGACGTAACAGACCCGTTCTTCCCCCGTCTGGCCATCCATAAATCCCCCTCCATCGGCT
TCCCTTTCTCAATCCAGCACCTGATTCCGATCGAAAAGTCCCCGCAAGAGCAAGCGACCGATCTCGTGAATCTCC
GTCAAGGTATGCAGCCTCGCTTCCTCCTCGCTACCGTTTCAATTCTGGAGTAGGTCGTAGAGGATACCATGTTGATT
TGACAGAGGGAGTAGATTAGATACTTGTAGATCGAAGTGCGGATGTTCCATGGTAGATGATACCATGTTGATTTTCA
TTAGATCGGATTAAATCTTTGTAGATCGAAGTGCGCATGTTCCATGAATTGCCTGTTACCAGTAGATTCAAGTTTTT
CTGTGTTATAGAGGTGGGATCTACTCGTTGAGATGATTAGCTCCTAGAGGACACCATGCCGTTTTGGAAAATAGATC
AGAACCGTGTAGATCGATGTGAGCATGTGTTCTGTAGATCCAAGTTCTTTTCGCATGTTACTAGTTGTGATCTATTG
TTTGTGTAATACGCTCTCGATCTATCCGTGTAGATTTCACTCGATTACTGTTACTGTGGCTTGATCGTTCATAGTTG
TTCGTTAGGTTTGATCGAACAGTGTCTGAACCTAATTGGATATGTATTCTTGATCTATCAACGTGTAGGTTTCAGTC
ATGTATTTATGTACTCCCTCCGTCCCAAATTAAGTACGTGGATTTTGTATAAGAATCTATACAAATCCATGTCAGT
TAATTCGGGATGGAGTACCATATTCAATAATTTGTTTATTGCTGTCCACTTATGTACCATATGTTTGTGTTCTCA
TGTGGATTCTACTAATTATCATTGATTGGTGATCTTCTATTTTGCTAGTTTCCTAGCTCAATCTGGTTATTCATGTA
GATGTGTTGTTGAAATCGGAGACCATGCTTGTTATTAGATAGTTTATTGCTTATCAGTTTCATGTTCTGGTTGATGC
AACACATATTGATGTTGCTATCTGGTTGCTGCTTGATATTCTCTGATTTACATTCATTATAAGAATATATTCTGCT
CTGGTTGTTGCTTCTCATGACTTTACCTACTCGGTAGGTGACTTACCTTTTGGTTTACAATTGTCAACTATGCAG
(SEQ ID NO:15)

[0188] 在一个实施方案中,与5'-UTR和内含子可操作连接的启动子可以是与5'-UTR和内含子可操作连接的二穗短柄草泛素1(Ubi1)启动子。在一个实施方案中,所述启动子包含或为SEQ ID NO:16:

[0189]

GGCGTCAGGACTGGCGAAGTCTGGACTCTGCAGGGCCGAAGTCTGAAGACGAAGCAGAGGAAGAGAAAGGGAAGTG
TTCGACTTGTAATTTGTAGGGGTTTTTTTTAGAGGAACTTGTAATTTGTAGGTGGGCTGGCCTCGTTGGAAAAACGA
TGCTGGCTGGTTGGGCTGGGCCGATGTACGCTTGCAAACAAGTGTGGCGGCCGTTCTGGACGAGCAGGAGTTTCT
TTTTTGTCTCACTTTTCTGGTCTTCTTTAGTTACGGAGTACCTTTTGTTTTTTAAAGGAGTTACCTTTTTTTTAGG
AATTCTTTAGTTACCTTTTCGCTTGTCTCAAAAAATATTTAACTTTTCGCTTTTTTTTCATTTTAATTTTGCACAT
TTACGAGTTTCATGAATGCTTATTTTCCAGCATATCATTATTTGCAAGTATTTTTATGCCGTATGTATTGGACGAGA
G C C A T C G G G A C T G T T C C A G A G A C T G C G T G G T G G
GGACGGCTCCCAACCGCCTTTTCTATCTCTGTTTCGCATCCGGTGGCCGACTTGGCTCGCGCGTGAGCCGTGACGTAA
CAGACTTGGTCTCTTCCCCATCTGGCCATCTATAAATTCCCCATCGATCGACCTCCCTTTCCCCAATCCAGCACC
CCCGATCCCGATCGAAAATTCTCCGCAACAGCAAGCGATCGATCTAGCGAATCCCCGTCAAGGTATGTAGCCTCTCG
ATTCTCCTCAGCCCTGCCCTCGATTTGGTGACGCGTTGAGATGATGATCTCGTAGATGTCTAGATGACACCATGT
CGATTTGAAATAGATCAGATCCGTGTAGATCGATGAGCTCCTGTGTACCTGTGGATTCAAGTTATTTTCGCATGCTA
TTGTTGTGATCTACTAGATCTAGTGTGTATTCTATGCTATCGATTTCTCCGTGTAGATTTCACTCGATTACTGTT
ACTGTGGCTTGATCGGCCATAGATGTTGGTTAAGGTTTGATCGGTTAGTGTGTTGAACCTGCGTGGATATCTAGCATC

CATCTATTATCGTGTAGGTTTCGAACAAACAAGCACTATTATTGTACTGATGGTTCGTCTATGGTTGGTTTTGACCG
TTTTAGTGTTGAACGAGCCTTCTGTATTTGTTTATTGCTGTCCAGTGATGTACCATGTTTCGTTGAGTGTGCGATTAT
ACTAATTATTGTTGATTGATAATCTTGTAGTTTGCTTTTCCTAATTTATTTATCGTAGTCCTGATTTGCCTCAGCTG
TGCCTCACCCGTGCGATGGTCAATCAACTTGTTAGCCCAATCTGCTTAATCATGTACATTTGTTGTTAGAATCAGAG
ATCAAGCCAATTAGCTATCTTATTGCTTATCTGTTCCATGTTCTGATCGATGTAACAGTCTACACTTTTGCTCTGTG
CTACTTGATTAAACATTCTGACTTAAATTCATGATTGGAAGTTTCAGATCTGATTGTTGCCTTACTTGACTAATAT
CTATTCATGTGACACCTCTCTGTCTTGGTAACTTACCGCTGTTTGTGTAATTTCTGACTATGCAG (SEQ ID
NO:16)

[0190] 在一个实施方案中,与5'-UTR和内含子可操作连接的启动子可以是与5'-UTR和内含子可操作连接的栗泛素2(Ubi2)启动子。在一个实施方案中,所述启动子包含或为SEQ ID NO:17:

[0191]

TGCGTCTGGACGCACAAGTCATAGCATTATCGGCTAAAATTTCTTAATTTCTAAATTAGTCATATCGGCTAAGAAAG
TGGGGAGCACTATCATTTTCGTAGAACAAGAACAAGGTATCATATATATATATATAATATTTAACTTTGTAA
GTGGAATCAAAGTGCTAGTATTAATGGAGTTTCATGTGCATTAAATTTTATGTCACATCAGCAATTTTGTGACTTG
GCAAGGTCATTTAGGGTGTGTTTGGAAAGACAGGGGCTATTAGGAGTATTAAACATAGTCTAATTACAAAATAATTG
CACAACCGCTAAGCTGAATCGCGAGATGGATCTATTAAGCTTAATTAGTCCATGATTTGACAATGTGGTGCTACAAT
AACCATTTGCTAATGATGGATTACTTAGGTTTAATAGATTTCGTCTCGTGATTTAGCCTATGGGTTCTGCTATTAATT
TTGTAATTAGCTCATATTTAGTTCTTATAATTAGTATCCGAACATCCAATGTGACATGCTAAAGTTTAACCCTGGTA
TCCAAATGAAGTCTTATGAGAGTTTCATCACTCCGGTGGTATATGTACTTAGGCTCCGTTTTCTCCACCGACTTAT
TTTTAGCACCCGTCACATTGAATGTTTAGATACTAATTAGAAGTATTAAACGTAGACTATTTACAAAATCCATTACA
TAAGACGAATCTAAACGGCGAGACGAATCTATTAAACCTAATTAGTCCATGATTTGACAATGTGTTGCTACAGTAAA
CATTGCTAATGATGGATTAATTAGGCTTAATAGATTTCGTCTCGCCGTTTAGCCTCCACTTATGTAATGGGTTTTCT
AAACAATCTACGTTTAACTCCTAATTAGTATCTAAATATTCAATGTGACACGTGCTAAAAATAAGTCAGTGGAAG
GAAGAGAACGTCCCTTAGTTTTCCATCTTATTAATTGTACGATGAACTGTGCAGCCAGATGATTGACAATCGCAA
TACTTCAACTAGTGGGCCATGCACATCAGCGACGTGTAACGTCGTGAGTTGCTGTTCCCGTAGAGAAATATCAACTG
GTGGGCCACGCACATCAGCGTCGTGTAACGTGGACGGAGGAGCCCCGTGACGGCGTCGACATCGAACGGCCACCAAC
CACGGAACCAACCGTCCCCACCTCTCGGAAGCTCCGCTCCACGGCGTCGACATCTAACGGCTACCAGCAGGCGTACG
GGTTGGAGTGGACTCCTTGCTCTTTGCGCTGGCGGCTTCCGAAATTGCGTGGCGGAGACGAGGCGGGCTCGTCTC
ACACGGCACGGAAGACGTACGGGTTCCCTCCCCACCTCTCCTTCCCCACCGCCATAAATAGCCGACCCCTCGC
CTTTCTCCCAATCTCATCTCGTCTCGTGTTGTTTCGGAGCACACCACCCGCCCCAAATCGTTCTTCCCGCAAGCCTC
GGCGATCCTTACCCGCTTCAAGGTACGGCGATCGTCTTCTCCTCTAGATCGGCGTGATCTGCAAGTAGTTGATTT
GGTAGATGGTTAGGATCTGTGCACTGAAGAAATCATGTTAGATCCGCGATGTTTCTGTTTCGTAGATGGCTGGGAGGT
GGAATTTTTGTGTAGATCTGATATGTTCTCCTGTTTATCTTGTACAGCTCCTGCGATTTGTGGGGATTTTAGGTCGT
TGATCTGGGAATCGTGGGGTTGCTTCTAGGCTGTTTCGTAGATGAGGTCGTTCTCACGGTTTACTGGATCATTGCCTA
GTAGATCAGCTCGGGCTTTCGTCTTTGTATATGGTGCCATACTGCATCTATGATCTGGTTCGGTGGTGTTACCTA
GGTTTCTGCGCCTGATTCGTCCGATCGATTTTGTAGCATGTGGTAAACGTTTGGTCATGGTCTGATTTAGATTAGA
GTCGAATAGGATGATCTCGATCTAGCTCTTGGGATTAATATGCATGTGTCACCAATCTGTTCCGTGGTTAAGATGAT
GAATCTATGCTTAGTTAATGGGTGTAGATATATATGCTGCTGTTCCCTCAATGATGCCGTAGCTTTTACCTGAGCAGC

ATGGATCCTCCTGTTACTTAGGTAGATGCACATGCTTATAGATCAAGATATGTACTGCTACTGTTGGAATTCTTTAG
TATACCTGATGATCATCCATGCTCTTGTTACTTGTGTTTGGTATACTTGGATGATGGCATGCTGCTGCTTTTTGTTGA
TTTGAGCCCATCCATATCTGCATATGTCACATGATTAAAGATGATTACGCTGTTTCTGTATGATGCCATAGCTTTTAT
GTGAGCAACATGCATCCTCCTGGTTATATGCATTAATAGATGGAAGATATCTATTGCTACAATTTGATGATTATTTT
GTACATACGATGATCAAGCATGCTCTTCATACTTTGTTGATATACTTGGATAATGAAATGCTGCTGCACGTTTCATTC
TATAGCACTAATGATGTGATGAACACGCACGACCTGTTTGTGGCATCTGTTTGAATGTGTTGTTGCTGTTCACTAGA
GACTGTTTTATTAACTACTGCTAGATACTTACCCTTCTGTCTGTTTATTCTTTTGCAG(SEQ ID NO:17)

[0192] 在一个实施方案中,与5'-UTR和内含子可操作连接的启动子可以是与5'-UTR和内含子可操作连接的柳枝稷(Switchgrass)泛素启动子。在一个实施方案中,所述启动子包含或为SEQ ID NO:39:

[0193]

TTGAATTTTAATTTCAAATTTTGCAGGGTAGTAGTGGACATCACAATACATATTTAGAAAAAGTTTTATAATTTTCC
TCCGTTAGTTTTTCAATAATTTTGAACCTCCAACGATTAATCTATTATTAATATCCCGATCTATCAAAATAATGATA
AAAATTTATGATTAATTTTTCTAACATGTGTTATGGTGTGTAATCTATCGTCTTATAAAATTTCAACTTAAACTCCAC
CTATACATGGAGAAATGAAAAAGACGAATTACAGTAGGGAGTAATTTGAACCAAATGGAATAGTTTGAGGGTAAAT
GAACTAAACAATAGTTTAGGAGGTTATTCAGATTTTAGTTATAGTTGAGAGGAGTAATTTAGACTTTTTCTATCTT
GAATTGTTGACGGCTCTCCTATCGGATATCGGATGGAGTCTTTCAGCCCAACATAACTTCATTGGGCCCCAACGTT
CGTCCATCCAGCCTAGGGAGAACATTTTGCCCATGATATCTGTTTTTCTTTTTTCTATTTTCACTGGTATTATAGG
AGGGAAATATACAACGTGTTACCTTTGGTTTCACTTCTGTTCCATCTGAATTTATCTAAACTGTGTTTGAACCTC
GTAAGAATTTTGTTCGATCTGTCCGGTACATCGTGTGATAGGTGGCCTCCGAGATTCTTCTTTTAAACGGCAAAG
TAAATAATCTCAGCTCCAGCCTAACGTCAATTATCAGAGAGAGAAAAAATATTTTTTTATGATTGATCGGAAACC
AACCGCCTTACGTGTCGATCCTGGTTCTTGCCCGGCACGGCGGAGGAAAGCGACCGACCTCGCAACGCCGGCGCACG
GCGCCGCCGTGTTGGACTTGGTCTCCCGCGACTCCGTGGGCCCTCGGCTTATCGCCGCCGTCCATCTCAACCGTCCG
CTTGACACGTGGAAGTTGATCCGTCGCGCACCAGCCTCGGAGGTAACCTAAGTCCCGTACTATAAATCCGGGATC
CGGCCTCTCCAATCCCCATCGCCACAAGTTTCGCGATCTCTCGATTTACAAATCGCCGAGAAGACCCGAGCAGAGAA
GTTCCCTCCGATCGCCTTGCCAAGGTACTCTACCTAATCCTCCTTAAGTATCTCTCTCTATCACGTTGGTAATC
TTCGAATGATCTGCTGCCTGGCTCGCTGTTCCCCCTCGTTATGCACTGTTTCCATCACGAGTTTTTTTTTTCATCAT
CTAATCTATGCGGTTGCGGAAGAATTGTGGCTAGTGGAGTAGTTTTCTGTGCTTGATCGGTAGATTCGATGTGTGGG
TGTATGGATGTTTTCTGAAAAGTTGCTGGATTAGTTTACGCTTTCAGGCCGAGGTCTGTTGCAAATTGATTATGAA
GTCTATATGCTTTGGATCTATCGATTTCCAGTTTATTCAGATGTAGGCCAAAAAATTGTCGGCATTGTGTGGAAT
TAGTTGGCCTTTAGGTCTGCACATTCATGGTGACGGCACAGTTGCTGCTGGCTGTTGCGTGGGACGAGTTATTATAG
TTGTTTTTGTGTTTTCCCTGATTGATTCACATTTCAATGATAACTAGCCTTTGTCACCTAACCAAGTCCAGGTTGAT
CCTATCTGTGTTCTTCAGCTACCAGTTTGCATAGATGATGGTGTATTGATTGCTTTAGTAGGCCTTCTGATTTTAC
ATCTAATTCTGTGATGAATATAGATAACTTTACATGCTTTTGATATACTTTATATTTGAACTGTTCACTGTCCAGCC
TATTTTGGATAATTGAGTGCATTGGCTTTTGATGCCTGAATTATTCACATGTTCCCTGGATAATTGACCTGTGTCACC
TAGTTGACTGTTTTTTGAGGTGCCACCCGTCTGTTACAGCTGATTTGTGTATTGCTCTAGTTAATCTTTTGAT
TATGCAGCTAGTGCTTTGTCATATGTAGCTTTATAGGCTTCTGATGTCCTGGATATAGTTCAGTCTACTTGTCAAG
TTGCTTTACAAGTAGTAGCTCTGATTCTATTTGGCTTCCTGAGTCAGAGCTTTGCAAATTGCTTGTTGTTACATTAC
ATCATATTACTTGAATTGCAGTTATTTAATGGTTGGATTGTTGCTGTTTACTTCTACATTTTTTGTGTTTTATATT

ATACTAAATGTTTGTGTTGCTGCTTTTCAG(SEQ ID NO:39)

[0194] 在一个实施方案中,提供了包含与内含子和5'-UTR可操作连接的启动子的核酸构建体。在一个实施方案中,所述构建体包含与SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、或SEQ ID NO:39至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同的启动子。在一个实施方案中,提供了包含与多接头可操作连接的泛素启动子序列的核酸构建体,所述泛素启动子序列包含或为与SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、或SEQ ID NO:39至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同的序列。任选地,所述构建体可进一步包含与多接头的3'端可操作连接的3'-UTR。在一个实施方案中,提供了包含与非泛素转基因可操作连接的泛素启动子序列的基因表达盒,其中所述启动子序列包含或为与SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、或SEQ ID NO:39至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同的序列。任选地,所述构建体可进一步包含与非泛素转基因的3'端可操作连接的3'-UTR。在一个实施方案中,3'-UTR序列由SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36组成。在说明性实施方案中,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。在一个实施方案中,所述转基因是除草剂抗性基因。在一个实施方案中,提供了包含1、2、3或4个启动子序列的载体,这些启动子序列独立地选自SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36。

[0195] 在一个实施方案中,基因表达盒包含泛素启动子、泛素5'-UTR、泛素内含子、和泛素3'-UTR。在一个实施方案中,泛素启动子、泛素5'-UTR、泛素内含子、和泛素3'-UTR可以各自独立地是:柳枝稷、二穗短柄草或粟泛素启动子;柳枝稷、二穗短柄草或粟泛素5'-UTR;柳枝稷、二穗短柄草或粟泛素内含子;以及柳枝稷、二穗短柄草或粟泛素3'-UTR。在一个实施方案中,基因表达盒包含:a)启动子,其中所述启动子与SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、或SEQ ID NO:36至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同;b)3'-UTR,其中所述3'-UTR与SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:37至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同;c)5'-UTR,其中所述5'-UTR与SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、或SEQ ID NO:38至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同;或d)内含子,其中所述内含子与SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、或SEQ ID NO:39至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同。

[0196] 例如,基因表达盒可包括启动子、内含子、和5'-UTR,其中所述启动子是多核苷酸SEQ ID NO:2,所述内含子是多核苷酸SEQ ID NO:8,所述5'-UTR是多核苷酸SEQ ID NO:12。同样,基因表达盒可包括启动子、内含子、和5'-UTR,其中所述启动子是多核苷酸SEQ ID NO:2,所述内含子是多核苷酸SEQ ID NO:8,所述5'-UTR是多核苷酸SEQ ID NO:12。此外,基因表达盒可包括启动子、内含子、和5'-UTR,其中所述启动子是多核苷酸SEQ ID NO:3,所述内含子是多核苷酸SEQ ID NO:9和/或SEQ ID NO:10,所述5'-UTR是多核苷酸SEQ ID NO:

13.另外,基因表达盒可包括启动子、内含子、和5'-UTR,其中所述启动子是多核苷酸SEQ ID NO:35,所述内含子是多核苷酸SEQ ID NO:37,所述5'-UTR是多核苷酸SEQ ID NO:38。

[0197] 例如,基因表达盒可包括启动子、内含子、5'-UTR、和3'-UTR,其中所述启动子是多核苷酸SEQ ID NO:2,所述内含子是多核苷酸SEQ ID NO:8,所述5'-UTR是多核苷酸SEQ ID NO:12,所述3'-UTR是多核苷酸SEQ ID NO:5。同样,基因表达盒可包括启动子、内含子、5'-UTR、和3'-UTR,其中所述启动子是多核苷酸SEQ ID NO:2,所述内含子是多核苷酸SEQ ID NO:8,所述5'-UTR是多核苷酸SEQ ID NO:12,所述3'-UTR是多核苷酸SEQ ID NO:5。此外,基因表达盒可包括启动子、内含子、5'-UTR、和3'-UTR,其中所述启动子是多核苷酸SEQ ID NO:3,所述内含子是多核苷酸SEQ ID NO:9和/或SEQ ID NO:10,所述5'-UTR是多核苷酸SEQ ID NO:13或14,所述3'-UTR是多核苷酸SEQ ID NO:6。另外,基因表达盒可包括启动子、内含子、5'-UTR、和3'-UTR,其中所述启动子是多核苷酸SEQ ID NO:35,所述内含子是多核苷酸SEQ ID NO:37,所述5'-UTR是多核苷酸SEQ ID NO:38,所述3'-UTR是多核苷酸SEQ ID NO:36。

[0198] 另外,基因表达盒可包括启动子、和3'-UTR,其中所述启动子是多核苷酸SEQ ID NO:2,且所述3'-UTR是SEQ ID NO:6。在一个实施方案中,基因表达盒可包括启动子和3'-UTR,其中所述启动子是多核苷酸SEQ ID NO:2而3'-UTR是SEQ ID NO:5。在一个实施方案中,基因表达盒可包括启动子和3'-UTR,其中所述启动子是多核苷酸SEQ ID NO:2,且3'-UTR是SEQ ID NO:5。在一个实施方案中,基因表达盒可包括启动子和3'-UTR,其中所述启动子是多核苷酸SEQ ID NO:35,且3'-UTR是SEQ ID NO:36。

[0199] 在一个实施方案中,基因表达盒包含与非泛素转基因可操作连接的泛素启动子、泛素5'-UTR、和泛素3'-UTR。在一个实施方案中,基因表达盒包含与非泛素转基因可操作连接的泛素启动子、泛素内含子、泛素5'-UTR、和泛素3'-UTR。

[0200] 当基因表达盒包含一个或多个转基因时,启动子、内含子、5'-UTR和3'-UTR可与基因表达盒之内的不同转基因可操作连接。在说明性实施方案中,基因表达盒包含与转基因可操作连接的泛素启动子,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。在说明性实施方案中,基因表达盒包含与转基因可操作连接的泛素启动子、内含子、和5'-UTR,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。在说明性实施方案中,基因表达盒包含与转基因可操作连接的泛素3'-UTR,其中所述转基因编码可提高杀虫剂抗性、除草剂耐受性、氮利用效率、水分利用效率、营养品质或它们的组合的基因产物。

[0201] 泛素内含子和5'-UTR可与基因表达盒内的不同启动子可操作连接。在说明性实施方案中,启动子来源于植物(例如,玉米泛素1启动子)、病毒(例如,木薯叶脉花叶病毒启动子)或细菌(例如,根癌土壤杆菌 δmas)。在说明性实施方案中,基因表达盒包含与转基因可操作连接的泛素启动子,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。

[0202] 在一个实施方案中,载体包含如本文中公开的基因表达盒。在一个实施方案中,载

体可以是质粒、粘粒、细菌人工染色体(BAC)、噬菌体、病毒、或在直接转化或基因打靶中使用的切割的多核苷酸片段,如供体DNA。

[0203] 根据一个实施方案,提供了包含重组基因盒的核酸载体,其中所述重组基因盒包含与多接头序列、非泛素转基因、或它们的组合可操作连接的基于泛素的启动子。在一个实施方案中,所述重组基因盒包含与非泛素转基因可操作连接的基于泛素的启动子。在一个实施方案中,所述重组基因盒包含与多接头序列可操作连接的如本文中公开的基于泛素的启动子。所述多接头以一定的方式与基于泛素的启动子可操作连接,使得编码序列插入到所述多接头的限制性位点之一中后会可操作地连接上,从而允许在载体转染到宿主细胞中时表达该编码序列。

[0204] 根据一个实施方案,基于泛素的启动子包含SEQ ID NO:2或与SEQ ID NO:2具有90%、95%或99%序列同一性的序列。根据一个实施方案,所述启动子序列的总长度不超过1.5、2、2.5、3或4kb。根据一个实施方案,基于泛素的启动子由SEQ ID NO:2或与SEQ ID NO:2具有90%、95%或99%序列同一性的636bp序列组成。

[0205] 根据一个实施方案,提供了包含基因盒的核酸载体,所述基因盒的组成为SEQ ID NO:16、非泛素转基因、和3'-UTR,其中SEQ ID NO:16与非泛素转基因的5'端可操作连接,而3'-UTR与非泛素转基因的3'端可操作连接。在另一个实施方案中,3'非翻译序列包含SEQ ID NO:5或与SEQ ID NO:5具有90%、95%、99%或100%序列同一性的序列。根据一个实施方案,提供了包含基因盒的核酸载体,所述基因盒的组成为SEQ ID NO:16或与SEQ ID NO:16具有90%、95%、或99%序列同一性的1563bp序列、非泛素转基因和3'-UTR,其中SEQ ID NO:16与非泛素转基因的5'端可操作连接,而3'-UTR与非泛素转基因的3'端可操作连接。在另一个实施方案中,3'非翻译序列包含SEQ ID NO:5或与SEQ ID NO:5具有90%、95%、99%或100%序列同一性的序列。在另一个实施方案中,3'非翻译序列由SEQ ID NO:5、或与SEQ ID NO:5具有90%、95%、或99%序列同一性的1020bp序列组成。

[0206] 根据一个实施方案,所述核酸载体进一步包含编码选择标志物的序列。根据一个实施方案,所述重组基因盒与土壤杆菌T-DNA边界可操作连接。根据一个实施方案,所述重组基因盒进一步包含第一和第二T-DNA边界,其中第一T-DNA边界与基因构建体的一端可操作连接,且所述第二T-DNA边界与基因构建体的另一端可操作连接。所述第一和第二土壤杆菌T-DNA边界可独立地选自来源于细菌菌株的T-DNA边界序列,所述T-DNA边界序列选自合成胭脂碱的土壤杆菌T-DNA边界、合成章鱼碱(ocotopine)的土壤杆菌T-DNA边界、合成琥珀碱的土壤杆菌T-DNA边界、或它们的任何组合。在一个实施方案中,提供了土壤杆菌菌株,其选自胭脂碱合成菌株、甘露碱合成菌株、琥珀碱合成菌株、或章鱼碱合成菌株,其中所述菌株包含质粒,其中所述质粒包含与选自SEQ ID NO:2、SEQ ID NO:16或与SEQ ID NO:2或SEQ ID NO:16具有90%、95%、或99%序列同一性的序列可操作连接的转基因。

[0207] 感兴趣的且适合用于当前公开的构建体中的转基因包括但不限于,赋予下列各项的编码序列:(1)对害虫或疾病的抗性,(2)对除草剂的抗性,和(3)增值性状,如公开于W02013116700(DGT-28)、US20110107455(DSM-2)、美国专利8,283,522(AAD-12);7,838,733(AAD-1);5,188,960;5,691,308;6,096,708;和6,573,240(Cry1F);美国专利6,114,138;5,710,020;和6,251,656(Cry1Ac);美国专利6,127,180;6,624,145和6,340,593(Cry34Ab1);美国专利6,083,499;6,548,291和6,340,593(Cry35Ab1)中,其公开内容并入本文。根据一

个实施方案,所述转基因编码选择标志物或赋予杀虫剂抗性、除草剂耐受性、氮利用效率、水分利用效率、或营养品质的基因产物。

[0208] 根据一个实施方案,提供了包含基因盒的核酸载体,其中所述基因盒包含与转基因的5'端可操作连接的启动子区,其中转基因的3'端与3'非翻译区连接。在一个实施方案中,所述启动子区包含SEQ ID NO:2或与SEQ ID NO:2具有90%、95%、或99%序列同一性的序列。根据一个实施方案,所述启动子区由SEQ ID NO:2或SEQ ID NO:16组成。在一个实施方案,3'非翻译序列包含SEQ ID NO:5或与SEQ ID NO:5具有90%、95%、99%或100%序列同一性的序列,并且在一个实施方案中,3'非翻译序列由SEQ ID NO:5、或与SEQ ID NO:5具有90%、95%、或99%序列同一性的1020bp序列组成。

[0209] 根据一个实施方案,提供了包含基因盒的核酸载体,其中所述基因盒包含与5'非翻译区的5'端可操作连接的启动子区,其中5'非翻译区的3'端与转基因的5'端可操作连接,其中转基因的3'端与3'非翻译区连接。在一个实施方案中,所述启动子区包含或为SEQ ID NO:2或与SEQ ID NO:2具有90%、95%或99%序列同一性的序列。在一个实施方案中,所述启动子区由SEQ ID NO:2或与SEQ ID NO:2具有90%、95%或99%序列同一性的636bp序列组成。根据一个实施方案,所述5'非翻译序列包含或为SEQ ID NO:38或与SEQ ID NO:12具有90%序列同一性的序列。根据一个实施方案,所述5'非翻译序列由SEQ ID NO:12或与SEQ ID NO:12具有90%序列同一性的75bp序列组成。在一个实施方案中,所述3'非翻译序列包含或为SEQ ID NO:5或与SEQ ID NO:5具有90%、95%或99%序列同一性的序列。在一个实施方案中,所述3'非翻译序列由SEQ ID NO:5或与SEQ ID NO:5具有90%、95%或99%序列同一性的1020bp序列组成。在另一个实施方案中,所述核酸载体进一步包含插入在所述5'非翻译区与所述转基因之间、并且与所述启动子和所述转基因可操作连接的泛素内含子。在一个实施方案中,所述泛素内含子包含或为SEQ ID NO:8或与SEQ ID NO:8具有90%、95%或99%序列同一性的序列。在一个实施方案中,所述泛素内含子由SEQ ID NO:8或与SEQ ID NO:8具有90%、95%或99%序列同一性的852bp序列组成。

[0210] 根据一个实施方案,提供了包含基因盒的核酸载体,其中所述基因盒包含与转基因的5'端可操作连接的启动子区,其中转基因的3'端与3'非翻译区连接。在一个实施方案中,所述启动子区包含SEQ ID NO:40或与SEQ ID NO:40具有90%、95%或99%序列同一性的序列。

[0211]

```
GGCGTCAGGACTGGCGAAGTCTGGACTCTGCAGGGCCGAAGTCTGAAGACGAAGCAGAGGAAGAGAAAGGGAAGTG
TTCGACTTGTAATTTGTAGGGGTTTTTTTTAGAGGAACTTGTAATTTGTAGGTGGGCTGGCCTCGTTGGAAAAACGA
TGCTGGCTGGTTGGGCTGGGCCGATGTACGCTTGCAAACAACCTGTGGCGGCCCGTTCTGGACGAGCAGGAGTTTCT
TTTTGTTCTCACTTTTCTGGTCTTCTTTAGTTACGGAGTACCTTTTGTTTTTTAAAGGAGTTACCTTTTTTTTTAGG
AATTCTTTAGTTACCTTTTCGCTTGCTCTCAAAAAATATTTAACTTTCGCTTTTTTTCATTTTAATTTTGCACATAT
TTACGAGTTTCATGAATGCTTATTTTCCAGCATATCATTATTTGCAAGTATTTTTATGCCGTATGTATTGGACGAGA
GCCATCGGGACTGTTCCAGAGACTGCGTGGTGCGGACGGCTCCCAACCGCCTTTTCTATCTCTGTTCGCATCCGGTG
GCCGACTTGGCTCGCGCGTGAGCCGTGACGTAACAGACTTGGTCTCTTCCCCATCTGGCCATCTATAAATTCCCCCA
TCGATCGACCCTCCCTTTCCCCAATCCAGCACCCCGATCCCGATCGAAAATTCTCCGCAACAGCAAGCGATCGATC
TAGCGAATCCCCGTCAAG(SEQ ID NO:40)
```

[0212] 根据一个实施方案,所述启动子区由SEQ ID NO:40或与SEQ ID NO:40具有90%、95%或99%序列同一性的711bp序列组成。根据一个实施方案,所述启动子区由SEQ ID NO:40组成。在一个实施方案中,所述3'非翻译序列的由SEQ ID NO:5或与SEQ ID NO:5具有90%、95%或99%序列同一性的1020bp序列组成,在一个实施方案中,所述3'非翻译序列由SEQ ID NO:5组成。

[0213] 在一个实施方案中,提供了包含如本文中公开的基因表达盒的细胞或植物。在一个实施方案中,细胞或植物包含含有如本文中公开的基因表达盒的载体。在一个实施方案中,载体可以是质粒、粘粒、细菌人工染色体(BAC)、噬菌体、或病毒。由此,包含如本文中公开的基因表达盒的细胞或植物分别是转基因细胞或转基因植物。在一个实施方案中,转基因植物可以是单子叶植物。在一个实施方案中,转基因单子叶植物可以是,但不限于,玉米、小麦、水稻、高粱、燕麦、黑麦、香蕉、甘蔗、和粟。在一个实施方案中,转基因植物可以是双子叶植物。在一个实施方案中,转基因双子叶植物可以是,但不限于,大豆、棉花、向日葵、和芸苔。一个实施方案还包括来自如本文中公开的转基因植物的转基因种子。

[0214] 在一个实施方案中,基因表达盒包含两个或更多个转基因。所述两个或更多个转基因可能不与相同的启动子、内含子、或如本文中公开的5'-UTR或3'-UTR可操作连接。在一个实施方案中,基因表达盒包含一个或多个转基因。在具有一个或多个转基因的实施方案中,至少一个转基因与主题公开的启动子、内含子、5'-UTR、或3'-UTR可操作连接。

[0215] 选择标志物

[0216] 可以将各种选择标志物(亦称报告基因)纳入选定的表达载体中,以允许鉴定并选择经转化的植物(转化体)。可以用来确认选择标志物在经转化植物中表达的方法有很多,包括例如DNA测序和PCR(聚合酶链式反应)、Southern印迹法、RNA印迹法、用于检测从载体中表达出来的蛋白质(例如介导膦丝菌素抗性的沉淀蛋白质)的免疫学方法,或其他蛋白质的视觉观察,所述其他蛋白质如报告基因编码的 β -葡萄糖醛酸糖苷酶(GUS)、萤光素酶、绿色荧光蛋白(GFP)、黄色荧光蛋白(YFP)、DsRed、 β -半乳糖苷酶、氯霉素乙酰转移酶(CAT)、碱性磷酸酶,等(参阅Sambrook等,Molecular Cloning:ALaboratory Manual,Third Edition,Cold Spring Harbor Press,N.Y.,2001,其全部内容通过提述并入本文)。

[0217] 利用选择标志物基因来选择转化的细胞或组织。选择标志物基因包括编码抗生素抗性的基因,例如编码新霉素磷酸转移酶II(NEO)和潮霉素磷酸转移酶(HPT)的那些基因,以及赋予对除草化合物的抗性的基因。除草剂抗性基因通常编码对该除草剂不敏感的修饰的靶蛋白,或编码在植物中在该除草剂起作用之前降解该除草剂或将该除草剂脱毒的酶。例如,已通过使用编码突变体靶酶5-烯醇式丙酮酰莽草酸-3-磷酸合酶(EPSPS)的基因而获得了对草甘膦的抗性。针对EPSPS的基因和突变体是熟知的,并且在下文进一步描述。已经通过使用编码使对应的除草剂脱毒的pat或DSM-2、腈水解酶、aad-1、或aad-12基因的细菌基因获得了对草铵膦、溴苯腈、以及2,4-二氯苯氧乙酸(2,4-D)的抗性。

[0218] 在一个实施方案中,除草剂,包括包括咪唑啉酮或磺酰脲,可以抑制生长点或分生组织,并且针对这些除草剂的抗性/耐受性的基因,乙酰羟酸合酶(AHAS)和乙酰乳酸合酶(ALS),是公知的。草甘膦抗性基因包括突变体5-烯醇式丙酮酰莽草酸-3-磷酸合酶(EPSP)和dgt-28基因(分别通过引入天然EPSP基因、aroA基因和草甘膦乙酰转移酶(GAT)基因的重组核酸和/或各种形式的体内诱变)。其他膦酰基化合物的抗性基因包括来自链霉菌属物种

(包括吸水链霉菌和绿色产色链霉菌)的bar基因、以及吡啶氧基或苯氧基丙酸和环己酮(ACC酶抑制剂编码基因)。赋予对环己二酮和/或芳氧基苯氧基丙酸(包括氟吡甲禾灵、禾草灵、精恶唑禾草灵、吡氟禾草灵、喹禾灵)抗性的示例性基因包括乙酰辅酶A羧化酶(ACC酶)基因,--Acc1-S1、Acc1-S2和Acc1-S3。在一个实施方案中,除草剂可以抑制光合作用,包括三嗪(psbA和1s+基因)或苄腈(腈水解酶基因)。

[0219] 在一个实施方案中,选择标志物基因包括但不限于,编码下列物质的基因:新霉素磷酸转移酶II;氰酰胺水合酶;天冬氨酸激酶;二氢吡啶二羧酸合成酶;色氨酸脱羧酶;二氢吡啶二羧酸合酶和脱敏型天冬氨酸激酶;bar基因;色氨酸脱羧酶;新霉素磷酸转移酶(NEO);潮霉素磷酸转移酶(HPT或HYG);二氢叶酸还原酶(DHFR);磷丝菌素乙酰转移酶;2,2-二氯丙酸脱卤酶;乙酰羟酸合酶;5-烯醇丙酮酰-莽草酸-磷酸合酶(aroA);卤代芳基腈水解酶;乙酰辅酶A羧化酶;二氢蝶酸合酶(sul I);和32kD光合体系II多肽(psbA)。

[0220] 一个实施方案还包括编码对下列物质的抗性的基因:氯霉素;甲氨蝶呤;潮霉素;大观霉素;溴苯腈;草甘膦;和磷丝菌素。

[0221] 以上列出的选择标志物基因并非意味着限制。本发明涵盖任何报告基因或选择标志物基因。

[0222] 合成选择标志物基因,以实现在植物中的最优表达。例如,在一个实施方案中,基因的编码序列已通过密码子优化进行修饰,以增强在植物中的表达。可以为了在特定的植物物种中表达而优化选择标志物基因,或者可以为了在双子叶或单子叶植物中最优表达而修饰选择标志物基因。可以从特定植物物种中表达量最大的蛋白质中频率最高的密码子确定出植物优选的密码子。在一个实施方案中,选择标志物基因被设计为以较高水平在植物中表达,从而导致较高的转化效率。用于基因的植物优化的方法是熟知的。关于合成的DNA序列的优化和生产的指导可例如在W02013016546、W02011146524、W01997013402、美国专利6166302、美国专利5380831中找到,将它们通过提述并入本文。

[0223] 转化

[0224] 用于植物转化的适合方法包括任何将DNA导入细胞中的方法,例如但不限于:电穿孔(参见,例如,美国专利5,384,253);微粒轰击(参见,例如,美国专利5,015,580,5,550,318,5,538,880,6,160,208,6,399,861、和6,403,865);土壤杆菌介导的转化(参见,例如,美国专利5,635,055,5,824,877,5,591,616;5,981,840、和6,384,301);和原生质体转化(参见,例如,美国专利5,508,184)。

[0225] 可以使用诸如利用碳化硅纤维搅动的技术(参见,例如,美国专利5,302,523和5,464,765)将DNA构建体直接导入到植物细胞的基因组DNA中,或者可以使用生物射弹法如DNA粒子轰击将DNA构建体直接导入到植物组织中(参见,例如,Klein等人(1987)Nature 327:70-73)。或者,可通过纳米颗粒转化将DNA构建体导入植物细胞中(参见,例如,美国专利公开20090104700,通过提述将其全文并入本文)。

[0226] 此外,基因转移可以使用非土壤杆菌的细菌或病毒,例如根瘤菌菌种NGR234,苜蓿中华根瘤菌、百脉根根瘤菌、马铃薯病毒X、花椰菜花叶病毒和木薯叶脉花叶病毒和/或烟草花叶病毒来实现,参见,例如Chung等人(2006)Trends Plant Sci.11(1):1-4。

[0227] 通过转化技术的应用,几乎任何植物物种的细胞均可稳定转化,并且可通过熟知的技术使这些细胞发育为转基因植物。例如,美国专利5,846,797、5,159,135、5,004,863、

和6,624,344中描述了在棉花转化背景下可能特别有用的技术;例如在美国专利5,750,871中描述了用于转化芸苔属植物的技术;用于转化大豆的技术例如描述于美国专利6,384,301中;用于转化玉米的技术例如描述于美国专利7,060,876和5,591,616、以及国际PCT公开WO 95/06722中。

[0228] 在完成外源核酸到受体细胞的递送之后,通常把转化的细胞鉴定出来用于进一步培养和植物再生。为了提高鉴定转化体的能力,可能期望与产生转化体用的转化载体一道使用选择标志物基因。在说明性实施方案中,可通过使细胞暴露于一种或多种选择剂来测定转化的细胞群,或者可筛选所述细胞的期望标志物基因性状。

[0229] 对于暴露于选择剂后存活的细胞、或者在筛选测定中已被评分为阳性的细胞,可以在支持植物再生的培养基中进行培养。在一个实施方案中,可通过加入其他物质,如生长调节剂来改良任何适合的植物组织培养基。可将组织维持在具有生长调节剂的基本培养基上,直到可得到足以启动植物再生工作的组织为止,或者进行重复多轮的手工选择,直到组织的形态适合于再生为止(例如,至少2周),然后转移到有助于芽形成的培养基中。定期转移培养物,直到已经发生充分的芽形成为止。一旦形成芽,将它们转移到有助于根形成的培养基中。一旦形成足够的根,可将植物转移到土壤中,以便进一步生长和成熟。

[0230] 为了证实提供在再生植物中的包含期望核酸的构建体的存在,可进行多种测定。这样的测定可包括:分子生物学测定,如Southern和northern印迹以及PCR;生物化学测定,如检测蛋白质产物的存在,例如通过免疫学手段(ELISA、western印迹、和/或LC-MS MS分光光度法)或借助于酶功能;植物部分测定,如叶或根测定;和/或再生的全植物的表型分析。

[0231] 可例如通过使用对感兴趣核酸分子特异的寡核苷酸引物进行PCR扩增来筛选转基因事件。PCR基因分型理解为包括但不限于:对来源于分离的宿主植物愈伤组织(预期该愈伤组织含有整合到基因组中的感兴趣核酸分子)的基因组DNA的聚合酶链式反应(PCR)扩增,然后进行PCR扩增产物的标准克隆和序列分析。PCR基因分型方法已有详尽的描述(参见,例如,Rios等(2002)Plant J.32:243-53),并可应用于来源于任何植物物种或组织类型(包括细胞培养物)的基因组DNA。在PCR扩增反应中可按顺序使用、或者复用与靶序列和导入序列两者都结合的寡核苷酸引物的组合。可产生旨在与靶位点、导入的核酸序列、和/或这两者的组合退火的寡核苷酸引物。因此,PCR基因分型策略可包括,例如但不限于:植物基因组中特异性序列的扩增;植物基因组中的多个特异性序列的扩增;植物基因组中的非特异性序列的扩增;以及任何前述项的组合。本领域技术人员可设计其他引物组合和扩增反应来探查基因组。例如,可以设计一组正向和反向寡核苷酸引物,令它们退火到对导入的核酸序列的边界之外的靶标而言为特异性的核酸序列。

[0232] 可设计特异性地退火到导入的核酸分子的正向和反向寡核苷酸引物,例如,退火到导入的核酸分子中包含的某个感兴趣核苷酸序列内的编码区的相应序列;或所述核酸分子的其他部分。引物可以结合本文中描述的引物使用。寡核苷酸引物可根据期望的序列来合成并且可商购(例如,来自Integrated DNA Technologies, Inc., Coralville, IA)。扩增之后可以进行扩增产物的克隆和测序、或直接序列分析。在一个实施方案中,PCR扩增中使用对基因靶标特异性的寡核苷酸引物。

[0233] 表达转基因的方法

[0234] 在一个实施方案中,在植物中表达至少一个转基因的方法包括种植这样的植物:

该植物包含与至少一个转基因可操作连接的泛素启动子。在一个实施方案中,在植物中表达至少一个转基因的方法包括种植这样的植物:该植物包含与至少一个转基因可操作连接的泛素5'-UTR。在一个实施方案中,在植物中表达至少一个转基因的方法包括种植这样的植物:该植物包含与至少一个转基因可操作连接的泛素内含子。在一个实施方案中,在植物中表达至少一个转基因的方法包括种植这样的植物:该植物包含与至少一个转基因可操作连接的泛素启动子、泛素5'-UTR、和泛素内含子。在一个实施方案中,在植物中表达至少一个转基因的方法包括种植这样的植物:该植物包含与至少一个转基因可操作连接的泛素3'-UTR。在一个实施方案中,在植物组织或植物细胞中表达至少一个转基因的方法包括培养这样的植物组织或植物细胞:所述植物组织或植物细胞包含与至少一个转基因可操作连接的泛素启动子。在一个实施方案中,在植物组织或植物细胞中表达至少一个转基因的方法包括培养这样的植物组织或植物细胞:所述植物组织或植物细胞包含与至少一个转基因可操作连接的泛素5'-UTR。在一个实施方案中,在植物组织或植物细胞中表达至少一个转基因的方法包括培养这样的植物组织或植物细胞:所述植物组织或植物细胞包含与至少一个转基因可操作连接的泛素内含子。在一个实施方案中,在植物组织或植物细胞中表达至少一个转基因的方法包括培养这样的植物组织或植物细胞,所述植物组织或植物细胞包含与至少一个转基因可操作连接的泛素启动子、泛素5'-UTR、和泛素内含子。在一个实施方案中,在植物组织或植物细胞中表达至少一个转基因的方法包括培养这样的植物组织或植物细胞,所述植物组织或植物细胞包含与至少一个转基因可操作连接的泛素3'-UTR。

[0235] 在一个实施方案中,在植物中表达至少一个转基因的方法包括种植包含含有与至少一个转基因可操作连接的泛素启动子的基因表达盒的植物。在一个实施方案中,泛素启动子由选自SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:35、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、和SEQ ID NO:39的序列或与选自SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:35、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、和SEQ ID NO:39的序列具有90%、95%或99%序列同一性的序列组成。在一个实施方案中,在植物中表达至少一个转基因的方法包括种植包含含有与至少一个转基因可操作连接的泛素内含子的基因表达盒的植物。在一个实施方案中,在植物中表达至少一个转基因的方法包括种植包含含有与至少一个转基因可操作连接的泛素5'-UTR的基因表达盒的植物。在一个实施方案中,在植物中表达至少一个转基因的方法包括种植包含含有与至少一个转基因可操作连接的泛素启动子、泛素5'-UTR、和泛素内含子的基因表达盒的植物。在一个实施方案中,在植物中表达至少一个转基因的方法包括种植包含含有与至少一个转基因可操作连接的泛素3'-UTR的基因表达盒的植物。在一个实施方案中,在植物组织或植物细胞中表达至少一个转基因的方法包括培养包含含有与至少一个转基因可操作连接的泛素启动子的基因表达盒的植物组织或植物细胞。在一个实施方案中,在植物组织或植物细胞中表达至少一个转基因的方法包括培养包含含有与至少一个转基因可操作连接的泛素内含子的基因表达盒的植物组织或植物细胞。在一个实施方案中,在植物组织或植物细胞中表达至少一个转基因的方法包括培养包含含有与至少一个转基因可操作连接的泛素5'-UTR的基因表达盒的植物组织或植物细胞。在一个实施方案中,在植物组织或植物细胞中表达至少一个转基因的方法包括培养包含含有与至少一个转基因可操作连接的泛素启动子、泛素5'-UTR、和泛素内含子的基因表达盒的植物组织或植物细胞。在一个实施方案中,在植物组织或植物细胞中

表达至少一个转基因的方法包括培养包含含有与至少一个转基因可操作连接的泛素3'-UTR的基因表达盒的植物组织或植物细胞。

[0236] 转基因植物

[0237] 在一个实施方案中,植物、植物组织、或植物细胞包含泛素启动子。在一个实施方案中,泛素启动子可以是柳枝稷、二穗短柄草或粟泛素启动子。在一个实施方案中,植物、植物组织、或植物细胞包含含有启动子的基因表达盒,其中所述启动子与SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3或SEQ ID NO:35具有至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同,其中所述启动子与非泛素转基因可操作连接。在一个实施方案中,植物、植物组织、或植物细胞包含含有选自SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:35、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、和SEQ ID NO:39的序列或与选自SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、SEQ ID NO:35、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:17、和SEQ ID NO:39的序列具有90、95或99.5序列同一性的序列的基因表达盒。在说明性实施方案中,植物、植物组织、或植物细胞包含含有与转基因可操作连接的泛素启动子的基因表达盒,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。

[0238] 在一个实施方案中,植物、植物组织、或植物细胞包含含有3'-UTR的基因表达盒。在一个实施方案中,植物、植物组织、或植物细胞包含含有泛素3'-UTR的基因表达盒。在一个实施方案中,所述泛素3'-UTR是柳枝稷、二穗短柄草或粟泛素3'-UTR。在一个实施方案中,3'-UTR可以是二穗短柄草泛素1C(Ubi1C)3'-UTR、二穗短柄草泛素1 3'-UTR、或粟泛素3'-UTR。

[0239] 在一个实施方案中,植物、植物组织、或植物细胞包含含有内含子的基因表达盒,其中所述内含子与SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、或SEQ ID NO:37具有至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同。在一个实施方案中,基因表达盒包含与启动子可操作连接的泛素内含子,其中所述启动子为柳枝稷、二穗短柄草或粟泛素启动子,或为源自植物(例如,玉米泛素1启动子)、病毒(例如,木薯叶脉花叶病毒启动子)或细菌(例如,根癌土壤杆菌 δ mas)的启动子。在一个实施方案中,植物、植物组织、或植物细胞包含与转基因可操作连接的泛素内含子的基因表达盒。在说明性实施方案中,植物、植物组织、或植物细胞包含含有与转基因可操作连接的泛素内含子的基因表达盒,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。

[0240] 在一个实施方案中,植物、植物组织、或植物细胞包含含有5'-UTR的基因表达盒,其中所述5'-UTR与SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14或SEQ ID NO:38至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同。在一个实施方案中,基因表达盒包含与启动子可操作连接的泛素内含子,其中所述启动子为柳枝稷、二穗短柄草或粟泛素启动子,或为源自植物(例如,玉米泛素1启动子)、病毒(例如,木薯叶脉花叶病毒启动子)或细菌(例如,根癌土壤杆菌 δ mas)的启动子。在一个实施方案中,植物、植物组织、或植物细胞包含与转基因可操作连接的泛素5'-

UTR的基因表达盒。在说明性实施方案中,植物、植物组织、或植物细胞包含含有与转基因可操作连接的泛素5'-UTR的基因表达盒,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。

[0241] 在一个实施方案中,植物、植物组织、或植物细胞包含含有泛素启动子和泛素3'-UTR的基因表达盒。在一个实施方案中,植物、植物组织、或植物细胞包含泛素启动子和3'-UTR,各自独立地可以是柳枝稷、二穗短柄草或粟泛素启动子以及柳枝稷、二穗短柄草或粟泛素3'-UTR。在一个实施方案中,植物、植物组织、或植物细胞包含基因表达盒,所述基因表达盒包含:a)启动子,其中所述启动子SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3或SEQ ID NO:35至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同;b)3'-UTR,其中所述3'-UTR与SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、或SEQ ID NO:36至少80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.5%、99.8%、或100%相同。

[0242] 在一个实施方案中,植物、植物组织、或植物细胞包含含有与转基因可操作连接的泛素启动子、泛素5'-UTR、泛素内含子、和泛素3'-UTR的基因表达盒。当基因表达盒包含两个或更多个转基因时,所述启动子、内含子、5'-UTR、和3'-UTR可与基因表达盒之内的不同转基因可操作连接。在说明性实施方案中,基因表达盒包含与转基因可操作连接的泛素启动子,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。在说明性实施方案中,基因表达盒包含与转基因可操作连接的泛素内含子,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。在一个实施方案中,基因表达盒包含与启动子可操作连接的泛素内含子,其中所述启动子为柳枝稷、二穗短柄草或粟泛素启动子,或为源自植物(例如,玉米泛素1启动子)、病毒(例如,木薯叶脉花叶病毒启动子)或细菌(例如,根癌土壤杆菌 δ mas)的启动子。在说明性实施方案中,基因表达盒包含与转基因可操作连接的泛素5'-UTR,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。在一个实施方案中,基因表达盒包含与启动子可操作连接的泛素5'-UTR,其中所述启动子为柳枝稷、二穗短柄草或粟泛素启动子,或为源自植物(例如,玉米泛素1启动子)、病毒(例如,木薯叶脉花叶病毒启动子)或细菌(例如,根癌土壤杆菌 δ mas)的启动子。在说明性实施方案中,基因表达盒包含与转基因可操作连接的泛素3'-UTR,其中所述转基因可以是杀虫剂抗性转基因、除草剂耐受性转基因、氮利用效率转基因、水分利用效率转基因、营养品质转基因、DNA结合转基因、选择标志物转基因、或它们的组合。

[0243] 在一个实施方案中,植物、植物组织、或植物细胞包含含有如本文中公开的泛素启动子、5'-UTR、内含子、和/或3'-UTR的载体。在一个实施方案中,植物、植物组织、或植物细胞包含含有与非泛素转基因可操作连接的如本文中公开的泛素启动子、5'-UTR、内含子、和/或3'-UTR的载体。在一个实施方案中,植物、植物组织、或植物细胞包含含有如本文中公开的基因表达盒的载体。在一个实施方案中,载体可以是质粒、粘粒、细菌人工染色体

(BAC)、噬菌体、或病毒。

[0244] 根据一个实施方案,提供了植物、植物组织、或植物细胞,其中所述植物、植物组织、或植物细胞包含与转基因可操作连接的非内源泛素来源的启动子序列,其中所述泛素来源的启动子序列包含序列SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、或SEQ ID NO:35、或与SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:3、或SEQ ID NO:35具有90%、95%、98%或99%序列同一性的序列。在一个实施方案中,提供了植物、植物组织、或植物细胞,其中所述植物、植物组织、或植物细胞包含与非泛素转基因可操作连接的SEQ ID NO:2或与SEQ ID NO:2具有90%序列同一性的序列。在一个实施方案中,所述植物、植物组织、或植物细胞为双子叶或单子叶植物或来源于双子叶或单子叶植物的细胞或组织。在一个实施方案中,所述植物选自玉米、小麦、水稻、高粱、燕麦、黑麦、香蕉、甘蔗、大豆、棉花、向日葵、和芸苔。在一个实施方案中,所述植物为玉米。根据一个实施方案,所述植物、植物组织、或植物细胞包含与非泛素转基因可操作连接的SEQ ID NO:2、SEQ ID NO:16、或与SEQ ID NO:2或SEQ ID NO:16具有90%、95%、98%或99%序列同一性的序列。在一个实施方案中,所述植物、植物组织、或植物细胞包含与转基因可操作连接的启动子,其中所述启动子由SEQ ID NO:2、SEQ ID NO:16、或与SEQ ID NO:2或SEQ ID NO:16具有90%、95%、98%或99%序列同一性的序列组成。根据一个实施方案,包含与转基因可操作连接的非内源泛素来源的启动子序列的基因构建体被纳入植物、植物组织、或植物细胞的基因组中。

[0245] 在一个实施方案中,提供了包含与转基因可操作连接的SEQ ID NO:2、或与SEQ ID NO:2具有90%、95%、98%或99%序列同一性的序列的非短柄草属植物、植物组织、或植物细胞。根据一个实施方案,所述非短柄草属植物、植物组织、或植物细胞为双子叶或单子叶植物,或来源于双子叶或单子叶植物的植物细胞或组织。在一个实施方案中,所述植物选自玉米、小麦、水稻、高粱、燕麦、黑麦、香蕉、甘蔗、大豆、棉花、向日葵、和芸苔。在一个实施方案中,所述植物为玉米。根据一个实施方案,与转基因可操作连接的启动子序列被纳入植物、植物组织、或植物细胞的基因组中。在一个实施方案中,所述植物、植物组织、或植物细胞进一步包含5'非翻译序列,所述5'非翻译序列包含SEQ ID NO:12或与SEQ ID NO:12具有90%序列同一性的序列,其中所述5'非翻译序列插入在所述启动子和所述转基因之间并且与它们可操作连接。在另一个实施方案中,所述植物、植物组织、或植物细胞进一步包含插入在所述5'非翻译序列之后的内含子序列。在一个实施方案中,所述内含子序列是分离自柳枝稷、二穗短柄草、或栗的泛素基因的内含子序列。在一个实施方案中,所述序列包含或为SEQ ID NO:8。

[0246] 在一个实施方案中,提供了非短柄草属植物、植物组织、或植物细胞,其包含与转基因的5'端可操作连接的SEQ ID NO:2、或与SEQ ID NO:2具有90%、95%、98%或99%序列同一性的序列、以及包含SEQ ID NO:5或与SEQ ID NO:5具有90%序列同一性的序列的3'非翻译序列,其中3'非翻译序列与所述转基因可操作连接。根据一个实施方案,所述非短柄草属植物、植物组织、或植物细胞是双子叶或单子叶植物或来源于双子叶或单子叶植物的植物组织或细胞。在一个实施方案中,所述植物选自玉米、小麦、水稻、高粱、燕麦、黑麦、香蕉、甘蔗、大豆、棉花、向日葵、和芸苔。在一个实施方案中,所述植物为玉米。根据一个实施方案,与转基因可操作连接的启动子序列被纳入植物、植物组织、或植物细胞的基因组中。在一个实施方案中,所述植物、植物组织、或植物细胞进一步包含5'非翻译序列,所述5'非翻

译序列包含SEQ ID NO:12或与SEQ ID NO:12具有90%序列同一性的序列,其中所述5'非翻译序列插入在所述启动子和所述转基因之间并且与它们可操作连接。在另一个实施方案中,所述植物、植物组织、或植物细胞进一步包含插入在所述5'非翻译序列之后的内含子序列。在一个实施方案中,所述内含子序列是分离自柳枝稷、二穗短柄草、或粟的泛素基因的内含子序列。在一个实施方案中,5'非翻译序列由SEQ ID NO:12组成。

[0247] 在一个实施方案中,提供了包含与转基因可操作连接的SEQ ID NO:16或与SEQ ID NO:16具有90%序列同一性的序列的非短柄草属植物、植物组织、或植物细胞。在一个实施方案中,提供了包含与转基因可操作连接的启动子的非短柄草属植物、植物组织、或植物细胞,其中所述启动子由SEQ ID NO:16、或与SEQ ID NO:16具有90%序列同一性的序列组成。在另一个实施方案中,非短柄草属植物、植物组织、或植物细胞进一步包含柳枝稷、二穗短柄草、或粟的泛素基因的3'非翻译序列。在一个实施方案中,所述3'非翻译序列包含或为SEQ ID NO:5或与SEQ ID NO:5具有90%序列同一性的序列,其中所述3'非翻译序列与转基因的3'端可操作连接。

[0248] 在一个实施方案中,根据本文中公开的方法的植物、植物组织、或植物细胞可以是单子叶植物。所述单子叶植物、植物组织、或植物细胞可以是,但不限于玉米、水稻、小麦、甘蔗、大麦、黑麦、高粱、兰花、竹子、香蕉、香蒲、百合、燕麦、洋葱、小米、和黑小麦。

[0249] 在一个实施方案中,根据本文中公开的方法的植物、植物组织、或植物细胞可以是双子叶植物。所述双子叶植物、植物组织、或植物细胞可以是,但不限于,油菜、芸苔、芥菜、埃塞俄比亚芥、大豆、向日葵、和棉花。

[0250] 关于基因修饰植物的生产,用于植物的基因工程的方法在本领域中是熟知的。例如,已经开发了用于植物转化的多种方法,包括双子叶植物以及单子叶植物的生物和物理转化方案(例如,Goto-Fumiyuki等人,Nature Biotech17:282-286(1999);Miki等人,Methods in Plant Molecular Biology and Biotechnology,Glick,B.R.和Thompson,J.E.编辑,CRC Press,Inc.,Boca Raton,pp.67-88(1993))。另外,有用于植物细胞或组织转化和植物再生的载体和体外培养方法,例如,见于Gruber等人,Methods in Plant Molecular Biology and Biotechnology,Glick,B.R.和Thompson,J.E.编辑,CRC Press,Inc.,Boca Raton,pp.89-119(1993)。

[0251] 本领域技术人员将认识到,在外源序列稳定纳入转基因植物中并证实其可操作之后,可通过有性杂交将其导入其他植物中。可以使用许多标准育种技术中的任何一种,这取决于被杂交的物种。

[0252] 对于工程化的植物材料选择或筛选转化DNA上存在的标志物基因所编码的性状,来鉴定和分离转化的植物细胞、愈伤组织、组织、或植物。例如,可以将工程化的植物材料在含有抑制量的抗生素或除草剂(转化基因构建体提供对该抗生素或除草剂的抗性)的培养基上培育来实施选择。此外,还可以通过筛选重组核酸构建体上可能存在的任何可见的标志物基因(例如yfp、gfp、 β 葡萄糖苷酸酶、荧光素酶、B或C1基因)的活性,来鉴定转化的细胞。这样的选择和筛选方法对于本领域技术人员是熟知的。

[0253] 还可以使用物理和生物化学方法来鉴定含有插入的基因构建体的植物或植物细胞转化体。这些方法包括但不限于:1)Southern分析或PCR扩增,用于检测和确定重组DNA插入物的结构;2)Northern印迹、S1RNA酶保护、引物延伸或逆转录酶PCR扩增,用于检测和检

查基因构建体的RNA转录本;3)酶学分析,用于检测酶或核酶活性,在这类基因产物由基因构建体编码的场合;4)新一代测序分析;5)蛋白凝胶电泳、Western印迹技术、免疫沉淀、或酶联免疫吸附测定(ELISA),当基因构建体产物是蛋白质时。其他技术,例如原位杂交、酶染色和免疫染色,也可以用于检测特定植物器官和组织中重组构建体的存在或表达。所有用于完成这些测定法的方法都是本领域技术人员熟知的。

[0254] 使用本文公开的方法进行基因操作的效果可以通过例如对从感兴趣组织分离的RNA(例如,mRNA)的northern印迹来观察。通常,如果mRNA出现或mRNA量增加,可以推定对应的转基因正在表达。可以使用其它测量基因和/或编码的多肽活性的方法。根据所使用的底物、以及检测反应产物或副产物的增加或减少的方法,可以使用不同类型的酶学测定。此外,表达的多肽的水平可以通过免疫化学检测,即ELISA、RIA、EIA和其他本领域技术人员所熟知的基于抗体的测定法,例如电泳检测测定法(可以使用染色或免疫印迹)。作为一个非限制性例子,使用ELISA测定法对AAD-1(芳氧基链烷酸酯双加氧酶;参见WO 2005/107437)和PAT(膦丝菌素乙酰转移酶),EC 2.3.1.183)蛋白的检测记载于美国专利公开20090093366,通过提述将其全文并入本文。转基因可以在植物的某些类型和组织中、或在不同的发育阶段选择性地表达,或者转基因可以在几乎所有植物组织中的整个生命周期内表达。然而,任何组合表达模式也是可应用的。

[0255] 本发明公开还包括如上面所记载的转基因植物的种子,其中种子具有转基因或基因构建体。本发明公开进一步包括如上面所述的转基因植物的后代、克隆、细胞系或细胞,其中所述后代、克隆、细胞系或细胞具有转基因或基因构建体。

[0256] 虽然已经参考特定的方法和实施方案描述了本发明,应当理解的是,可以做出各种不偏离本发明的修改和变化。

[0257] 实施例1

[0258] 根癌土壤杆菌的转化

[0259] 将二元表达载体转化到根癌土壤杆菌菌株DAt13192(RecA阴性三重菌株)(国际专利公开No.WO2012016222)中。分离细菌菌落,分离二元质粒DNA并经由限制酶消化进行证实。

[0260] 玉米转化

[0261] 土壤杆菌培养启动。从甘油储液将土壤杆菌培养物划线接种在土壤杆菌(AB)基本培养基(如公开于WO 2013090734中,其公开内容通过提述并入本文)上并在20℃黑暗中温育3天。然后将土壤杆菌培养物划线接种在YEP(参见WO 2013090734)培养基的平板上并在20℃黑暗中温育1天。

[0262] 在实验当天,制备接种培养基(参见WO 2013090734)与乙酰丁香酮的混合物,混合物的体积与实验中包含植物转化构建体的细菌菌株的数目相适应。将接种培养基吸移到一次性250ml无菌烧瓶中。然后,向含有接种培养基的烧瓶中加入适宜体积的100%二甲亚砜中的1M乙酰丁香酮储液,以产生200μM的乙酰丁香酮终浓度。所需的接种培养基和1M乙酰丁香酮储液的体积在表1中列出。

[0263] 表1:根据制备的构建体数目制备接种培养基/乙酰丁香酮混合物的量

[0264]

制备构建体的数目	接种培养基(mL)	1M 乙酰丁香酮储备溶液(μ L)
1	50	10
2	100	20
3	150	30
4	200	40
5	250	50

[0265] 对于每种构建体,从YEP平板取1-2环土壤杆菌悬浮在一次性50ml无菌离心管内的15ml的接种培养基/乙酰丁香酮混合物中,在分光光度计中在600nm(OD_{600})测量溶液的光密度。然后使用额外的接种培养基/乙酰丁香酮混合物将悬液稀释到0.25-0.35 OD_{600} 。再将土壤杆菌悬液的管水平放置在平台摇床上,平台摇床设置为室温下约75rpm,在使用之前温育1到4小时。

[0266] 穗消毒和胚分离.玉米栽培种B104授粉10-12天后收获穗。将收获的穗去壳,浸没含20%商用漂白剂(UltraClorox[®] Germicidal Bleach,6.15%次氯酸钠)和两滴Tween[®] 20的溶液中持续20分钟进行表面消毒,随后在层流罩内在无菌去离子水中冲洗三次。从每个穗无菌地切下未成熟合子胚(1.8-2.2mm长)并分配到一个或多个含有2.0ml的土壤杆菌悬液的微量离心管中,土壤杆菌悬液中已添加了2 μ l的10% Break-Thru[®] S233表面活性剂。

[0267] 土壤杆菌共培养.在完成胚分离动作后,将装有胚的管封闭,放置在摇摆式平台上持续5分钟。然后将管的内容物倾倒入共培养培养基平板上,用一次性无菌移液管移除液态的土壤杆菌悬浮物。将含有胚的共培养平板放置在层流罩的后部,盖子半开,历时30分钟;此后使用显微镜使胚盾片朝上定向。然后使带有胚的共培养平板放回到层流罩的后部,盖子半开,再经过15分钟。然后盖上平板,用3M[®] Micropore[®] 胶带密封,放置在25 $^{\circ}$ C、每天24小时光照、光强度大约60 μ mol $m^{-2}s^{-1}$ 的培养箱中。

[0268] 愈伤组织选择和转基因事件的再生.在共培养阶段后,将胚转移到静息培养基(参见W0 2013090734)上。将不超过36个胚移动到每个平板上。将平板放置在透明的箱中,在27 $^{\circ}$ C下、每天24小时大约50 μ mol $m^{-2}s^{-1}$ 光强度的光照下温育7-10天。然后将愈伤胚转移到选择I培养基(参见W0 2013090734)上。移到每个选择I平板上的愈伤胚不超过18个。将平板放置在透明的箱中,在27 $^{\circ}$ C、每天24小时大约50 μ mol $m^{-2}s^{-1}$ 光强度的光照下温育7天。然后将愈伤胚转移到选择II培养基(参见W0 2013090734)上。移到每个选择II培养基平板上的愈伤胚不超过12个。将平板放置在透明的箱中,在27 $^{\circ}$ C、每天24小时大约50 μ mol $m^{-2}s^{-1}$ 光强度的光照下温育14天。

[0269] 在这个阶段,将抗性愈伤组织移动到预再生培养基(参见W0 2013090734)上。移动到每个预再生培养基平板上的愈伤组织不超过9个。将平板放置在透明的箱中,在27 $^{\circ}$ C下、每天24小时大约50 μ mol $m^{-2}s^{-1}$ 光强度的光照下温育7天。然后将再生愈伤组织转移到Phytatrays[™]中的再生培养基上(参见W0 2013090734)。在28 $^{\circ}$ C、每天16小时光照/8小时黑

暗、大约 $150\mu\text{mol m}^{-2}\text{s}^{-1}$ 光强度的光照下温育7-14天,直到发出芽时为止。每个Phytatray™中放置的愈伤组织不超过5个。然后分离具有主根的小芽并转移到芽伸长培养基(参见WO 2013090734)上。将约6cm或更高的生根的小植株移栽到土壤中,并移出到生长室炼苗。

[0270] YFP瞬时表达.在转化的胚中,以及在与土壤杆菌共培养3天之后观察到瞬时YFP表达。使用YFP滤光片和500nm光源在立体显微镜(Leica Microsystems, Buffalo Grove, IL)下观察胚。

[0271] T₀植物在温室中的转移和建立.定期将转基因植物转移到温室。将植物从Phytatrays™移栽到充填有生长培养基(Premier Tech Horticulture, ProMix BX, 0581P)的小盆(T.O. Plastics, 3.5" SVD, 700022C)中,用湿度控制圆顶(humidome)覆盖,以帮助驯化植物。将植物放置在Conviron生长室中(28°C/24°C, 16小时光周期, 50-70% RH, $200\mu\text{mol m}^{-2}\text{s}^{-1}$ 光强度),直到达到V3-V4阶段。这有助于使植物适应土壤和更苛刻的温度。然后将植物移动到温室(光暴露类型:光合(Photo)或同化(Assimilation);高光限: $1200\mu\text{mol m}^{-2}\text{s}^{-1}$ 光合活性照射(PAR);16小时昼长;27°C昼/24°C夜),并从小盆移栽到5.5英寸盆中。在移栽到更大的盆中之后约1-2周,对植物取样进行生物测定。对每个事件测定一株植物。

[0272] 实施例2:启动子的鉴定

[0273] 使用二穗短柄草和粟作为靶基因组,在Phytozome(Goodstein et al., 2012)数据库中对玉米泛素编码序列进行BLASTx搜索。

[0274] 玉米泛素(ZM Ubi1)编码序列

[0275]

ATGCAGATCTTTGTGAAAACCTGACTGGCAAGACTATCACCCCTCGAGGTGGAGTCGTCTGACACCATTGACAACGT
TAAGGCCAAGATCCAGGACAAGGAGGGCATCCCCCAGACCAGCAGCGGCTCATCTTTGCTGGCAAACAGCTTGAGG
ACGGGCGCACGCTTGCTGACTACAACATCCAGAAGGAGAGCACCCCTCCACCTTGTGCTCCGTCTCAGGGGAGGCATG
CAGATCTTTGTGAAAACCTGACCGGCAAGACTATCACCCCTCGAGGTGGAGTCCTCTGACACCATTGACAACGTCAA
GGCCAAGATCCAGGACAAGGAGGGCATCCCTCCAGACCAGCAGCGGCTCATCTTTGCTGGGAAGCAGCTTGAGGACG
GGCGCACGCTTGCCGACTACAACATCCAGAAGGAGAGCACCCCTCCACTTGGTGCTGCGCCTCAGGGGAGGCATGCAG
ATCTTCGTGAAGACCCTGACCGGCAAGACTATCACCCCTCGAGGTGGAGTCTTCAGACACCATCGACAACGTCAAGGC
CAAGATCCAGGACAAGGAGGGCATTCCCCCAGACCAGCAGCGGCTCATCTTTGCTGGAAAGCAGCTTGAGGACGGGC
GCACGCTTGCCGACTACAACATCCAGAAGGAGAGCACCCCTCCACTTGGTGCTGCGCCTCAGGGGAGGCATGCAGATC
TTCGTGAAGACCCTGACCGGCAAGACTATCACCCCTCGAGGTGGAGTCTTCAGACACCATCGACAATGTCAAGGCCAA
GATCCAGGACAAGGAGGGCATCCCACCGACCAGCAGCGTTTGATCTTCGCTGGCAAGCAGCTGGAGGATGGCCGCA
CCCTTGCGGATTACAACATCCAGAAGGAGAGCACCCCTCCACCTGGTGCTCCGTCTCAGGGGTGGTATGCAGATCTTT
GTGAAGACACTCACTGGCAAGACAATCACCCCTTGAGGTGGAGTCTTCGGATACCAATTGACAATGTCAAGGCCAAGAT
CCAGGACAAGGAGGGCATCCCACCGACCAGCAGCGCTCATCTTCGCCGGCAAGCAGCTGGAGGATGGCCGCACCC
TGGCGGATTACAACATCCAGAAGGAGAGCACTCTCCACCTGGTGCTCCGCCTCAGGGGTGGCATGCAGATTTTTGTG
AAGACATTGACTGGCAAGACCATCACCTTGAGGTGGAGAGCTCTGACACCATTGACAATGTGAAGGCCAAGATCCA
GGACAAGGAGGGCATTCCCCCAGACCAGCAGCGTCTGATCTTTGCGGGCAAGCAGCTGGAGGATGGCCGCACTCTCG
CGGACTACAACATCCAGAAGGAGAGCACCCCTTCACCTGTTCTCCGCCTCAGAGGTGGTATGCAGATCTTTGTAAAG
ACCCTGACTGGAAAAACCATTAACCTGGAGGTTGAGAGCTCGGACACCATCGACAATGTGAAGGCCAAGATCCAGGA
CAAGGAGGGCATCCCCCGGACCAGCAGCGTCTGATCTTCGCCGGCAAACAGCTGGAGGATGGCCGCACCCCTAGCAG

ACTACAACATCCAAAAGGAGAGCACCTCCACCTTGTGCTCCGTCTCCGTGGTGGTCAGTAA(SEQ ID NO:18)

[0276] 蛋白质比对显示在图1中。从二穗短柄草鉴定出两条与玉米泛素1蛋白对齐的序列。从粟和柳枝稷分别仅鉴定出一条与玉米泛素1蛋白对齐的序列。在预期的翻译起始位点(ATG)上游有一大约2kb的DNA序列,将其确定为推定的启动子序列的起点,并用于表达的表征。将分离自柳枝稷、二穗短柄草和粟的新启动子的多核苷酸序列比对与ZM Ubi1启动子进行比对,发现它们在所述2kb DNA区上享有低水平的序列相似性(图2A-C)。

[0277] 柳枝稷、二穗短柄草和粟泛素基因的UBI编码序列以及推定启动子在图35-38中有示明。

[0278] 实施例3:载体构建

[0279] 商业合成四个启动子序列并将其纳入质粒载体中,这些质粒载体描绘于图3(pDAB113091)、图4(pDAB113092)、图5(pDAB113066)和图22(pDAB118238)中。类似地,商业合成四个3'UTR/转录终止序列并将其纳入质粒载体中,所述质粒载体描绘于图23(pDAB118237)、图24(pDAB118207)、图25(pDAB118208)和图26(pDAB118209)中。所述序列的两端侧翼有用无缝克隆(GeneArt[®] Seamless Cloning and Assembly Kit, Invitrogen, Carlsbad, CA)的15-18个核苷酸的同源片段,以及为了分离启动子片段而插入的II型限制性酶位点。利用PCR或II型限制性酶,获得了无缝克隆相容性的玉米Ubi1启动子(Christensen and Quail(1996)Transgenic Research.5;213-218;Christensen et al., (1992)Plant Molecular Biology.18;675-689)或水稻肌动蛋白启动子(McElroy et al., (1990)Plant Cell.2;163-71)、和包含ST-LS1内含子的PhiYFP(Shagin et al., (2004)Mol Biol Evol.21;841-50)编码序列(Vancanneyt等人, (1990)Mol Gen Genet.220;245-50)、以及St PinII或天然3'-UTR(An et al., 1989Plant Cell.1;115-22.)片段。最后,利用无缝克隆将启动子::PhiYFP::St PinII 3'-UTR片段组装而产生用于瞬时表达测试的瞬时表达载体(图6, pDAB113103;图7, pDAB113104;图8, pDAB113105;图9, pDAB113106;以及,图10, pDAB113107;图27, pDAB120403;图28, pDAB118234, 图29, pDAB118235;和图30, pDAB118236)。将这些瞬时表达载体整合到含有Zm Ubi1启动子和AAD-1编码序列(国际专利公开2005107437)和Zm Lip3'UTR的二元载体中(Paek等人, (1998)Molecules and Cells, 8(3):336-342)。通过限制酶消化和测序反应确证了生成的二元载体(图12, pDAB113117;图13, pDAB113118;图14, pDAB113119;图15, pDAB113120;图16, pDAB113121;图31, pDAB120400;图32, pDAB120404;图33, pDAB120401;以及,图34, pDAB120402)。

[0280] 实施例4:瞬时表达测试

[0281] 利用不成熟玉米(B104)胚的粒子轰击测试了瞬时表达。对于每次处理,在一块培养皿中使用四十个胚供轰击。在粒子轰击的过夜温育之后进行YFP图像分析。图19显示了从新启动子获得的YFP表达水平。数据表明,从新启动子(pDAB113103、pDAB113104、和pDAB113105)获得的YFP表达水平与从ZM Ubi1启动子(pDAB113106)和OS Act1启动子(pDAB113107)获得的YFP表达水平相当,如在显微镜下目测的。在Leica EL6000-mercury metal halide[™]显微镜上将植物组织成像。使用Chroma 42003-ZsYellow 1[™]滤光片捕获共聚焦和微分干涉差(DIC)图像。

[0282] 实施例5:利用实时TaqMan[®] PCR估计转基因拷贝数

[0283] 通过水解探针测定确证了yfp转基因在转基因玉米植物的基因组内的稳定整合。

获得从愈伤组织发育的稳定转化的转基因玉米小植株,并分析以鉴定含有低拷贝数(1-2个拷贝)的全长T链插入物的事件。将鉴定的小植株移动到温室中培育。

[0284] 使用Roche Light Cycler480™系统来确定转基因拷贝数。该方法利用了二重TaqMan®反应,在单一测定中采用对yfp基因和对内源玉米参考基因(转化酶,Genbank登录号:U16123.1)特异的寡核苷酸。通过测量yfp特异性荧光相对于转化酶特异性荧光的相对强度,与已知的拷贝数标准进行比较,确定了拷贝数和接合性。

[0285] 用一个含有FAM™荧光染料标记探针的TaqMan®引物/探针组扩增了yfp基因特异性DNA片段,并且用含有HEX™荧光标记探针的第二引物/探针组扩增了转化酶(表2)。如表3所示制备了PCR反应混合物,并根据表4中所示的条件扩增了基因特异性DNA片段。通过测量报告基因特异性荧光相对于参考基因转化酶特异性荧光的相对强度,并与已知的拷贝数标准比较,确定了样品的拷贝数和接合性。

[0286] 表2:正向和反向核苷酸引物和荧光探针。

[0287]

	引物/探针	序列
PhiYFP v3	正向引物	(SEQ ID NO:28) CGTGTTGGGAAAGAACTTGGA
PhiYFP v3	反向引物	(SEQ ID NO:29) CCGTGGTTGGCTTGGTCT
PhiYFP v3	探针	(SEQ ID NO:30) 5'FAM/ CACTCCCCACTGCCT /MGB_BHQ_1/3'
转化酶	正向引物	(SEQ ID NO:31) TGGCGGACGACGACTTGT
转化酶	反向引物	(SEQ ID NO:32) AAAGTTTGGAGGCTGCCGT
转化酶	探针	(SEQ ID NO:33) 5'HEX/ CGAGCAGACCGCCGTGTACTT /3BHQ_1/3'

[0288] (由Integrated DNA Technologies合成,Coralville,IA)。

[0289] 表3: Taqman®PCR反应混合物。

	成分	工作浓度	终浓度	体积(μ l)
	水	-	-	0.5
	Roche LightCycler 480探针预混液	2X	1X	5
	PhiYFP v3 F	10 μ M	400nM	0.4
	PhiYFP v3 R	10 μ M	400nM	0.4
	PhiYFP v3 探针 -FAM	5 μ M	200nM	0.4
[0290]	转化酶F	10 μ M	400nM	0.4
	转化酶R	10 μ M	400nM	0.4
	转化酶探针 -Hex	5 μ M	200nM	0.4
	聚乙烯吡咯烷酮 (PVP)	10%	0.1%	0.1
	基因组DNA模板	稀释的 BioCel DNA (~5ng/ μ l)	~10ng/ μ L	2
	总反应体积	-	-	10.0

[0291] 表4:用于PCR扩增的热循环仪条件。

	PCR步骤	温度($^{\circ}$ C)	时间	循环数
	步骤1	95	10分钟	1
[0292]	步骤2	95	10秒	40
		58	35秒	
		72	1秒	
	步骤3	40	秒	1

[0293] 如下建立标准:将载体pDAB108706稀释到玉米B104基因组DNA(gDNA)中,获得关于pDAB108706:gDNA关系已知的标准。例如,制备了每一个拷贝玉米B104gDNA具有一个、两个、四个拷贝载体DNA的样品。针对已知为半合的对照玉米事件、以及已知为纯合的对照玉米事件(玉米事件278;参见PCT国际专利公开No.WO 2011/022469A2),验证了pDAB108706与玉米B104gDNA标准品混合的一个和两个拷贝稀释物。实施TaqMan[®]二重测定,该测定使用对AAD1基因特异的寡核苷酸和对内源玉米参考基因转化酶特异的寡核苷酸,用一个含有FAM荧光染料标记的TaqMan[®]引物/探针组扩增并检测AAD1的基因特异性DNA片段,用第二个含有HEX[™]荧光标记探针的TaqMan[®]引物/探针组扩增并检测转化酶的基因特异性DNA片段(表2)。如表3中所示制备了AAD1 TaqMan[®]反应混合物,并根据表4中所示的条件扩增了特异性片段。

[0294] 根据制造商的指导,使用RocheLightCycler[®] 480热循环仪分析了每个反应产生的荧光水平。在465/510nm光密度激发FAM[™]荧光部分,在533/580nm光密度激发HEX[™]荧光部分。通过将未知样品的目标/参考值(由LightCycler[®] 480输出)与四个已知拷贝数的标准品[空、1拷贝(半合)、2拷贝(纯合)和4拷贝]的目标/参考值进行比较,确定了拷贝数。用不同的启动子构建体转化而获得的转基因植物的转基因拷贝数分析的结果显示在表5中。仅将具有1-2个拷贝的yfp转基因的植物转移到温室中用于进一步表达分析。

[0295] 表5:从本文中描述的启动子构建体和对照构建体获得的转基因植物的转基因拷贝数估计。

[0296]

构建体	阳性事件数	1-2个拷贝的yfp
pDAB113117	32	17
pDAB113118	26	13
pDAB113119	30	16
pDAB113120	43	10
pDAB113121	36	19

[0297] 实施例6:与泛素启动子可操作连接的基因的表达

[0298] 蛋白质提取

[0299] 利用叶子的ELISA测定,在V4-5对T₀植物取样。将样品收集在96孔收集管板中,对每个样品取4个叶盘(纸打孔器大小)。向每个管中加入两个4.5mm BBs和200μL提取缓冲液[1x PBS,补充有0.05% Tween[®]-20和0.05% BSA(Millipore Probumin[®], EMD Millipore Corp., Billerica, MA)]。为了AAD1提取,将BSA的浓度增加到0.5%。在KLECO珠磨机中以全速处理平板3分钟。向每个管添加另外200μL提取缓冲液,随后倒置混合。使平板在3000rpm旋转5分钟。上清液转移到保存在冰上的96深孔板中的相应孔中。

[0300] YFP和AAD1ELISA程序

[0301] 使用Nunc[®] 96孔Maxi-Sorp板(Thermo Fisher Scientific Inc., Rockford, IL)进行ELISA。用小鼠单克隆抗YFP捕捉抗体(OriGene Technologies Inc., Rockville, MD)包被平板。该抗体稀释在PBS中(1μg/mL),且每个孔加入150μL稀释的PBS。平板在4℃温育过夜。将过夜平板在室温下保持20-30分钟,然后用350μL洗涤缓冲液[1x PBS,补充有0.05% Tween[®]-20(Sigma-Aldrich, St. Louis, MO)]洗涤4次。在37℃用每孔200μL封闭缓冲液[1x PBS,补充有0.05% Tween[®]-20和0.5% BSA(Millipore Probumin[®])]封闭平板至少1小时,随后用350μL洗涤缓冲液(Tomtec QuadraWash[™]2, Tomtec, Inc., Hamden, CT)洗涤4次。

[0302] 对于YFP ELISA,使用Evrogen重组Phi-YFP 1mg/mL(Axxora LLC, Farmingdale, NY)作为标准。使用5参数拟合标准曲线(在1ng/ml和0.125ng/ml标准之间)来确保所有数据落入曲线的线性部分中。向孔中加入100μL的标准品或样品。使用样品在测定缓冲液中的至少1:4稀释物。将平板在室温下在平板摇床(250rpm; Titer Plate shaker)上温育1小时,随后用350μL洗涤缓冲液(Tomtec QuadraWash[™]2)洗涤4次。向每个孔中加入约100μL的1μg/mL Evrogen兔多克隆抗PhiYFP第一抗体(Axxora)。将平板在室温下在平板摇床上250rpm温育1

小时,随后用350 μ L洗涤缓冲液(Tomtec QuadraWash™2)洗涤4次。然后,向每个孔中加入以1:5000稀释在封闭/测定缓冲液中的100 μ L的抗兔IgG HRP第二抗体(Thermo Scientific)。在室温下在平板摇床上250rpm温育平板1小时,随后用350 μ L洗涤缓冲液(Tomtec QuadraWash™2)洗涤4次。将100 μ L的Pierce 1Step Ultra TMB ELISA(Thermo Scientific)底物加入孔中,轻轻振摇10分钟。通过加入50 μ L的0.4N H₂SO₄停止反应。在450nm处读取吸光度,使用650nm参考滤光片。

[0303] 使用来自Acadia BioSciences(Portland,ME)的试剂盒通过ELISA确定AAD1表达水平。使用提取物的多个稀释度,且使用由供应商提供的试剂和说明进行ELISA。利用可溶性总蛋白测定将蛋白质水平归一化,蛋白质水平归一化使用由Thermo Scientific提供的660nm蛋白测定试剂遵照供应商的说明进行。

[0304] 实施例7:全植物YFP图像分析,例示与泛素启动子可操作连接的基因的稳定表达

[0305] 在温室中培育含有低拷贝数的二元质粒的全植物。在Leica EL6000-mercury metal halide™显微镜上将植物组织成像。使用Chroma 42003-ZsYellow 1™滤光片捕获共聚焦和微分干涉差(DIC)图像。图20到图21分别呈现了用本文中描述的二穗短柄草泛素1C、二穗短柄草泛素1及粟泛素2启动子转化玉米胚获得的转基因T₀玉米植物的愈伤组织和根组织中YFP稳定表达的代表性实例。上述启动子驱动yfp编码序列在愈伤组织(图20)和根(图21)植物组织中强劲表达。

[0306] 实施例8:与泛素启动子可操作连接的基因的全植物T₀稳定表达

[0307] 对表达的YFP蛋白进行ELISA分析产生了额外的数据。ELISA分析进一步证实,新启动子驱动了转基因的强劲表达。图17和表6中显示了从包含新启动子构建体的转基因植物获得的YFP蛋白的定量测量。数据表明,YFP蛋白在含有新启动子(pDAB113117、pDAB113118、和pDAB113119)的植物中的表达比从Os Act1(水稻肌动蛋白1)启动子(pDAB113120)获得的YFP表达高几倍。相比而言,图18和表7显示,从所有构建体获得的AAD1表达水平相似。这是预料之中的,因为所有构建体的AAD1都是由Zm Ubi1启动子驱动的。

[0308] 表6:跨物种泛素启动子T₀叶的YFP表达

[0309]

构建体	平均值(ng/mg TSP)	统计显著性
pDAB113121	144.00173	A
pDAB113118	92.37256	AB
pDAB113119	65.30393	B
pDAB113117	55.24345	B
pDAB113120	12.77181	B

[0310] 不是由相同字母连接的水平是显著不同的。

[0311] 表7:跨物种泛素启动子T₀叶的AAD1表达

[0312]

构建体	平均值(ng/mg TSP)	统计显著性
pDAB113121	119.06932	A
pDAB113118	109.19796	A
pDAB113119	96.29021	A

pDAB113117	85.40412	A
pDAB113120	83.81594	A

[0313] 不是由相同字母连接的水平是显著不同的。

[0314] 实施例9:与泛素启动子和3'UTR可操作连接的基因的全植物T₁稳定表达

[0315] T₀单转基因拷贝植物与野生型B104玉米植物回交,以获得T₁种子。将半合T₁植物用于分析。V4和V12叶表达使用每个构建体五个事件、每个事件5-10株植物。其他组织类型表达使用每个构建体三个事件、每个事件3株植物。对AAD1/YFP进行了接合性分析。

[0316] 从包含新启动子构建体的T₁转基因植物的叶组织获得的YFP蛋白的定量测量显示在表8中。数据证实了T₀叶表达结果,并且进一步表明,在含有新启动子(pDAB113117、pDAB113118、和pDAB113119)的植物的V4、V12和R3叶组织中获得了YFP蛋白的一致高表达。表8还表明,当这种新启动子与它们的天然3'UTR(pDAB120400、pDAB120401、和pDAB120402)而不是PinII3'UTR(pDAB113117、pDAB113118、和pDAB113119)结合使用时,YFP蛋白的表达增加若干倍。从含有构建体pDAB120404的植物中检测到YFP蛋白表达,证实这种构建体中所使用的新启动子和3'UTR驱动了转基因的表达。

表8: 跨物种泛素启动子和3'UTR T₁叶表达

		平均YFP (ng/mg TSP)			
构建体	事件	V4叶	V12叶	R3叶	
[0317]	pDAB113117	pDAB113117[1]-006	44.0	169.4	2108.3
	pDAB113117	pDAB113117[1]-007	44.8	181.4	2582.7
	pDAB113117	pDAB113117[1]-008	79.6	322.8	4096.3
	pDAB113117	pDAB113117[1]-019	74.3	369.1	3420.3
	pDAB113117	pDAB113117[1]-028	34.2	168.2	2164.1
	pDAB113118	pDAB113118[1]-005	33.6	148.8	2094.7
	pDAB113118	pDAB113118[1]-007	54.9	180.1	2171.4
	pDAB113118	pDAB113118[1]-010		138.0	2748.2
	pDAB113118	pDAB113118[1]-023	46.2	156.6	2216.8
	pDAB113118	pDAB113118[1]-025	41.7	132.9	2071.4
	pDAB113119	pDAB113119[1]-001	133.1	436.0	6744.0
	pDAB113119	pDAB113119[1]-005	49.2	138.6	1772.9
	pDAB113119	pDAB113119[1]-011	54.6	133.9	1415.5
	pDAB113119	pDAB113119[1]-013		129.1	1807.7
	pDAB113119	pDAB113119[1]-028	38.5	129.9	1632.8
	pDAB113120	pDAB113120[1]-005	9.8	69.6	493.5
	pDAB113120	pDAB113120[1]-010	24.3	74.5	638.3
	pDAB113120	pDAB113120[1]-014	17.2	79.7	552.4
	pDAB113120	pDAB113120[1]-023	13.2	55.4	372.2
	pDAB113120	pDAB113120[1]-032	12.5	69.6	233.7
	pDAB113121	pDAB113121[1]-008	327.9		
	pDAB113121	pDAB113121[1]-011	166.2	271.2	4472.6
	pDAB113121	pDAB113121[1]-018	128.2	362.0	7116.3
	pDAB113121	pDAB113121[1]-023	112.2	309.1	6813.7
	pDAB113121	pDAB113121[1]-026	118.7	311.7	6300.7
	pDAB120400	pDAB120400[1]-001	640.8182		
	pDAB120400	pDAB120400[1]-002	339.24463		

	pDAB120400	pDAB120400[1]-004	943.96511
	pDAB120400	pDAB120400[1]-007	1653.7402
	pDAB120400	pDAB120400[1]-024	466.01906
	pDAB120401	pDAB120401[1]-001	833.04373
	pDAB120401	pDAB120401[1]-011	471.9103
	pDAB120401	pDAB120401[1]-019	795.08285
	pDAB120401	pDAB120401[1]-022	721.58288
	pDAB120401	pDAB120401[1]-025	696.94286
[0318]	pDAB120402	pDAB120402[1]-010	750.82185
	pDAB120402	pDAB120402[1]-011	619.38603
	pDAB120402	pDAB120402[1]-014	618.98144
	pDAB120402	pDAB120402[1]-030	625.84385
	pDAB120404	pDAB120404[1]-003	44.088479
	pDAB120404	pDAB120404[1]-013	47.464389
	pDAB120404	pDAB120404[1]-014	52.204801
	pDAB120404	pDAB120404[1]-016	45.397854
	pDAB120404	pDAB120404[1]-020	46.913279

[0319] 从含有驱动YFP的新泛素启动子的转基因玉米植物取样的不同组织类型(包括穗轴、壳、粒、花粉、根、须和茎)中发现了高的YFP蛋白表达(表9)。这些数据证明,本申请请求保护的新启动子和3'UTR可驱动转基因在植物中组成性高表达,并且在生物技术应用中是有用的。

[0320] 表9:在不同组织类型中的跨物种泛素启动子T1表达

[0321]

构建体	事件	平均YFP (ng/mg TSP)						
		穗轴	壳	粒	花粉	V12根	须	茎
pDAB113117	pDAB113117[1]-006	3452.1	1164.3	1341.2	397.1	2292.7	1405.0	7279.8
pDAB113117	pDAB113117[1]-007	2519.6	954.7	1410.9	414.3	2245.6	1974.3	6179.0
pDAB113117	pDAB113117[1]-019	8362.3	2280.8	2829.6	749.5	7112.4	4790.2	13044.1
pDAB113118	pDAB113118[1]-005	2801.6	620.9	886.3	782.2	1136.1	636.7	1953.6

[0322]

pDAB113118	pDAB113118[1]-007	2339.2	524.9	725.1	376.1	1495.6	1271.0	2806.9
pDAB113118	pDAB113118[1]-023	1302.1	491.8	716.8	435.3	1193.0	829.1	1522.9
pDAB113119	pDAB113119[1]-011				399.7	1025.9	942.2	2475.6
pDAB113119	pDAB113119[1]-013	2238.1	572.0	1050.5	438.1	1311.2	539.7	2235.2
pDAB113119	pDAB113119[1]-028	2013.9	536.4	1061.4	450.0	1166.3	826.9	1912.5
pDAB113120	pDAB113120[1]-005	1166.8	310.5	514.0	1704.1	169.7	322.1	739.3
pDAB113120	pDAB113120[1]-023	1096.4	531.9	845.8	1433.9	268.4	572.2	877.9
pDAB113120	pDAB113120[1]-032	1344.1	587.8	985.1	1252.3	187.6	472.4	694.0
pDAB113121	pDAB113121[1]-011	6779.1	2942.3	3452.6	2022.6	5834.0	2881.7	7445.5
pDAB113121	pDAB113121[1]-023	4830.0	2689.8	1913.7	1641.8	2547.7	2453.8	8295.9
pDAB113121	pDAB113121[1]-026	8186.2	3889.3	4432.3	1432.0	2521.9	2182.5	7760.7

[0323] 本文中引用的所有参考文献(包括出版物、专利、和专利申请)通过提述将其并入本文,其程度不与本公开的明确细节相冲突,并入的程度如同每篇参考文献被单独地并且明确地指示通过提述并入,并且在本文中以其整体进行阐述。提供在本文中论述的参考文献只不过是它们的公开在本申请的申请日之前。在本文中不应被解释为承认本发明人因现有发明而无权先于这些公开。提供以下实施例以展示某些具体特征和/或实施方案。不应将这些实施例解释为将本公开限制到所例示的具体特征或实施方案。

序列表

<110> S·库马尔
J·贝林格

<120> 利用来自短柄草属泛素基因的调节元件的转基因表达用构建体

<130> 14754-231172

<150> 61/872,134

<151> 2013-08-30

<160> 40

<170> PatentIn version 3.5

<210> 1

<211> 1029

<212> DNA

<213> 二穗短柄草

<400> 1

```

ctgctcgttc agccacacgt aacacgccg gcgacatgca gaigccctcc accacgccga      60
ccaaccccaa giccgcgcgc ctgctccacg gcgccatccg catccgcgcg tcaacgtcat      120
ccggaggagg cgagcgcgat gtcgacggcc acggcggcgg cggacacgac ggcgacgccc      180
cgactccgcg cgcgcgtcaa ggctgcagtg gcgtcgttgt ggccgtccgc ctgcacgaga      240
tccccgcgig gacgagcgcc gccctcaccg agccccata tcgagaaatc aacggtgggc      300
tcgagctcct cagcaacctc cccacccccc ctcccgacca cgtcccttc ccccggtccc      360
ctcttcctcc taaacccgag ccgcgcgaga caacaccaac gaaagggcga agagaatcgc      420
[0001] catagagagg agatgggcgg aggcggatag tticagccat tcacggagaa atggggagga      480
gagaacacga catcatacgg acgcgacccf ctagctggct ggctgtccta aagaatcgaa      540
cggaatcgct cgcgcaggag aaacacgaac gtccctgaagc atgctgcgcc gattcttcca      600
aaacacttat cttaagatt gaagtatgat atatgactga aatttttaca aggtttttcc      660
ccataaaaca ggtgagctta tcctatccit ttgttiagga tgiacgtati atatatgact      720
gaatattttt tattttcatt gaatgaagat tticgacccc ccaaaaaaia aaacggagg      780
gagtaccrit ggcgcgtgta tatggacfat agccaicggg acgtttccgg agactgcgig      840
gtggggcgca tggacgcaca acgaaccgat tticggttgc cgaactgcgc ttgcacttg      900
gtaggcacga ctgcicgggt tcggctcttg cgtgagccgt gacgtaacag acccgttcic      960
ttcccccgic tgcccatcca taatccccc ctccatcggc ttccctttcc tcaatccagc      1020
accctgatt                                     1029

```

<210> 2

<211> 636

<212> DNA

<213> 二穗短柄草

<400> 2

```

ggcgtcagga ctggogaagt ctggactctg cagggccgaa ctgctgaaga cgaagcagag      60
gaagagaaag ggaagtgttc gacttgtaa ttgtagggtt tttttttaga ggaacttgta      120
attttaggtt ggctggccct cgttggaana acgatctgg ctggttgggc tgggccgaig      180
tacgottgca aacaacttgt ggccggccgt tcggacagag caggagtctc tttttgttc      240

```

	tcacitititc ggctitititc agttacggag tacctititgt ittttaaagg agitacctit	300
	tttttaggaa ticttiaggt accititcgct tgcctcmeta aaatatttaa ctitcgctit	360
	tttcatitit aattititga actaittacg agtticacga atgcttatit tccagcatai	420
	cattatitgc aagtaititit atgocgtatg taitggacga gagccatcgg gactgttcca	480
	gagactgcgt ggtagggacg gctcccaacc gccititcta tctctgttcg catcgggtgg	540
	ccgacttgge tgcgcgiga gccgigacgt aacagactig gictcttccc catcggcca	600
	tctataaait ccccatcga tegacccicc ctitcc	636
	<210> 3	
	<211> 1064	
	<212> DNA	
	<213> 粟	
	<400> 3	
	tgcgtctgga cgcacaagtc atagcattat cggctaaaai tctttaattt ctaaattagi	60
	cataicggct aagaaagigg ggagcactat catticgtag aacaagaaca aggtatcata	120
	tatataata tatataatit ttaaacittg itaagiggaa tcaaagtgtt agtattaatg	180
	gagtticagt tgcattaaat tttatgtcac atcagcaait ttgttgactt ggcaaggta	240
	tttaggggtg gtttgggaag caggggctat taggagiatt aaacatagtc taattacaaa	300
	actaatigca caaccgttaa gctgaatcgc gagatggatc tattaagctt aattagacca	360
	tgattigaca atgttggtgt acaataacca ttigtctaat atggattact taggtttaat	420
[0002]	agattcgcti cgtgatttag cctaiggtt ctgctattaa ttigttaatt agctcatait	480
	tagttcttat aattagtatc cgaacatcca atgtgacatg cttaaagtta accctggat	540
	ccaaatgaag tcttaigaga gtttcatcac tccggiggta taigtactta ggctcgtit	600
	tcttccaccg acttattitit agcacccgtc acattgaatg tttagatact aattagaagi	660
	attaaacgtt gactatttac aaaatccait acataagacg aatctaaacg gcgagacgaa	720
	tctattaaac ctaattagtc catgatttga caatgtgttg ctacagttaa catitgctaa	780
	tgatggatta attaggctta atagattcgt ctgcgcgtit agccctccact tatgtaatgg	840
	gttttctaaa caatctacgt ttaatactcc taattagiat cttaatatit aatgtgacac	900
	gtgtctctga taagtcagtg gaaggagag aacgtccctt tagtttcca tctattatit	960
	tgtacgatga aactgtgcag ccagatgatt gacaatcga atacttcaac tagtgggcca	1020
	tgcacalcag cgacgtgtaa cgtcgtgagt tgcgttccc gtag	1064
	<210> 4	
	<211> 1026	
	<212> DNA	
	<213> 二穗短柄草	
	<400> 4	
	gtttgtctaa aactggccta cagctcgtg cccctgttgg tctgcccctt ggaagtatgc	60
	gtgtctatgg ttatgtgaga agtcgtgtgt ttctttctaa tcccgtactg ttgtgtgaa	120
	catctgctgc tgcgttatgt catcgtgaag aaacctgtta tgaataagtg aacaigaacc	180
	ttgtctgtg attacggctt cgtgttatgt cgaacgttct tacaacgca atgtcacctg	240

	aigtaaaaac gtttttgcia gcgtatgga acaagtgctc atgatttca igcaagatgc	300
	aattccagct ttgttgggtt igtcatcttt giactgtgct taccgcacat aaagattgca	360
	tcttgcctat tgccttgttg ctttgggtgt cgtccgcttc tecttgcacc ttatcaaacc	420
	ttgttttaga ttctcttctt atagcacttg gtaaccttca gccttacaac gccagtactg	480
	ttcttgaaat ttcatgactg ataaagctga tagatggagt actaataatg gacatcttcc	540
	cafaaatgtt cgggtgcaga gataaggagg ccccgaggat ctatttacag gatgaacctg	600
	ccctggcccg tgtacgcatt acatcccgga gcaagtctga ggctctcaat gtacacatga	660
	aattgatttt tgcctcggtt ggcttggctg atcgttgcct ttgttctgat tcatcagagt	720
	taaataacgg atataacagc aaataacgc agcatccaca ccgaccacac gtccggtaaa	780
	cagagtcctc cgtcttctgt ttaattatga cggagtactc cgtattaat ccttagatat	840
	gtttcgaagg aactcaaacc ttctctcctc tgcatacttc agtcttcaa aactggaaat	900
	agataattga aaccttcatt cgggtgcaat tcacaactgc aaattgaaca gcactgtcaa	960
	tttcaatttc ggttcaaga ttccaccgat aggttgacat gatccatgat ccacctatg	1020
	tacaaac	1026
	<210> 5	
	<211> 1020	
	<212> DNA	
	<213> 二穗短柄草	
[0003]	<400> 5	
	gtcttgcctg aactgggtca cagtctgtg ccttgggtgg tcgcccctt agtggtcatg	60
	ccctttgtta tgggtcttgc gtcccaatcc tgtatcgttt gtgtgaacat ctctgtgtgt	120
	gtatagcagc ttgaatcctg ttatgaattt gigaacctga acctgttccc gtgaatcatg	180
	ttatgaataa gigaacctga acctgttccc gtgattattg ttacaatctg ttgtgccgta	240
	tgggtgtctg tgggtgattt atgtgaact ggagaaccaa gtctgttcca ggacataatg	300
	caacctgaag taacctatgt agaactactt gttctgggag acataaaacg tcatttttat	360
	gcattcgtaa catataagca taclacaata attgatattg ccttttccca ctcatccctg	420
	aaaccataig cctcttctca ggcctctac atgcagtgig ctacagaaca acagggccctg	480
	ccagctgttt ttcaattttc caattataaa ccacantagt cggactatgg catctgtggg	540
	tgactaigca agatgttgt gtcaggcttc tgaacctttt cccatgtatc tgttgaattt	600
	accagtaaa ttcatgcctc tatttaattt ggcatgggtg attttcaaac agaattgttt	660
	tttttttgtt ctggaagcta tatgtgtaaa taaatacaaa gctggagtgt gattatattt	720
	ccaacagata ttcaagaaaa tctcagtga ttattttact actgtagtat atataatata	780
	cttacagtgt acctctcata ttcaaacga catgtgagca catgtttcag ttctttagga	840
	tgtgttgtgt gctcaagggt gtaattttgc attctgccc cagagttaac actacacgta	900
	tttttttgag tggcagtgtc ttgtattaca aggcacaaac aacaaaaacc tatggcaaga	960
	tatccttctt agaggctgcc aggatcattt tgactgaact atgtaaggct gaagaaaagg	1020
	<210> 6	

<211> 1032
 <212> DNA
 <213> 粟

<400> 6
 gcccatcggt catgagctt tctactgtac ctgggicgtc tggctctcgc ctgtgtcacc 60
 ttggaagtac ctgtgtcggg aattggtttg gtcataaact gcagtttgtc ttgatgttc 120
 ttgtgtcggg tctatgaac tgggtgtatc tgaatgtt cgttaaactg ttgtgtcggg 180
 gcagcagtat ggcatccgaa tgaataaatg atgtttggac ttaaatctgt actctgtttg 240
 ttgtcgttgc tgcctgttct atattgtctg agatcagaat gtttgccttc tgcctgttgc 300
 ttggcgttgc gtcactcct gtttcttact tgaggcgtaa ctctgttctg gcaaacctaa 360
 atgtctaacg gaattgttgc ggacataatt gttggacaga ttaacgtgtt tgggtgttgc 420
 ctgattgtgc attcgggaag ctgtgtgttt gttggaatcaa ggagagcagc taggtctgtg 480
 cagaacgtta ttgtggaatt aagccttctc agattatgcc attactctaa acctaatgat 540
 atcatatttc actcggggat gttggagtag tcttttcttt ctctgcaga caaaatgatt 600
 ttgtttctgt gtgtgtacat gattttgtgc aactgttgc acaactgaag tagacaagtt 660
 ttgacctcac cagaagaatg aaaaagaatt tggattttgt taccatgaca aacctgtgta 720
 acttggccca tcagaatgca cagaagagcg gctacaaatt gacatgcgtt gcaaaccttg 780
 caatagttag tgcacatgtt tgcctatgcc tgcagctctt aggaaaagtg tgtgtgtcga 840
 gaaatctaag catagtgtgt ctgtctacat tgcgtggaac ccacacagct ttgtcacact 900
 [0004] ctgttccact ccagaagtc ttcctggtgc tgtttacccc tggtaaaagg taaccgaaaa 960
 ctctcaagg ctgtacccaa aactggaagg aaatttggag gaaatctttg ctttgatctg 1020
 gctcactctt tc 1032

<210> 7
 <211> 993
 <212> DNA
 <213> 二穗短柄草

<400> 7
 gtatgcagcc tgccttctc ctctctaccg ttcaattctt ggagtaggtc gtagaggata 60
 ccaattgat ttgacagagg gattagatta gatacttgta gatcgaagtg cgaatttcc 120
 atggtatagc ataccatgtt gatttgcatt agatcggatt aaatctttgt agatcgaagt 180
 ggcacgttc catgaattgc ctgttaccag tagattcaag ttttctgtg ttatagaggi 240
 gggatctact cgttgagatg attagctctt agaggacacc atgccttttt ggaaaataga 300
 tcagaacctg gtatgcgat gtgagcagt gtctctgtag atccaagtgc ttctgcattg 360
 tactagtgtg gatctattgt ttgtgttaata cgtctctgat ctatccgtgt agatttctat 420
 cgattactgt tactgtggtc tgcctgttca tagttgttgc ttagggttga tgaacagtg 480
 tcgaacctta atgggatag tattcttgat ctatcaacgt gtagggttca gtcattgtat 540
 tatgtactcc ctctgtccca aattaacgt cgtggatttt gtataagaat ctatacaaat 600
 ccaattcagc taattcggga tggagtacca tattcaataa ttgttttatt gcgttccact 660
 tatgtacctt atgtttgttg ttctctatgt ggattctact aattatcatt gattgtgat 720

	cttctatitit gctagtittcc tagctcaatc tggtiattca tctagatgig ttgttgaaat	780
	cggagaccaa gcttgttatt agatagtita ttgcttatca gttaicatgt cttgttgatg	840
	caacacatat tcaigticgc taictgggtg ctgcttgata ttctctgatt tacattcaat	900
	ataagaaiat attctgcict ggttgttgc tctcatgact ttacctactc ggtagggtgac	960
	ttaccttttg gtttacaatt gtcaactatg cag	993
	<210> 8	
	<211> 852	
	<212> DNA	
	<213> 二穗短柄草	
	<400> 8	
	gtaigttagcc tctcgattcc tcttcagecc tgcctctgat ttgggtacg cgttgagaig	60
	atgatctcgt agatgtctag atgacaccaa gtctgattga aatagatcag atccgtgtag	120
	atcgatgagc tctgtgttac ctgttgatc aagttatitl cgcattctat tttgttgatc	180
	tactagatct agtgtgtgta ttctatgcta tctgattctc cgtgttagat tcaactgatt	240
	actgttactg tggcttgatc ggccatagat gtgtgttaag gttgatcgg ttagtgttg	300
	aaactgcgtg gatatctagc atccatctat tatcgtgtag gtttgaaca aacaagcact	360
	attattgtac tgaatgttgc tctatgtgtg gttttgaccg tttagtgtt gaacagccct	420
	tctgtatttg ttattgtctg tccagtgtg tactaatttc gttgagtgtc ggtattatct	480
	aattattgtt gattgataat ctgttagttt gcttttctta atttatttat cgtagtctctg	540
[0005]	atttgcctca gctgtgcctc acccgtgcga tggicaatca actgtttagc ccaatctct	600
	taatcatgta catttgttgt tagaatcaga gatcaagcca attagctatc ttattgttta	660
	tctgttccat gttctgatcg atgtaacagt ctacactttt gctctgtgct acttgattaa	720
	aaattctga cttaaatcca tgaatggaag ttccagatct gattgttgc ttacttgact	780
	aaatctatct catgtgacac ctctctgtct tggtaactta cgcctgttg ttgttaattl	840
	cigactaigc ag	852
	<210> 9	
	<211> 48	
	<212> DNA	
	<213> 粟	
	<400> 9	
	gtcacgggtt ccttccccac ctctctctct cccacacgcc ataaatag	48
	<210> 10	
	<211> 1114	
	<212> DNA	
	<213> 粟	
	<400> 10	
	gtacggcgat cgtcttctct ctctagatcg gcgtgatctg caagtagttg atttggtaga	60
	tggtaggat ctgtgcactg aagaaatcat gttagatccg cgatgtttct gttctgiatg	120
	ggctgggagg tggaaatitl gtgtagatct gatagttct cctgtttatc ttgtcacgt	180
	cttgcgattl gttgggaatt taggtctgtg atctgggaat cgtgggttg ctcttaggt	240
	gttcttagat gaggctgttc tcacggitta ctggatcaat gccatgtaga tcagctcggg	300

	ctttcgtctt tgaataaggt gccatactt gcataatga tctggtccg tgggttacc	360
	taggttctcg cggcigattc giccgaicga tlllgttagc atgigglaaa cglltgglica	420
	tggctcgatt tagattagag tggaaatagga tgaatcgaat ctacgtcttg ggattaatat	480
	gcattggtca ccaactgttt ccggtgttaa gatgatgaat ctatgcttag ttaatgggtg	540
	tagatataa tgcctcgttt cctcaatgat gccgiagctt ttacctgagc agcatggatc	600
	ctcctgttac ttaggiagat gcacatgctt atagatcaag ataigtactg ctactgttgg	660
	aattcttttag tatacctgat gatcatecat gctcttgtaa ctgttttgg tatactigga	720
	tgatggcatg ctgctgcttt ttgttgatt gagccatcc atatctgcat atgtcacatg	780
	attaagaaga ttacgtgttt tctgtatgat gccatagctt ttaigtgagc aacatgcatc	840
	ctcctggtta tatgcattaa tagatggaag atacttatig ctacaatttg atgattattt	900
	tgtacatagc atgacaaagc atgccttcca tactttgttg atatacttg ataatgaaat	960
	gcctgcgcaac gttcattcta tagcactaat gatgtgatga acacgcacga cctgtttgtg	1020
	gcattctgtt gaattgttg ttgtgttca cttagagctg ttttattaac ctactgctag	1080
	atacttaacc ttctgtctgt ttattctttt gcag	1114
[0006]	<210> 11	
	<211> 55	
	<212> DNA	
	<213> 二穗短柄草	
	<400> 11	
	ccgatcgaaa agtccccga agagcaagcg accgatctcg tgaatctccg tcaag	55
	<210> 12	
	<211> 75	
	<212> DNA	
	<213> 二穗短柄草	
	<400> 12	
	ccaatccagc acccccgatc ccgatcgaaa attctccga acagcaagcg atcgatctag	60
	cgaatccccg tcaag	75
	<210> 13	
	<211> 261	
	<212> DNA	
	<213> 粟	
	<400> 13	
	agaaaatca actggigggc cagccacatc agcgtcgtgt aacgtggacg gaggagcccc	60
	gtgacggcgt cgacatcgaa cggccaccaaa ccacggaacc acccgtcccc acctctcgga	120
	agctccgcic caegcgctcg acatcilaag gclaccagca ggcgtacggg ttgagagga	180
	ctccttgcct ctltgcgcig gcggcttccg gaaatlgcgt ggcggagacg aggcgggcic	240
	gtctcacacg gcacggaaga c	261
	<210> 14	
	<211> 113	
	<212> DNA	
	<213> 粟	
	<400> 14	

<400> 14	ccgacccctt cgctttcttc cccaatcica tctcgtctcg tgtgttcgg agcacaccac	60
	ccgccccaaa tegtcttcc cgaagcctc ggcatctt caccgcttc aag	113
<210> 15		
<211> 2077		
<212> DNA		
<213> 二穗短柄草		
[0007]	<400> 15	
	ctcgtcttc agccacagt aacacgcgt gcgacatga gatgccctcc accacgccga	60
	ccaaccccaa gtccgcgcg ctcgtccacg gcgccatccg caccgcgcg tcaacgtcat	120
	ccggaggagg cgagcgcgt gtgcacggcc acggcggcgg cggacacgac ggcaegccc	180
	cgactccgcg cgcgcgtcaa ggcgtcagt gggtcgtggg ggccgtccg ctgcacgaga	240
	tccccgcgtg gacgagcgc gccccacccc agccctata tcgagaaac aacggtgggc	300
	tcgagctctt cagcaacctc cccaccccc ctccgaeca cgtcccttc cccgtgccc	360
	ctctctccg taaacccgag ccgccgagaa caacaccaac gaaaggcgga agagaatccg	420
	catagagagg agatgggcgg aggcggatag tticagccat tcacggagaa atggggagga	480
	gagaacacga catcatcgg acccgacctt ctacgtggct ggctgtctta aagaatcgaa	540
	cggaaatcgt gcgccaggag aaaacgaacg gtccgaagc atgtgcgcgc ggttcttcca	600
	aaacacttat cttaagatt gaagtatgt atatgactga aatttttaca aggtttttcc	660
	ccataaaaca ggtgagctta tcatactct tttgttagga gtacgtatt atatatgact	720
	gaataitttt tatttttatt gaaigaagt ttccgacccc ccaaaaataa aaacggagg	780
	gagtaccttt gtgcgtgtga tatggactag agccatcggg acgtttccgg agacgtcgig	840
	gtggggcgga tggacgcaca accaccgcat ttccgttgc cgactcgcgc ttccgactcg	900
	gtaggcagca ctctcgggt tgggtcttg cgtgagccgt gacglaacag accgttctc	960
	ttccccctgc tggccatcca taaaccccc ctccatcggc ttcccttccc tcaatccagc	1020
	acctgtatcc cgatcgaaaa gtccccgaa gagcaagcga ccgatctcgt gaatctccgt	1080
	caaggatgc agcctcgtt cctctcgtt accgtttcaa ttctggagia ggtcgtagag	1140
	gataccatgt tgattigaca gagggagtag attagatact gtagatcga agtgcggatg	1200
	ttccatggta gatgatacca tttgatitc gattagatcg gattaaatct ttgtagatcg	1260
	aaagtcgcat gtccatgaa ttgcctgita ccagtagatt caagtttttc tgtgttatag	1320
	aggtgggata tactcgttga gatgattagc tcttagagga caccatgccg ttttgaaaa	1380
	tatgtcagaa ccgtgtatgt cgtgtgagc atgtgttctt gtagatccaa gtcttttcgc	1440
	atgttactag ttgtatctta ttgtttgtgt aatacgtctt cgatctatcc gtgtagattt	1500
	cactcgatta ctgttactgt ggcttgatcg ttcatagttg ttctgtaggt ttgatcgaac	1560
	aggtgtcgaa cctaatggga tatgtattct tgatctatca acgtgtaggt ttcatgcatg	1620
	tatttatgta ctccctccgt ccaaaataa ctgacgtgga ttttgtataa gaatctatac	1680
	aaatccatgt cagttaattc gggatggagt accatattca ataatttgtt taattcgttc	1740
	cacttatgta ccatatgttt gtgtttcttc atgtggatcc tactaattat catgtatgg	1800

	tgatcttcta ttttctagt ttctagctc aatctggta tcatgtaga tgtgtgtg	1860
	aaatcggaga ccatgcttgi taitagatag ttiattgcit atcagtttca tgttciggt	1920
	gaigcaacac atattcaigt tegtatctg gtgtctgtt gatattctct gattfacatt	1980
	caitataaga atatatctg ctctgggtgt tgcctctcat gacittacct actcggtagg	2040
	tgactfacct ttgggtttac aatgtcaac tatgcag	2077
	<210> 16	
	<211> 1563	
	<212> DNA	
	<213> 二穗短柄草	
	<400> 16	
	ggcgtcagga cggcggaagt ctggactctg caggccgaa ctgtgaaga cgaagcagag	60
	gaagagaaag ggaagigtic gacttgtaai tigtagggtt tttttttaga ggaacttgta	120
	alltgtaggt gggctggcct cgttggaaaa acgatgctgg ctggttgggc tggccgatg	180
	taagcttgca aacaacttgi ggcggccgtt tctggacgag caggagtctt tttttgtt	240
	tcacttttct ggtcttcttt agttacggag tacatttgtt tttttaagg agttaccttt	300
	tttttaggaa ttcttagtt acccttctgt tgcctctaaa aaatatttaa ctctcgttt	360
	tttctatttt aatttttgca actatttacg agtttcatga atgcttattt tccagcaat	420
	cattatttgc aagtattttt atgccgtatg tattggacga gagccatcgg gactgttcca	480
	gagactgcgt ggtggggacg gctcccaacc gccttttcta tctctgttct catecgggtg	540
[0008]	ccgacttggc tgcgcgtiga gccgtgacgt aacagaattg gtctcttccc catctggcca	600
	tcataaatt ccccatcga tgaaccttcc ctctcccaa tccagcaacc ccgacctcga	660
	tcaaaaattc tccgaacag caagcgatcg atctagcgaa tccccgtcaa ggtatgiacg	720
	ctctcgatc ctctcagcc ctgccctcga ttgggtgtac gcgttagat gatgatctcg	780
	tagatgtcta gatgacacca tgtcgatttg aaatagatca gatccgtgta gatcgatgag	840
	ctctgttgta cctgttgatt caagttattt tgcattgcta ttgttgtat ctactagaic	900
	tagtgtgtgt attctatgct atcgatttct cctgttagat ttactctgat tactgttact	960
	gtggcttgat cggccataga tgttggttaa ggtttgatcg gttagtgttt gaacctgcgt	1020
	ggatacttag catccattcta ttaacgtgta ggtttcgaac aaacaagcac tattatigta	1080
	ctgatggctt gctatgggtt ggttttgacc gtttttaggt tgaacgagcc ttctgtattt	1140
	gtttattgct gtccagtgai gtacctgtt cgttgagtgt cggattatct taattattgt	1200
	tgattgataa tcttgiagtt tgccttctct aattatttta tctagttctt gatttgcctc	1260
	agcgtgcctt caccctggcg atgttcaatc aactgttag cccaatctgc ttaatcatgt	1320
	acatttgttg ttagaatcag agatcaagcc aattagctat cttattgtct atctgttcca	1380
	tgtcttgatc gatgtaacag tctacatttt tgcctgtgct tacttgatta aaacattctg	1440
	acttaaatc atgatiggaa gtttcagatc tgattgttgc cttacttgac taatactat	1500
	tcatgtgaca cctctctgtc ttgtaactt acccgtgttt gtttgttaatt tctgactatg	1560
	cag	1563

<210> 17		
<211> 2600		
<212> DNA		
<213> 粟		
<400> 17		
tgctctgga cgcacaagtc atagcatat cggcfaaaat tcttaattt cttaaatiagt	60	
catatcggct aagaaagtg ggagcactat caittcgtag aacaagaaca aggtatcata	120	
tatatataa tatataatat ttaaccttg ttaagiggaa tcaaatgct agtattaatg	180	
gagtticag tgcattaaat ttaatgac atcagcaatt ttgttactt ggcaaggta	240	
tttaggggt gtttgaaga caggggctat taggagtatt aaacatagtc taattacaaa	300	
actaatgca caaccgcta gctgaatgc gagatggatc tattaagctt aattagacca	360	
tgattigaca atgttggtct acaataacca ttgtctaag atggattact taggtttaat	420	
agattcgtct cgtgaattag cctaagggtt ctgctatata ttigttaatt agctcatatt	480	
tagttcttat aattagtatc cgaacatcca atgtgacatg cttaaagtta accctggtat	540	
ccaaatgaag tcttaagaga gttaacacac tccggtggta taigtacata ggctccgttt	600	
tctccaccg acttattttt agcacccgic acattgaatg tttagatatt aattagaagt	660	
attaaacgta gactatttac aaaatccatt acataagacg aatctaaacg gcgagacgaa	720	
tctatlaaac ctatagagtc catgattga caatggttg ctacaglaaa catttgctaa	780	
tgaaggatta attaggttta atagattcgt ctcccggttt agcctccact taigtatgg	840	
gttttctaaa caatctacgt ttaatacacc taattagat ctatattatc aaigtgacac	900	
gtgctaaaaa taagtacgtg gaaggagag aacgtccctt tagttttcca tcttatiaat	960	
tgtacgatga aactgtgcag ccagatgatt gacaatcgca atacttcaac tagtgggcca	1020	
tgcacatcag cgacgtgtaa cgtctgtagt tgcgttccc gtagagaaat atcaactggt	1080	
gggccacgca catcagcgtc gtgtaacgtg gacggaggag ccccgtagc gctcgcacat	1140	
cgaacggcca ccaaccacgg aaccacccgt cccaccctt cggagctcc gctccacggc	1200	
gtcgacatct aacggctacc agcagggcta cgggttggag lggactcctt gctcttttgc	1260	
gttggcgggt tccggaattt gcgtggcgga gacgaggcgt gctcgtctca cagggcacgg	1320	
aagacgtcac ggttcccttc cccacctctc ctcttcccca ccgccataaa tagccgaccc	1380	
ctctgccttt ctccccaate tcatctcgtc tegtgttgtt cggagcacac caecgcctcc	1440	
aaatcgttct tcccgcaagc ctggcgatc ctacacccgc ttcaaggtae ggcgatcgtc	1500	
trectctct agatcggcgt gatctgcaag tagttgattt ggtagatggt taggatctgt	1560	
gcacigaaga aatcatgta gatcccgcat gtctctgttc gtagatggct gggagggtga	1620	
atttttgtgt agatctgata tgtctcctg ttatctttgt cacgtctctg cgatttgttg	1680	
ggattttagg tcttgatct gggaaatcgt ggttgccttc tagctgttc gtagatgagg	1740	
tctttctcac ggtttacttg atcatgacct agtagatcag ctccggcttt cgtctttgta	1800	
tatgtgtccc atacttgcat ctatgatctg gtcccggtt gttacctagg ttctctgccc	1860	
tgattcgtcc gatcgatttt gttagcagt ggtaaacgtt tggatcgtt ctgatttga	1920	
ttagagtcga ataggatgat ctgatctag ctcttgggtt taatatgcat gtgtcaccaa	1980	

[0009]

	tctgttccgt ggtaagatg atgaatctat gcttagttaa tgggtgtaga tatataigct	2040
	gcgtttccctc aatgatgccg tagcttttac ctagcagca tggatccctcc tgttacttag	2100
	gtagatgcac atgcttatag atcaagatat gtacigctac tgttgaati ctttagtata	2160
	ccgatgac atccatgctc ttgttacttg ttgtgtata ctggagat ggcatgcgc	2220
	tgtttttgt tgaattgagc ccatccatat ctcatatgt cacatgatta agatgattac	2280
	gcgtttctg tatgatgcca tagcttttat gtgagcaaca tgcacccctcc tggttatafg	2340
	cattaaaga tggagaat atattgctac aatttga ttttttgta catacgaaga	2400
	tcaagcatgc tcttcatact ttgttgatat acttggataa tgaatgcgc ctcacgttc	2460
	attctatagc actaatgatg tgaagaacac gcacgacctg ttgttgcat ctgttgaat	2520
	ggtgtgtgc tgttcactag agactgtttt attaacctac tgcagatac ttaccctct	2580
	gtcgtttat tcttttgcag	2600
	<210> 18	
	<211> 1602	
	<212> DNA	
	<213> 玉米	
	<400> 18	
[0010]	atgcagatct ttgtgaaac cctgactggc aagactatca ccccgaggc ggagtcgtc	60
	gacaccattg acaacgttaa ggccaagatc caggacaagg agggcatccc ccagaccag	120
	cagcggctca tcttgcctgg caaacagctt gaggacgggc gcacgcttgc tgaactaac	180
	atccagaagg agagcaccct ccaccttgig ctccgctca ggggaggcat gcagacttt	240
	gtgaaacccc tgaccggcaa gactatcacc ctgagggtg agtccctga caccattgac	300
	aacgtcaagg ccaagatcca ggacaaggag ggcatccctc cagaccagca gggcctc	360
	ttgtctgga agcagctga ggacgggggc acgcttgcg actacaacat ccagaaggag	420
	agcaccctcc acttggctgc gcgcctcagg ggaggcatgc agatcttctg gaagaccctg	480
	accggcaaga ctatcaccct cgaggiggag tcttcagaca ccaatgaca cgtaaggcc	540
	aagaaccagg acaaggaggg catcccccga gaccagcagc ggcctcctt tgcggaaag	600
	cagcttgagg acggcgccac gcttgcggac tacaacatcc agaaggagag caccctccac	660
	ttgtgtctgc gccctcaggg aggcattgag atcttctga agacctgac cggcaagact	720
	atcaccctcg aggtggagtc ttcagacacc atcgacaatg tcaaggccaa gatccaggac	780
	aaggagggca tcccaccgga ccagcagcgt ttgatcttgc ctggcaagca gcaggaggat	840
	ggccgcaccc ttgcggatta caacatccag aaggagagca cctccacctt ggtgtctcgt	900
	ctcaggggtg gtaicagat ctttgaag acactcacig gcaagacaat cacccttgag	960
	gtggagtctt cggataccat tgacaatgtc aaggccaaga tccaggacaa ggagggcac	1020
	ccaccgacc agcagcgctt catcttgcg ggcaagcagc tggaggatgg ccgcaccccg	1080
	gcggattaca acaaccagaa ggagagcact ctccacctgg tgcctcgctt caggggtggc	1140
	atgcagattt ttgtgaagac attgacttgc aagaccaica ccttggaggc ggagagctct	1200
	gacaccattg acaatgtaa ggccaagatc caggacaagg agggcatccc ccagaccag	1260
	cagcgtctga tcttgcggg caagcagcgt gaggatggcc gcactctcgc ggactacaac	1320

```

aiccagaagg agagaccct tcacctgtt ctccgcctca gaggtggtat gcagatcttt 1380
gtaaagacct tgactggaaa aaccataacc ctggagggtg agagctcggg caccatcgac 1440
aatgtgaagg cgaagatcca ggacaaggag ggcaaccccc cggaccagca gcgtctgalc 1500
ttcgccggca aacagctgga ggaaggccgc acctagcag actacaacat ccaaaaggag 1560
agcaccctcc accttggtct ccgtctccgt ggtggtcagt aa 1602

```

<210> 19

<400> 19
000

<210> 20

<400> 20
000

<210> 21

<400> 21
000

<210> 22

<211> 685

<212> PRT

<213> 玉米

[0011]

<400> 22

```

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1           5           10           15

```

```

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp
20           25           30

```

```

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
35           40           45

```

```

Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu
50           55           60

```

```

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe
65           70           75           80

```

```

Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser
85           90           95

```

```

Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile
100          105          110

```

```

Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp
115          120          125

```

```

Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His
130          135          140

```

```

Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu
145              150              155              160

Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp
              165              170              175

Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln
              180              185              190

Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu
              195              200              205

Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg
              210              215              220

Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr
225              230              235              240

Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala
              245              250              255

Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile
              260              265              270

Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn
              275              280              285

Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly
290              295              300

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
305              310              315              320

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp
              325              330              335

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
              340              345              350

Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu
              355              360              365

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe
              370              375              380

Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser
385              390              395              400

Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile
              405              410              415

Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp
              420              425              430

```

[0012]

Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His
 435 440 445
 Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu
 450 455 460
 Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp
 465 470 475 480
 Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln
 485 490 495
 Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu
 500 505 510
 Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg
 515 520 525
 Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr
 530 535 540
 Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala
 545 550 555 560
 Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile
 565 570 575
 Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn
 580 585 590
 Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly
 595 600 605
 Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
 610 615 620
 Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp
 625 630 635 640
 Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
 645 650 655
 Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu
 660 665 670
 Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Gln
 675 680 685

[0013]

<210> 23
 <211> 685
 <212> PRT
 <213> 粟
 <400> 23

```

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1           5           10           15

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp
20           25           30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
35           40           45

Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu
50           55           60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe
65           70           75           80

Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser
85           90           95

Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile
100          105          110

Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp
115          120          125

Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His
130          135          140

Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu
145          150          155          160

Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp
165          170          175

Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln
180          185          190

Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu
195          200          205

Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg
210          215          220

Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr
225          230          235          240

Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala
245          250          255

Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile
260          265          270

Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn
275          280          285

```

[0014]

```

Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly
290                295                300

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
305                310                315                320

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp
325                330                335

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
340                345                350

Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu
355                360                365

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe
370                375                380

Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser
385                390                395                400

Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile
405                410                415

Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp
420                425                430

Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His
435                440                445

Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu
450                455                460

Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp
465                470                475                480

Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln
485                490                495

Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu
500                505                510

Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg
515                520                525

Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr
530                535                540

Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala
545                550                555                560

Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile
565                570                575

```

[0015]

Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn
580 585 590

Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly
595 600 605

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
610 615 620

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp
625 630 635 640

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
645 650 655

Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu
660 665 670

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Gln
675 680 685

<210> 24

<211> 533

<212> PRT

<213> 二穗短柄草

<400> 24

[0016]

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1 5 10 15

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp
20 25 30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
35 40 45

Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu
50 55 60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe
65 70 75 80

Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser
85 90 95

Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile
100 105 110

Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp
115 120 125

Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His
130 135 140

Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu

	145	150	155	160
	Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp	165	170	175
	Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln	180	185	190
	Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu	195	200	205
	Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg	210	215	220
	Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr	225	230	235
	Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala	245	250	255
	Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile	260	265	270
	Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn	275	280	285
[0017]	Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly	290	295	300
	Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu	305	310	315
	Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp	325	330	335
	Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys	340	345	350
	Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu	355	360	365
	Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe	370	375	380
	Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser	385	390	395
	Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile	405	410	415
	Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp	420	425	430
	Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His			

```

435          440          445
Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu
450          455          460

Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp
465          470          475          480

Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln
485          490          495

Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu
500          505          510

Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg
515          520          525

Leu Arg Gly Gly Gln
530

<210> 25
<211> 533
<212> PRT
<213> 二穗短柄草

<400> 25

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1          5          10          15

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp
20          25          30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
35          40          45

Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu
50          55          60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe
65          70          75          80

Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser
85          90          95

Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile
100         105         110

Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp
115         120         125

Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His
130         135         140

Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu
145         150         155         160

```

[0018]

Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp
 165 170 175
 Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln
 180 185 190
 Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu
 195 200 205
 Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg
 210 215 220
 Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr
 225 230 235 240
 Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala
 245 250 255
 Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile
 260 265 270
 Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn
 275 280 285
 Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly
 290 295 300
 Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
 305 310 315 320
 Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp
 325 330 335
 Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
 340 345 350
 Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu
 355 360 365
 Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe
 370 375 380
 Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser
 385 390 395 400
 Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile
 405 410 415
 Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp
 420 425 430
 Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His
 435 440 445

[0019]

Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu
450 455 460

Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp
465 470 475 480

Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln
485 490 495

Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu
500 505 510

Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg
515 520 525

Leu Arg Gly Gly Gln
530

<210> 26

<211> 685

<212> PRT

<213> 人工序列

<220>

<223> 共有序列

<400> 26

[0020]

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
1 5 10 15

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp
20 25 30

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
35 40 45

Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu
50 55 60

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe
65 70 75 80

Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser
85 90 95

Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile
100 105 110

Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp
115 120 125

Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His
130 135 140

```

Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu
145                      150                      155                      160

Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp
                      165                      170                      175

Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln
                      180                      185                      190

Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu
                      195                      200                      205

Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg
                      210                      215                      220

Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr
225                      230                      235                      240

Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala
                      245                      250                      255

Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile
                      260                      265                      270

Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn
                      275                      280                      285

Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly
290                      295                      300

Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
305                      310                      315                      320

Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp
                      325                      330                      335

Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
                      340                      345                      350

Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu
                      355                      360                      365

Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe
                      370                      375                      380

Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser
385                      390                      395                      400

Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile
                      405                      410                      415

Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp
                      420                      425                      430

```

[0021]

Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His
 435 440 445
 Leu Val Leu Arg Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu
 450 455 460
 Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp
 465 470 475 480
 Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln
 485 490 495
 Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu
 500 505 510
 Ala Asp Tyr Asn Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg
 515 520 525
 Leu Arg Gly Gly Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr
 530 535 540
 Ile Thr Leu Glu Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala
 545 550 555 560
 Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile
 565 570 575
 Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn
 580 585 590
 Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly
 595 600 605
 Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
 610 615 620
 Val Glu Ser Ser Asp Thr Ile Asp Asn Val Lys Ala Lys Ile Gln Asp
 625 630 635 640
 Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
 645 650 655
 Gln Leu Glu Asp Gly Arg Thr Leu Ala Asp Tyr Asn Ile Gln Lys Glu
 660 665 670
 Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly Gln
 675 680 685

[0022]

<210> 27
 <211> 1986
 <212> DNA
 <213> 玉米
 <400> 27
 ggcagcgig acccggctgt gcccctctct agagataatg agcatigcat gctaaagtia.

60

[0023]

```

taaaaaatta ccacatattt tttttgtcac acctgtttga agtgcagitt atctatcttt 120
atacatatat ttaaacttta ctctacgaat aataataatc ataglaactac aataatafca 180
gigtittaga gaaicatata aatgaacagt tagacaiggt ctaaaggaca atigagiatf 240
ttgacaacag gactctacag ttttatcttt tttagigtga tgtgtctccc ttttttttg 300
caaatagctt caccatataa atacttcac caitttatta gtacatccat ttagggttta 360
gggttaatgg tttttataga ctaatttttt tagtacatct attttattct attttagcct 420
ctaaatfaag aaaaactfaaa ctctatttta gtttttttat ttaataattt agatataaaa 480
tagaatfaaa taaagtgcct aaaaattfaa caaataccct ttaagaaatt aaaaaaacta 540
aggaaacatt tttctgtgtt cgagtagata atgccagcct gtfaaacgcc gtgcagcagf 600
ctaaccgaca ccaaccagcg aaccagcagc gtctgcgtgg gccaaagcaa gcagacggca 660
cggcatctct gtctcgtcct ctggacccct ctgcagaggt ccgtctccac gtggacttg 720
ctccgtctgc ggcatccaga aattgcgtgg cggagcggca gacgtgagcc ggcacggcag 780
ggggctctct cctcctctca cggcacggca gctacggggg attccttccc caccgtctct 840
tcgtcttccc tctctgccc gccglaaiaa atagacaccc ctccacaccc ctcttctccc 900
aacctcgtgt tgttcggagc gcacacacac acaaccagat ctcccccaaa tccaccgtc 960
ggcacctcgg ctcaaggia cgcgcgtcgt cctccccccc cccccctctc taccitctct 1020
agatcggcgt tccggctcat ggttagggcc cggtagttct acttctgttc atgtttgtgt 1080
tagatccgtg ttgtgttag atccgtgtgt ctacgttctg tacacggaig cgaacctgac 1140
gtcagacacg tctctattgc taacttgcca gtgtttctct ttggggaatc ctgggaigcc 1200
tctagccgtt ccgcagacgg gatcgatttc atgatttttt ttgtttcgtt gcataggggt 1260
tggtttgccc ttttcttfaa ttcaatata tgcctgtcac ttgtttgtcg ggcatcttt 1320
tcatgttttt tttgtcttg gttgtgaiga ttgtgtctgg ttggcgggtc gttctagatc 1380
ggagtagaat tctgtttcaa actaccgtgt ggatttatta attttggatc tctatgtgtg 1440
tgcatacat atcatagtt acgaattga gaatgtgga ggaaatacgc atctaggata 1500
ggatatactg ttgatgcggg ttttactgat gcatafacag agatgtcttt tgttcgtttg 1560
gttgtgaiga tgtgtgtgtg ttggcgggtc gttaattctt tctagatcgg agtagaatac 1620
tgtttcaaac taccgtgtgt atttattaat ttggaaactg tatgtgtgtg tcatacatct 1680
tcatagttao gatttfaaga tggatggaaa taicgatcta ggatagggtt acatgttgat 1740
gggggtttta ctgatcata tcatgatgg catafgcagc atctattcat atgtctaac 1800
cttagtacc tctctattat aataaacaag taigtittat aattattttg atctigatat 1860
acttgatga tggcatafge agcagctata tgggtatttt tttagccctg ccttcatacg 1920
ctatttattt gcttggtaact gtttcttttg tcaatgtca cctgtttgtt tgggttaact 1980
tctgca 1986

```

<210> 28
<211> 21
<212> DNA
<213> 人工序列

[0024]

<220>
<223> 核苷酸引物

<400> 28
cggtgttggga aagaacttgg a 21

<210> 29
<211> 18
<212> DNA
<213> 人工序列

<220>
<223> 核苷酸引物

<400> 29
ccgtggttgg ctgggtct 18

<210> 30
<211> 15
<212> DNA
<213> 人工序列

<220>
<223> 核苷酸引物

<400> 30
cactcctccac tgcct 15

<210> 31
<211> 18
<212> DNA
<213> 人工序列

<220>
<223> 核苷酸引物

<400> 31
tggcggacga cgacttgt 18

<210> 32
<211> 19
<212> DNA
<213> 人工序列

<220>
<223> 核苷酸引物

<400> 32
aaagtttggg ggcgtccgt 19

<210> 33
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 核苷酸引物

<400> 33
cgagcagacc gccgtgtact t 21

<210> 34
<400> 34
000

<210>	35	
<211>	1025	
<212>	DNA	
<213>	柳枝稷	
<400>	35	
	ttgaatttta atticaaatt ttgcagggia giagtgaca tcacaataca taittagaaa	60
	aagtttata attitctctc gttagtittc atataatttt gaactccaac gattaatcta	120
	ttatfaata tcccgacta tcaaaataat gatanaaatt tatgallaat tittctaaca	180
	tgtgtatgg tgtgtactat cgtcttataa aatticaact taaaactcca cctatcatg	240
	gagaaatgaa aaagacgaat tacagtaggg agtantiiga accaaatgga atagtittgag	300
	ggtaaaaiga actaaacaat agtttaggag gttatcaga tttagttaa agttgagagg	360
	aglaatttag acttttctt atcttgaatt gttagcggct ctcttatcgg atatcggaig	420
	gagctttica gcccaacata acttcaatcg gccccaaaag ttctgtccatc cagcctaggg	480
	agaacatttt gcccatgata tctgtttttc ttttttcta ttctacttgg tattatagga	540
	gggaataata caacgtgttc accttgggtt tcatcttgtt tccatctgaa tttatctaaa	600
	actgtgttgg aactctgtaa gaattttgtt cgtatcttgc ggtacatcgt gtgataggt	660
	ggctctcgag attcttcttt ttaaccggca aagtaanaata atctcagctc cagcciaacg	720
	tcaattatca gagagagaaa aaaatatitt ttatgatgtt atcggaacc aaccgccita	780
	cgtgtcgaic ctggttcttg gccggcacgg cggaggaaag cgaccgacct cgcaacgcgc	840
	gcgcacggcg ccgcctgttt ggacttggtc tcccgcgact ccggtggcct cggcttatcg	900
[0025]	ccgcctctcc atctcaaccg tccgtttgga cactgtgaag ttgatccgtc gcgcaccage	960
	ctcggaggta acctaaactgc cgtactata aatccgggat ccggcctctc caatcccat	1020
	cgcca	1025
<210>	36	
<211>	1000	
<212>	DNA	
<213>	柳枝稷	
<400>	36	
	gcctagtgtc cctagattgc ctttgtctgt tatgttcaac ctctgttta agtctgtiga	60
	actctctgca ttgcgttctt agtgtctgtt tgtggttga ataagaacat gaagaacatg	120
	tgtctgtgga tcacatgact ttttttttg aaccggaaga tcacatgact ttcatggctt	180
	taagttcttg aactctgaaa tctggacccc tttttaagct ctgaacicat cttcttgc	240
	tttacctctg gtgttgatct taitgatgtt atgcagtcct gctgaaatag tcaatgtaga	300
	ttcatgactg actgatgtcg tttatgtgtt gtatgttgtt aacaagctga aggtctgtgt	360
	gtgtcttctc agttagacga agtgtcttt attgtagcgt gtatgtctgc tggatgattg	420
	atgaactgaa acattctgca tttagcact agcgagccaa aggigaagac tgagtttctg	480
	tagaccigtg ttttatgccc catgtctgtt cttaaatgc acttgatttt cacattagct	540
	ggatcataat ctgagcagac tactcaaaag tacaagtttc atctctgcta tgacgtttg	600
	ccactaggat ttcttttga tgatttgtt acaaaacctg taatctagtc aaaagaaaag	660
	ccaaaatttt tcttgtatg atttgtttac aaatcctcta atctagtcaa agaaaagcca	720

	aaattiatccc tccgggtccc ctacatcacg tagctatgig gcccgcaagc agaigaaagc	780
	agccccgtca gccgacgcgc agcccgacgc caacacatcc igctcctccc tgcggggcgc	840
	cggcgccggc gagggcgcac cggcgctgcc ccgiggccgc aggcacacgg tgcgcacig	900
	ccggcgcccc gggcgcgag gcacacggig ccgcactgcc gccgcctccc ctccggcat	960
	tgcgggacgg ctgggctact gtcccgccg ccttcccaat	1000
	<210> 37	
	<211> 1085	
	<212> DNA	
	<213> 柳枝稷	
	<400> 37	
	gtactcciac ctaatccccc ttaactgacg tctcctctat caggttgga atcttcgaat	60
	gatctgctgc ctggctcgtc gtccccctc gtaatgcaat gttccatca cgagttttt	120
	tttcatcat ctaatctatg cgggtgcgga agaattgtgg ctagtggagt agttttctgt	180
	gcttgatcgg tagattgat gtgtgggigt atggatgttt tctgaaaagt tcttggaat	240
	gtttacgctt tcaggccgca ggtcgttgc aaattgaita tgaagctat atgctttgga	300
	tctatcgatt tccagtttta ttcagatga gcccaaaaaa ttgtcgcat ttgtgtgaa	360
	ttagttggcc tttaggtctg cacattcatg gtgacggcac agttgctgct ggtgttgcg	420
	tgggacgagt tattatagtt gttttgttt ttccctgatt gattcacatt ttcaatgata	480
	actagccctt gtcacctaac caagtcagg ttgatccat ctgtgttctt cagctaccag	540
[0026]	tttgcataga tgatggtgta ttcatgtct ttagtaggac ttctgattc acatctaatt	600
	ctgtcatgaa tatagataac ttacatgct ttgatatac ttatatttg aactgttca	660
	tgtccagcct attttgata attgagtga ttggcttttg atgcttgaat tattcacatg	720
	ttcttgata attgacctgt gtcacctagt tgaatgttt ttgaggigcc accgtctgt	780
	tcagctgatt ttgtatttcg attgctctag ttaactcttt gattatgcag ctagtgttt	840
	gtcatatgta gctttatagg ctctgatgt ccttgatat agttcagict acitgtcaag	900
	ttgctttaca agtagtagct ctgattctat ttggcttccc gaticagagc ttigcaaatt	960
	gcttgttgtt acattacatc atattacttg aattgcagtt attaatggt tggattgtg	1020
	ctgtttactt ctacattttt tgcgttttta tattatacta aaatgtttgt gtgtctgtt	1080
	ttcag	1085
	<210> 38	
	<211> 77	
	<212> DNA	
	<213> 柳枝稷	
	<400> 38	
	caagttcgcg atctctcga ttacaaaac gccgagaaga cccgagcaga gaagttccct	60
	ccgacgcct tgcgaag	77
	<210> 39	
	<211> 2187	
	<212> DNA	
	<213> 柳枝稷	

[0027]	<400> 39	
	ttgaatttta atttcaaatt ttcagggtta gtagtgaca tcacaataca taittagaaa	60
	aagtttiata attttctctc gttagtttic atataattit gaactccaac gattaatcta	120
	ttattaaata tcccgatcta tcaaaataat gataaaaatt taigattaat ttttctaaca	180
	tgtgttaagg tgtgtactat cgtctiataa aatttcaact taaaactcca cctatataig	240
	gagaaatgaa aaagacgaat tacagtaggg agtaattiga accaaatgga atagittgag	300
	ggtaaaatga actaaacaat agttiaggag gttattcaga ttttagttat agttgagagg	360
	agtaatttag actttttctc atcttgaatt gttagcggct ctcctatcgg atatcggatg	420
	gagttttta gcccaacata acttcatcgg ggcctaaacg ttcttctcgc cagcttaggg	480
	agaacatttt gcccaigata tctgttttcc tttttttcia ttttactgg tattatagga	540
	gggaataata caacgtgttc acctttgggt tcatcttgt tccatctgaa ttatctaaa	600
	actgtgtttg aacttcgtaa gaattttgtt cgtatcttcc ggtacatcgt gttgataggt	660
	ggcctccgag attcttcttt ttaaccggca aagtaaaata atctcagctc cagcctaacg	720
	tcaattatca gagagagaaa aaaatatttt ttatgatig atcggaacc aaccgctta	780
	cgtgtcgtc ctgttctcgg gccggcaccg cggaggaaag cgaccgacct cgcaacgccg	840
	gcgcacggcg ccgccgtgtt ggacttggtc tcccgcgact ccgtgggcct cggcitatcg	900
	ccgccgtccc atctcaaccg tccgtttgga cacttggaag ttgatccgtc gcgcaccagc	960
	ctcggaggta accttaactgc ccgtactata aatccgggat ccggcctctc caatcccat	1020
	cgcacaaagt tccggtctc tcgatttcac aaatcgccga gaagaccga gcagagaagt	1080
	tccctccgat cgccttgcca aggtactcct acctaatcct ccttaactga tctctctct	1140
	atcacgttgg taatcttcca atgactcgtt gccgtggctg ctgttcccc tcgtatgca	1200
	ctgtttccat cagcagtttt ttttttctc atctaataa tgcggttgcg gaagaaifgt	1260
	ggctagtggg gtagtttctt gtccttgcg ggtagattcg atgtgtgggt gtatggatgt	1320
	tttctgaaaa gtgtctggat tagtttacgc ttccaggccg cagggtctgtt cgaattgat	1380
	taigaagctt atatgctttg gatctatcga ttccagttt tatcagaig taggccaaaa	1440
	aattgtcggc atttgttgg aattagtgg cccttaggtc tgcacattca tggtagcggc	1500
	acagtgtcgt ctggcgtttg cgtgggacga gttattatag ttgtttttgt ttttccctga	1560
	ttagttcaca ttttcaatga taactagcct ttgtcaccta accaagtcga ggttgatcct	1620
	atctgtgttc ttcagctacc agttgcata gatgatgggt tattcgattg ctttagtagg	1680
	ccctctgatt tcacatctaa ttctgtcatg aataatgata actttacatg cttttgatat	1740
	actttatatt tgaactgttc actgtccagc ctattttgga taattgagtg cattggcctt	1800
	tgatgcctga attattcaca tgttcttgga taattgacct gtgtcaccta gttgactgtt	1860
	ttttgagggt ccaccgtctt gttcagctga ttgtgtatc cgtttgtctt agttaatctt	1920
	ttgattatgc agctagtgtt ttgtcatatg tagctttata ggcttctgat gtccttgat	1980
	atagttcagt ctacttgica agttgcctta caagtagtag ctctgattct atttggcttc	2040
	ctgagtcaga gctttgcaa ttgtttgttg ttacattaca tcataattact tgaattgcag	2100

	ttatittaatg gtiggatitg tgcgttttac ticiacattt ttgctgttt tataattatac	2160
	taaaatgttt ggttgcctgc ttttcag	2187
	<210> 40	
	<211> 711	
	<212> DNA	
	<213> 二穗短柄草	
	<400> 40	
	ggcgtcagga ctggcgaagt ctggactctg cagggccgaa ctgctgaaga cgaagcagag	60
	gaagagaaag ggaagtgttc gacttgtaat ttgtaggggt tttttttaga ggaacttgta	120
	atttgttaggt gggctggcct cgttggaaaa acgatgcctg ctggttgggc tgggccgatg	180
[0028]	tacgtttgca aacaacttgi ggcggccctg tctggacgag caggagtttc tttttgttc	240
	tcactttict ggcttcttt agttacggag tacctttgti tttttaaagg agttaccitt	300
	tttttaggaa ttctttagt accittcgti tgcctcaaa aaatatttaa ctttcgttt	360
	tttcatittt aatttttgca actatttacg agtttcatga atgcttattt tccagcatat	420
	cattatttgc aagtattttt atgccgtatg tatggacga gagccatcgg gactgttcca	480
	gagactgcgt ggtggggacg gctcccaacc gccitttcta tctcgttctg catccggigg	540
	ccgacttggc tcgcgcgiga gccgtgacgt aacagacttg gtctcttccc catctggcca	600
	tctataaatt ccccatcga tcgaccttc ctctcccaa tccagcacc ccatccga	660
	tcgaaaattc tccgcaacag caagcgtatg atctagcgaa tccccgtcaa g	711

										节段1							
2341801	(1)	1	15	25	35	45	55	65	75	88							
2341802	(1)																
2341803	(1)																
二模制钢带 1801	(1)																
二模制钢带 1802	(1)																
共有	(1)																
										节段2							
2341801	(88)	25	105	115	125	135	145	155	165	175							
2341802	(88)																
二模制钢带 1801	(1)																
二模制钢带 1802	(1)																
共有	(88)																
										节段3							
2341801	(177)	177	190	205	215	225	235	245	255	264							
2341802	(177)																
二模制钢带 1801	(1)																
二模制钢带 1802	(1)																
共有	(177)																
										节段4							
2341801	(265)	265	275	290	295	305	315	325	335	342							
2341802	(265)																
二模制钢带 1801	(1)																
二模制钢带 1802	(1)																
共有	(265)																
										节段5							
2341801	(353)	353	365	375	385	395	405	415	425	430							
2341802	(353)																
二模制钢带 1801	(1)																
二模制钢带 1802	(1)																
共有	(353)																
										节段6							
2341801	(441)	441	450	465	475	485	495	505	515	528							
2341802	(441)																
二模制钢带 1801	(1)																
二模制钢带 1802	(1)																
共有	(441)																
										节段7							
2341801	(528)	528	534														
2341802	(528)																
二模制钢带 1801	(1)																
二模制钢带 1802	(1)																
共有	(528)																

图1

图 2

											第段:11
		(871)	871	880	890	900	910	920	930	940	950
二端部部 1801	(1801)	 第段:11									
二端部部 1802	(1802)	 第段:11									
二端部部 1803	(1803)	 第段:11									
二端部部 1804	(1804)	 第段:11									
二端部部 1805	(1805)	 第段:11									
二端部部 1806	(1806)	 第段:11									
二端部部 1807	(1807)	 第段:11									
二端部部 1808	(1808)	 第段:11									
二端部部 1809	(1809)	 第段:11									
二端部部 1810	(1810)	 第段:11									
二端部部 1811	(1811)	 第段:11									
二端部部 1812	(1812)	 第段:11									
二端部部 1813	(1813)	 第段:11									
二端部部 1814	(1814)	 第段:11									
二端部部 1815	(1815)	 第段:11									
二端部部 1816	(1816)	 第段:11									
二端部部 1817	(1817)	 第段:11									
二端部部 1818	(1818)	 第段:11									
二端部部 1819	(1819)	 第段:11									
二端部部 1820	(1820)	 第段:11									
二端部部 1821	(1821)	 第段:11									
二端部部 1822	(1822)	 第段:11									
二端部部 1823	(1823)	 第段:11									
二端部部 1824	(1824)	 第段:11									
二端部部 1825	(1825)	 第段:11									
二端部部 1826	(1826)	 第段:11									
二端部部 1827	(1827)	 第段:11									
二端部部 1828	(1828)	 第段:11									
二端部部 1829	(1829)	 第段:11									
二端部部 1830	(1830)	 第段:11									
二端部部 1831	(1831)	 第段:11									
二端部部 1832	(1832)	 第段:11									
二端部部 1833	(1833)	 第段:11									
二端部部 1834	(1834)	 第段:11									
二端部部 1835	(1835)	 第段:11									
二端部部 1836	(1836)	 第段:11									
二端部部 1837	(1837)	 第段:11									
二端部部 1838	(1838)	 第段:11									
二端部部 1839	(1839)	 第段:11									
二端部部 1840	(1840)	 第段:11									
二端部部 1841	(1841)	 第段:11									
二端部部 1842	(1842)	 第段:11									
二端部部 1843	(1843)	 第段:11									
二端部部 1844	(1844)	 第段:11									
二端部部 1845	(1845)	 第段:11									
二端部部 1846	(1846)	 第段:11									
二端部部 1847	(1847)	 第段:11									
二端部部 1848	(1848)	 第段:11									
二端部部 1849	(1849)	 第段:11									
二端部部 1850	(1850)	 第段:11									
二端部部 1851	(1851)	 第段:11									
二端部部 1852	(1852)	 第段:11									
二端部部 1853	(1853)	 第段:11									
二端部部 1854	(1854)	 第段:11									
二端部部 1855	(1855)	 第段:11									
二端部部 1856	(1856)	 第段:11									
二端部部 1857	(1857)	 第段:11									
二端部部 1858	(1858)	 第段:11									
二端部部 1859	(1859)	 第段:11									
二端部部 1860	(1860)	 第段:11									
二端部部 1861	(1861)	 第段:11									
二端部部 1862	(1862)	 第段:11									
二端部部 1863	(1863)	 第段:11									
二端部部 1864	(1864)	 第段:11									
二端部部 1865	(1865)	 第段:11									
二端部部 1866	(1866)	 第段:11									
二端部部 1867	(1867)	 第段:11									
二端部部 1868	(1868)	 第段:11									
二端部部 1869	(1869)	 第段:11									
二端部部 1870	(1870)	 第段:11									
二端部部 1871	(1871)	 第段:11									
二端部部 1872	(1872)	 第段:11									
二端部部 1873	(1873)	 第段:11									
二端部部 1874	(1874)	 第段:11									
二端部部 1875	(1875)	 第段:11									
二端部部 1876	(1876)	 第段:11									
二端部部 1877	(1877)	 第段:11									
二端部部 1878	(1878)	 第段:11									
二端部部 1879	(1879)	 第段:11									
二端部部 1880	(1880)	 第段:11									
二端部部 1881	(1881)	 第段:11									
二端部部 1882	(1882)	 第段:11									
二端部部 1883	(1883)	 第段:11									
二端部部 1884	(1884)	 第段:11									
二端部部 1885	(1885)	 第段:11									
二端部部 1886	(1886)	 第段:11									
二端部部 1887	(1887)	 第段:11									
二端部部 1888	(1888)	 第段:11									
二端部部 1889	(1889)	 第段:11									
二端部部 1890	(1890)	 第段:11									
二端部部 1891	(1891)	 第段:11									
二端部部 1892	(1892)	 第段:11									
二端部部 1893	(1893)	 第段:11									
二端部部 1894	(1894)	 第段:11									
二端部部 1895	(1895)	 第段:11									
二端部部 1896	(1896)	 第段:11									
二端部部 1897	(1897)	 第段:11									
二端部部 1898	(1898)	 第段:11									
二端部部 1899	(1899)	 第段:11									
二端部部 1900	(1900)	 第段:11									
二端部部 1901	(1901)	 第段:11									
二端部部 1902	(1902)	 第段:11									
二端部部 1903	(1903)	 第段:11									
二端部部 1904	(1904)	 第段:11									
二端部部 1905	(1905)	 第段:11									
二端部部 1906	(1906)	 第段:11									
二端部部 1907	(1907)	 第段:11									
二端部部 1908	(1908)	 第段:11									
二端部部 1909	(1909)	 第段:11									
二端部部 1910	(1910)	 第段:11									
二端部部 1911	(1911)	 第段:11									
二端部部 1912	(1912)	 第段:11									
二端部部 1913	(1913)	 第段:11									
二端部部 1914	(1914)	 第段:11									
二端部部 1915	(1915)	 第段:11									
二端部部 1916	(1916)	 第段:11									
二端部部 1917	(1917)	 第段:11									
二端部部 1918	(1918)	 第段:11									
二端部部 1919	(1919)	 第段:11									
二端部部 1920	(1920)	 第段:11									
二端部部 1921	(1921)	 第段:11									
二端部部 1922	(1922)	 第段:11									
二端部部 1923	(1923)	 第段:11									
二端部部 1924	(1924)	 第段:11									
二端部部 1925	(1925)	 第段:11									
二端部部 1926	(1926)	 第段:11									
二端部部 1927	(1927)	 第段:11									
二端部部 1928	(1928)	 第段:11									
二端部部 1929	(1929)	 第段:11									
二端部部 1930	(1930)	 第段:11									
二端部部 1931	(1931)	 第段:11									
二端部部 1932	(1932)	 第段:11									
二端部部 1933	(1933)	 第段:11									
二端部部 1934	(1934)	 第段:11									
二端部部 1935	(1935)	 第段:11									
二端部部 1936	(1936)	 第段:11									
二端部部 1937	(1937)	 第段:11									
二端部部 1938	(1938)	 第段:11									
二端部部 1939	(1939)	 第段:11									
二端部部 1940	(1940)	 第段:11									
二端部部 1941	(1941)	 第段:11									
二端部部 1942	(1942)	 第段:11									
二端部部 1943	(1943)	 第段:11									
二端部部 1944	(1944)	 第段:11									
二端部部 1945	(1945)	 第段:11									
二端部部 1946	(1946)	 第段:11									
二端部部 1947	(1947)	 第段:11									
二端部部 1948	(1948)	 第段:11									
二端部部 1949	(1949)	 第段:11									
二端部部 1950	(1950)	 第段:11									
二端部部 1951	(1951)	 第段:11									
二端部部 1952	(1952)	 第段:11									
二端部部 1953	(1953)	 第段:11									
二端部部 1954	(1954)	 第段:11									
二端部部 1955	(1955)	 第段:11									
二端部部 1956	(1956)	 第段:11									
二端部部 1957	(1957)	 第段:11									
二端部部 1958	(1958)	 第段:11									
二端部部 1959	(1959)	 第段:11									
二端部部 1960	(1960)	 第段:11									
二端部部 1961	(1961)	 第段:11									
二端部部 1962	(1962)	 第段:11									
二端部部 1963	(1963)	 第段:11									
二端部部 1964	(1964)	 第段:11									
二端部部 1965	(1965)	 第段:11									
二端部部 1966	(1966)	 第段:11									
二端部部 1967	(1967)	 第段:11									
二端部部 1968	(1968)	 第段:11									
二端部部 1969	(1969)	 第段:11									
二端部部 1970	(1970)	 第段:11									
二端部部 1971	(1971)	 第段:11									
二端部部 1972	(1972)	 第段:11									
二端部部 1973	(1973)	 第段:11									
二端部部 1974	(1974)	 第段:11									
二端部部 1975	(1975)	 第段:11									
二端部部 1976	(1976)	 第段:11									
二端部部 1977	(1977)	 第段:11									
二端部部 1978	(1978)	 第段:11									
二端部部 1979	(1979)	 第段:11									
二端部部 1980	(1980)	 第段:11									
二端部部 1981	(1981)	 第段:11									
二端部部 1982	(1982)	 第段:11									
二端部部 1983	(1983)	 第段:11									
二端部部 1984	(1984)	 第段:11									
二端部部 1985	(1985)	 第段:11									
二端部部 1986	(1986)	 第段:11									
二端部部 1987	(1987)	 第段:11									
二端部部 1988	(1988)	 第段:11									
二端部部 1989	(1989)	 第段:11									
二端部部 1990	(1990)	 第段:11									
二端部部 1991	(1991)	 第段:11									
二端部部 1992	(1992)	 第段:11									
二端部部 1993	(1993)	 第段:11									
二端部部 1994	(1994)	 第段:11									
二端部部 1995	(1995)	 第段:11									
二端部部 1996	(1996)	 第段:11									
二端部部 1997	(1997)	 第段:11									
二端部部 1998	(1998)	 第段:11									
二端部部 1999	(1999)	 第段:11									
二端部部 2000	(2000)	 第段:11									
二端部部 2001	(2001)	 第段:11									
二端部部 2002	(2002)	 第段:11									
二端部部 2003	(2003)	 第段:11									
二端部部 2004	(2004)	 第段:11									
二端部部 2005	(2005)	 第段:11									
二端部部 2006	(2006)	 第段:11									
二端部部 2007	(2007)	 第段:11									
二端部部 2008	(2008)	 第段:11									
二端部部 2009	(2009)	 第段:11									
二端部部 2010	(2010)	 第段:11									
二端部部 2011	(2011)	 第段:11									
二端部部 2012	(2012)	 第段:11									
二端部部 2013	(2013)	 第段:11									
二端部部 2014	(2014)	 第段:11									
二端部部 2015	(2015)	 第段:11									
二端部部 2016	(2016)	 第段:11									
二端部部 2017	(2017)	 第段:11									
二端部部 2018	(2018)	 第段:11									
二端部部 2019	(2019)	 第段:11									
二端部部 2020	(2020)	 第段:11									
二端部部 2021	(2021)	 第段:11									
二端部部 2022	(2022)	 第段:11									
二端部部 2023	(2023)	 第段:11									
二端部部 2024	(2024)	 第段:11									
二端部部 2025	(2025)	 第段:11									
二端部部 2026	(2026)	 第段:11									
二端部部 2027	(2027)	 第段:11									
二端部部 2028	(2028)	 第段:11									
二端部部 2029	(2029)	 第段:11									
二端部部 2030	(2030)	 第段:11									
二端部部 2031	(2031)	 第段:11									
二端部部 2032	(2032)	 第段:11									
二端部部 2033	(2033)	 第段:11									
二端部部 2034	(2034)	 第段:11									
二端部部 2035	(2035)	 第段:11									
二端部部 2036	(2036)	 第段:11									
二端部部 2037	(2037)	 第段:11									
二端部部 2038	(2038)	 第段:11									
二端部部 2039	(2039)	 第段:11									
二端部部 2040	(2040)	 第段:11									
二端部部 2041	(2041)	 第段:11									
二端部部 2042	(2042)	 第段:11									
二端部部 2043	(2043)	 第段:11									
二端部部 2044	(2044)	 第段:11									
二端部部 2045	(2045)	 第段:11									
二端部部 2046	(2046)	 第段:11									
二端部部 2047	(2047)	 第段:11									
二端部部 2048	(2048)	 第段:11									
二端部部 2049	(2049)	 第段:11									
二端部部 2050	(2050)	 第段:11									
二端部部 2051	(2051)	 第段:11									
二端部部 2052	(2052)	 第段:11									
二端部部 2053	(2053)	 第段:11									
二端部部 2054	(2054)	 第段:11									
二端部部 2055	(2055)	 第段:11									
二端部部 2056	(2056)	 第段:11									
二端部部 2057	(2057)	 第段:11									
二端部部 2058	(2058)	 第段:11									
二端部部 2059	(2059)	 第段:11									
二端部部 2060	(2060)	 第段:11									
二端部部 2061	(2061)	 第段:11									
二端部部 2062	(2062)	 第段:11									
二端部部 2063	(2063)	 第段:11									
二端部部 2064	(2064)	 第段:11									
二端部部 2065	(2065)	 第段:11									
二端部部 2066	(2066)	 第段:11									
二端部部 2067	(2067)	 第段:11									
二端部部 2068	(2068)	 第段:11									
二端部部 2069	(2069)	 第段:11									
二端部部 2070	(2070)	 第段:11									
二端部部 2071	(2071)	 第段:11									
二端部部 2072	(2072)	 第段:11									
二端部部 2073	(2073)	 第段:11									
二端部部 2074	(2074)	 第段:11									
二端部部 2075	(2075)	 第段:11									
二端部部 2076	(2076)	 第段:11									
二端部部 2077	(2077)	 第段:11									
二端部部 2078	(2078)	 第段:11									
二端部部 2079	(2079)	 第段:11									
二端部部 2080	(2080)	 第段:11									
二端部部 2081	(2081)	 第段:11									
二端部部 2082	(2082)	 第段:11									
二端部部 2083	(2083)	 第段:11									
二端部部 2084	(2084)	 第段:11									
二端部部 2085	(2085)	 第段:11									
二端部部 2086											

图2续

图2续

图2续

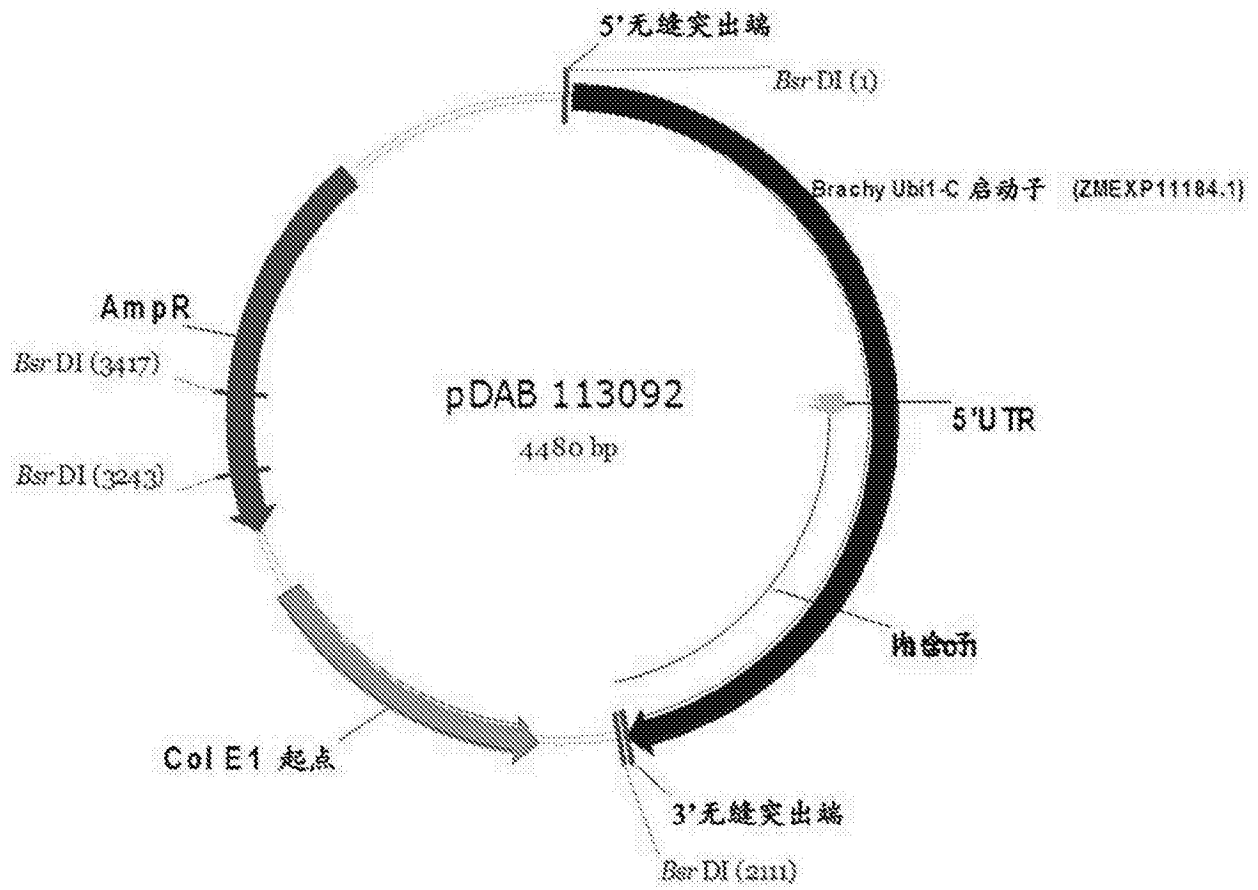


图4

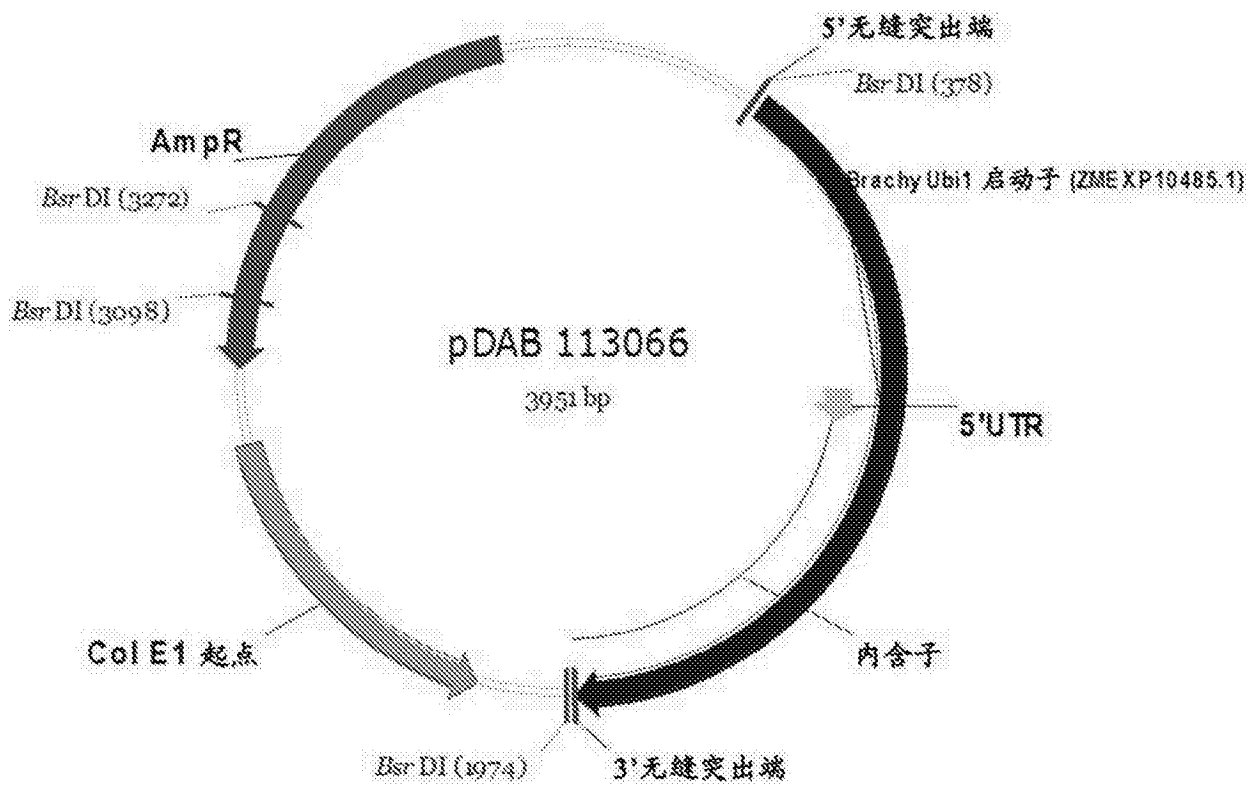


图5

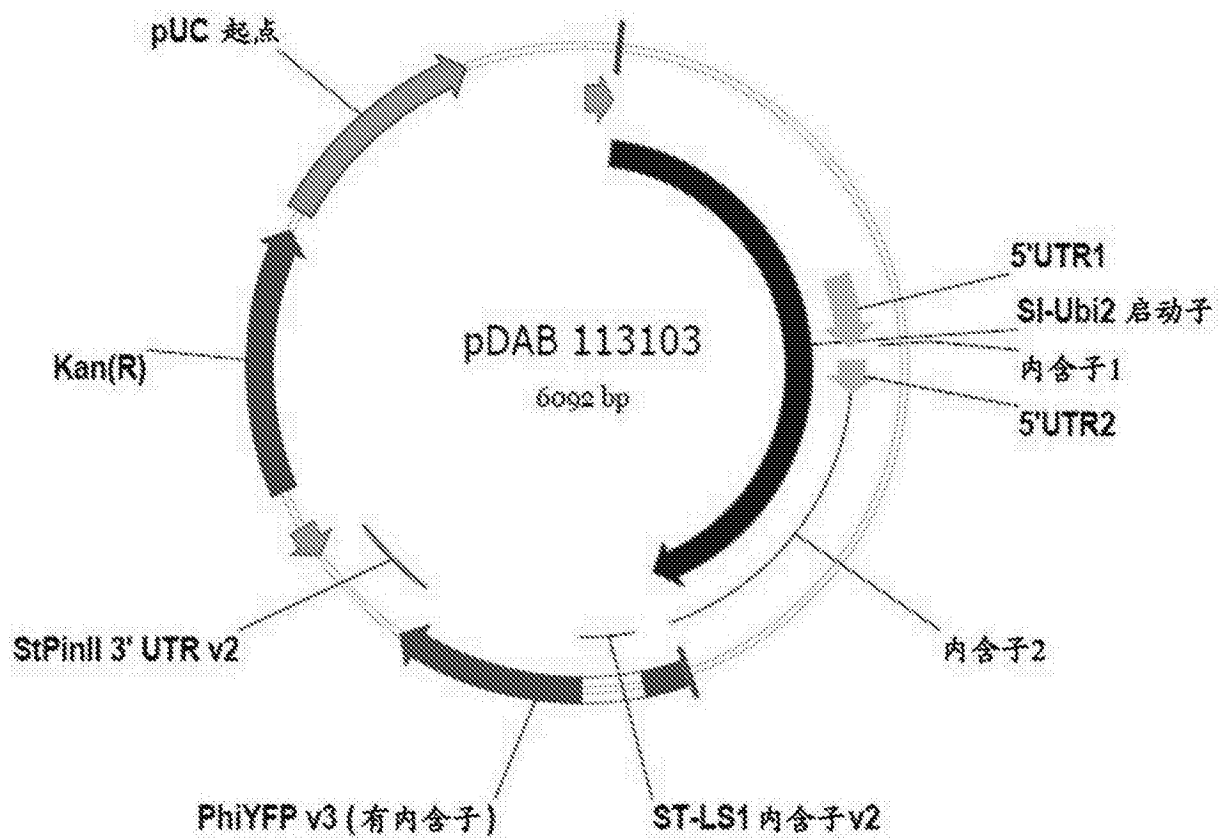


图6

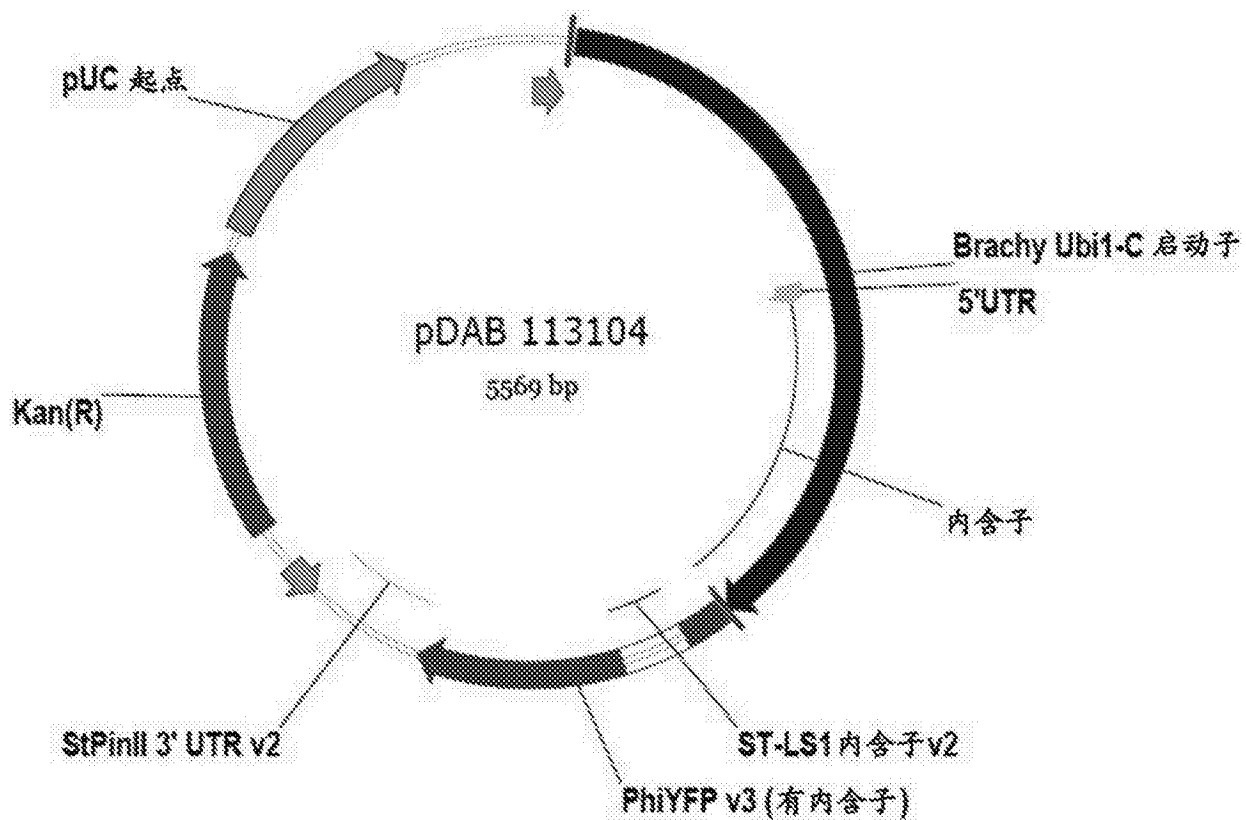


图7

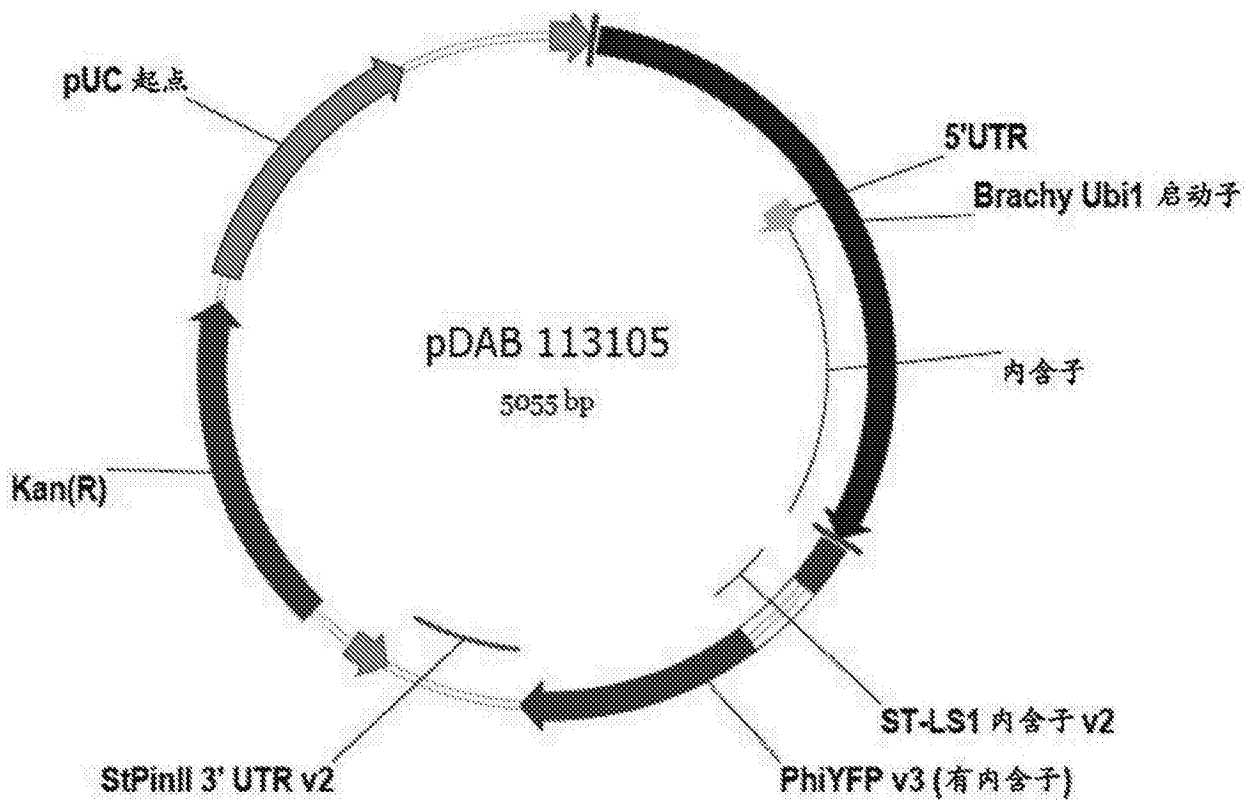


图8

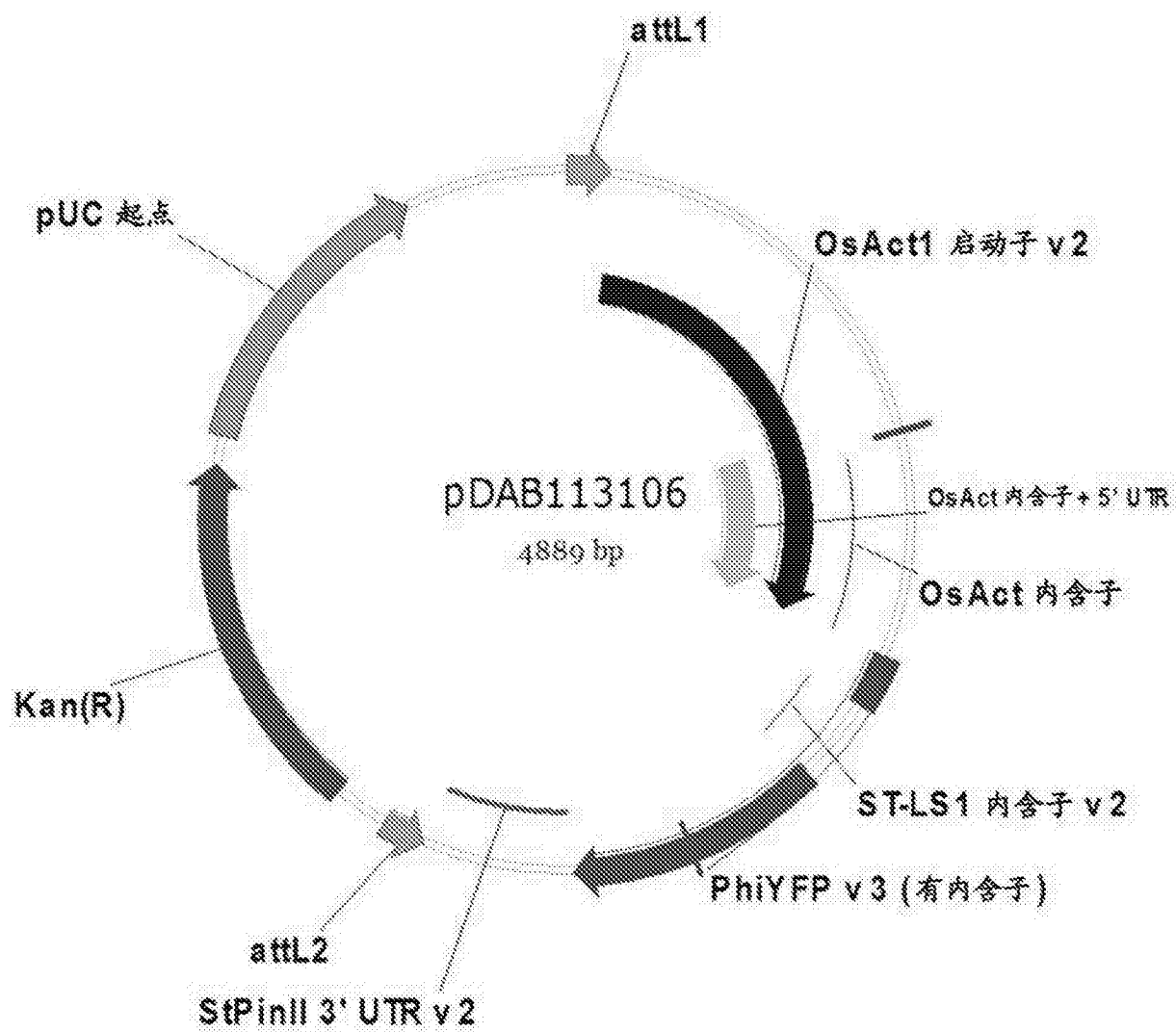


图9

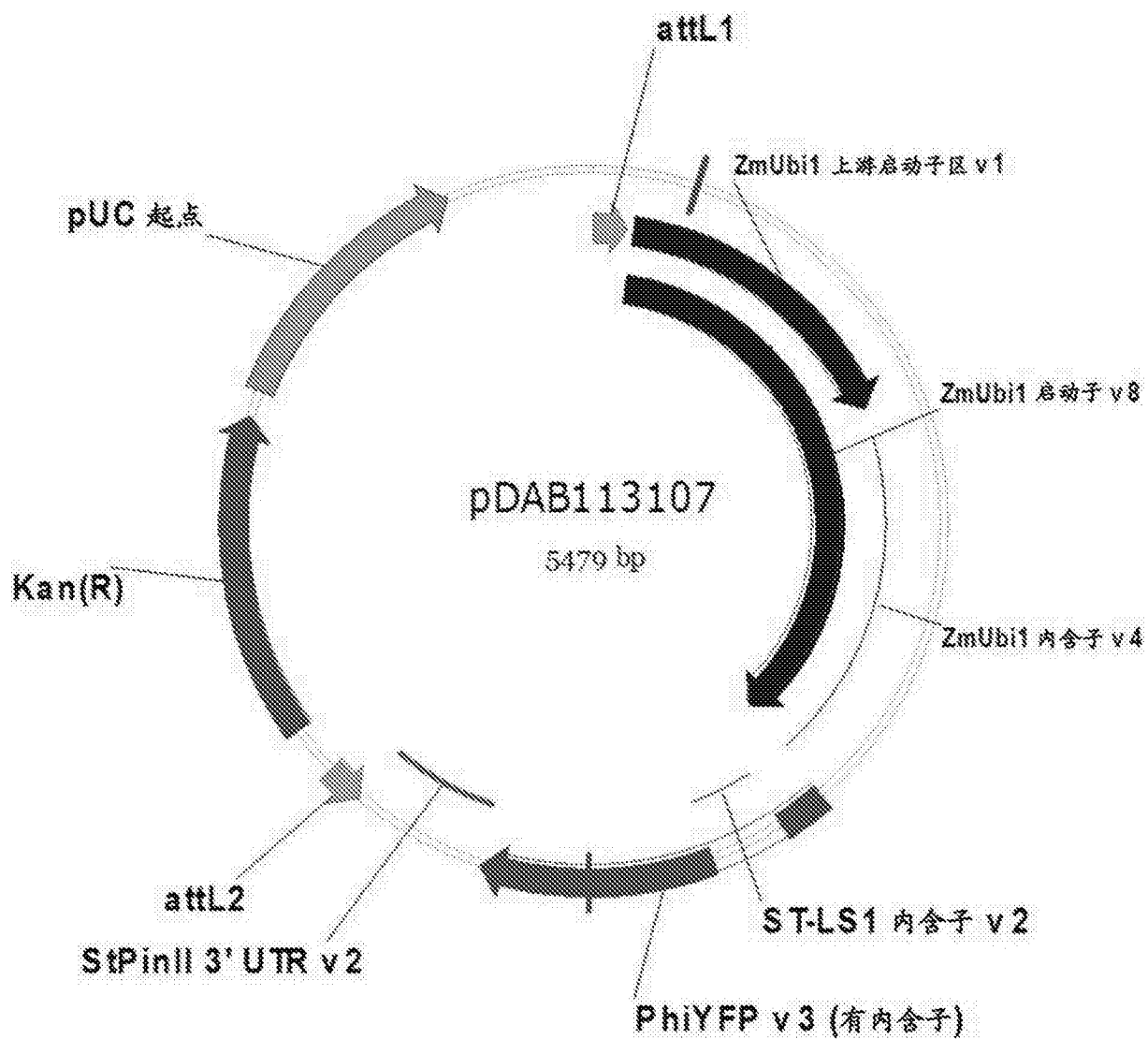


图10

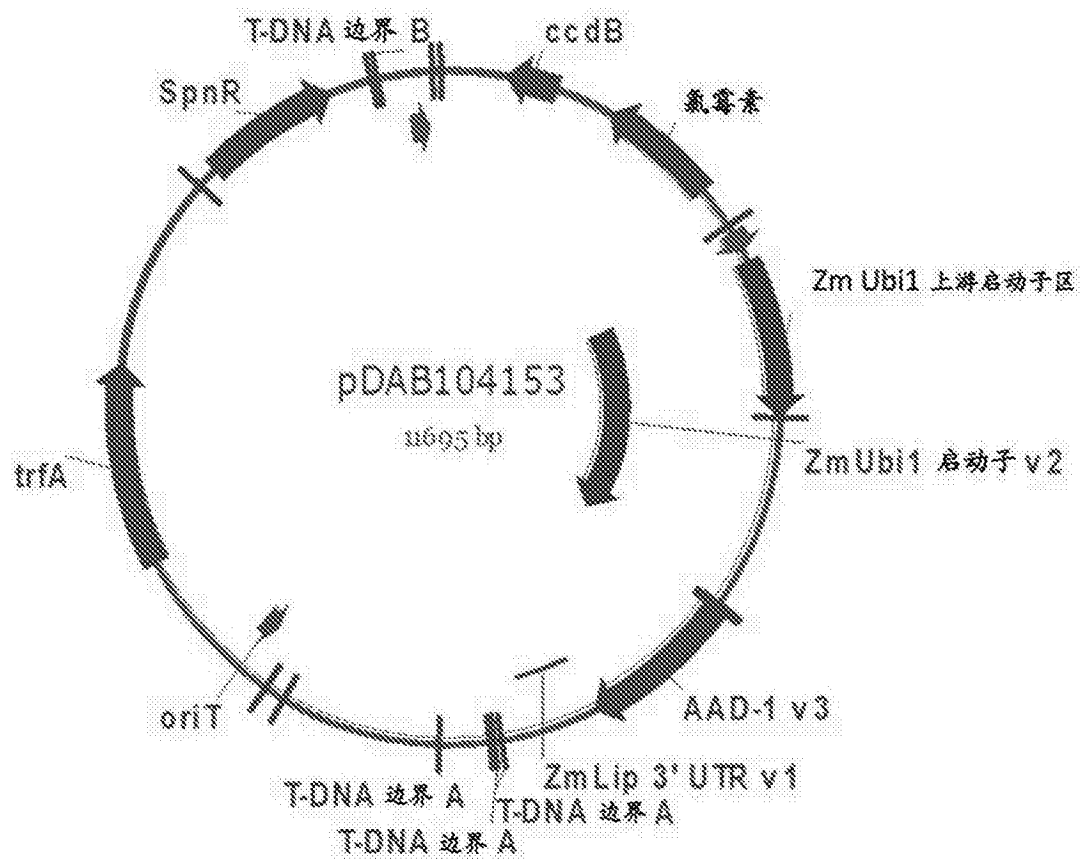


图11

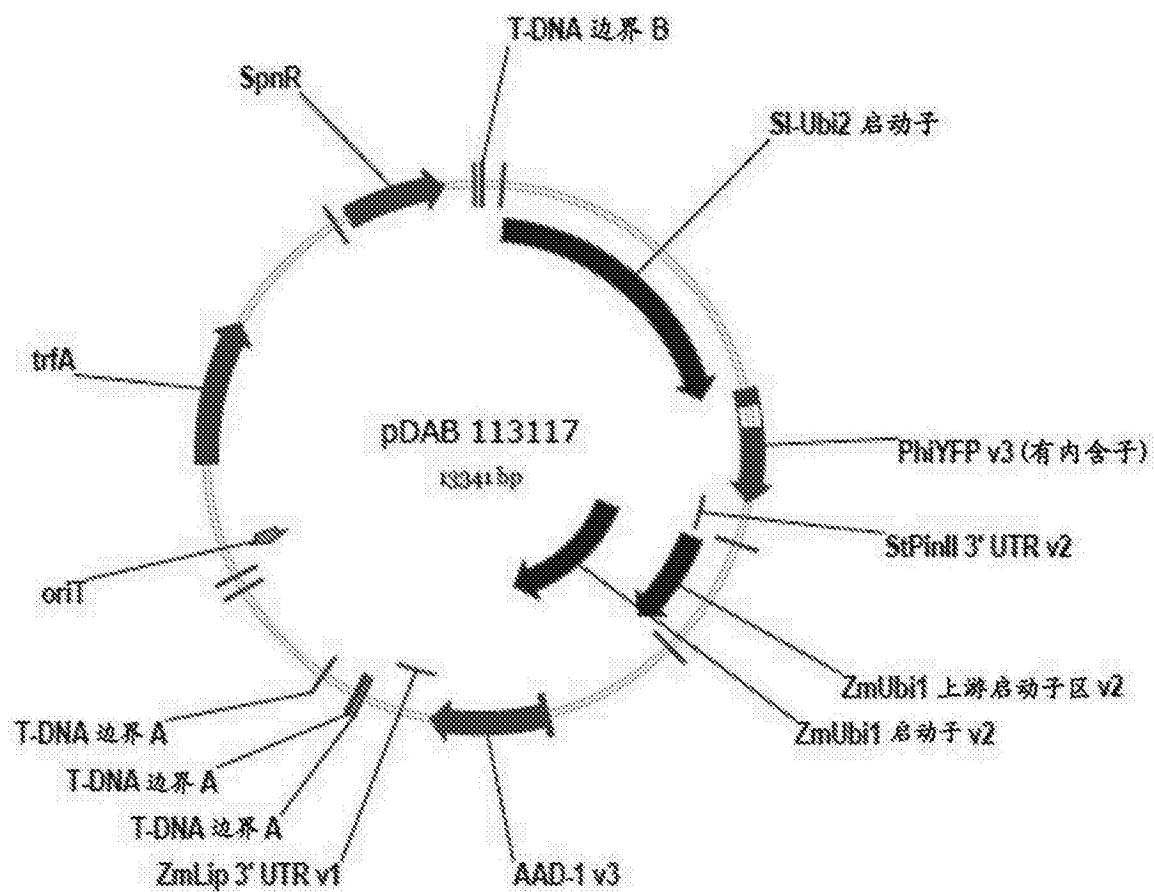


图12

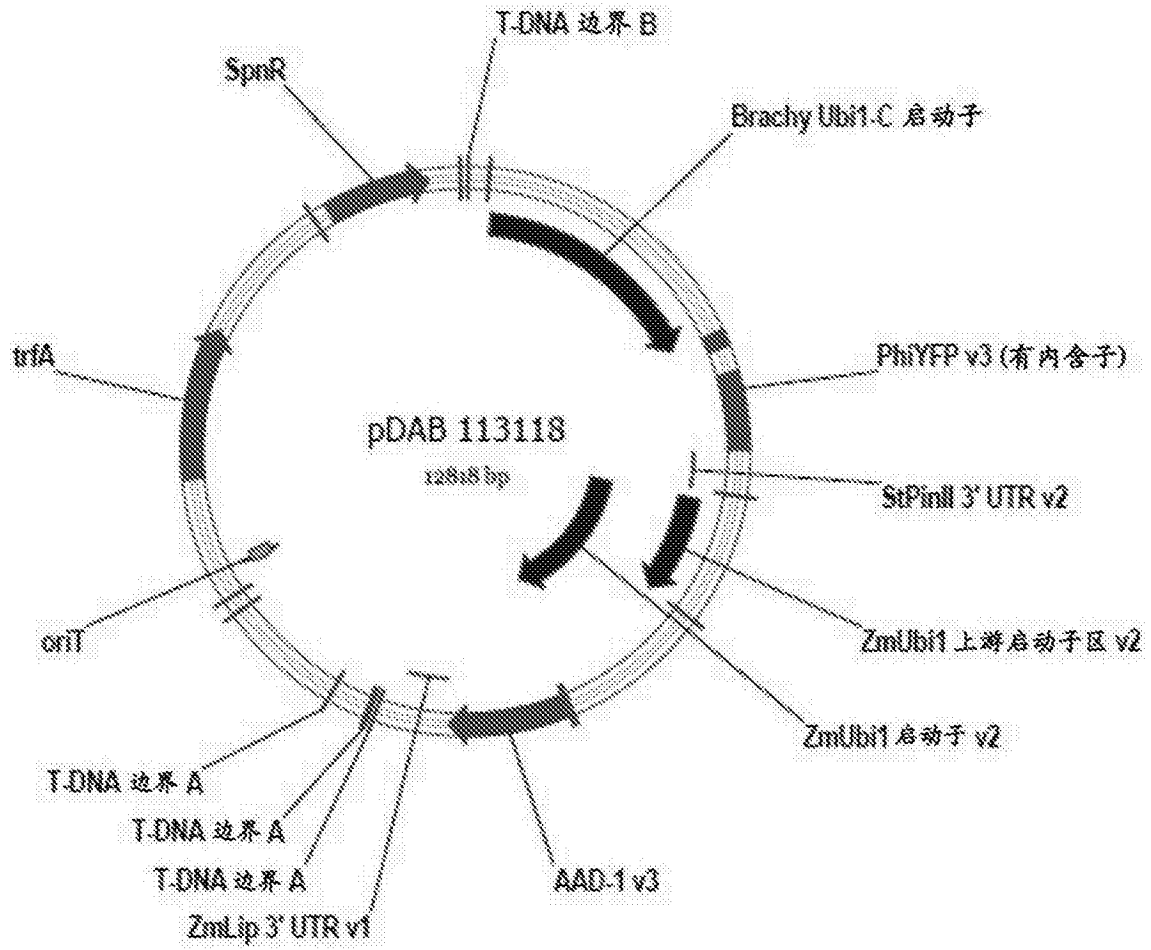


图13

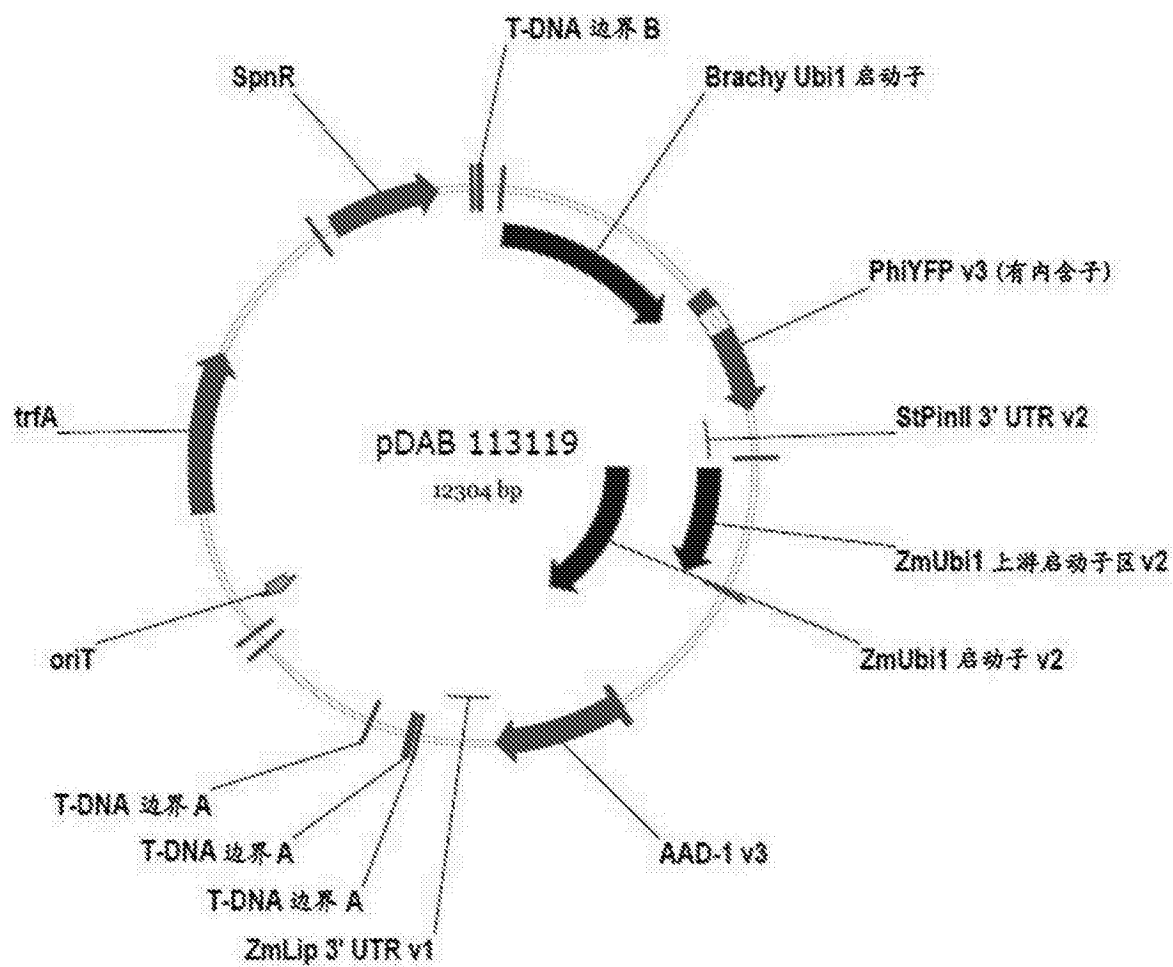


图14

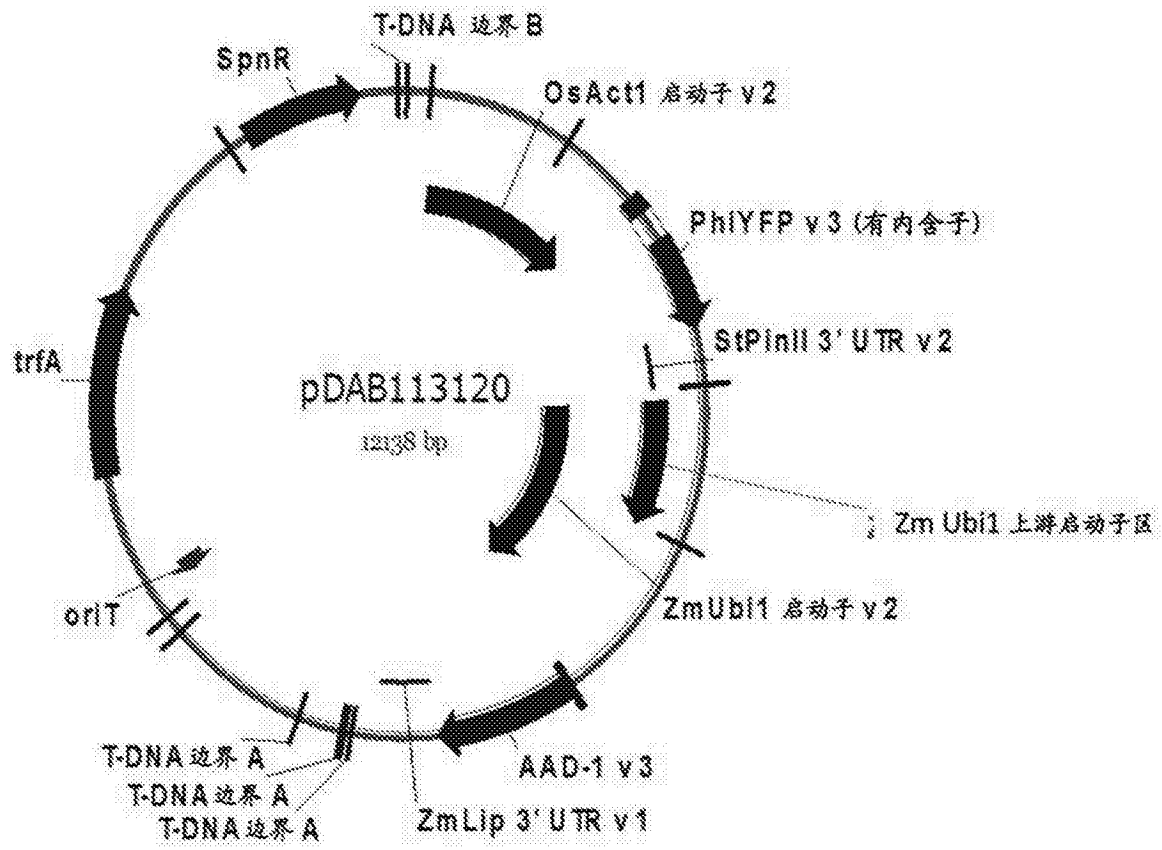


图15

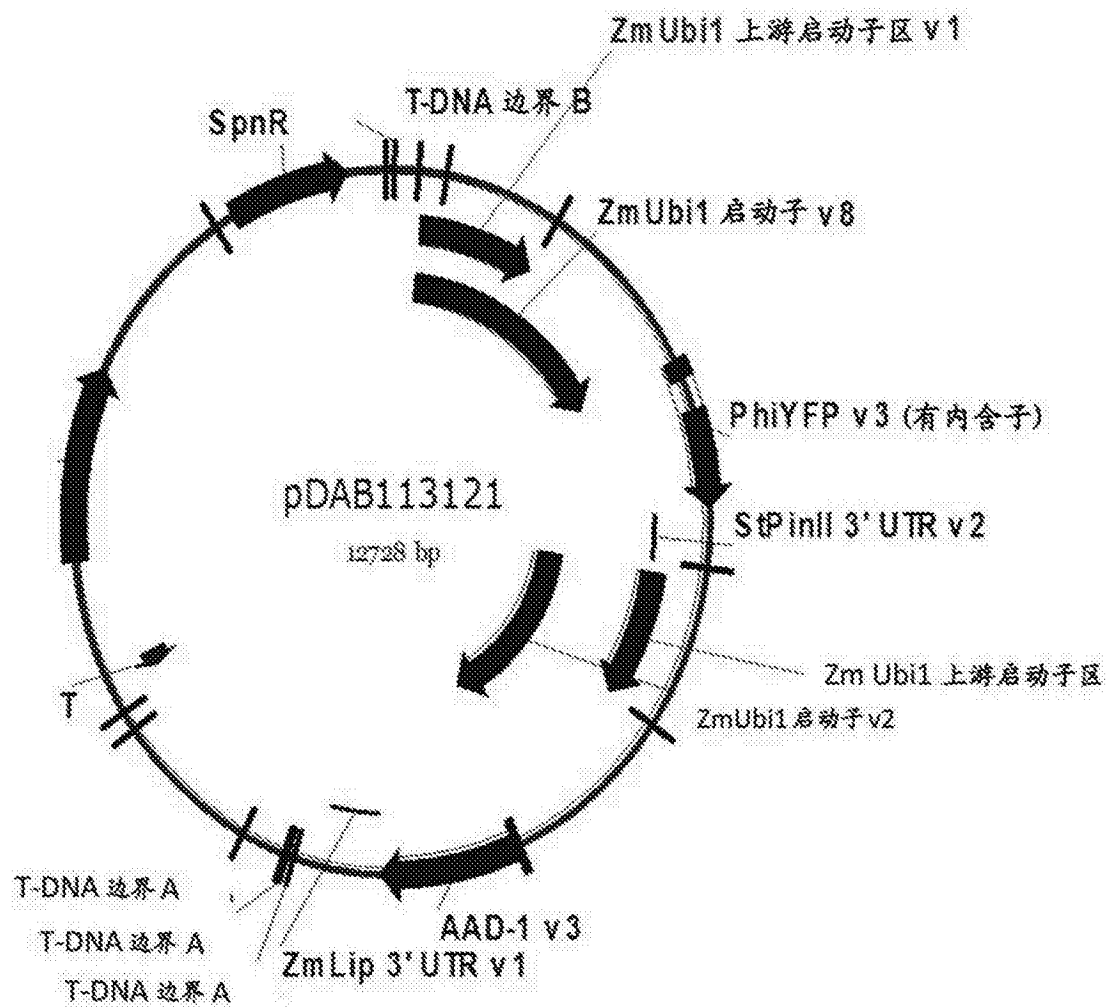


图16

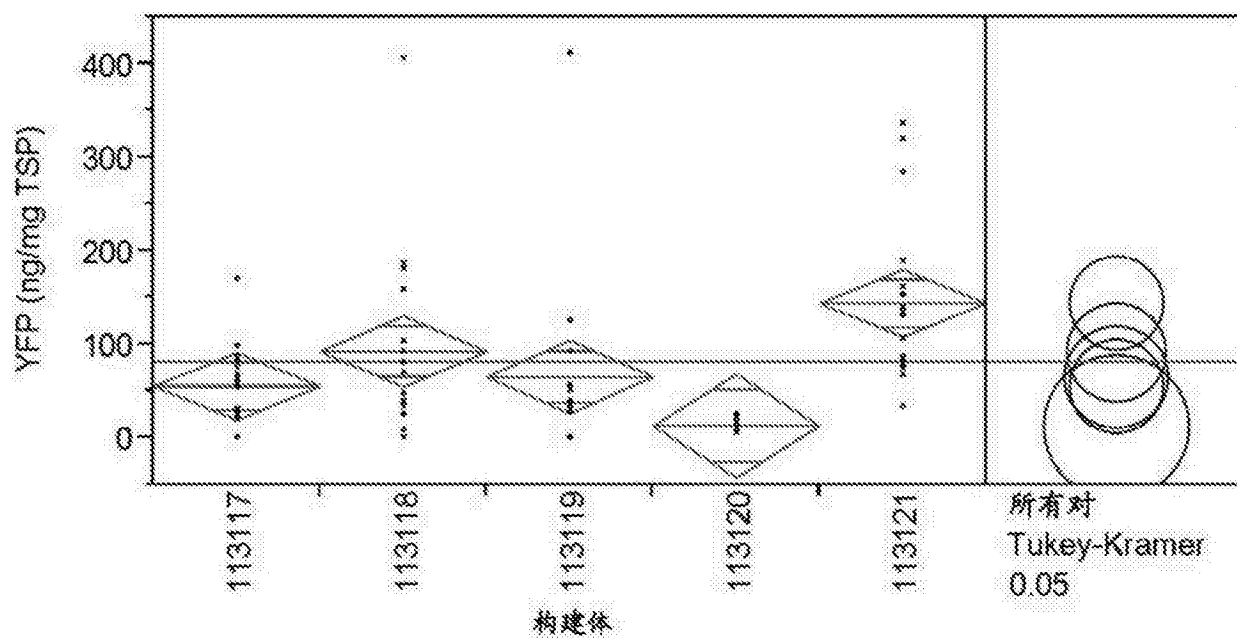


图17

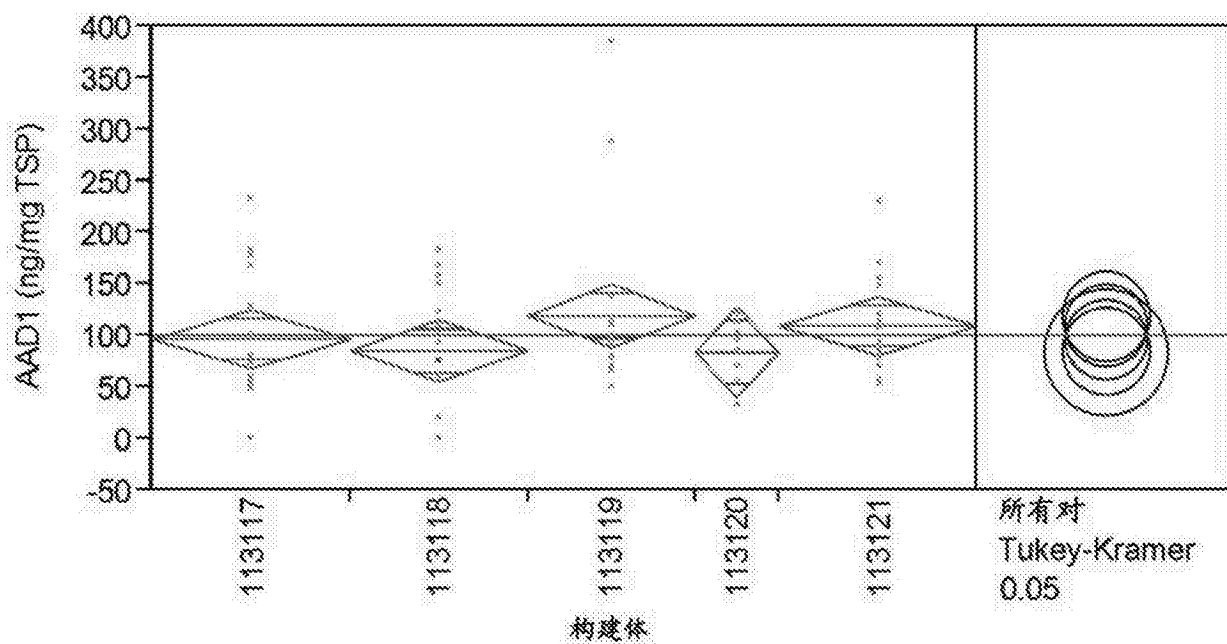


图18

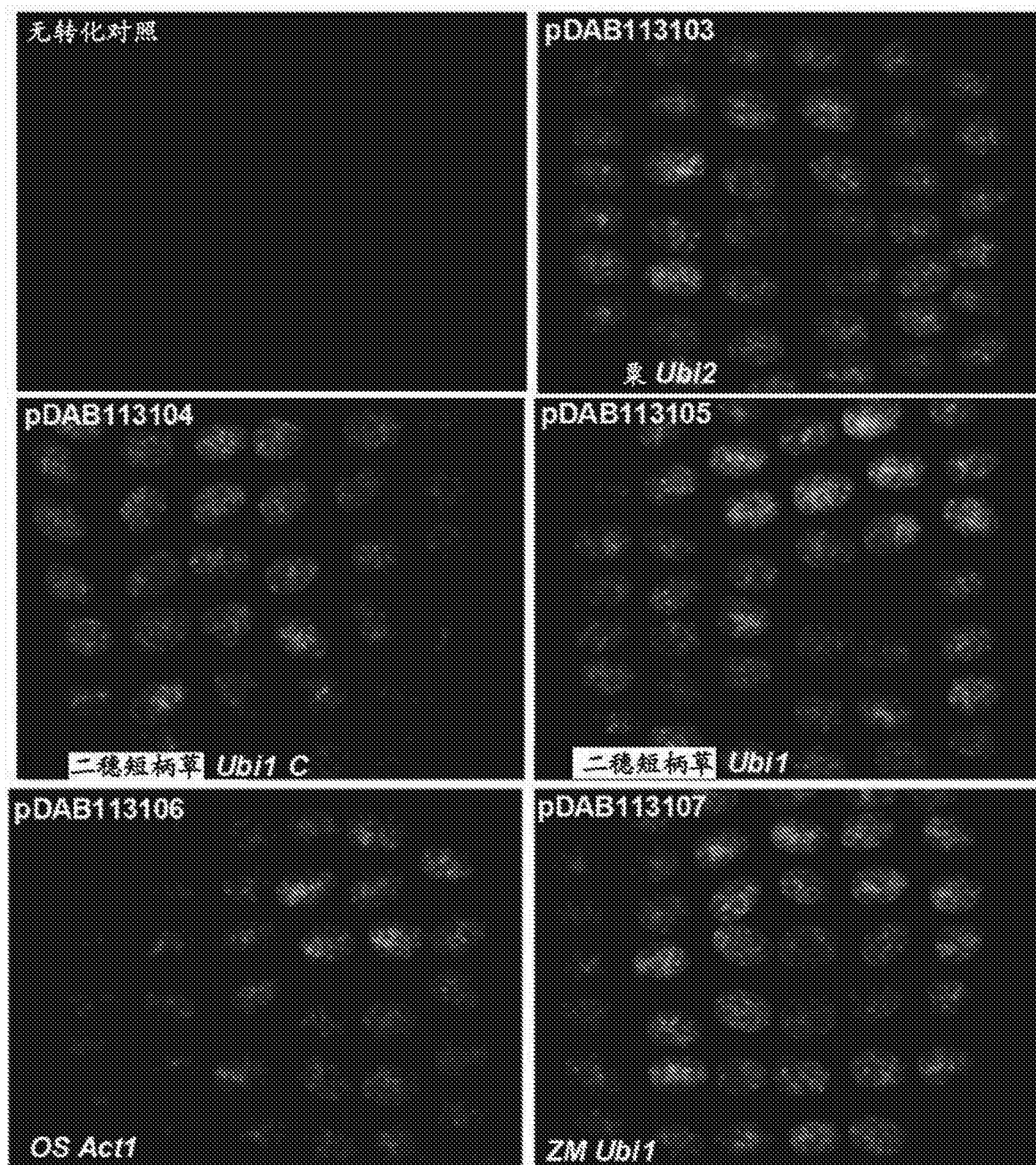


图19

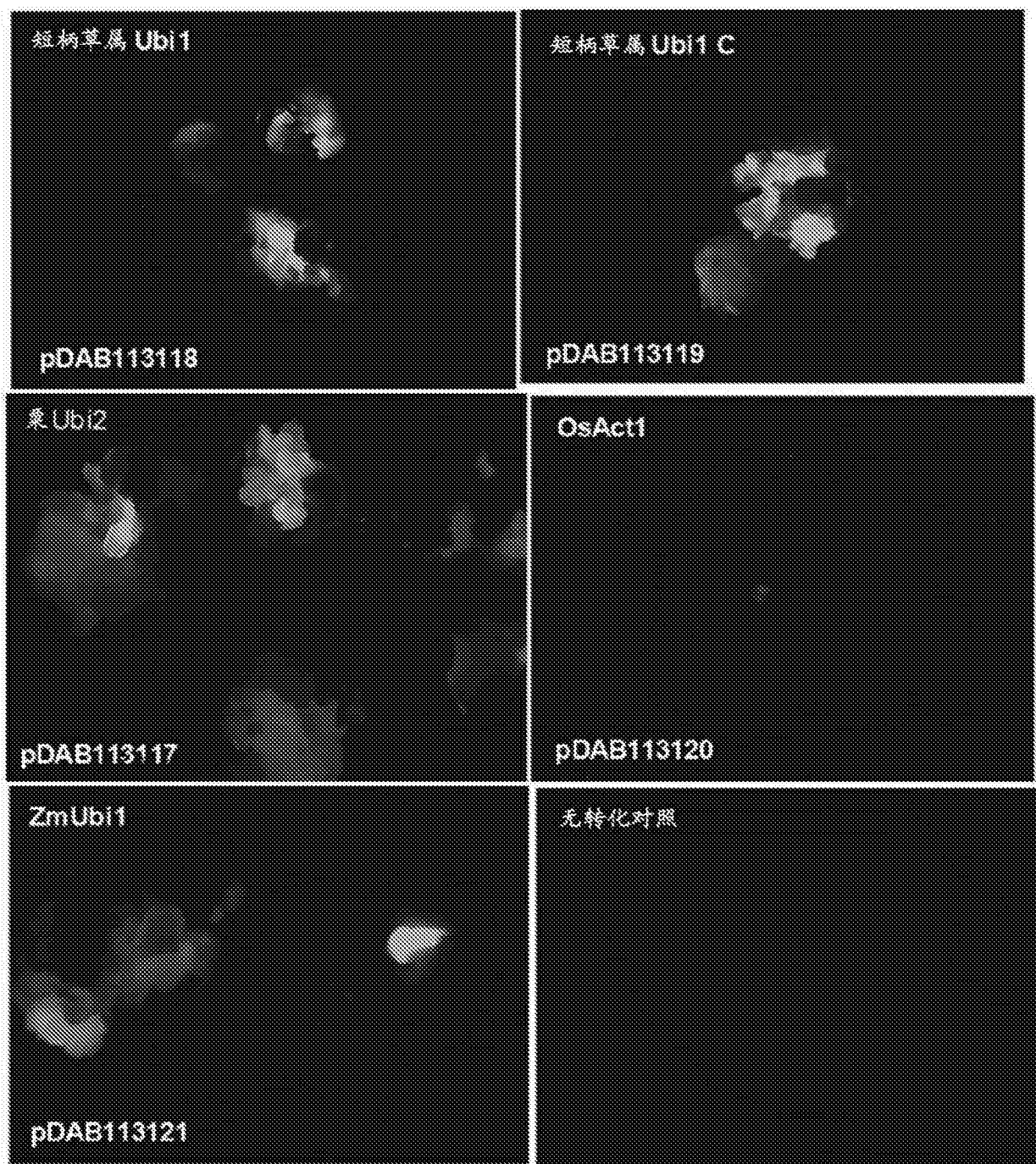


图20

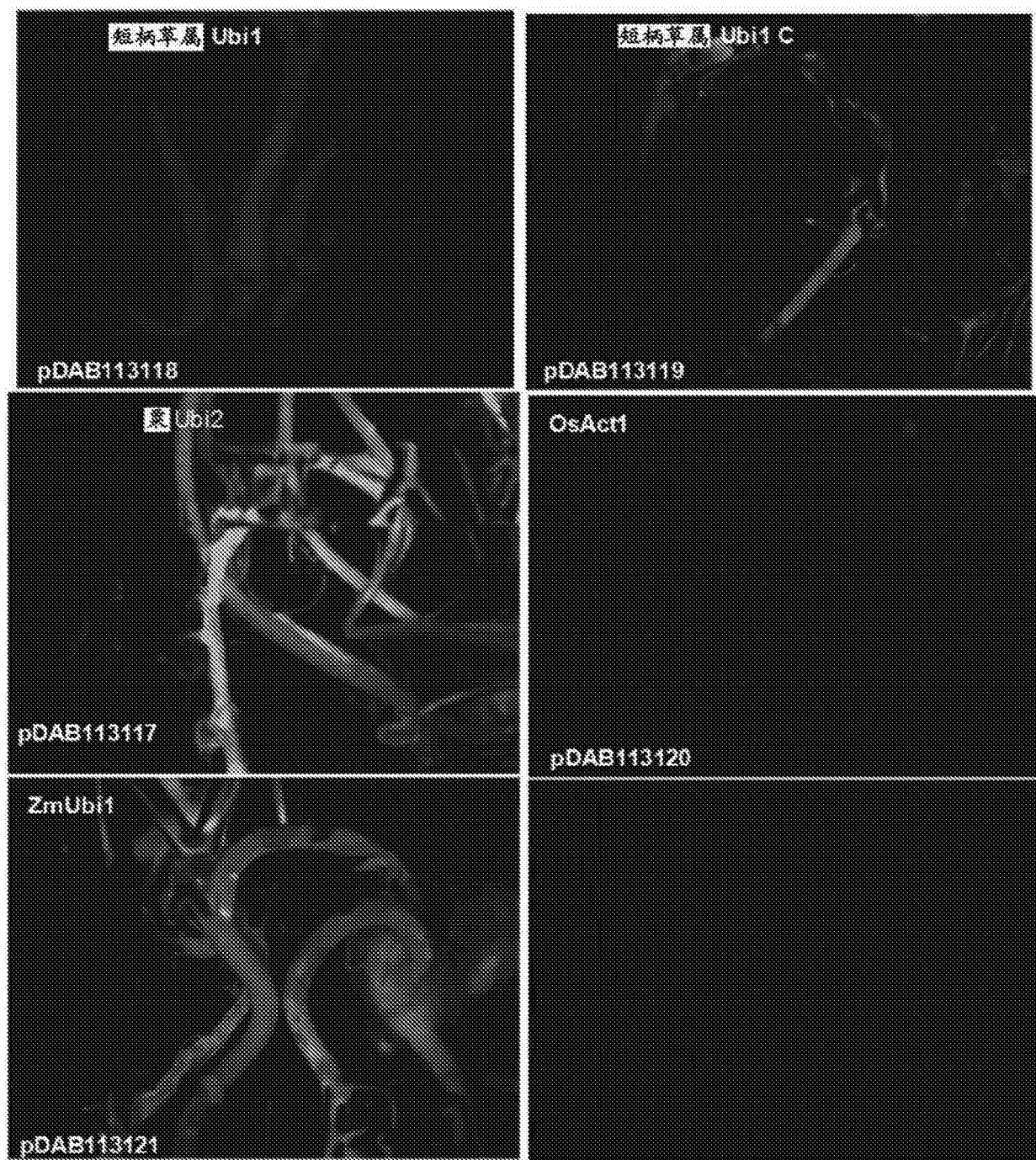


图21

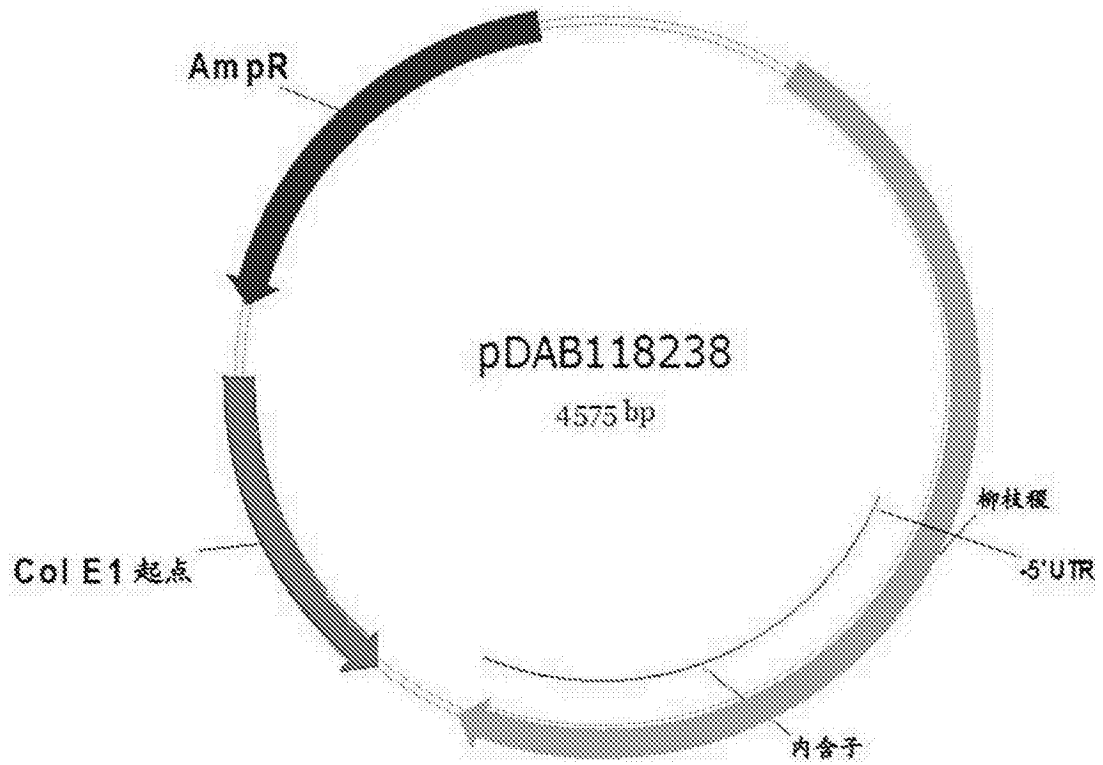


图22

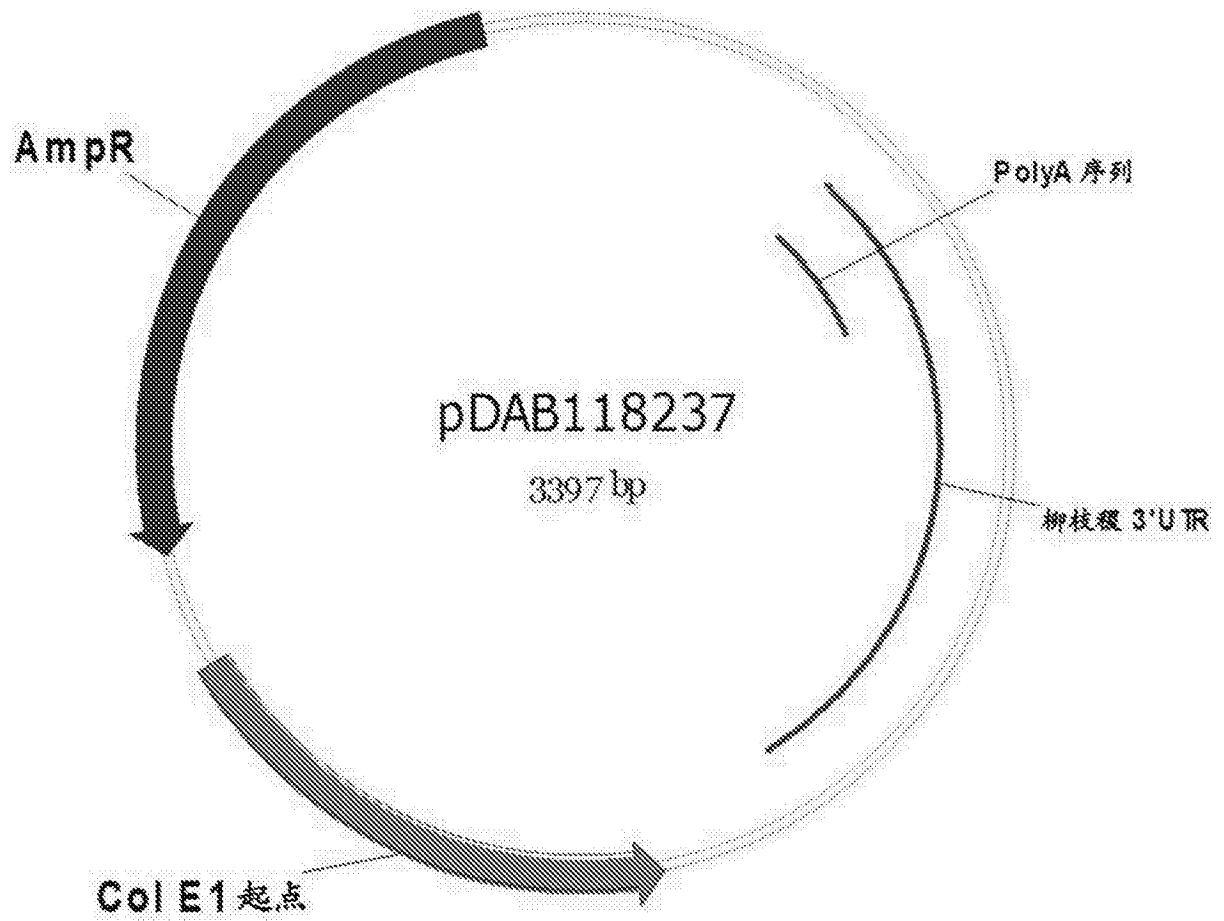


图23

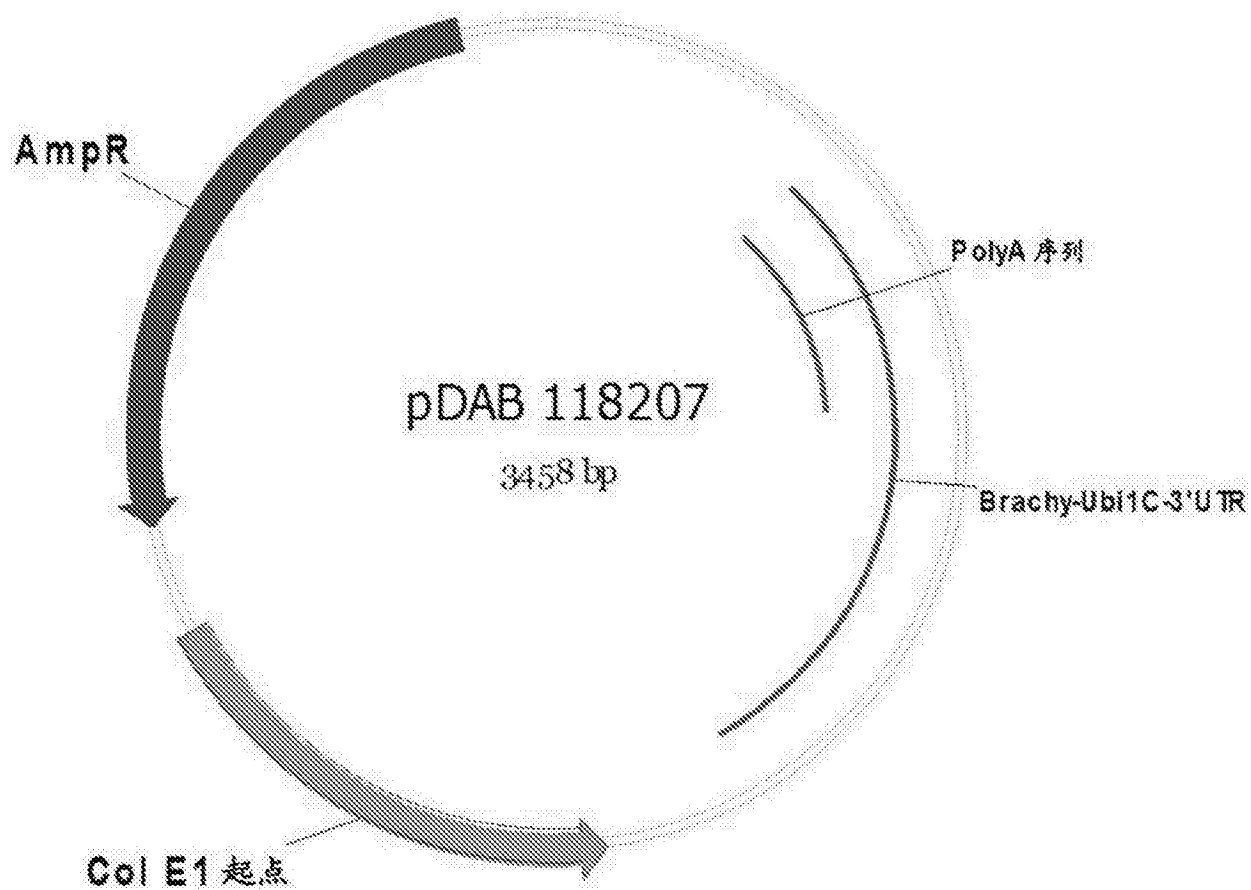


图24

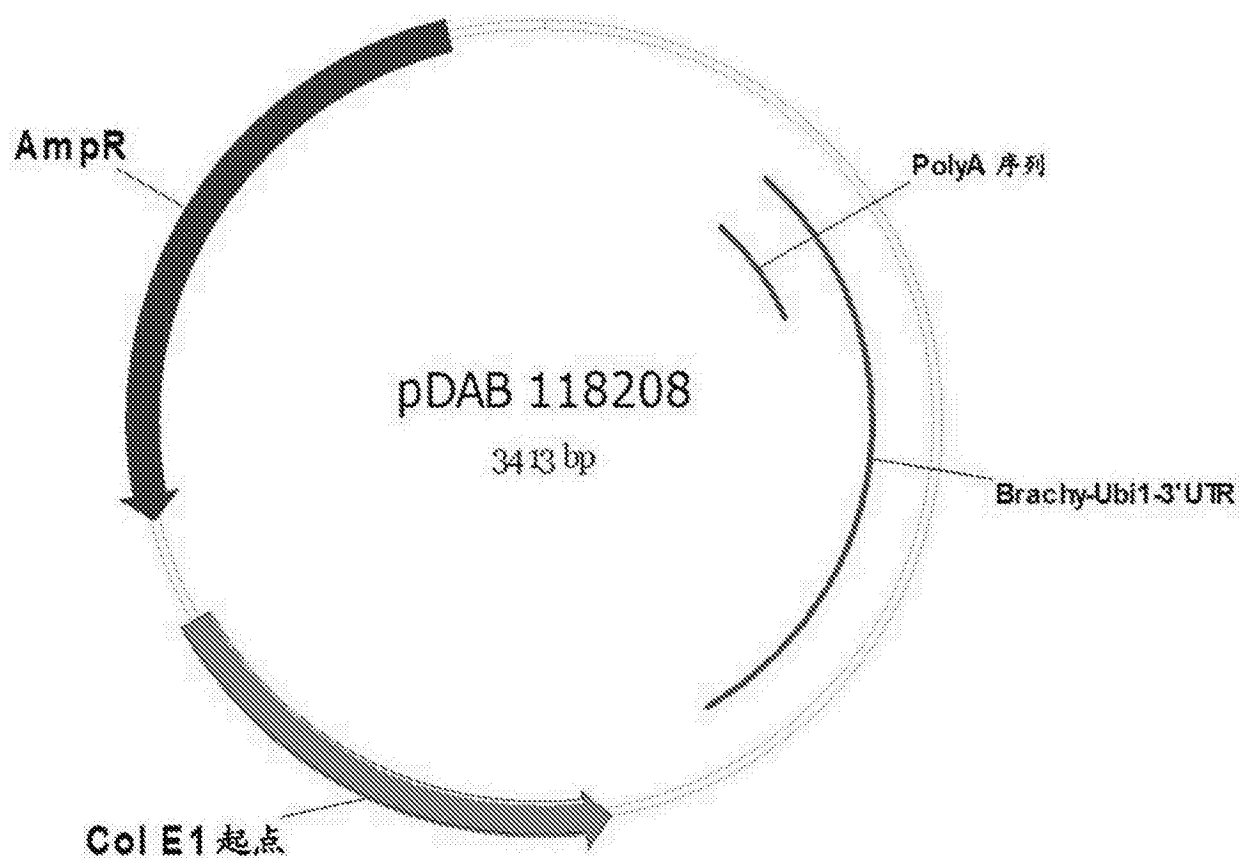


图25

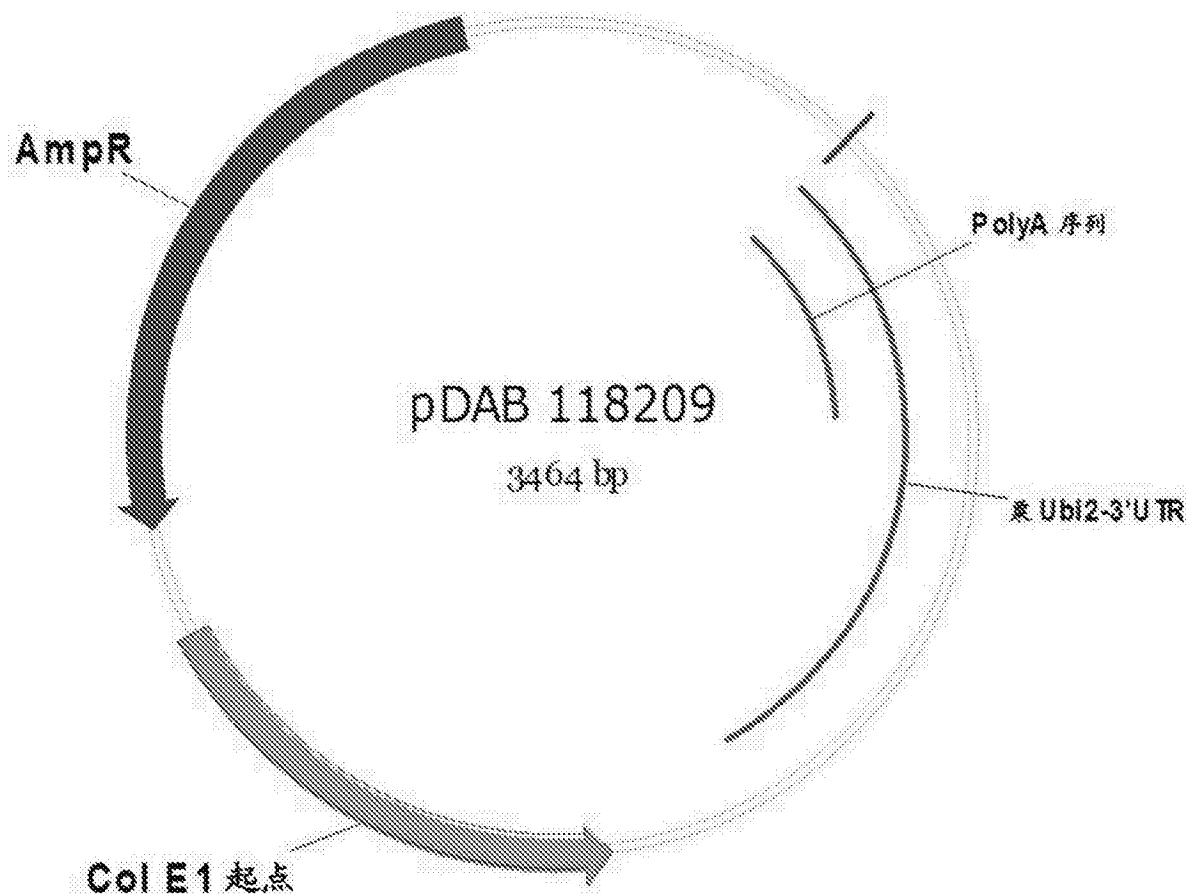


图26

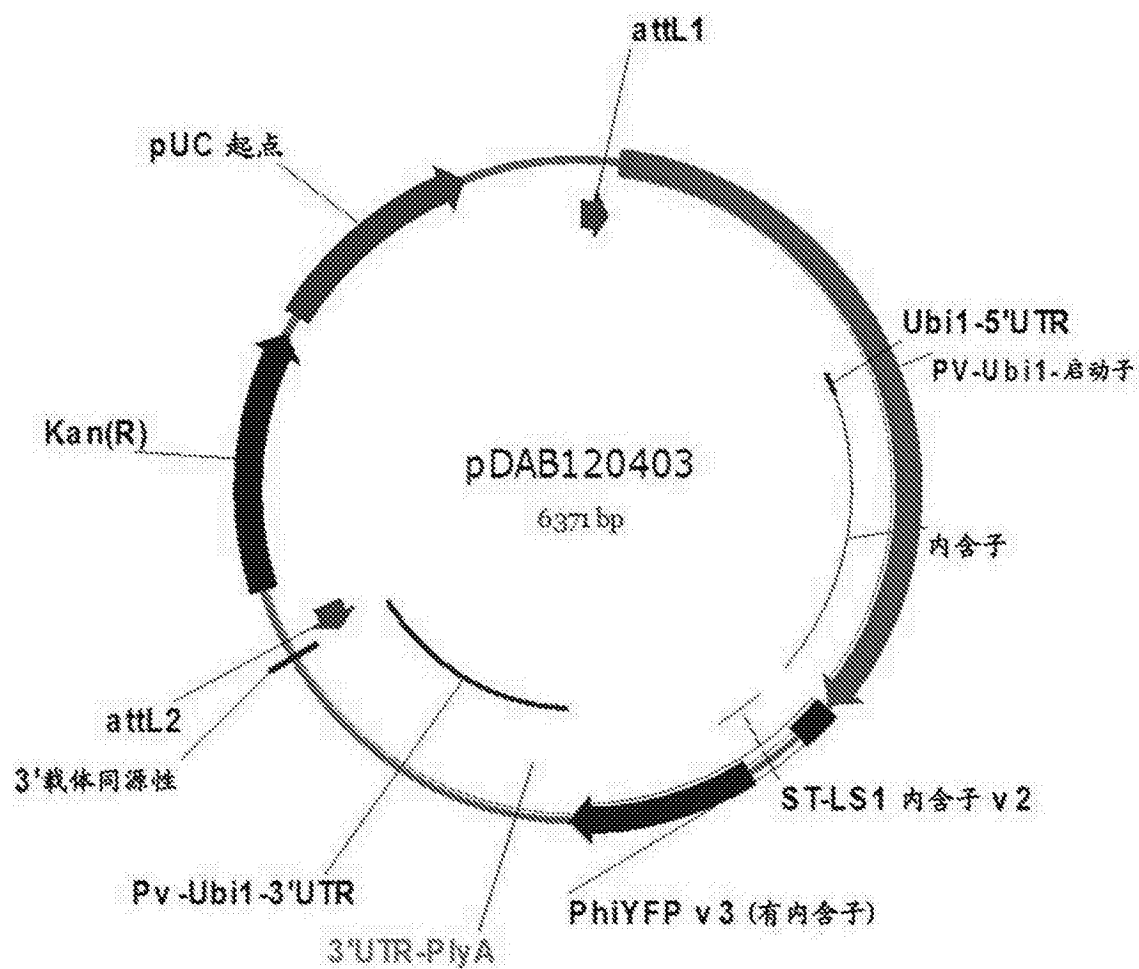


图27

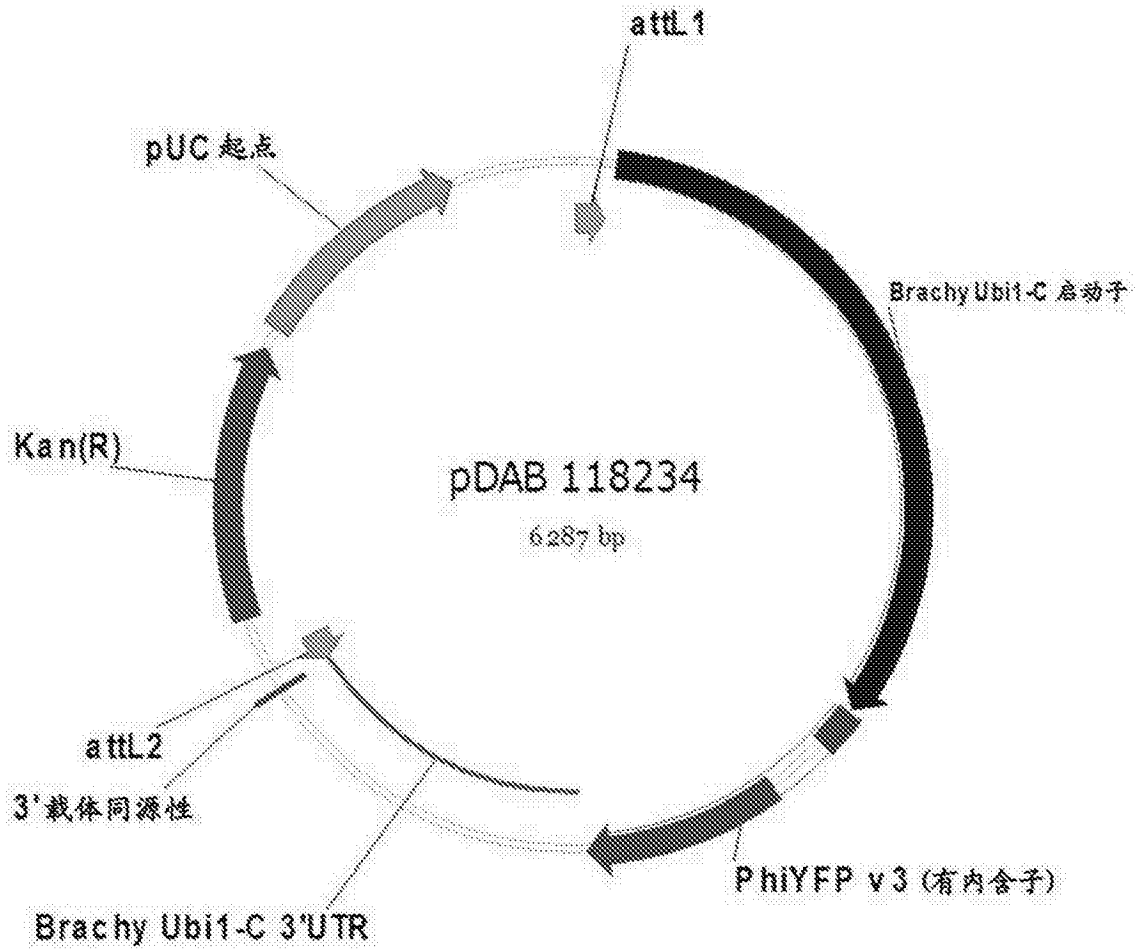


图28

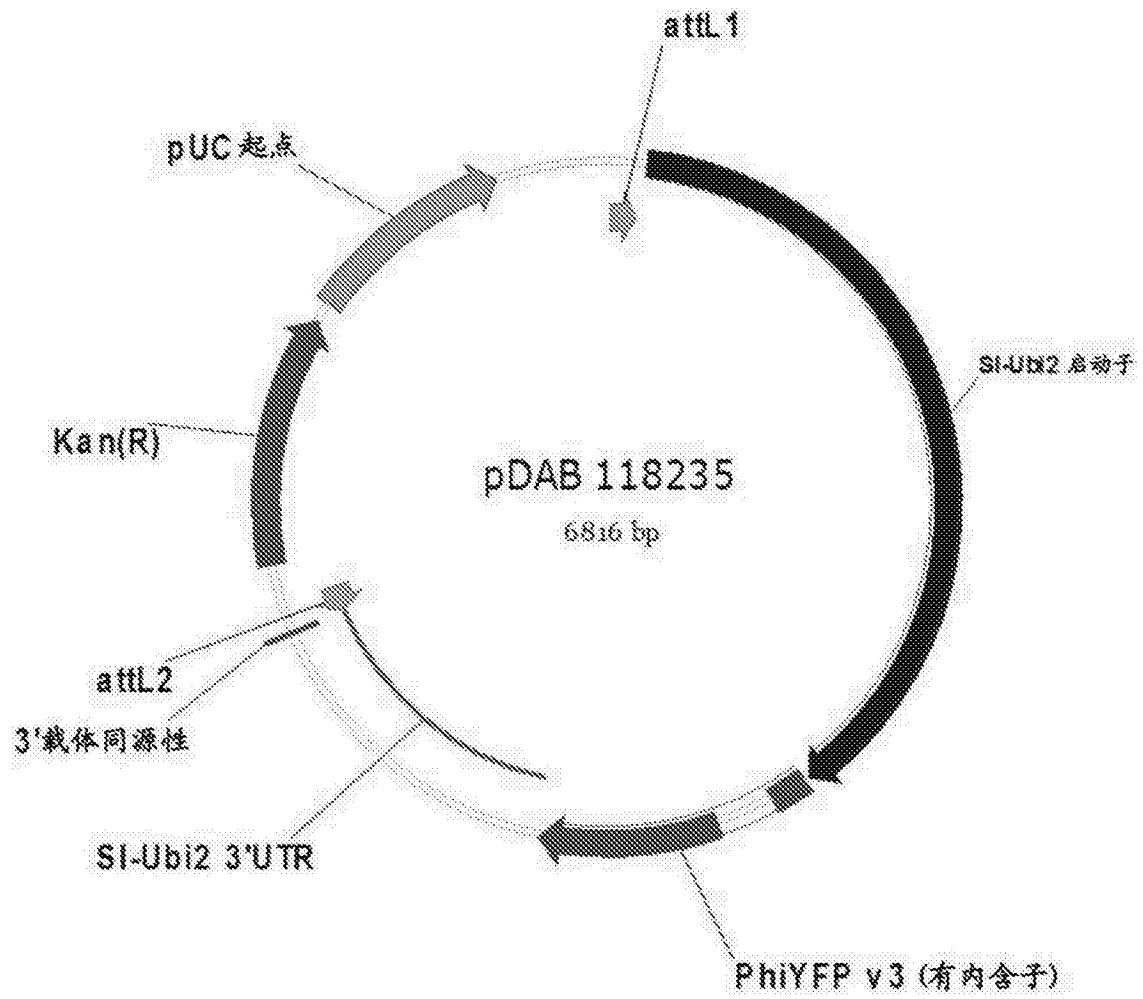


图29

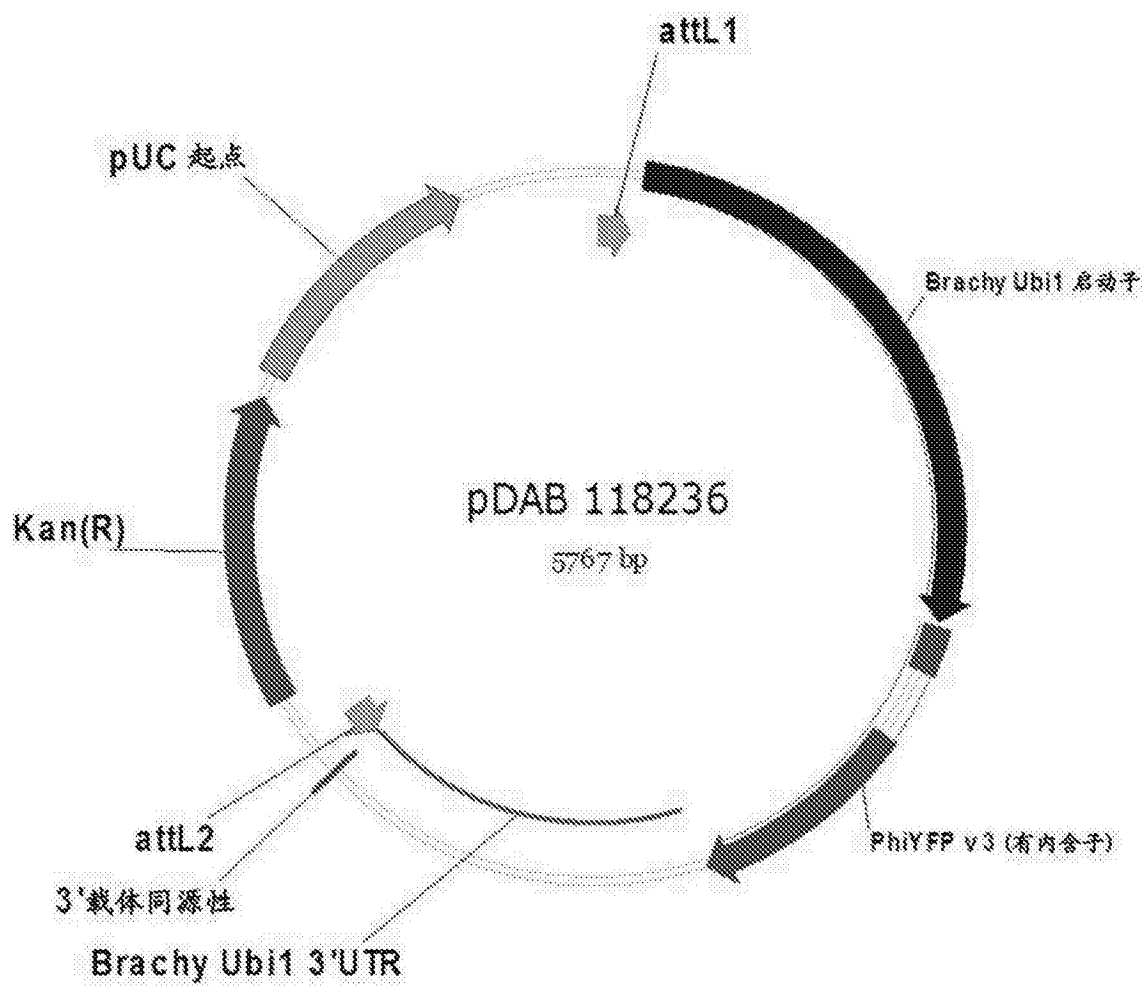


图30

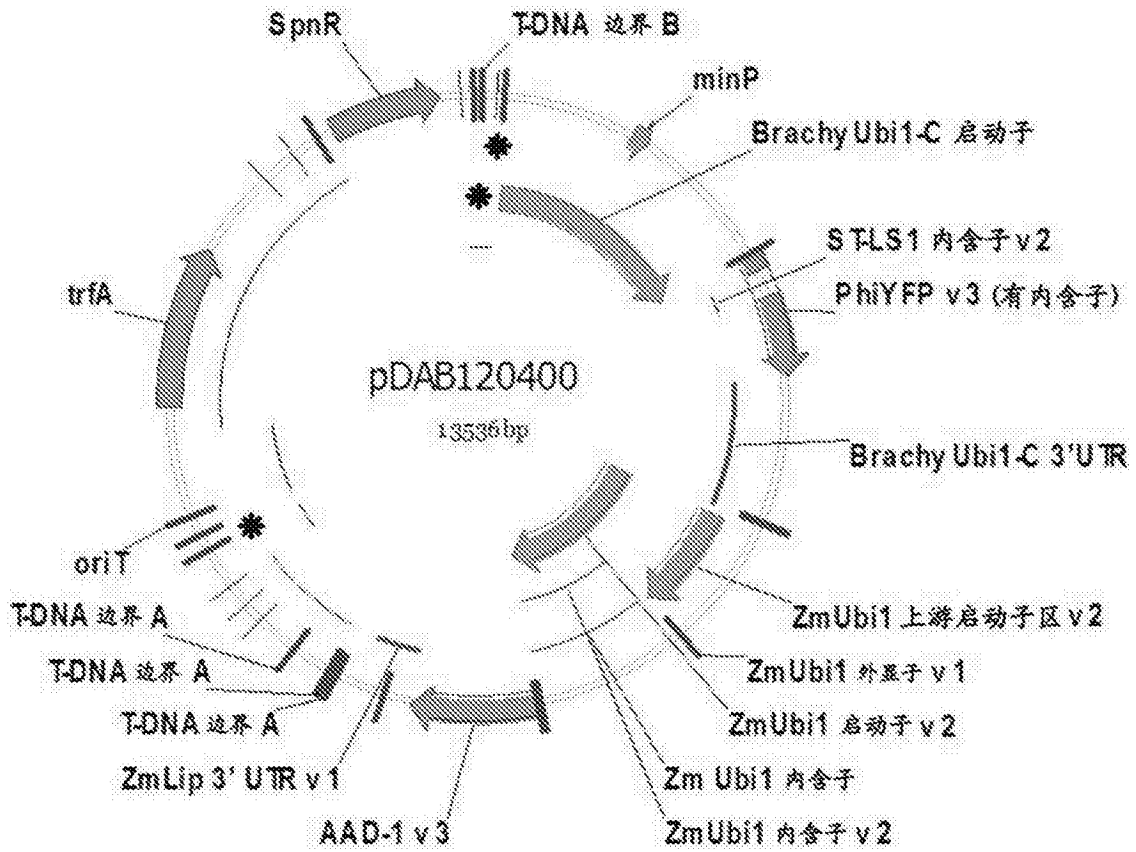


图31

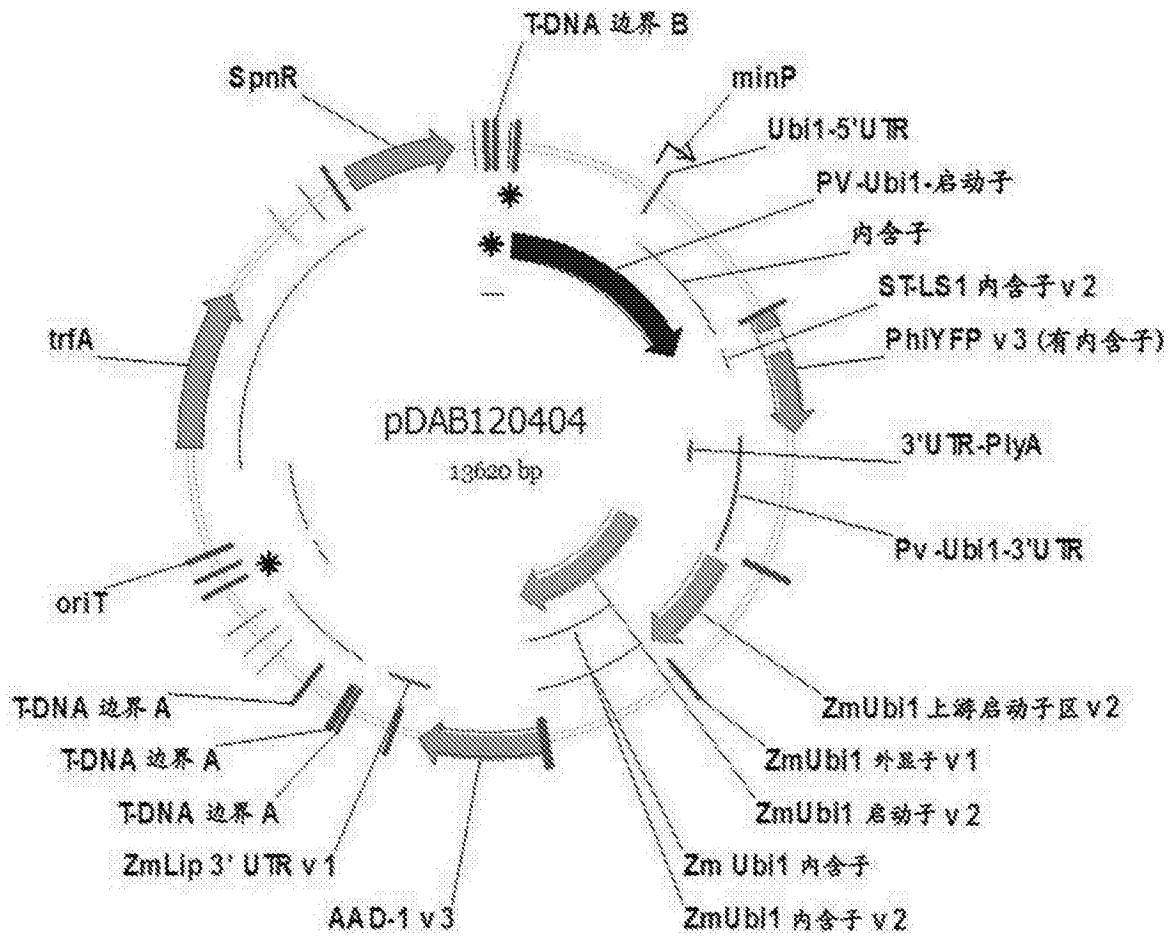


图32

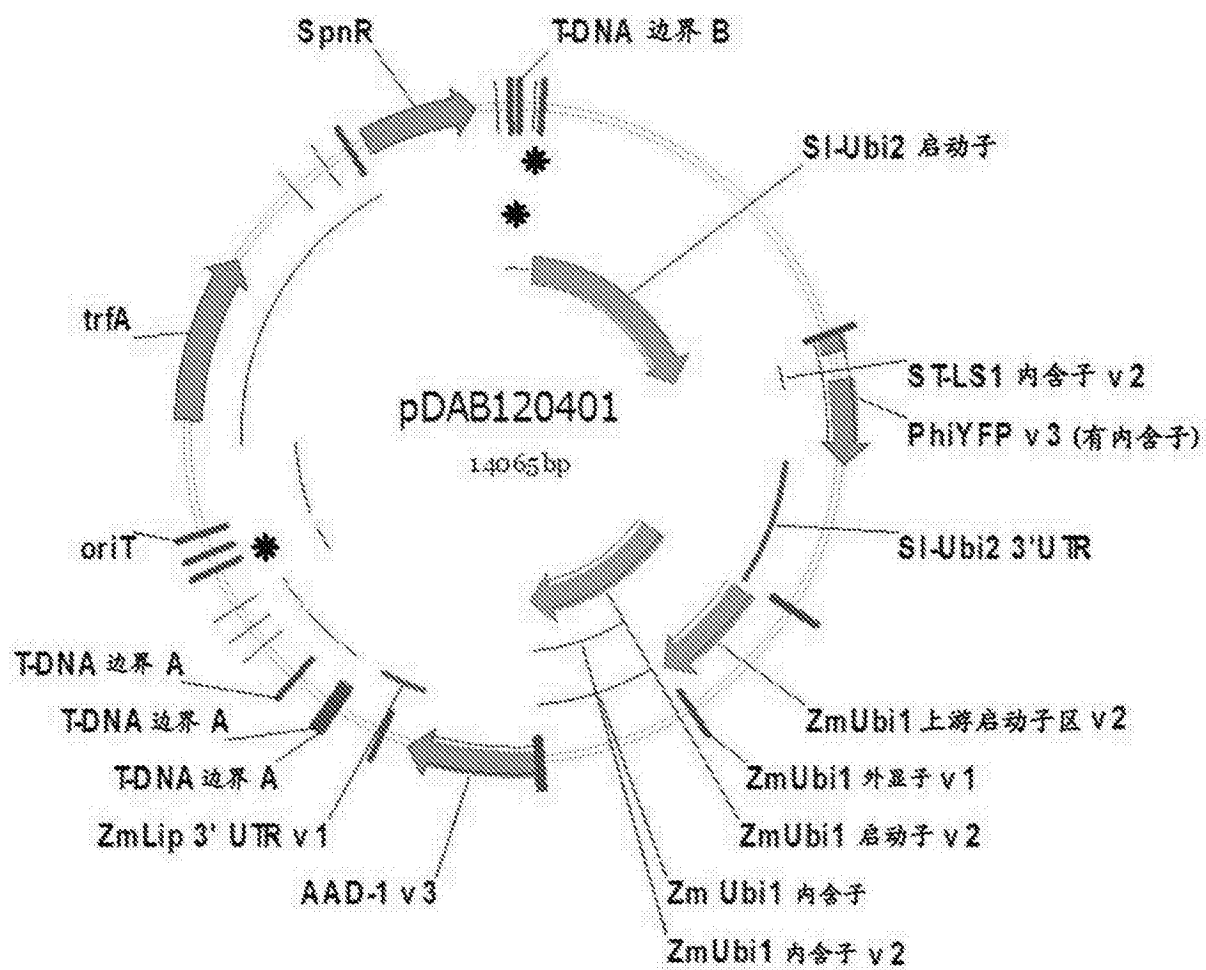


图33

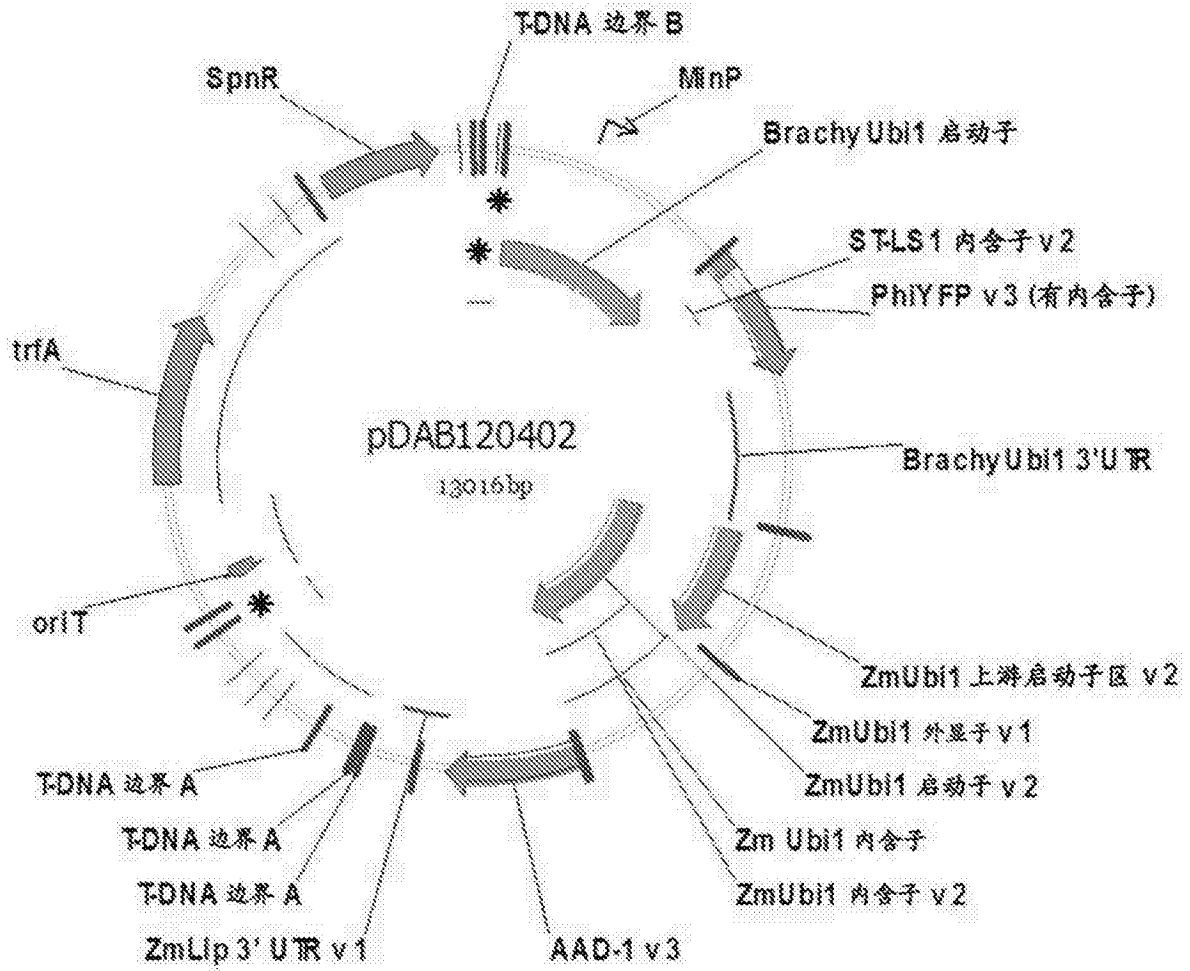


图34

ctgctcgttcagccacagtaacacgcgcgtggacatgcagatgccctccaccacgcgcgaccaaccccaagt
cgcgcgcgcctcgtccacggcgccatccgcctccgcgcgtcaacgtccatccggaggaggcgagcgcgatgtcg
acggccacggcgggcgggacacgacggcgacgcccgcactccgcgcgcgcgtcaaggctgcagtgccgtcg
tggtggcgcgtccgcctgcacgagatccccgcgtggacgagcgccgcctccaccacgcccctatatcgagaaa
tcaacggtgggctcgaggtcctcagcaacctcccccacccccttcgcgaccacgtcccttccccgcgtgccc
cttttctccgtaaaccccgagcgcgcgcgagaacaacaccaacgaaaggcggaagagaatcgccatagagaggag
atggggcgaggcgagatagtttcagccattcacggagaaaatggggaggagagaacacgacatcatacggacgc
gaccctctagctggctggctgtcctaaagaatcgaaacggaatcgctgcgcaggagaaaaacgaacggctcctg
aagcatgtgcgcgcggttcttccaaaaacacttatctttaagattgaagtagtatatatgactgaaattttta
caagggttttccccataaaacaggtgagcttatctcatccttttgttttaggatgtacgtatttatatatgact
gaatatttttttattttcattgaatgaagattttcagcccccaaaaaataaaaaacggaggggagtacctttgt
gcctgttatatggactagagccatcgggacggtttccggagactgcgtgggtggggcgatggacgcacaaacga
ccgcattttcgggtgcgcactgcgcgttcgcactctggtaggcacgactcgtcgggttcgggtcttgcgtgag
cgtgacgtaacagaccgcgttctctccccgcgtctggccatccataaatccccctccatcggttccccctt
cctcaatccagcaccctgattCCGATCGAAAAGTCCCGCAAGAGCAAGCGACCGATCTCGTGAATCTCCGT
CAAGgtatgcagccctcgttccctccctcgcctacggtttcaattctggagtaggctcgtagaggataccaatgttg
atgtgacagaggaggtagattagatacttgttagatcgaagtgcggatgttccatggtagatgataccatgtt
gatttcgattagatcggattaaatctttgttagatcgaagtgcgcactgttccatgaattgcctgttaccagta
gattcaagtttttctgtgttatagagggtgggatctactcgttgagatgattagctcctagaggacaccatgc
cgttttgaaaaatagatcagaaccgtgtagatcgatgtgagcatgtgttccctgtagatccaagttctttcgc
atgttactagttgtgatctattgtttgtgtaatacgcctctcgatctatccgtgtagatttcactcgattact
gttactgtggtttgatcgttccatagttgttcgttaggtttgatcgaacagtgctgaacctaaattggatatg
tattcttgatctatcaacgtgtagggttcagtcactgtatttatgtactccctccgtcccaaatctaactgacg
tggtattttgtataagaatctatacaaatccatgtcagtttaattcgggatggagtaccatattcaataatttg
tttattgctgtccacttatgtaccatattgtttgtttgttccctcatgtggattctactaattatcattgattgg
tgatcttctattttgcctagtttccctagctcaatctgggttattcatgtagatgtgtttgttgaatccggagacc
atgcttgttattagatagtttattgcttatcagtttccatgttctgggttgatgcaacacatattcatgttccg
tatctggttgctgcttgatattctctgatttacattcattataagaatatattctgctctgggtgttgccttc
tcatgactttacctactcggtaggtgacttaccttttgggtttacaattgtcaactatgcagATGcagatctt
tgtgaagaccctccacggcaagaccatcaccttgagggtcgagtccttctgatacgatcgacaatgtcaaggc
aaagatccaggacaaggaggggcatcccccgagaccagcagcgtctcatcttcgcccgggaagcagctggagga
tgcccgccaccctggcagattacaacatccagaaggagtcacccctcatctggtgctcaggctcaggggtgg
catgcaaatctttgtgaagacccttactggcaagaccatcacactcgagggtcgagtcgtctgacacgatoga
caatgtgaaggcaagatccaggacaaggaggggcatccccccagaccagcagcgcctcatctttgctggcaa
gcaactggaagacggtcgcaaccctggcagattacaatatccagaaggagtcacacttgcaccttgtgcttcg
cctccgtgggtggcatgcaaatctttgtcaagacccttgacaggcaagaccattactctggagggtcgagtcgtc
tgacacaatcgataatgtgaaggcgaagatccaggacaaggagggaattccacgggaccagcagcgcctaat
ctttgctggcaaacagcttgaggatggcgcacccctggcagattacaacatccagaagaatccactctgca
cctgggtgcttcgcctccgtgggtggcatgcagatctttgtcaagacccttgacagggaagacaatcacactgga
ggtcgagtcgtctgacacaatcgataatgtgaaggcgaagatccaggacaaggaggggcattccacgggacca
gcagcgccttatcttcgcccggcaagcagcttgaggatggccgcaccccttgctgattacaatatccagaagga

图35

atccaccctgcacctggtgcttcgcctccgtggtggcatgcagatctttgtcaagactttgaccgggaagac
cattacactggaggttgaatcttcagacacnctcgacaacgtgaaggcgaagatccaggacaaggagggcat
ccccccagaccagcagcgctgatctttgctggttaagcagcttgaggatggacgcactctggcggattataa
catccagaaggaggtctacccctacacctggtgctccgcctccgtggtggccagTAAgtttgtcaaaaactggc
ctacagtctgctgcccctgttgggtctgccccttggaagtagtctgtctatggttatgtgagaagtcgttgt
gttctttctaatcccgtaactgtttgtgtgaacatctgctgctgctgctattgcatcgtgaagaatcctgttat
gaataagtgaacatgaaccttgttctgtgattacggcttcgtggttatgccaacgttcttataaaacgcaatt
gcacctgatgtaaaatcgtttttgcctagctgtatggacaagtgctcatgatgttcatgcaagatgcaattc
cagcttttgttgggttgcctctttgtactgtgcttaccgcacataaagattgcatcttgccttattgctttg
ttgctttggtgctcgtccgcttctccttgcaaccttatcaaacctttgttttagattctcttcttatagcaact
ggtaactctcagctttacaacgccagtaactgtttctgaaatttcatgactgataaagctgatagatggagta
ctaataatataacatctttccataaatgttcgggtgcagagatatggaggcccccaggatccctatttacaggat
gaacctacctgggcccgtgtacgcctgacatccgcgagcaagtcctgaggttctcaatgtacacatgaaattg
atttttgcctgcgttttggcttggctgatcgttgcatttgttctgattcatcagagttaaataacggatatatc
agcaaatatccgcagcatccacacccgaaccacacgtccgggttaacagagtcctccctgccttgccttaattatt
acggagtactccgctattaatccttagatatgtttcgaaggaaactcaaaccttccctccatctgcaaatctca
gtgcttcaaaaactggaattagataattgaaaaccttcattcgggttgcaattcacaaactgcaaatgaaacagca
ctgtcaatttcaatttcgggttcacgattccaccgataggttgacatgatccatgatccaccattgtacaa
c (SEQ ID NO:19)

图35续

ggcgctcaggactggcgaagtcttgactctgcagggccgaactgctgaagacgaagcagaggaagagaaaagg
 aagtgttcgacttgtaattttagtaggggttttttttagaggaacttgtaattttagtaggtgggtggcctcgtt
 ggaaaaacgatgctgggtggttgggtgggtggcggatgtacgcttgcacaaacttggtggcggccggttctgga
 cgagcaggagtttttttttggttctcacttttttgggtcttcttttagttacggagtaccttttgttttttaa
 ggagttaccttttttttaggaattcttttagttaccttttgccttgccttcaaaaaatatitaaactttcgtttt
 ttttcatttttaatttttgcactattttagcagtttcaatgaatgcttattttccagcatalcattatttgcac
 gtaftttttatgcccgtatgtatttggacgagagccatcggaactgttccagagactgcgtgggtggggacggctc
 ccaacccgctttttctatctctgttgcacccgggtggcggacttggctcggcgtgagccgtgacgtaacaga
 ettgggtctcttcccatctggccatctataaattcccccatcgatcgacccctccctttccCCAATCCAGCAC
 CCCCATCCCGATCGAAAAATTCCTCCGCAACAGCAAGCGATCGATCTAGCGAATCCCGCTCAAGgtatgtagc
 ctctcgatttccctcctcagccctgcctcogatttgggtgtacgcttgagatgatgatctcgtagatgtctaga
 tgacaccatgtcgatttgaatatagatcagatccgtgttagatcgatgagctcctgtgtacctgttgattcaag
 ttatttttgcacgtctattgttgtgtatctactagatctagtgtgtgtattctatgctatogattttctcgtgt
 agatttccactcgattactgttactgtggcttgcacggccatagatgttgggttaaggtttgacgggttagtgt
 ttgaacctgcgtggatcttagcatccatctattatcgtgttaggtttgcacacaaacaaaggcactattattgta
 ctgatgggttgcctatggttgggttttgacgggttttaggttgaacgagccttctgtattttgtttattgtctgt
 ccagtgatgtacatgttctgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgc
 ctctcctaattttattttatcgttagtctctgatttgcctcagctgtgcctcaccggtgcgttgcgttgcgttgc
 tttagcccaatctgcttaaatcatgtacatttgttgttgaatacagagatcaagccaattagctatcttattgc
 ctatctgttccatgttctgatcgatgtacacagctctacacttttgcctctgtgcctacttgattaaaacattctg
 acttaaatctcatgatgtgaagtctcagatctgattgttgccttacttgactaatatctatctcatgtgacacc
 ctctctgtcttggtaacttaacgctgttctgttctgtaatttctgactatgcagATGcagatcttctgtgaagacc
 ctactggcaaaaaccatcacctttaggttcgagtcgttcgacacgacatcgacacacgtcaaggcaagatccag
 gacaaggagggcatctcctccagaccagcagcgcctcatctttgctggaaagcagcttgaggacggccgcacc
 ctgcgcgactacaacatccagaaggagtccacccctccacctgggtcctgaggctccgtgggtggcatgcagatc
 ttctgcaagaccccttaacggcaagaccatcacgctggaggttcgagtcctctgacacgacatcgacaatgtgaag
 gcgaagatccaggataaggagggcatccccccggaccagcagcgcctcatctttgcgggcaagcagcttgag
 gacggcgttacccctgcgcgactacaacatccagaaggaaatccacactccatctggtgctcaggtgcgtggt
 ggcattgcagatcttctgcaagaccctaaacggcaagaccatcacctctggaggttgagtcctctgacacgac
 gacaatgtgaaggcaagatccaggataaggagggcatccccccggaccagcagcgcctcatctttgcgggcaagcagcttgag
 aagcagcttgaggatggccgcacccctggcagattacaatatccagaaggaaatccacacttgcaactgggtgctt
 cgcctccgtgggtggcatgcagatctttgtaaagaccttgactggcaagacaattaccctggaggttgagtcg
 tccgacacaattgacaatgtcaaggcgaagatccaggacaaggagggcatccacccggaccagcagcgcctc
 atcttgcgggcaagcagcttgaggatgggtgcaccccttgacagattacaatatccagaaggaaatccactctg
 catctgggtgcttctcctcgggtggaatgcagatcttctgttaagacgttgacaggggaagaccatcacactg
 gaggttgaaatcttgcgacaccattgacaacgtgaaggcaagatccaggacaaggagggcatccccccagac
 cagcagcgcctcatctttgctggtaagcagcttgaggatggccgcaccccttgacagattacaacattcagaag
 gagtccacccctgcacactgggtgctcctcctcgggtgggtggcagTAAgcttctgcggaactgggttcacagctg
 ctgccttgggtgggtctgccccttagtggtcatgcttttgttatgtgtcttgcgtcccaatcctgtatcgtt
 tgtgtgaacatctctgctgctgtatagcagcttgaatcctgttatgaatttgtgaacctgaaccttgttccg
 tgaatcatgttatgaataagtgaacctgaaccttgttccgtgattattgttacaatctgttgtgcgtatgg

图36

ttggtcgtgtgtgatttatgttgaactggagaaccaagttcgttccaggacatatattgcaacctaagctaacc
catgtagaactacttgttctgggagacataaaaagtcatttttatgcattcgtaacatttaagcataactaca
ataattgtattgtccttttctactcatccttgaaaaccatatgcctcttctcagcgccctctacatgcagtgt
gctcagaacaaaacaggccctgccagctgcttttcaattttccaattaataaccacaatagtcggactatggc
atctgtgggtgactatgcaagatgttgctgtcaggtctcttgaaaacttttcccatgtatctgttgaaattacc
cagtaaatcatgcctctatttaaatctggcatgggttgattttcaaacagaatgtgttttttttgttctgga
agctatatattggtaaaataaatacaaaagctggagtggtgattatatattccaacagatattcaagaaaatctcagt
tgatttatttactactgtagtatatatatatatcttacagttgactttctcatatttcaaacgacatgtgagc
acattgttcagtttcttaggatgtgttggtgtgctcaaagggtgtaattttgcattctgcctccgagtaaca
ctacaogtattttttgagtggcagtgcatttgattacaaggcaacaacaacaaaacctaaggcaagatat
ccttcttagaggctgccaggatcattttgactgaactatgtaaggctgaagaaaagg (SEQ ID NO:20)

图36续

tgcgtctggacgcacaagtcattatcggtctaaaattttcttaattttctaaattagtcattatcggtctaa
gaaagtggggagcactatcatttcgtagaacaagaacaagggtatcatatatatatatatataatatttaa
actttgttaagtgggaatcaaagtgcattgatttaattggagtttcatgtgcattaaattttatgtcacatcagc
aattttgttgacttggcaagggtcatttaggtgtgtgtttggaagacaggggtatttaggaggtattaaacatag
tctaattacaasactaattgcacaacogctsaagctgaatcgcgagatggatctattaagcttaattagttca
tgatttgacaatgtggtgctacaataaccatttgcataatgatggattacttaggtttaatagattcgtctcg
tgatttagcctatgggttctgctattaattttgttaattagctcatatttagttcttataaattagtafcgaa
catccaatgtgacatgctaaagttaaaccctggtatccaatgaagtcttatgagagtttcatcactccggt
ggtatatgtacttaggctccgttttcttccaccgacttatttttagcaccggtcacattgaatgttttagata
ctaattagaagtattaaacgtagactatittacaaaatccattacataagacgaatctaaacggcgagacgaa
tctattaaacctaatttagtccatgatttgacaatgtgttgctacagtaaacattttgctaattgatggattaat
taggcttaattagattcgtctcgcggttttagcctccacttatgtaattgggttttctaaacaattctacgttta
tactcctaatttagtatctaaatattcaatgtgacacgtgctaaaaataaagtcagtggaagggaagagaacgtc
ccttagttttccatcttattaattgtacgatgaaactgtgcagccagatgattgacaatcgcaatacttca
actagtggggccatgcacatcagcgacgtgtaacgtcgtgaggttgcgtgttcccgtagAGAAATATCAACTGGT
GGGCCACGCACATCAGCCTCGTGTAAACGTGGACGGAGGAGCCCCGTGACGGCGTCGACATCTAACGGCTACCAGCA
AACCACGGAACCACCGCTCCCCACCTCTCGEAAGCTCCGCTCCACGGCGTCGACATCTAACGGCTACCAGCA
GGCGTACGGGTTGGASTGSACTCCTTGCCCTCTTTGCGCTGGCGGCTTCCGGAAATTGCGTGGCGGAGACCAG
GCGGCTCGTCTCACACGGCACGGAAGACgtcacgggttcccttccccacctcLccLcttccccaccgccata
aatagCCGACCCCTCGCCTTTCTCCCCAATCTCATCTCGTCTCGTGTGTTGTTCCGAGCACACCACCCGCCCC
AAATCGTTCTTCCCGCAAGCCTCGGCGATCCTTCACCCGCTTCAAGgtacggcgatcgtcttctctctctag
atcggggtgatctgcaagtagttgattttggttagatgggttaggactctgtgcaactgaagaatcatgttagatc
cgogatgttctgttctgtagatggctgggaggtggaatttttgtgtagatctgatatgttctctctgtttatc
ttgtcacgctcctgogattttgtggggatttttaggtcgttgatctgggaatcgtggggttgcctctaggtctgt
tcgttagatgaggtcgttctcacgggttactggatcattgcctagtagatcagctcgggctttcgtctttgtat
tatggtgcccatacttgcactctatgatctggttccgtgggtgttacctaggtttctgogcctgatttgcctga
tcgattttgttagcatgtggttaaacgttttggtcatggtctgatttagattagagtcgaataggatgatctcg
atctagctcttgggattaatatgcatgtgtcaccaatctgttccgtggttaagatgatgaatctatgcttag
ctaattgggtgtagatatatatgctgctgttctcaatgatgcctagcttttacctgagcagcatggatcct
cctgttacttaggttagatgcacatgcttatagatcaagatatgtactgctactgttggaaattcttttagtata
cctgatgatcatcatgctcttctgttacttgttttggtaacttgatgatggcatgctgctgctttttgttg
atttgagcccatccatatctgcataatgtcacatgattaagatgattacgctgtttctgtatgatgccatagc
ttttatgtgagcaacatgcactcctcctggttatatgcattaaatagatggaagatactattgctacaatttg
atgattattttgtacatacgtatgaagcatgctcttcatactttgttgatatacttggataatgaaatgc
tgctgcacgttcatctatagcactaatgatgtgataaacacgcacgacctgtttgtggcatctgtttgaat
gtgttgttgcgtgttcaactagagactgttttattaacctactgctagatacttacccttctgctgtttatto
ttttgcagATGcagatctttgtcaagacctcaccggcaagaccatcaccctcgaggtggagctctctgaca
ccattgacaacgtcaaggccaagatccaggcaaggaaggcatcccccggatcagcagcggtcatctttg
ccggcaagcagcttgaggatgggcccacctggctgactacaacatccagaaggagagcaccctccacctgg
tgctcgtctcagggggaggcatgcagatctttgtgaagaccttgactggcaagaccatcacccttgagggtg

图37

agtcttccgacacccatcgacaacgtcaaggccaagatccaggacaaggagggcatccccccggaccagcaga
ggctcatctttgccccgaagcagcttgaggatggacgcacccctggctgactataacatccagaaggagagca
ccctccatctgggtgctccgtctcaggggaggcatgcagatcttcgtgaagactctcactggcaagaccatca
ccctcgaggtggagtcttccgacaacatcgacaacgtcaaggccaagatccaggacaaggaggggcatcccc
cagaccagcagagggctcatctttgctggcaagcagcttgaggacggacgcacccctggctgactataacatcc
agaaggagagcaccctccacctgggtgctccgcctgaggggtgggatgcagatctttgtgaagactttgactg
gcaagaccattactttggaggttgagagctccgacacccatcgacaacgtgaaggccaagatccaggacaagg
aaggcatccccccggaccagcagagggctcatcttcgccggcaagcagcttgaggacggacgcacccctggctg
actataacatccagaaggagagcaccctccacctgggtgctccgtctcaggggaggcatgcagatcttcgtga
agaccctcactggcaagaccatcaccccttgaggtggagtcttcgcgacacccatcgacaatgtcaaggccaaga
tccaggacaaggagggcatccccccagaccagcagagactcatctttgcaggcaagcagcttgaggacggac
gcacccctggctgactacaacatccagaaggagagcaccctccacctgggtgctccgtctcaggggaggcatgc
agatcttcgtgaagaccctcactggcaagaccatcacccctcgaggtggagtcttctgacaaccatcgacaacg
tcaaggccaagatccaggacaaggagggcatccccccggaccagcagcgtcttatctttgccggcaagcagc
tgaggatggcgccaccccttgccgattacaatatccagaaggagagcaccctccatctgggtgctccgtctga
gggtgggatgcagatattcgtgaagactttgacggccaagaccatcacctttggaggttgagagctccgaca
ccattgacaatgtgaaggccaagatccaggacaaggaggggtatccccccggaccagcagcgtctgatcttcg
ccggcaagcaactggaggatggccgcacccctggcggactacaatatccagaaggagtcacccctccacctgg
tgctccgcctccgtgggtggcagTAAgcccacatcggtcatggatgcttctactgtacctgggtcgtctggct
ctgcctgtgtcacctttgaagtacctgtgtcgggattgtgtttgggtcatgaactgcagtttgtctttgatgt
tcttttgtctggctcttatgaactgggtgtgtatctgtatgtttactgtaaactgttgttgoggtgcagcagtat
ggcatccgaatgaataaatgatgtttggacttaaatctgtactctgtttgttttcgggttatgccagttctat
attgcctgagatcagaatgttttagcttttgagttctgtttggcttggtgcactcctgtttcttaacttgag
gcgtaactctgttctggcaaaactcaaatgtctaaactgaatgttttaggacttaattgttggacagattaacg
tgtttggttttgtttctagattgtgattcggaaggcttgttagttgtggaatcaaggagagcagctagggtctg
tgcagaacgttattttggatttaagccttctcagattatgccattactctaaacctaataatgatcatatcttc
actcggggatgttggagtagtcttttctttctcctgcagacaaaatgattttgctttcgtgtgtgtacatga
ttttgtgcaaotgttgcacaactgaagtagacaagttttgacctcaccagaagaatgaaaaagattttgga
atttgttacatcgacaaaacattgttaacttggcccatcagaatgcacagaagagcggctacaaattgacatg
cgttgcaaaactttgcaatagttgatgcacatgtttgccattgcctgccagtccttaggaaaagtgtgtggttc
gagaaatctaagcatatgtgtctctgctcacattgcgtggaacccacacagctttgtcacactcttgtccact
ccagaagtcattcctggcgctgtttacccctggtaaaaggtaaccgaaaacttctcaaggctgtacccaaaa
ctggaaggaaaatttggaggaaatctttgcttttgatcggtcactctttc (SEQ ID NO:21)

图37续

ttgaattttaatttcaaatftttgcagggttagtagtggacatcacaaatcacatatttagaaaaagttttataat
tttccotccgttagttttcatatatttttgaactccaacgattaatctattatttaaataatcccgatctatcaa
aataatgataaaaaatttatgattaatttttctaacaatgtgttatgggtgtgtactatcgtctttataaaaatttc
aacttaaaaactccacctatacstggagaaatgaaaaagacgaattacagtagggagtaattttgaaccaatg
gaatagtttgagggttaaaaatgaaactaascaatagtttagggaggttatttcagatttttagtttatagttgagagg
agtaattttagacittttccatctttgaattgttgaaggctctccatccggatatccggatggagcttttcagc
ccaacataaacttcattccgggcccaaacgttccgtccatccagccatagggagaaacattttgcccattgatatctc
ttttttttttttttctatftttcactgggtatttatagggaaggaaatatacaacgtgttcacctttgggtttcatt
ttgttccatctgaattttatctaaaactgtgtttgaacttcgtaagaattttgttcgatctgtccgggtacate
gtgttgataggttggcctccgagattcttctttttaacccggcaaaagttaaaataatctcagctccagccataacg
tcaattatcagagagagaaaaaaatatttttttatgattgatcggaaaccaaccgccttacgtgttcgatctct
ggttccctggccggccacgggaggagaaagcgacccgaacctogcaacgcggcgccagggcgccgcgtgtttggac
ttggctctcccggaactccgttgggtccctcggcttatccggccggctccatctcaaccgtccgcttggagacagtg
gaagttgatccgtccgcaccagcctccggaggttaacctaaactgcccgtactataaatccgggatccggcctc
tccaatcccatccgccaCAAGTTCGCGATCTCTCGATTTCACAAATCGCGGAGAAGACCCGAGCAGAGCAAGT
TCCCTCCGATCGCCTTGCCAAAGgtactccacctaatactcccttaactgatctctctctatccagttggta
atcttcgaatgatctgctgctgggtccgtgttcccccctcggttatgcaactgtttccateaccgagtttttttt
ttcatcatctaatctatgcgggttgcgggaagaattgtggttagtggagtagttttctgtgcttgatccgtaga
ttcgatgtgtgggtgtatggatgtttttctgaaaaagtgtcgtggattagtttacgctttcaggccgcaggctctg
ttcgaaaattgattatgaagctctatatgtctttggatctatcgatttcagttttatccagatgtaggccaaaa
aattgtccgcatttctgtggaattagttggccttttaggtctgcacattcatggtgaaggccacagtttgcctgct
ggctgtttgcgtgggacgagttattatagttgtttttgtttttccctgtattgatattacattttcaatgataac
tagcctttgtccactaaaccaagtcacaggttgatccatctctgtgttcttcagctaccagttttgcatagatgat
ggtgtattcgattgcttttagtaggcttctgatttcacatctaatctgtcatgaatatagataactttaca
tgcttttgatataactttatatttgaactgttccactgtccagcctattttggataaattgagtgcaattggcttt
tgatgcctgaattattcaaatgttccctggataattgaacctgtgtccactagttgaactgttttttgaggtgcc
accgctctgttccagctgaattgtgttaticgattgctctagttaatcttttgattatgcagctagtgctttgt
catatgtagctttataggcttctgatgtccctggatatagttccagctctactgtcaagttgctttacaagta
gtagctctgattctatttggcttccctgagtcagagccttgcaaaattgcctgttggttacattacatcatatta
cttgaattgcagttattttaatgggttggattgttgcctgtttacttctacatttttttgcctgttttatattatac
taaaaatgttttgttggctgtcttttcagATGcagatcttttgtgaagacaactcaactggcaagactatcaccctt
gaggtggagctctctgacacaattgacaatgtcaaggcaagatccaggacaaggaagggtattcctccagac
cagcagcgccttatcttccgtggcaagcagcttgaggatggccgtacacttgcagattacaacattcagaag
gagtcacacactgcaaccttgtccctcaggctcgcgtggaggcatgcagattttcgtgaagacccctcaactggcaag
acgatcaccctagaggttggagtcctctgacaccatcgacaattgtgaaggcaagatccaggacaaggagggc
atccccctgaccagcagcgcctcatctttgcaggcaagcagttggaggatgggcgaactctggctgactat
aatatccagaaaagagtcaccccttcacctcgtcctccgcctgcgaggtggcatgcagatctttgtgaagacg
ctgacaggcaagaccatcacarttgaagttgagtcctccgatataatcgacaattgtgaaggccaagatccag
gataaggaggggtatccccccgggaaccagcagcgcctcatcttgcgggcaagcagctcgaggatggccgcact
ctcgtctactacaaacatccagaaggaatcaaccctgcacctgtgtgtctccgctcgtgtgttggaaatgcagat

图 38

图38续