



(12)发明专利申请

(10)申请公布号 CN 107075526 A

(43)申请公布日 2017. 08. 18

(21)申请号 201580052482.8

(74)专利代理机构 北京润平知识产权代理有限公司 11283

(22)申请日 2015.09.29

代理人 刘依云 严政

(30)优先权数据

62/056,852 2014.09.29 US

(51)Int.Cl.

C12N 15/82(2006.01)

(85)PCT国际申请进入国家阶段日

A23K 10/00(2006.01)

2017.03.28

(86)PCT国际申请的申请数据

PCT/US2015/052940 2015.09.29

(87)PCT国际申请的公布数据

W02016/054039 EN 2016.04.07

(71)申请人 谷万达公司

地址 美国马萨诸塞州

(72)发明人 R·M·拉布 M·拉纳汉 C·博宁

O·布格里

权利要求书5页 说明书63页

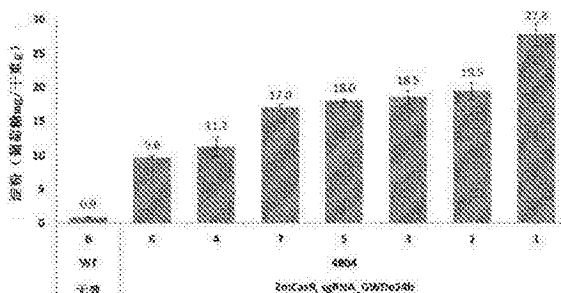
序列表89页 附图13页

(54)发明名称

具有工程内源基因的植物

(57)摘要

本发明提供了表达改变的葡聚糖水合二激酶且具有升高的淀粉水平的基因工程植物。本发明提供了基因工程改造植物以表达改变的葡聚糖水合二激酶的方法和基因构建体。本文描述了育种对于编码改变的葡聚糖水合二激酶的突变基因是纯合的基因工程植物的方法。本发明还提供了使用基因工程植物进行农业加工和制备动物饲料的方法。



1. 一种含有编码改变的葡聚糖水合二激酶的工程核酸的基因工程植物,其特征在于,与含有野生型葡聚糖水合二激酶的具有相同遗传背景的植物相比,所述植物具有升高的淀粉水平。

2. 根据权利要求1所述的基因工程植物,其中,与所述野生型葡聚糖水合二激酶的活性相比,所述改变的葡聚糖水合二激酶的活性降低。

3. 根据权利要求1所述的基因工程植物,其中,所述改变的葡聚糖水合二激酶是无活性的。

4. 根据权利要求1所述的基因工程植物,其中,与编码野生型葡聚糖水合二激酶的内源核酸相比,所述工程核酸为编码野生型葡聚糖水合二激酶的基因的至少一个等位基因的修饰序列。

5. 根据权利要求1所述的基因工程植物,其中,所述内源核酸含有与SEQ ID NO:1 (Zm GWD编码序列)或SEQ ID NO:2 (Sb GWD编码序列)所示的参考序列具有至少90%同一性的序列。

6. 根据权利要求4所述的基因工程植物,其中,所述修饰序列含有选自内源核酸中至少一个核苷酸的插入、缺失或取代中的至少一种的突变。

7. 根据权利要求6所述的基因工程植物,其中,所述突变在与参考序列具有至少90%同一性的靶序列内,所述参考序列选自由SEQ ID NO:3 (Zm GWD外显子24+内含子)、SEQ ID NO:4 (SbGWD外显子24+内含子)、SEQ ID NO:182 (ZmGWD外显子24无内含子)、SEQ ID NO:183 (Sb GWD外显子24)和SEQ ID NO:184 (SbGWD外显子7)所组成的组。

8. 根据权利要求6所述的基因工程植物,其中,所述突变在与参考序列具有至少90%同一性的靶序列内,所述参考序列选自由SEQ ID NO:91 (GWDe1a)、SEQ ID NO:92 (GWDe24b)、SEQ ID NO:93 (GWDe24c)和SEQ ID NO:94 (GWDe25a)所组成的组。

9. 根据权利要求1所述的基因工程植物,其中,所述工程核酸含有与选自由SEQ ID NO:12-40、114-120、131-146和188 (Zm GWD突变-外显子24)和SEQ ID NO:119和120 (Zm GWD突变-外显子25)所组成的组的序列具有至少90%同一性的多核苷酸。

10. 根据权利要求1所述的基因工程植物,其中,所述工程核酸含有与选自由SEQ ID NO:106 (Sb4715_1 (WT+ins) _外显子24)和SEQ ID NO:107 (Sb4715_2 (WT+ins) _外显子24)所组成的组的序列具有至少90%同一性的多核苷酸。

11. 根据权利要求1所述的基因工程植物,其中,所述改变的葡聚糖水合二激酶含有与选自由SEQ ID NO:45-73 (Zm GWD突变蛋白M1-M29)、121-124 (Zm GWD突变体M32-M36)、126-127 (Zm GWD突变体M38-M39)和147-162 (Zm GWD突变体M40-M55)所组成的组的序列具有至少90%同一性的氨基酸序列。

12. 根据权利要求1所述的基因工程植物,其中,所述改变的葡聚糖水合二激酶含有与SEQ ID NO:194 (Sb GWD突变蛋白Sb4715_1WT+ins)或SEQ ID NO:195 (Sb GWD突变蛋白Sb4715_2WT+del)所示的序列具有至少90%同一性的氨基酸序列。

13. 一种用于基因工程改造含有改变的葡聚糖水合二激酶的植物,该方法包括:
将至少一个植物细胞与含有第一核酸的载体接触,所述至少一个植物细胞含有存在于编码葡聚糖水合二激酶的内源基因中的靶序列,所述第一核酸编码能够诱导靶序列上的双链断裂的核酸酶;

- 选择在所述靶序列中具有改变的植物细胞;和
再生含有来自植物细胞的改变的基因工程植物。
14. 根据权利要求13所述的方法,其中,所述基因工程植物对于所述改变是纯合的。
15. 根据权利要求13所述的方法,其中,所述基因工程植物对于所述改变是杂合的。
16. 根据权利要求15所述的方法,其中,该方法还包括将杂合的基因工程植物自交,或与对于相同改变是杂合的另外的基因工程植物杂交,且选择改变是纯合的第一子代植物。
17. 根据权利要求15所述的方法,其中,该方法还包括将所述基因工程植物与具有相同遗传背景的野生型植物杂交,且选择改变是杂合的第一子代植物。
18. 根据权利要求17所述的方法,其中,该方法还包括将第一杂合的子代植物自交并选择改变是纯合的第二子代植物。
19. 根据权利要求13所述的方法,其中,所述改变包括突变。
20. 根据权利要求19所述的方法,其中,所述突变包括靶序列中至少一个核苷酸的插入、缺失或取代中的至少一种。
21. 根据权利要求20所述的方法,其中,所述突变是无效突变。
22. 根据权利要求13所述的方法,其中,与具有相同遗传背景的非基因工程植物相比,包含所述改变的基因工程植物或其子代具有升高的淀粉水平。
23. 根据权利要求13所述的方法,其中,所述核酸酶选自由大范围核酸酶、Cas9核酸酶、锌指核酸酶和转录激活子样效应物核酸酶所组成的组。
24. 根据权利要求23所述的方法,其中,所述核酸酶是大范围核酸酶,且由与参考序列具有至少90%同一性的序列编码,所述参考序列选自由SEQ ID NO:108(4715_大范围核酸酶)和SEQ ID NO:109(4716_大范围核酸酶)所组成的组。
25. 根据权利要求24所述的方法,其中,所述大范围核酸酶能够切割靶序列,且所述靶序列含有SEQ ID NO:41(4715_WWD-9/10x.272的靶标)或SEQ ID NO:42(4716_3e GWD-7/8x276的靶标)所示的多核苷酸。
26. 根据权利要求23所述的方法,其中,所述核酸酶是Cas9核酸酶。
27. 根据权利要求26所述的方法,其中,所述Cas9核酸酶由与SEQ ID NO:74(Cas9核酸酶)或SEQ ID NO:75(ZmCas9)所示的参考序列具有至少90%同一性的核酸编码。
28. 根据权利要求13所述的方法,其中,所述载体还含有编码sgRNA的第二核酸。
29. 根据权利要求28所述的方法,其中,所述sgRNA能够与靶序列结合,所述靶序列选自由SEQ ID NO:91(GWDe1a)、SEQ ID NO:92(GWDe24b)、SEQ ID NO:93(GWDe24c)和SEQ ID NO:94(GWDe25a)所组成的组。
30. 根据权利要求13所述的方法,其中,所述载体还含有与所述第一核酸可操作地连接的第一核酸启动子。
31. 根据权利要求28所述的方法,其中,所述载体还含有与所述第二核酸可操作地连接的第二核酸启动子。
32. 根据权利要求30或31所述的方法,其中,所述第一核酸启动子或所述第二核酸启动子含有与参考序列具有至少90%同一性的序列,所述参考序列选自由SEQ ID NO:78(MzU3.8)、SEQ ID NO:79(ZmU3)、SEQ ID NO:82(ZmU3P1)、SEQ ID NO:84(ZmU3P2)和SEQ ID NO:86(MzU3.8)所组成的组。

33. 根据权利要求13所述的方法,其中,所述第二核酸含有与参考序列具有至少90%同一性的序列,所述参考序列选自SEQ ID NO:95 (ZmU3P1:sgRNA_GWDe24b)、SEQ ID NO:96 (ZmU3P2:sgRNA_GWDe24b)、SEQ ID NO:97 (ZmU3.8P:sgRNA_GWDe24b)、SEQ ID NO:98 (ZmU3P2:sgRNA_GWDe24c)、SEQ ID NO:99 (ZmU3P2:sgRNA_GWDe25a)和SEQ ID NO:100 (ZmU3P2:sgRNA_GWDe1a)所组成的组。

34. 由权利要求13-31和33中任意一项所述的方法产生的基因工程植物、或其子代或其后代,其中,所述植物、其子代或其后代含有所述改变。

35. 根据权利要求34所述的基因工程植物,其中,与含有野生型葡聚糖水合二激酶的具有相同遗传背景的植物相比,所述基因工程植物具有升高的淀粉水平。

36. 一种提高植物中淀粉水平的方法,该方法包括表达编码能够诱导靶序列上的双链断裂的核酸酶的核酸,并选择与含有野生型序列的具有相同遗传背景的植物相比,在靶序列中具有改变且具有升高的淀粉水平的纯合植物。

37. 一种农业加工的方法,该方法包括:

表达编码能够诱导靶序列上的双链断裂的核酸酶的核酸,其中,所述靶序列被包含在编码葡聚糖水合二激酶的内源基因中;

选择与含有野生型序列的具有相同遗传背景的植物相比,在靶序列中具有改变且具有升高的淀粉水平的纯合植物;和

加工所述纯合植物,其中,所述加工包括选自收获、排水、粉碎、干燥、发酵、使用化学品水解、使用外源酶水解和与植物生物质组合中的一个或多个步骤。

38. 一种制备动物饲料的方法,该方法包括:

表达编码能够诱导靶序列上的双链断裂的核酸酶的核酸,其中,所述靶序列被包含在编码葡聚糖水合二激酶的内源基因中;

选择与含有野生型序列的具有相同遗传背景的植物相比,在靶序列中具有改变且具有升高的淀粉水平的纯合植物;和

进行选自以下所组成的组中的至少一个步骤:收获、排水、粉碎、干燥、青贮、造粒、与可食用纤维素源组合和与植物生物质组合。

39. 一种合成的核酸启动子,该核酸启动子含有与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列,所述参考序列选自SEQ ID NO:78 (MzU3.8)、SEQ ID NO:79 (ZmU3)、SEQ ID NO:82 (ZmU3P1)、SEQ ID NO:84 (ZmU3P2)和SEQ ID NO:86 (MzU3.8P)所组成的组。

40. 一种基因构建体,该基因构建体含有编码Cas9核酸酶的第一工程核酸序列,其中,所述Cas9核酸酶能够切割植物中被包含在编码葡聚糖水合二激酶的内源核酸中的靶序列。

41. 根据权利要求40所述的基因构建体,其中,所述第一工程核酸序列与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性,所述参考序列如SEQ ID NO:74 (Cas9核酸酶)或SEQ ID NO:75 (ZmCas9)所示。

42. 根据权利要求40所述的基因构建体,其中,所述第一核酸被融合至编码至少一个核定位信号(NLS)的多核苷酸序列。

43. 根据权利要求42所述的基因构建体,其中,所述多核苷酸序列选自SEQ ID NO:163-168。

44. 根据权利要求40所述的基因构建体,其中,所述基因构建体还含有编码sgRNA的第二工程核酸序列。

45. 根据权利要求44所述的基因构建体,其中,所述sgRNA能够结合靶序列。

46. 根据权利要求44所述的基因构建体,其中,所述第二工程核酸含有与如SEQ ID NO:135 (ZmU3P1:sgRNA_GWDe24b)、SEQ ID NO:136 (ZmU3P2:sgRNA_GWDe24b)、SEQ ID NO:137 (ZmU3.8P:sgRNA_GWDe24b)、SEQ ID NO:138 (ZmU3P2:sgRNA_GWDe24c)、SEQ ID NO:139 (ZmU3P2:sgRNA_GWDe25a)和SEQ ID NO:40 (ZmU3P2:sgRNA_GWDe1a)中任意一项所示的序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列。

47. 根据权利要求40所述的基因构建体,其中,所述靶序列与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性,所述参考序列选自由SEQ ID NO:91 (GWDe1a)、SEQ ID NO:92 (GWDe24b)、SEQ ID NO:93 (GWDe24c)和SEQ ID NO:94 (GWDe25a)所组成的组。

48. 根据权利要求40所述的基因构建体,其中,所述基因构建体还含有与第一工程核酸可操作地连接的第一核酸启动子。

49. 根据权利要求40所述的基因构建体,其中,所述基因构建体还含有与第二工程核酸可操作地连接的第二核酸启动子。

50. 根据权利要求48或49所述的基因构建体,其中,所述第一核酸启动子或所述第二核酸启动子为权利要求39所述的合成的核酸启动子。

51. 一种用于鉴定样品中编码葡聚糖水合二激酶的内源基因的修饰序列的试剂盒,其中,所述试剂盒含有第一引物和第二引物,其中,所述第一引物和所述第二引物能够扩增被包含在编码葡聚糖水合二激酶的内源基因中的靶序列,且所述靶序列含有与选自如SEQ ID NO:1-4、75、170-184、186、187、189-193所示的参考序列具有至少90%同一性的核酸序列。

52. 根据权利要求51所述的试剂盒,该试剂盒还含有用于检测所述靶序列的扩增区域中的修饰的一个或多个组件。

53. 根据权利要求52所述的试剂盒,其中,所述扩增区域中的修饰含有如SEQ ID NO:12-40、106-107、114-120、131-146和188所示的序列。

54. 根据权利要求51所述的试剂盒,其中,所述第一引物含有选自如SEQ ID NO:6、7、9、11、101、103、105、110和111所示的核酸序列。

55. 根据权利要求51所述的试剂盒,其中,所述第二引物含有选自如SEQ ID NO:5、8、10、102和104所示的核酸序列。

56. 根据权利要求51所述的试剂盒,其中,所述第一引物和所述第二引物能够扩增靶序列,以产生含有修饰的靶序列的扩增产物。

57. 根据权利要求51所述的试剂盒,其中,所述修饰的靶序列能够在高严格度的条件下,与含有选自如SEQ ID NO:12-40、106-107、114-120、131-146和188所示的序列的核酸序列杂交。

58. 根据权利要求51所述的试剂盒,其中,所述样品含有来自于基因工程植物的植物物质,所述基因工程植物在编码葡聚糖水合二激酶的内源基因中具有至少一个突变。

59. 一种鉴定样品中编码葡聚糖水合二激酶的内源基因的修饰序列的方法,该方法包括:

将样品与第一引物和第二引物接触；

扩增被包含在编码葡聚糖水合二激酶的内源基因中的靶序列，且所述靶序列含有与选自如SEQ ID NO:1-4、75、170-184、186、187、189-193所示的参考序列具有至少90%同一性的核酸序列；和

检测靶序列中的修饰。

60. 根据权利要求59所述的方法，其中，所述靶序列中的修饰含有选自如SEQ ID NO:12-40、106-107、114-120、131-146和188所示的序列。

具有工程内源基因的植物

[0001] 相关申请的交叉引用

[0002] 本申请要求于2014年9月29日提交的申请号为62/056,852的美国临时申请的权利,这些都通过引用并入本文,如同完整阐述。

[0003] 与本申请一起以电子方式提交的标题为“Sequence Listing”的序列列表创建于2015年9月29日,具有159,500字节大小,也通过引用并入本文,如同完整阐述。

[0004] 政府支持声明

[0005] 本发明是在由高级能源研究计划署ARPA-E颁发的奖励号为DE-AR0000042的政府支持下完成的。政府对本发明具有一定的权利。

技术领域

[0006] 本文的公开内容涉及具有优化的编码改变的葡聚糖水合二激酶(glucan water dikinase)的内源核酸序列和具有升高的淀粉水平的基因改良植物。本发明还涉及优化的编码改变的葡聚糖水合二激酶的核酸序列,优化内源核酸的方法,提高植物中的淀粉水平的方法以及制备和繁殖基因改良植物的方法。

背景技术

[0007] 在白天,植物在营养组织中合成淀粉并在晚上降解淀粉来动员产生的糖以支持植物的能量需求。在夜间,营养型植物细胞表达一系列酶以启动中间淀粉的动员。使淀粉磷酸化的葡聚糖水合二激酶(“GWD”)是这些酶之一。GWD的转录水平显示出经历了昼夜波动(Smith et al.Plant Phys.Preview, April 29, 2014)。增加生物质的淀粉含量可以增加动物饲料中的能量含量(卡路里)或提高从生物质中提取葡萄糖以用于乙醇或其他生化制品的生产。

[0008] 存在不同的用于操纵植物特性的分子方法。几乎所有的方法都依赖于通过转化过程将新的、合成的或重组的核酸插入至植物中。因此,插入的核酸可以编码核糖核酸(RNA)或蛋白质,它们可以被转化的植物表达从而改变植物表型。在许多情况下,核酸可以编码异源蛋白或产生更多的内源蛋白。类似地,转化的核酸可以产生通过多种机制(例如,RNA干扰、反义RNA等)降低内源基因的表达从而“沉默”基因并产生其产物的RNA。在所有情况下,插入植物中的核酸以显性方式表达,即,它的存在对植物的特性具有直接的影响。最近,已经证明通过在生物体中表达编码改变蛋白质(例如核酸酶)的脱氧核糖核酸(DNA)的核酸,生物体的基因组可以被永久地改变,即使在插入的核酸已经被除去以及内源基因优化之后。以这种方式,可能不仅产生有益的显性性状,而且产生具有非常特异性的靶向突变作为产生有益的隐性性状(recessive trait)的基础,而这本来是非常难以用于商业应用的发现和发展。目前,在行栽作物(row crops)的商业用途中尚没有使用核酸酶产生的隐性性状。使用核酸酶产生的隐性性状先前已在植物和植物细胞中证明,但是在完全发育的多细胞玉米和高粱植物(包括杂交玉米和高粱)中未见。像显性性状(dominant traits)一样,隐性性状可能具有商业价值,并且可能具有优于显性性状的特定商业优势(特别是安全和监

管方面的益处)。这种隐性性状将需要新的繁殖、跟踪和传递性状的方法,特别是在杂交作物中。

[0009] 显性性状特别是在杂交和交叉授粉作物如玉米中的一个问题是它们很容易地转移到相同物种的其他品系。在世界上的一些地区中,农民至少生产部分他们自己的用于种植的种子,这为将显性性状培育至农民现有的品系中提供了机会,而无需向技术所有者付费。已建立的性状商业模式目前需要种子和性状购买者向性状提供者支付特许权使用费,并且许可证通常将性状的使用限于单一种植并且禁止育种。对于很多性状,监测无授权的育种几乎是不可能的,并且在世界上的一些地方发生大量的性状的未经授权的性状转移(盗用,pirating)。根据性状,盗用或将性状转移到可用品系中而不向技术所有者付费是一项容易的工作,并且对于技术提供者来说难以检测。例如,抗虫性或农艺性状,它们的使用不需要任何其他材料,例如,除草剂抗性 or 特定肥料,如果它们已经转移到不同品系中,是几乎不可能检测的。可以产生后代,并且育种者可以使用商业上可以获得的测试条对其进行跟踪,或者如果该性状赋予了易获得的表型,可以通过表型进行跟踪。因为性状是显性的,所以对于农民使用而言在子代中它可能不需要是纯合的(homozygous),从而能够很容易继续育种和在技术许可者不知晓的情况下使用。

[0010] 与显性性状相反,隐性性状在作物中需要是纯合的,以便进行表型观察或易于评分。简单的测试条可能不可用于跟踪性状的分子基础,并且通过使用核酸酶产生的隐性性状的精准育种可能至少需要使用聚合酶链反应(PCR)进行检测。在这种情况下,携带该性状的纯合亲本的异交(outcross)所产生的后代将不显示该性状,并且在扩大育种、跟踪和一些情况下,杂交组合(hybrid crosses)将被要求使用该性状。这使得技术的盗用比使用显性性状显得更昂贵和困难。制备、维持和提供隐性性状的过程需要在显性性状的生产中所不必要的额外的步骤,因此,在种子和性状的生产中需要使用新的方法。

[0011] 基于含有特定基因突变的优化基因的隐性性状也可以具有超过使用转基因技术制备的显性性状的监管优势。因为这样的隐性性状可能不包含任何新引入的异源DNA,在世界的许多地区,它可能不会被作为转基因作物而被监管。

发明内容

[0012] 一方面,本发明涉及一种基因工程植物(genetically engineered plant),该基因工程植物含有编码改变的葡聚糖水合二激酶的工程核酸(engineered nucleic acid),其中,与含有野生型葡聚糖水合二激酶的具有相同遗传背景的植物相比,该植物具有升高的淀粉水平。

[0013] 一方面,本发明涉及一种基因工程改造植物的方法,以使其含有改变的葡聚糖水合二激酶。该方法包括将含有编码葡聚糖水合二激酶的内源基因中的靶序列的至少一个植物细胞与含有编码能够诱导靶序列上的双链断裂的核酸酶的第一核酸的载体接触。该方法还包括选择在靶序列中具有改变的植物细胞。该方法还包括从植物细胞再生包含改变的基因工程植物。

[0014] 一方面,本发明涉及一种提高植物中淀粉水平的方法。该方法包括表达编码能够诱导靶序列上双链断裂的核酸酶的核酸,其中,所述靶序列是编码葡聚糖水合二激酶的内源基因的序列。该方法还包括选择与含有野生型葡聚糖水合二激酶的具有相同遗传背景的

植物相比,在靶序列中具有改变并且具有升高的淀粉水平的纯合植物。

[0015] 一方面,本发明涉及一种农业加工的方法。该方法包括表达编码能够诱导靶序列上的双链断裂的核酸酶的核酸,其中,所述靶序列为编码葡聚糖水合二激酶的内源基因中的序列。该方法可以包括选择与含有野生型葡聚糖水合二激酶的具有相同遗传背景的植物相比,在靶序列中具有改变并且具有升高的淀粉水平的纯合植物。该方法还可以包括对所述纯合植物进行加工。

[0016] 一方面,本发明涉及一种制备动物饲料的方法。该方法包括表达编码能够诱导靶序列上的双链断裂的核酸酶和核酸,其中,所述靶序列是编码葡聚糖水合二激酶的内源基因的序列。该方法可以包括与含有野生型葡聚糖水合二激酶的具有相同遗传背景的植物相比,选择在靶序列中具有改变并且具有升高的淀粉水平的纯合植物。该方法还可以包括进行选自以下所组成的组中的至少一个步骤:收获(harvesting)、排水(bailing)、粉碎(shredding)、干燥(drying)、青贮(ensiling)、造粒(pelletizing)、与可食用的纤维素源结合和与植物生物质结合。

[0017] 一方面,本发明涉及一种制备对于工程核酸纯合的基因工程植物的方法,该工程核酸编码改变的葡聚糖水合二激酶,所述方法包括进行任意一种本文所述的对含有改变的葡聚糖水合二激酶的植物进行基因工程的方法。

[0018] 一方面,本发明涉及一种合成的核酸启动子(nucleic acid promoter),所述核酸启动子与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性(identity)的序列,所述参考序列选自:SEQ ID NO:78(MzU3.8)、SEQ ID NO:79(ZmU3)、SEQ ID NO:82(ZmU3P1)、SEQ ID NO:84(ZmU3P2)和SEQ ID NO:86(MzU3.8P)所组成的组中。

[0019] 一方面,本发明涉及一种含有编码Cas9核酸酶的第一工程核酸序列的基因构建体。所述Cas9核酸酶能够切割包含在编码植物中葡聚糖水合二激酶的内源核酸中的靶序列。

[0020] 一方面,本发明涉及一种鉴定样品中编码葡聚糖水合二激酶的内源基因的修饰序列的试剂盒。所述试剂盒含有第一引物和第二引物。所述第一引物和所述第二引物能够扩增包含在编码葡聚糖水合二激酶的内源基因中的靶序列。所述靶序列含有与选自SEQ ID NO:1-4、75、170-184、186、187、189-193的参考序列具有至少90%同一性的核酸序列。所述试剂盒还可以含有一种或多种组件用于检测靶序列扩增区域的修饰。所述修饰(modification)可以是编码本文所述的任意基因工程植物中的葡聚糖水合二激酶的内源基因的修饰序列(modified sequence)。

[0021] 一方面,本发明涉及一种鉴定样品中编码葡聚糖水合二激酶的内源基因的修饰序列的方法。该方法包括将样品与第一引物和第二引物进行接触。该方法包括扩增包含在编码葡聚糖水合二激酶的内源基因中的靶序列。所述靶序列含有与选自SEQ ID NO:1-4、75、170-184、186、187、189-193的参考序列具有至少90%同一性的核酸序列。该方法还包括检测靶序列的修饰。所述修饰可以是编码本文所述的任意基因工程植物中的葡聚糖水合二激酶的内源基因的修饰序列。

附图说明

[0022] 结合附图阅读将会更好地理解以下对本发明的实施方式的详细描述。出于说明的目的,图中显示的是具体的实施方式。然而,应当理解的是,本发明并不局限于显示出的具体设置和方法。在附图中:

[0023] 图1显示了用于表达大范围核酸酶(meganuclease)的载体pAG4715。

[0024] 图2显示了用于表达大范围核酸酶的载体pAG4716。

[0025] 图3显示了突变体的PCR检测。

[0026] 图4显示了描述通过使用载体pAG4715和pAG4716产生的突变的纯合子、杂合子和半合子玉米植物群之间的淀粉含量(mg淀粉/g干重)的图表。品系195、20、19、18、9和6为对照植物。

[0027] 图5显示了gwd敲除的玉米突变体的绿色组织中的淀粉含量:柱1是野生型(WT)植物,柱2是M17,柱3是M18,柱4是M1,柱5是M20,柱6是M13/M12,柱7是M9,柱8是M7/M11,柱9是M4/M14,柱10是M11/M12,柱11是M15,柱12是M14,柱13是M11,柱14是M11/M10,柱15是M13。

[0028] 图6显示了gwd敲除的(GWDko)大范围核酸酶穗轴(cobs)中的淀粉染色。

[0029] 图7显示了用于表达ZmCas9的载体pAG4800。

[0030] 图8显示了用于表达sgRNA支架(scaffold)和ZmCas9的载体pAG4804。

[0031] 图9显示了pAG4804玉米事件(maize events)中的淀粉积累。

[0032] 图10显示了pAG4806玉米事件中的淀粉积累。

[0033] 图11显示了源自玉米事件4716_164的靶向突变M20的自交和异交的示意图。

[0034] 图12显示了来自自交T0 4716_164 M20植株的T1子代的基因分型。

[0035] 图13显示了来自异交T0 4716_164 M20植株的T1子代的基因分型。

[0036] 图14显示了来自异交4716_164 M20植株的T1子代的基因分型。

具体实施方式

[0037] 以下描述中使用的特定术语仅为了方便而并非限制。

[0038] 本文使用的“工程核酸序列”、“工程多核苷酸”、“工程寡核苷酸”、“工程DNA”或“工程RNA”是指与自然界中的任意一种均不同的核酸、多核苷酸(polynucleotide)、寡核苷酸(oligonucleotide)、DNA或RNA,其具有与自然界中任意一种所不同的序列或者自然界中尚未发现的化学修饰。所述“工程核酸序列”、“工程多核苷酸”、“工程寡核苷酸”、“工程DNA”或“工程RNA”可以是合成的核酸序列、合成的多核苷酸、合成的寡核苷酸、合成的DNA或合成的RNA。工程核酸的定义包括但不限于使用生物技术工具产生的DNA序列。这些工具包括但不限于重组DNA技术、化学合成或核酸酶的定向使用(所谓的“基因组编辑”或“基因优化”技术)。

[0039] 本文所用的“内源核酸”是指天然存在于生物体或基因组中的核酸、多核苷酸、寡核苷酸、DNA或RNA。内源核酸可以是内源基因。

[0040] 本文所用的“改变的蛋白质”是指与天然存在的生物体,例如亲本生物体中所含的氨基酸序列相比含有至少一个氨基酸发生变化或缺失的蛋白质、多肽、寡肽或肽。改变的蛋白质可以保留或缺乏原始序列的生物活性。

[0041] 本文所用的“可操作地连接(operably linked)”是指两种或多种生物分子以相对于彼此的方式结合,以使所述生物分子能够进行正常功能。对于两个或多个核酸序列,“可

操作地连接”是指所述核酸序列以相对于彼此的方式结合,以使所述序列能够进行正常功能。例如,对于多肽,如果表达为参与多肽分泌的前蛋白(preprotein),编码前序列或分泌型前导体的核苷酸(nucleotide)序列与用于多肽的核苷酸序列可操作地连接;如果其影响编码序列的转录,启动子或增强子序列与编码序列可操作地连接;如果其被定位以促进核糖体与核酸结合,核酸核糖体结合位点与编码序列可操作地连接。

[0042] 本文所用的遗传背景定义为植物中所有基因或特定基因(例如,除基因工程修饰之外的所有基因)集合的总和。相同物种的植物可以被称为具有相同基因或相同遗传背景的植物。基因工程植物可以包括本文所述的工程核酸或多核苷酸,但是另外还具有与相同遗传背景的非基因工程植物相同的基因。

[0043] 用于权利要求和说明书相应部分的词“一种”(a)和“一个”(one),除特别声明外,定义为包括一个或多个引用项。该术语包括以上特别提到的词、其派生词和类似含义的词。短语“至少一种”之后列出的两项或多项,如“A、B或C”,意思为A、B或C中的任意单独的一个以及它们任意的组合。

[0044] 一种实施方式包括基因工程植物,该基因工程植物含有编码改变的葡聚糖水合二激酶的工程核酸。与具有相同遗传背景但含有野生型(wt)GWD的植物相比,基因工程植物具有升高的淀粉水平。与包含在具有相同遗传背景的植物中的野生型(wt)GWD相比,改变的GWD的活性可以被降低。基于wt GWD的水平,降低的水平可以是20、30、40、50、60、70、80、90、91、92、93、94、95、96、97、98、99或100%。可以通过监测植物中的淀粉含量,例如,通过使用本文实施例3中所述的傅里叶变换近红外(FT-NIR)技术来检测GWD的活性。改变的GWD可能是无活性的。增加的淀粉水平表明GWD的活性降低。

[0045] 在一种实施方式中,基因工程植物中的工程核酸可以含有内源核酸,所述内源核酸包括编码GWD蛋白但与野生型植物相比具有一种或多种修饰的gwd基因的至少一个等位基因。所述修饰可以通过对植物或其原种(ancestors)进行基因工程改造实现。所述内源核酸可以是工程植物中gwd基因的一个或多个等位基因。所述修饰可以位于gwd的编码序列中。所述内源核酸可以含有以下序列、基本上由以下序列组成或由以下序列组成:与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列,所述参考序列选自SEQ ID NO:1(Zm GWD编码序列)或SEQ ID NO:2(Sb GWD编码序列)。所述工程核酸可以包括相对于内源核酸的至少一个突变。与内源核酸相比,所述突变可以包括一个或多个核苷酸的插入。与天然核酸相比,突变可以包括核苷酸的缺失。与内源核酸相比,突变可以包括一个或多个核苷酸的取代。突变可以是几个突变的组合。所述至少一个突变可以在靶序列内与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性,所述参考序列选自SEQ ID NO:1(Zm GWD编码序列)、SEQ ID NO:2(Sb GWD编码序列)、SEQ ID NO:3(Zm GWD外显子24+内含子)、SEQ ID NO:4(SbGWD外显子24+内含子)、SEQ ID NO:91(GWDe1a)、SEQ ID NO:92(GWDe24b)、SEQ ID NO:93(GWDe24c)、SEQ ID NO:94(GWDe25a)、SEQ ID NO:182(ZmGWD外显子24)、SEQ ID NO:183(Sb GWD外显子24)、SEQ ID NO:184(SbGWD外显子7)和SEQ ID NO:189(Zm GWD外显子25)。

[0046] 在一种实施方式中,基因工程植物中的工程核酸可以是内源核酸,所述内源核酸包括编码GWD蛋白但具有通过对植物或其原种进行基因工程改造制备得到的具有一个或多个修饰的gwd基因的至少一个等位基因。内源基因可以是工程植物中gwd基因的一个或多个

等位基因。修饰可以位于gwd的编码序列中。内源核酸可以含有以下序列、基本上由以下序列组成或由以下序列组成：与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列，所述参考序列选自SEQ ID NO:1 (Zm GWD编码序列) 或SEQ ID NO:2 (Sb GWD编码序列)。工程核酸可以包括相对于内源核酸的至少一个突变。与内源核酸相比，突变可以包括一个或多个核苷酸的插入。与天然核酸相比，突变可以包括核苷酸的缺失。与内源核酸相比，突变可以包括一个或多个核苷酸的取代。突变可以是几个突变的组合。所述至少一个突变可以在与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性的靶序列内，所述参考序列选自SEQ ID NO:1 (Zm GWD编码序列)、SEQ ID NO:2 (Sb GWD编码序列)、SEQ ID NO:3 (Zm GWD外显子24+内含子)、SEQ ID NO:4 (SbGWD外显子24+内含子)、SEQ ID NO:91 (GWDe1a)、SEQ ID NO:92 (GWDe24b)、SEQ ID NO:93 (GWDe24c)、和SEQ ID NO:94 (GWDe25a)、SEQ ID NO:182 (ZmGWD外显子4)、SEQ ID NO:183 (Sb GWD外显子24)、SEQ ID NO:184 (SbGWD外显子7) 和SEQ ID NO:189 (Zm GWD外显子25)。

[0047] 在一种实施方式中，基因工程植物中的工程核酸含有玉米gwd基因的外显子24经过修饰的序列。工程核酸可以含有以下序列、基本上由以下序列组成或由以下序列组成：与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的多核苷酸，所述参考序列选自SEQ ID NO:12-40、114-118、119-120和131-146，其包含玉米gwd基因中的突变。

[0048] 在一种实施方式中，基因工程植物中的工程核酸可以含有在野生型Zm GWD基因 (SEQ ID NO:1) 的区域内具有一个或多个修饰的修饰序列，所述修饰序列位于3030核苷酸 (nt, 碱基) 至3243nt。在一种实施方式中，基因工程植物中的工程核酸可以含有在野生型Zm GWD基因 (SEQ ID NO:1) 的从3157nt至3213nt的位置内具有一个或多个修饰的修饰序列。在一种实施方式中，基因工程植物中的工程核酸可以含有在野生型Zm GWD基因外显子24 (SEQ ID NO:3) 的从81nt至160nt的位置内具有一个或多个修饰的修饰序列。

[0049] 在一种实施方式中，基因工程植物中的Zm GWD基因可以含有相对于野生型Zm GWD基因 (SEQ ID NO:1) 具有序列变化的修饰序列，其选自SEQ ID NO:12-40、114-118、119-120和131-146中的一个。具有位于变化位置之外的变化的Zm GWD序列可以与SEQ ID NO:1的相应区域具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性。所述变化可以与SEQ ID NO:12-40、114-118、119-120和131-146中的一个相同或不同。

[0050] 在一种实施方式中，基因工程植物中的工程核酸可以含有高粱gwd基因的外显子24的经修饰的序列。工程核酸可以含有以下序列、基本上由以下序列组成或由以下序列组成：与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列，所述参考序列选自SEQ ID NO:106 (Sb4715_1 (WT+ins) _外显子24) 和SEQ ID NO:107 (Sb4715_2 (WT+del) _外显子24)，这些是在高粱gwd基因的外显子24上的突变体。在一种实施方式中，基因工程植物的工程核酸可以含有在野生型Sb GWD基因 (SEQ ID NO:2) 的3030nt至3243nt位置内具有一个或多个修饰的修饰序列。基因工程植物的工程核酸可以含有在野生型Sb GWD基因 (SEQ ID NO:2) 的736nt至969nt位置内具有一个或多个修饰的修饰序列。改变的GWD可以由本文中任意一种工程核酸编码。

[0051] 在一种实施方式中，基因工程植物可以含有改变的玉米 (*Zea mays*) GWD (Zm GWD)。改变的Zm GWD可以含有以下氨基酸序列、基本上由以下氨基酸序列组成或由以下氨基酸序

列组成:与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性,所述参考序列选自SEQ ID NO:45-73 (Zm GWD突变蛋白M1-M29)、121-125 (Zm GWD突变蛋白M32-M36)、126-127 (Zm GWD突变蛋白M38-M39)和147-162 (Zm GWD突变蛋白M40-M55)。

[0052] 在一种实施方式中,基因工程植物中的改变的ZmGWD蛋白可以含有在野生型Zm GWD蛋白 (SEQ ID NO:43)的1040氨基酸(aa)至1120aa的位置内具有一个或多个修饰的修饰序列。基因工程植物中的改变的ZmGWD蛋白可以含有在野生型Zm GWD蛋白 (SEQ ID NO:43)的1054aa至1081aa的位置内具有一个或多个修饰的修饰序列。基因工程植物中的改变的ZmGWD蛋白可以含有在野生型Zm GWD蛋白 (SEQ ID NO:43)的1011aa至1057aa的位置内具有一个或多个修饰的修饰序列。基因工程植物中的改变的ZmGWD蛋白可以含有在野生型Zm GWD蛋白 (SEQ ID NO:43)的1082aa至1116aa的位置内具有一个或多个修饰的修饰序列。

[0053] 在一种实施方式中,基因工程植物中的Zm GWD蛋白可以含有相对于野生型Zm GWD蛋白 (SEQ ID NO:43)具有序列变化的的修饰序列,其选自SEQ ID NO:45-73 (Zm GWD突变蛋白M1-M29)、121-125 (Zm GWD突变蛋白M32-M36)、126-127 (Zm GWD突变蛋白M38-M39)和147-162 (Zm GWD突变蛋白M40-M55)中的一个。具有位于变化位置之外的变化的Zm GWD序列可以与SEQ ID NO:43的相应区域具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性。所述变化可以与SEQ ID NO:45-73 (Zm GWD突变蛋白M1-M29)、121-125 (Zm GWD突变蛋白M32-M36)、126-127 (Zm GWD突变蛋白M38-M39)和147-162 (Zm GWD突变蛋白M40-M55)中的一个相同或不同。

[0054] 在一种实施方式中,基因工程植物可以含有改变的高粱 (*Sorghum bicolor*) GWD (Sb GWD)。改变的Sb GWD可以包含以下氨基酸序列、基本上由以下氨基酸序列组成或由以下氨基酸序列组成:与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性,所述参考序列选自SEQ ID NO:194 (Sb GWD突变蛋白Sb4715_1WT+ins)和SEQ ID NO:195 (Sb GWD突变蛋白Sb4715_2WT+del)。本文的核酸、核苷酸序列、蛋白或氨基酸序列可以被分离、纯化、化学合成或通过重组DNA技术产生。所有这些方法是本领域公知的。

[0055] 在一种实施方式中,基因工程植物可以是任何类型的植物。基因工程植物可以是但不限于单子叶植物 (monocotyledonous plant)、双子叶植物 (dicotyledonous plant)、C4植物、C3植物、玉米 (corn)、大豆 (soybean)、水稻 (rice)、甘蔗 (sugar cane)、甜菜 (sugar beet)、高粱 (sorghum)、柳枝稷 (switchgrass)、芒草 (miscanthus)、桉树 (eucalyptus)、小麦、苜蓿 (alfalfa)、柳树 (willow) 或杨树 (poplar)。基因工程植物可以来自能源作物 (energy crop plant)、饲料作物 (forage crop plant) 或食物作物 (food crop plant)。所述能源作物植物可以是玉米植物、柳枝稷植物、高粱植物、杨树植物或芒草植物。所述饲料作物植物可以是玉米植物、苜蓿植物、高粱植物或大豆植物。所述食物作物植物可以是玉米植物、小麦植物、大豆植物、水稻植物或番茄植物。

[0056] 基因工程植物可以是转基因植物或突变植物。基因工程植物可以是转基因植物或突变植物的子代 (progeny), 或转基因植物或突变植物的后代 (descendant)。

[0057] 基因工程植物可以是在编码GWD的基因的核酸序列中具有一个或多个突变的常规突变体,其导致GWD的表达受抑制或GWD的活性降低。所述突变可以是GWD编码基因序列中核

酸的缺失、插入或取代。与具有相同遗传背景但表达野生型GWD的非突变植物相比，常规突变体可以具有改变的营养性淀粉水平。

[0058] 如本文所用，基因工程植物可以指完整的转基因植物或突变植物或其部分。该部分可以是但不限于叶、茎、花、芽、花瓣、子房、果实或种子中的一种或多种。该部分可以是来自转基因植物或突变植物的愈伤组织 (callus)。基因工程植物可以从转基因植物或突变植物的部分或者从植株再生。基因工程植物可以是第一转基因植物和第二转基因植物或非转基因植物的有性杂交的产物，其中，产物植物保留了被引入至第一转基因植物的工程核酸。基因工程植物可以是第一突变植物和第二非突变植物的有性杂交的产物，其中，产物植物保留了被引入第一突变植物的突变。转基因植物或突变植物可以是本文所述的转基因植物或突变植物中的任意一种。

[0059] 在一种实施方式中，提供了含有改变的葡聚糖水合二激酶的植物的基因工程方法。所述方法可以包括将至少一种含有编码葡聚糖水合二激酶的内源基因中的靶序列的植物细胞与载体接触。所述载体可以含有编码能够诱导靶序列上的单链断裂或双链断裂的核酸酶的第一核酸。载体可以通过转化或以其他基因工程的方式被引入植物。转化可以为使用含有编码核酸酶的第一核酸的载体的农杆菌 (*Agrobacterium*) 介导的转化。如前所述，核酸酶可以切割靶序列 (Puchta et al., 1993; Wright et al., 2005; Wehrkamp-Richter et al., 2009; Cong et al., 2013; Belhaj et al., 2013, 全部通过引用并入本文，如同完整阐述)。核酸酶可以是但不限于大范围核酸酶、Cas9核酸酶、锌指核酸酶 (zinc finger nuclease) 或转录激活子样效应物核酸酶 (transcription activator-like effector nuclease)。

[0060] 如本文所述，所述核酸酶可以是大范围核酸酶。大范围核酸酶可以引入单链或双链DNA断裂并且具有长度为14至40个核苷酸的识别位点，提供良好的特异性。对于大范围核酸酶用于靶向修饰的用途，参见Rosen et al., 2006; Wehrkamp-Richter et al., 2009; Djukanovic et al., 2013, 全部通过引用并入本文，如同完整阐述。大范围核酸酶可以是LAGLIDADG归巢核酸内切酶 (LHE)。LAGLIDADG归巢核酸内切酶 (LHEs) 是天然的基因靶向蛋白，其编码序列存在于内含子或内含肽 (intein) 中。参见Arnould et al., 2011, 全部通过引用并入本文，如同完整阐述。大范围核酸酶可以是I-CreI归巢核酸内切酶。如本文所用，I-CreI归巢核酸内切酶是天然存在于莱茵衣藻 (*Chlamydomonas reinhardtii*) 的叶绿体中的大范围核酸酶，并且是含有对于核酸酶活性很重要的单一序列基序 (motif) 的被充分表征的蛋白质。参见Heath et al., 1997, 全部通过引用并入本文，如同完整阐述。I-CreI核酸内切酶适用于蛋白质工程，并且被用于包括植物在内的若干物种中的靶向基因组修饰。参见Rosen et al., 2006; Arnould et al., 2007; Djukanovic et al., 2013, 全部通过引用并入本文，如同完整阐述。大范围核酸酶可以是I-DmoI、I-SceI、E-DmeI或DmoCre。可以使用其他大范围核酸酶。

[0061] 大范围核酸酶可以由与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列编码，所述参考序列选自自由SEQ ID NO:108 (4715_大范围核酸酶) 和SEQ ID NO:109 (4716_大范围核酸酶) 所组成的组。

[0062] 在一种实施方式中，所述核酸酶可以是Cas9核酸酶。所述Cas9核酸酶是在规律成簇间隔短回文重复 (Clustered Regularly Interspaced Short Palindromic Repeats,

CRISPR)/CRISPR相关蛋白9 (Cas9) 系统中使用的核酸酶 (Cong et al., 2013; Belhaj et al., 2013, 两者均通过引用并入本文, 如同完整阐述)。CRISPR/Cas9是基因组编辑技术, 由于它的低成本、高效率和对技术人员来说相对简单, 所以该技术有潜力成为各种物种基因组编辑的可选技术, 但是其尚未被证明通过使用稳定转化可以在多细胞植物中起作用。CRISPR/Cas9系统可以包括Cas9核酸酶和单引导RNA (single guide RNA, sg RNA)。本文的Cas9核酸酶可以由与SEQ ID NO:74 (Cas9核酸酶) 或SEQ ID NO:75 (ZmCas9) 的参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列编码。所述核酸酶可以对能够使核酸酶切割靶序列的序列具有亲和力, 或者可以通过使用sgRNA将其引导至靶序列。本文的载体可以进一步含有编码sgRNA的第二核酸序列。内源基因的靶向修饰可以通过在植物细胞中表达Cas9和sgRNA来进行。sgRNA嵌合体分子可以含有非翻译的CRISPR RNA (crRNA), 一个与具有3bp前间区序列邻近基序 (protospacer adjacent motif, PAM) 序列的靶基因组DNA序列互补的20bp间隔序列 (Jinek et al., 2012, 通过引入并入本文, 如同完整阐述)。Cas9核酸酶可以从植物例如拟南芥 (*Arabidopsis*)、玉米、烟草、水稻、小麦和高粱中的PPDK、CaMV 35S、肌动蛋白或泛素启动子表达。sgRNA可以主要从RNA聚合酶III启动子U6或U3和从RNA聚合酶II启动子CaMV E35S表达 (Belhaj et al., 2013; Upadhyay et al., 2014, 二者通过引用并入本文, 如同完整阐述)。sgRNA可以从本文所述的SEQ ID NO:78 (MzU3.8)、SEQ ID NO:79 (ZmU3)、ZmU3P1 (SEQ ID NO:82)、ZmU3P2 (SEQ ID NO:84) 或ZmU3.8启动子 (SEQ ID NO:86) 表达。所述启动子可以与SEQ ID NO:78 (MzU3.8)、SEQ ID NO:79 (ZmU3)、ZmU3P1 (SEQ ID NO:82)、ZmU3P2 (SEQ ID NO:84) 或ZmU3.8启动子 (SEQ ID NO:86) 中的一种具有70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性。所述启动子可以具有等于SEQ ID NO:78 (MzU3.8)、SEQ ID NO:79 (ZmU3)、ZmU3P1 (SEQ ID NO:82)、ZmU3P2 (SEQ ID NO:84) 或ZmU3.8启动子 (SEQ ID NO:86) 中的一种的核苷酸长度的长度。比SEQ ID NO:78 (MzU3.8)、SEQ ID NO:79 (ZmU3)、ZmU3P1 (SEQ ID NO:82)、ZmU3P2 (SEQ ID NO:84) 或ZmU3.8启动子 (SEQ ID NO:86) 更短的启动子的同一性百分数可以如上所述沿着较短启动子的长度。Cas9核酸酶可以将单链断裂或双链DNA断裂引入包含在基因组DNA中的内源核酸中。随后, 由Cas9在基因组DNA中引入的断裂可以通过两种不同的机制NHEJ (非同源末端连接) 和HR (同源重组) (Symington and Gautier, 2011, 通过引用并入本文, 如同完整阐述) 进行修复。

[0063] 在一种实施方式中, 所述核酸酶可以是转录激活子样效应物核酸酶 (transcription activator-like effector nucleases, TALEN)。本文所用的TALEN是指来自黄单胞菌属 (*Xanthomonas*) 的蛋白质。TALEN是可程序化运行的融合蛋白, 其包含与FokI核酸内切酶的DNA切割结构域融合的TAL效应子的工程DNA结合结构域 (Boch and Bonas, 2010; Christian et al., 2010; Jung and Sander, 2013; Li et al., 所有这些文献通过引用并入本文, 如同完整阐述)。这些嵌合蛋白可以以成对的两个单体起作用, 以将FokI核酸内切酶靶向作用于基因组内用于DNA切割的特定DNA序列。可以修饰TAL DNA结合结构域以识别不同的序列 (Cermak et al., 2011, 其通过引用并入本文, 如同完整阐述)。

[0064] 在一种实施方式中, 所述核酸酶可以是锌指核酸酶。(Wright et al., 2005; Shukla et al., 2009, 二者通过引用并入本文, 如同完整阐述)。

[0065] 在一种实施方式中, 所述核酸酶可以是任何其他的适用于靶序列的靶向修饰的核

酸酶。

[0066] 所述靶序列可以为靶基因。所述靶基因可以是植物天然的内源基因。所述靶序列可以是植物的gwd基因。所述靶序列可以被包含在SEQ ID NO:1 (Zm GWD编码序列)或SEQ ID NO:2 (Sb GWD编码序列)内。靶序列可以是包含在编码GWD的内源核酸的外显子中的任何核酸序列。所述靶序列可以被包含在编码玉米GWD的内源核酸的外显子中。所述靶序列可以包含在编码高粱GWD的内源核酸的外显子中。靶序列可以包含在编码GWD的内源核酸的外显子1、外显子7、外显子24或外显子25中。所述靶序列可以是大范围核酸酶的靶序列。所述靶序列可以含有以下序列、基本上由以下序列组成或由以下序列组成:与SEQ ID NO:41 (大范围核酸酶GWD-9/10x.272靶序列 (pAG4715))或SEQ ID NO:42 (大范围核酸酶GWD-7/8x靶序列 (pAG4716))具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列。所述靶序列可以是Cas9核酸酶的靶序列。所述靶序列可以含有以下序列、基本上由以下序列组成或由以下序列组成:与SEQ ID NO:91 (GWDe1a)、SEQ ID NO:92 (GWDe24b)、SEQ ID NO:93 (GWDe24c)或SEQ ID NO:94 (GWDe25a)具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列。所述sgRNA能够结合靶序列,所述靶序列选自自由SEQ ID NO:91 (GWDe1a)、SEQ ID NO:92 (GWDe24b)、SEQ ID NO:93 (GWDe24c)和SEQ ID NO:94 (GWDe25a)所组成的组。所述靶序列可以是与sgRNA杂交的任何序列。所述核酸酶可以对能够使核酸酶切割靶序列的序列具有亲和力,或者可以通过使用sgRNA将其引导至靶序列。

[0067] 一旦表达,核酸酶将在靶序列中引起单链或双链DNA断裂。例如,核酸酶可以切断短片段,然后通过细胞的DNA修复机制进行部分修复,但是在靶序列内留下损伤。修复的靶序列可以包括改变。所述改变可以包括突变。所述突变可以是靶序列中一个或多个核苷酸的插入、缺失或取代中的至少一种。所述突变可以是无效突变(null mutation)。如本文所用,术语“无效突变”是指基因中的突变,能够导致其不被转录成RNA或翻译成功能性蛋白质。由于靶序列中的突变,天然核酸序列可以编码改变的GWD。改变的GWD的活性可以被降低。降低的水平可以是野生型GWD活性水平的20、30、40、50、60、70、80、90、91、92、93、94、95、96、97、98、99或100%,其可以通过使用本文实施例3中所述的傅里叶变换近红外(FT-NIR)技术来监测植物中的淀粉含量进行检测。改变的GWD可能是无活性的。与具有相同遗传背景的非基因工程植物相比,基因工程植物具有升高的淀粉水平。

[0068] 所述方法可以包括选择含有靶序列的改变的植物细胞。所述方法可以包括从植物细胞再生含有所述改变的植物。所述基因工程植物对于该改变可以是纯合的。

[0069] 所述基因工程植物对于所述改变可以是杂合的。本文中的基因工程植物对于含有突变的基因可以是杂合的。所述基因可以包括编码改变的GWD的工程核酸。杂合植物可以含有编码野生型、未改变的GWD的内源基因的等位基因。当编码GWD的基因的至少一个等位基因缺失时,所述杂合植物还可以包括半合子植物。杂合植物可以在表型上与野生型植物无法区分,并且可以不具有升高的淀粉水平。为了产生具有升高的淀粉水平的纯合植物,杂合的基因工程植物可以进行自交。可以从这样的杂交获得子代。所述子代可以包括纯合子、杂合子和野生型植物。杂合植物可以在表型上与野生型植物无法区分。所述方法可以包括分析所述子代是否存在所述改变,并选择含有所述改变的子代植物。

[0070] 在一种实施方式中,所述方法可以进一步包括为了相同的改变将杂合的基因工程植物与另一个杂合的基因工程植物杂交。该方法可以包括选择对于该改变是纯合的第一子

代植物。所述方法可以进一步包括将基因工程植物与相同遗传背景的野生型植物杂交。可以从这样的杂交获得子代。子代可以包括杂合植物和野生型植物。所述方法可以包括选择对于所述改变是杂合的第一子代植物。所述方法可以进一步包括使第一杂合子代植物自交和选择对于该改变是纯合的第二子代植物。

[0071] 本文的基因工程植物对于含有突变的基因来说可以是纯合的或杂合的,并且可以含有编码核酸酶的转基因。所述编码核酸酶的转基因可以在上述杂交过程中分离。

[0072] 一种实施方式包括生产对于编码蛋白质的工程核酸是纯合的基因工程植物的方法。工程核酸可以编码隐性性状。所述隐性性状可以包括基因的切割的内源靶序列。所述隐性性状仅可以在不含有基因的未改变的野生型等位基因的植物中观察到。该方法可以包括通过修饰内源核酸的序列以制备工程核酸。该方法还可以包括将隐性性状育种至其他作物品系中。该方法可以包括保持作物品系的性状。该方法可以包括产生纯合子代。该方法可以包括制备具有隐性性状的杂种种子。

[0073] 一种实施方式包括通过本文所述的任意一种方法获得的基因工程植物。

[0074] 一种实施方式包括提高植物中淀粉水平的方法。该方法可以包括在植物中表达编码大范围核酸酶的核酸。该方法可以包括在植物中表达编码TALEN的核酸。该方法可以包括表达编码Cas9核酸酶的第一核酸和编码靶向特定序列的所需引导RNA的第二核酸。核酸在植物中的表达可以改变内源DNA序列的功能或编码。核酸在植物中的表达可以改变植物中GWD的活性和淀粉代谢。所述植物可以是本文中的任意转基因或突变植物。所述植物可以是转基因植物或突变植物的子代。核酸可以包含在基因构建体中。该方法可以包括制备本文中的任意基因工程植物。所述基因工程植物或其子代可以通过本文的方法得到的淀粉水平升高的植物。

[0075] 具有编码大范围核酸酶的核酸的基因构建体可以在所述方法中的任意点被表达,所述大范围核酸酶可以灭活或抑制参与植物中淀粉动员的GWD蛋白的表达。所述核酸可以在处理植物的步骤之前表达。所述核酸可以在处理植物的步骤期间表达。所述表达可以被诱导。在表达核酸时,与具有相同遗传背景但缺少一个或多个基因构建体的非基因工程植物中的淀粉水平相比,所述基因工程植物可以具有改变的营养性淀粉水平。

[0076] 本文的任意基因工程植物可以由农业加工方法、制备动物饲料的方法或饲喂动物的方法提供。提供基因工程植物的步骤可以包括从生产它的另一方获得。提供的步骤可以包括制备基因工程植物。所述基因工程植物可以是转基因植物或突变植物。提供的步骤可以包括通过将植物与本文的任意一种基因构建体接触以进行转化。提供的步骤可以包括通过本文所述的任何方法或已知方法稳定转化植物。提供的步骤可以包括将植物与含有编码核酸酶的多核酸的基因构建体接触之后,通过在植物中瞬时表达的核酸酶识别的切割位点切割编码参与淀粉代谢的蛋白的基因以对植物进行基因工程改造。提供的步骤还可以包括从具有改变的营养性淀粉水平的基因工程植物的组织再生植物。提供的步骤可以包括获得来自自花授粉或基因工程植物与非基因工程植物之间的异花授粉产生的基因工程植物的子代。提供的步骤可以包括获得纯合的子代。所述纯合的子代可以是近交植物。所述纯合的子代可以是杂交植物。基因工程植物可以被用于多种后续方法或用途。提供的步骤可以包括购得所述基因工程植物。提供的步骤可以包括使基因工程植物可用于进一步加工步骤。提供步骤可以包括使基因工程植物可用于作为动物饮食的部分。

[0077] 在农业加工方法中,基因工程植物可以是具有升高的淀粉水平经改造的原料和/或表达一种或多种多糖降解酶的原料。所述原料可以包括单独的或与其他组分组合的任何基因工程植物。所述其他组分可以包括其他植物材料。农业加工可以包括操作或转换任何农业原料,包括用于特定产品或用途的基因工程植物。农业加工可以包括干燥基因工程植物。农业加工可以包括发酵基因工程植物。农业加工可以包括使用一种或多种外源酶水解基因工程植物以获得生化产物。所述外源酶可以是木质素降解酶、纤维素降解酶或半纤维素降解酶。所述外源酶可以是糖苷酶、木聚糖酶、纤维素酶、内切葡聚糖酶、外切葡聚糖酶、纤维二糖水解酶、 β -木糖苷酶、阿魏酸酯酶(feruloyl esterases)、 β -葡糖苷酶和淀粉酶。所述外源酶可以从供应商处购买,并且可以包括可以从杰能科国际公司(Genencor International, 罗切斯特, 纽约)获得的Accellerase 1000、Accellerase 1500、Accelerase TRIO™和Accellerase XY。所述外源酶可以包括来自诺维信公司(Novozymes, 丹麦)的Cellic、CTEC、HTEC。所述外源酶可以包括淀粉降解酶。所述外源酶可以包括淀粉酶或转化酶。所述农业加工方法可以包括同时糖化和发酵可溶性糖以生产乙醇。

[0078] 本文所述的农业加工方法可以包括收获具有升高的淀粉水平的基因工程植物,以用作农业加工的原料。该方法可以包括将基因工程植物与植物生物质相结合。所述植物生物质可以包括非基因工程植物。所述植物生物质可以是基因工程植物生物质。所述基因工程植物生物质可以表达多糖降解酶。通过将基因工程植物与表达多糖降解酶的植物生物质组合,本文所述的方法可以不需要苛刻的预处理,以提高纤维素细胞壁对外源酶的易接近性。本文的方法可以利用用于表达细胞壁降解酶的植物生物质的固结预处理和水解的任何方法和组合物,所述细胞壁降解酶于2012年3月7日提交的申请号为13/414,627的美国专利申请和2012年3月7日提交的申请号为PCT/US2012/028132的国际专利申请中进行描述,均通过引用并入本文,如同完整阐述。具有改变的升高的淀粉水平的植物在2011年6月27日提交的申请号为PCT/US2011/041991的国际专利申请、2013年3月19日提交的申请号为13/806,654的美国专利和2013年3月11日提交的申请号为13/793,078的美国专利申请中进行了描述,全部通过引用并入本文,如同完整阐述。

[0079] 基因工程植物可以以制备动物饲料的方法提供。制备动物饲料可以包括将基因工程植物与动物饲料(包括但不限于玉米、谷物、大豆和/或其他饲料)相结合。制备动物饲料可以包括青贮基因工程植物,以制备青贮饲料。制备动物饲料可以包括将基因工程植物与酒糟(distillers' grains)组合。制备动物饲料可以包括将基因工程植物造粒成饲料颗粒。制备动物饲料可以包括将基因工程植物与可食用纤维源组合。制备动物饲料可以包括将基因工程植物与蛋白质源组合。制备动物饲料可以包括将基因工程植物与一种或多种作为能量源的碳水化合物组合。制备动物饲料可以包括将基因工程植物与本文所述的一种或多种外源酶组合。

[0080] 农业加工方法或制备动物饲料的方法还可以包括收获(harvesting)、打包(baling)、研磨(grinding)、碾磨(milling)、切碎(chopping)、减小尺寸、粉碎(crushing)、从原料中提取组分、纯化原料的组分或部分、提取或纯化淀粉、将多糖水解成寡糖或单糖、原料的化学转化或化学催化。

[0081] 在一种实施方式中,提供了在营养组织中含有升高的淀粉水平的动物饲料制剂。动物饲料制剂可以被用于通过饲喂动物具有升高的淀粉水平的植物材料以提高牛奶和牛

肉产量。发酵过程中可用的易发酵的糖可以由本文的实施方式提供。生物燃料的生产可以通过提供易发酵的糖来加强。提供易发酵的糖的方法和加强生物燃料的生产的方法作为本文的实施方式提供。动物饲料制剂可以含有一种或多种本文中的基因工程植物。动物饲料制剂可以包括本文的制备动物饲料的方法的产物。

[0082] 具有升高的营养性淀粉水平的作物可以具有多种用途和效用。在一种实施方式中,提供了来自相对于野生型植物积累营养性淀粉升高水平的植物中的生物质。所述生物质可以来自本文中的任何基因工程植物或其子代。这些植物可以具有作为用于发酵过程或动物饲料应用的原料的附加值。例如,在典型的纤维素工艺中,存在于生物质中的多糖,例如,纤维素和半纤维素,被水解成单糖,然后通过微生物将其发酵成乙醇、丁醇、异丁醇、脂肪酸或其他烃类。由于生物质的难降解性,从聚合物如纤维素和半纤维素中释放单糖通常需要使用苛刻的预处理条件和使用相对昂贵的酶混合物进行水解。类似的情况发生在食用饲料,包括青贮玉米,作为营养和能量来源的反刍动物中。在反刍动物中,饲料被咀嚼并移动至瘤胃中,在其中纤维多糖,如纤维素和半纤维素,被瘤胃菌群中的微生物水解和发酵。这些有机物产生被动物吸收并代谢的脂肪酸,为动物提供营养。在反刍动物消化或生物燃料处理中,存在于生物质中的任何淀粉代表易于发酵的糖(即葡萄糖)的额外来源,其对于水解较不耐受并且可以通过淀粉酶或温和的化学处理非常容易地释放。结果生物质中存在的淀粉量的任何增加将同时增加可回收的可发酵糖的量。含有升高水平的淀粉的生物质在饲料应用中可能具有更大的价值,其中,将植物材料喂养给家畜或奶畜。同样,存在于该材料中的过量淀粉比纤维素材料更容易被大多数动物消化,与具有普通水平的淀粉的生物质相比,每单位生物质提供更多的能量。实施方式包括将本文所述的植物用于这些方法中的任意一种。

[0083] 本文中的方法,包括前面段落中的那些,可以包括下述内容中的至少一种:改造植物以产生基因工程植物、种植基因工程植物、收获基因工程植物、按照其他饲料作物将其加工(例如,减小饲料的尺寸、青贮、使用接种剂处理、与其他饲料组分组合,或造粒)用于动物饲料应用、或以与用于纤维素加工中使用的处理方式类似的方式发酵基因工程植物。所使用的纤维素加工步骤可以包括通过酶促或化学水解或消化来预处理和水解多糖成其组分糖。本文的方法中可以提供本段落中阐述的任何一个步骤,一组步骤或所有步骤。

[0084] 一种实施方式包括基因构建体,旨在实施改变植物中营养性淀粉水平的策略。所述基因构建体可以含有编码能够切割编码GWD的内源核酸中的靶序列的核酸酶的第一工程核酸序列。所述第一工程核酸可以编码本文所述的任何一种核酸酶。所述基因构建体还可以含有编码sgRNA的第二工程核酸序列。所述第二工程核酸可以编码本文所述的任何一种sgRNA。所述基因构建体可以含有与第一工程核酸序列或第二工程核酸可操作地连接的启动子。所述可操作地连接的启动子可以允许编码核酸酶的第一工程核酸序列或编码sgRNA的第二工程核酸序列的转录。第一工程核酸序列的转录和翻译可以被称为核酸酶的表达。一旦表达,所述核酸酶可以切割内源核酸的靶序列。所述内源核酸可以编码GWD。所述第二工程核酸序列的转录可以导致识别内源核酸内的靶序列的sgRNA的产生,并将Cas9核酸酶引导至靶标以使其断裂。

[0085] 所述基因构建体可以含有编码核定位信号的区域。本文所使用的SEQ ID NO:135核定位信号(nuclear localization signal,NLS)是指核蛋白内的碱性氨基酸序列的短基

序。某些蛋白质从细胞质转运到核仁 (nucleolus) 中以发挥其特定功能,通过核包膜发生,并涉及核孔复合物 (NPC) (Wagner et al.,1990,通过引用并入本文,如同完整阐述)。在该过程中,核定位信号 (NLS) 发挥重要的作用,因为它们被认为被NPC受体识别,随后通过核孔复合体转移蛋白质。NLS属于若干定义的类别之一 (Garcia-Bustos et al.,1991,通过引用并入本文,如同完整阐述)。NLS可以是来自猿猴病毒40大T抗原的SV40NLS,由于其在各种包括植物在内的生物体中的活性而在靶向基因组修饰的实验中被集中使用 (Kalderon et al.,1984;Raikhel,1992,这些都通过引用并入本文,如同完整阐述)。SV40NLS可以由与SEQ ID NO:163的参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性的核酸序列编码。SV NLS可以含有以下序列、基本上由以下序列组成或由以下序列组成:与SEQ ID NO:196的参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的氨基酸序列。所述NLS可以是植物特异性NLS序列。还描述了所述植物特异性NLS序列,例如,在玉米调节蛋白opaque-2和R (Varagona et al.,1992; Shieh et al.,1993,两者通过引用并入本文,如同完整阐述)。所述植物特异性NLS可以由与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的核酸序列编码,所述参考序列选自自由SEQ ID NO:164 (NLS1)、165 (NLS3)、166 (NLS4)、167 (NLS5) 和168 (NLS6) 所组成的组。所述植物特异性NLS可以含有以下序列、基本上由以下序列组成或由以下序列组成:与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的氨基酸序列,所述参考序列选自自由SEQ ID NO:128 (NLS1)、129 (NLS3)、130 (NLS4)、169 (NLS5) 和170 (NLS6) 所组成的组。所述NLS序列可以是衍生的NLS序列。NLS序列或其衍生物可以被用于将大范围核酸酶、ZFN、TALEN或Cas9蛋白靶向作用于植物细胞核中以用于靶向基因组修饰。一个或多个NLS序列可以与所述核酸酶的氨基酸序列相融合。

[0086] 所述基因构建体可以进一步含有与编码核酸酶的核酸可操作地连接的一个或多个调节序列(也称为调节元件)。启动子可以是任何种类的启动子。所述启动子可以是诱导型启动子。所述启动子可以是组成型启动子。所述启动子可以是诱导型启动子,其仅当暴露于特定的化学或环境刺激时才启动编码核酸酶的核酸的转录。所述诱导型启动子的实例包括但不限于醇诱导型启动子、四环素诱导型启动子、类固醇诱导型启动子或激素诱导型启动子。所述启动子可以是组成型启动子,其在大多数细胞、组织和器官中以及在许多但不一定是所有发育阶段期间在整个植物中提供核酸或多核酸序列的转录。所述启动子可以对特定发育阶段、器官或组织具有特异性。组织特异性启动子可以能够在特定植物组织中启动转录。可以被组织特异性启动子靶向作用的植物组织可以为但不限于茎、叶、毛状体 (trichomes)、花药 (anther) 或种子。本文的组成型启动子可以是水稻泛素3启动子 (OsUbi3P) 或玉米泛素启动子 (ZmUbi1)。其他已知的组成型启动子可以是本文的基因构建体的一部分,并且包括但不限于花椰菜花叶病毒 (Cauliflower Mosaic Virus,CAMV) 35S启动子、夜香树黄化卷曲病毒启动子 (Cestrum Yellow Leaf Curling Virus,CMP) 或CMP短型 (CMPS)、二磷酸核酮糖羧化酶小亚基启动子 (Rubisco small subunit promoter)、水稻肌动蛋白启动子 (rice actin promoter,OsAct1P) 和玉米磷酸烯醇丙酮酸羧化酶启动子 (maize phosphoenolpyruvate carboxylase promoter,ZmPepCP)。所述启动子可以是来自玉米的合成的核酸启动子。所述来自玉米的合成的核酸启动子可以含有以下序列、基本上

由以下序列组成或由以下序列组成:与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列,所述参考序列选自SEQ ID NO:78 (MzU3.8)、SEQ ID NO:79 (ZmU3)、ZmU3P1 (SEQ ID NO:82)、ZmU3P2 (SEQ ID NO:84) 或ZmU3.8启动子 (SEQ ID NO:86) 所组成的组。所述合成的核酸启动子可以具有等于SEQ ID NO:78 (MzU3.8)、SEQ ID NO:79 (ZmU3)、ZmU3P1 (SEQ ID NO:82)、ZmU3P2 (SEQ ID NO:84) 或ZmU3.8启动子SEQ ID NO:86) 中的一种的核苷酸长度的长度。比SEQ ID NO:78 (MzU3.8)、SEQ ID NO:79 (ZmU3)、ZmU3P1 (SEQ ID NO:82)、ZmU3P2 (SEQ ID NO:84) 或ZmU3.8启动子SEQ ID NO:86) 短的启动子的同一性百分数可以如上所述沿着较短启动子的长度。一种实施方式包括本文所述的任何一种合成的核酸启动子。所述合成的核酸启动子可以与第一工程核酸或第二改造核酸分子可操作地连接,并且可以转录激活所述第一或第二工程核酸。作为转录激活的结果,所述第一或第二工程核酸可以在植物中组成型表达。

[0087] 本文的基因构建体中的调节元件可以是终止子。终止子能够终止转录。终止子序列可以被包括在表达盒 (expression cassette) 的转录单元的3'端。所述转录单元可以编码核酸酶。所述终止子可以是来源于在多种植物基因中发现的终止子。所述终止子可以是来自根癌土壤杆菌 (*Agrobacterium tumefaciens*) 的胭脂碱合酶 (NOS) 或章鱼碱合酶 (OCS) 基因的终止子序列。所述终止子可以是酿脓链球菌 (*S. pyogenes*) Cas9终止子 (SEQ ID NO:88)。所述终止子可以是ZmU3T终止子 (SEQ ID NO:89)。所述终止子序列可以是来自CaMV的CaMV 35S终止子,或显示出可以终止植物中转基因转录的任何3'UTR序列。例如,所述终止子可以是玉米PepC终止子 (3'UTR)。所述基因构建体可以被包括在载体中。所述基因构建体可以被整合到基因工程植物的基因组中。所述基因构建体可以在基因工程植物中瞬时表达。

[0088] 所述基因构建体可以被用于植物的转化。所述基因构建体可以被用于农杆菌 (*Agrobacterium*) 介导的植物转化。所述基因构建体可以通过任何已知的方法用于转化植物,例如,粒子轰击或直接DNA摄取。所述基因构建体可以被克隆并被并入载体中。

[0089] 一种实施方式包括含有本文的基因构建体并适用于基因工程改造植物的载体。所述载体可以是中间载体。所述载体可以是转化载体。并入了本文的基因构建体的载体还可以含有另外的基因元件,例如,促进分子克隆的多克隆位点和一个或多个选择性标记物以促进选择。包含在载体中的选择性标记物可以是来自大肠杆菌 (*Escherichia coli*) 的磷酸甘露糖异构酶 (PMI) 基因,其赋予转化细胞利用甘露糖生长的能力。包含在载体中的选择性标记物包括但不限于赋予卡那霉素抗性的新霉素磷酸转移酶 (npt) 基因、赋予潮霉素抗性的潮霉素磷酸转移酶 (hpt) 基因或赋予草甘膦抗性的烯醇丙酮酰莽草酸-3-磷酸合酶 (enolpyruvylshikimate-3-phosphate synthase gene) 基因。所述载体可以是在2013年3月11日提交的申请号为13/793,078的美国专利申请中描述的任何载体。全部通过引用并入本文,如同完整阐述。所述载体可以包括编码本文所述的任何一种核酸酶的基因构建体。所述载体可以含有与SEQ ID NO:108 (大范围核酸酶4715) 或SEQ ID NO:109 (大范围核酸酶4716) 的参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的核酸序列。所述载体可以含有与SEQ ID NO:75 (Zm Cas9) 的参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的核酸序列。所述载体可以含有与参考序列具有与至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同

一性的序列,所述参考序列选自由SEQ ID NO:95 (ZmU3P1:sgRNA_GWDe24b)、SEQ ID NO:96 (ZmU3P2:sgRNA_GWDe24b)、SEQ ID NO:97 (ZmU3.8P:sgRNA_GWDe24b)、SEQ ID NO:98 (ZmU3P2:sgRNA_GWDe24c)、SEQ ID NO:99 (ZmU3P2:sgRNA_GWDe25a)和SEQ ID NO:100 (ZmU3P2:sgRNA_GWDe1a)所组成的组。本文所述的载体或基因构建体可以含有工程核酸。所述载体可以是pAG4715(图1)、pAG4716(图2)或其修饰,使用本文另外描述的对应用物代替任何一处注释标志。在图1和2中注释的常规载体元件可以被本文中描述的或本领域已知的对应用物替代。

[0090] 一种实施方式包括工程核酸,该工程核酸含有以下序列、基本上由以下序列组成或由以下序列组成:与SEQ ID NO:41(大范围核酸酶GWD-9/10x.272靶序列(pAG4715))或SEQ ID NO:42(大范围核酸酶GWD-7/8x靶序列(pAG4716))具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列。

[0091] 一种实施方式包括工程核酸序列,该工程核酸序列含有以下序列、基本上由以下序列组成或由以下序列组成:与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列,所述参考序列选自由SEQ ID NO:91(GWDe1a)、SEQ ID NO:92(GWDe24b)、SEQ ID NO:93(GWDe24c)和SEQ ID NO:94(GWDe25a)所组成的组。

[0092] 一种实施方式包括具有如本文所列的任意一个工程核酸或其互补序列中如前所述序列的工程核酸。在一种实施方式中,提供了具有与本文所列任何核酸的序列或其互补序列的核酸杂交的序列的工程核酸。在一种实施方式中,杂交条件是低严格度条件。在一种实施方式中,杂交条件是中严格度条件。在一种实施方式中,杂交条件是高严格度条件。用于优化杂交方案的杂交方案和方法的实例如下列书籍所述:Molecular Cloning, T.Maniatis,E.F.Fritsch,and J.Sambrook,Cold Spring Harbor Laboratory,1982;和 Current Protocols in Molecular Biology,F.M.Ausubel,R.Brent,R.E.Kingston, D.D.Moore,J.G.Seidman,J.A.Smith,K.Struhl,Volume 1,John Wiley&Sons,2000,其全部内容通过引用并入本文,如同完整阐述。中等条件包括以下内容:在含有6×柠檬酸盐缓冲盐水(SSC;Amresco公司,梭伦,俄亥俄州)、0.5%十二烷基硫酸钠(SDS;Amresco公司,梭伦,俄亥俄州)、5xDenhardt溶液(Amresco公司,梭伦,俄亥俄州)和变性鲑精DNA(Invitrogen Life Technologies公司,卡尔斯巴德,加拿大)的溶液中,在68℃下预处理装载有DNA样品的滤器2-4小时。杂交在具有以下变化的相同溶液中进行:0.01M EDTA(Amresco公司,梭伦,俄亥俄州)、100μg/ml鲑鱼精DNA和5-20×10⁶cpm的³²P-标记的或荧光标记的探针。将滤膜在杂交混合物中孵育16-20小时,然后在含有2xSSC和0.1%SDS的溶液中洗涤15分钟。使用含有0.1×SSC和0.5%SDS的溶液代替洗涤溶液以进行第二次洗涤,并在低于T_m(熔融温度,℃)的20℃至29℃下再孵育2小时。 $T_m = 81.5 + 16.61 \log_{10}([\text{Na}^+] / (1.0 + 0.7[\text{Na}^+])) + 0.41 (\% [\text{G} + \text{C}]) - (500/n) - P - F$ 。[Na⁺] = 钠离子的摩尔浓度。% [G+C] = DNA序列中G+C碱基的百分比。N = 以碱基计的DNA序列长度。P = 对于错配的碱基对%的温度校正(~1℃/1%错配)。F = 甲酰胺浓度校正(=0.63℃/1%甲酰胺)。过滤器暴露于成像器或通过放射自显影进行显影。低严格度条件是指在37℃至60℃之间的低温下的杂交条件,并且在更高的[Na⁺] (高达0.825M)和低于T_m 40℃至48℃的温度下进行的第二次洗涤。高严格度是指在高于68℃的高温下的杂交条件,并且在低于T_m 5℃至10℃的温度下在[Na⁺] = 0.0165至0.0330M进行第二次洗涤。

[0093] 在一种实施方式中,提供了工程核酸,该工程核酸具有沿着其长度与具有本文所述的任何一种序列的核酸或其补体的连续部分(contiguous portion)具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列。所述连续部分可以为本文所述序列或其补体的全长。

[0094] 确定两种氨基酸序列或两种核酸序列的同一性百分数可以包括比对和比较两种序列相应位置的氨基酸残基或核苷酸。如果两种序列的所有位置被完全相同的氨基酸残基或核苷酸占据,则所述序列被认为具有100%同一性。同一性百分数可以使用Smith Waterman算法测量(Smith TF,Waterman MS 1981“Identification of Common Molecular Subsequences,”J Mol Biol 147:195-197,其通过引用并入本文,如同完整阐述)。

[0095] 一种实施方式包括工程核酸、工程多核苷酸或工程寡核苷酸,其具有如本文所列的任意一种核酸或其互补序列中所示的序列的一部分。这些工程核酸、工程多核苷酸或工程寡核苷酸可以具有10至全长、10至5000、10至4900、10至4800、10至4700、10至4600、10至4500、10至4400、10至4300、10至4200、10至4100、10至4000、10至3900、10至3800、10至3700、10至3600、10至3500、10至3400、10至3300、10至3200、10至3100、10至3000、10至2900、10至2800、10至2700、10至2600、10至2500、10至2400、10至2300、10至2200、10至2100、10至2000、10至1900、10至1800、10至1700、10至1600、10至1500、10至1400、10至1300、10至1200、10至1100、10至1000、10至900、10至800、10至700、10至600、10至500、10至400、10至300、10至200、10至100、10至90、10至80、10至70、10至60、10至50、10至40、10至35、10至30、10至25、10至20、10至15、或20至30个核苷酸、或10、15、20或25个核苷酸。具有在上述范围之一内的长度的工程核酸、工程多核苷酸或工程寡核苷酸可以具有在所述范围内(包括端点)的任何特定长度。所述核苷酸长度可以起始于参考序列(即,本文中的任何一个核酸)中的任何单一位置,其中,足够的核苷酸跟随在单一位置之后以适应所列出的长度。在一种实施方式中,杂交探针或引物与核酸具有85至100%、90至100%、91至100%、92至100%、93至100%、94至100%、95至100%、96至100%、97至100%、98至100%、99至100%或100%互补,所述核酸与探针和引物的长度相同,并且具有选自在本文所列的任何一种核酸中所述部分序列内的核苷酸的长度相当于探针或引物长度的序列。在一种实施方式中,杂交探针或引物沿着其长度与相应长度的具有本文所列的任意一种核酸中所述序列的核酸杂交。在一种实施方式中,杂交条件是低严格度的。在一种实施方式中,杂交条件是中严格度的。在一种实施方式中,杂交条件是高严格度的。

[0096] 一种实施方式包括用于鉴定样品中编码葡聚糖水合二激酶的内源基因的修饰序列的试剂盒。所述试剂盒可以含有第一引物和第二引物。所述第一引物和第二引物能够扩增包含在编码葡聚糖水合二激酶的内源基因中的靶序列。所述靶序列可以含有与选自SEQ ID NO:1-4、75、170-184、186、187、189-193的参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的核酸序列。所述试剂盒可以进一步含有一个或多个用于检测靶序列的扩增区域中的修饰的组件。所述试剂盒可以包含含有选自SEQ ID NO:6、7、9、11、101、103、105、110和111的核酸序列的第一引物。所述试剂盒可以包含含有选自SEQ ID NO:5、8、10、102和104的核酸序列的第二引物。所述试剂盒可以包含含有SEQ ID NO:6的核酸序列的第一引物和含有SEQ ID NO:5的核酸序列的第二引物。所述试剂盒可以包含含有SEQ ID NO:7的核酸序列的第一引物和含有SEQ ID NO:8的核酸序列的第二引物。

所述试剂盒可以包含含有SEQ ID NO:9的核酸序列的第一引物和含有SEQ ID NO:10的核酸序列的第二引物。所述试剂盒可以包含含有SEQ ID NO:11的核酸序列的第一引物和含有SEQ ID NO:13的核酸序列的第二引物。所述试剂盒可以包含含有SEQ ID NO:110的核酸序列的第一引物和含有SEQ ID NO:13的核酸序列的第二引物。所述试剂盒可以包含含有SEQ ID NO:111的核酸序列的第一引物和含有SEQ ID NO:112的核酸序列的第二引物。所述试剂盒可以包含含有SEQ ID NO:105的核酸序列的第一引物和含有SEQ ID NO:13的核酸序列的第二引物。所述第一引物和第二引物能够扩增靶序列以产生扩增产物。所述扩增产物可以含有修饰的靶序列。所述修饰的靶序列能够在高严格度条件下与含有选自SEQ ID NO:12-40、114-116、188-189、19-120和131-162的序列的核酸序列杂交。所述修饰的靶序列可以被用于诊断在编码葡聚糖水合二激酶的内源基因中具有突变的基因工程植物的探针。样品可以包括来自目前存在的植物物质 (plant matter) 的核酸的任何样品。所述样品可以包括任何植物物质。所述植物物质可以来自植物或其部分。所述植物材料可以来自动物饲料或食物。

[0097] 一种实施方式提供了鉴定样品中编码葡聚糖水合二激酶的内源基因的修饰的序列的方法。该方法可以包括将样品与第一引物和第二引物接触。该方法可以包括扩增含有被包含在编码葡聚糖水合二激酶的内源基因中的靶序列的合成的多核苷酸。所述靶序列可以是包含在编码本文所述的葡聚糖水合二激酶的内源基因中的任何靶序列。所述第一引物和第二引物能够扩增靶序列以产生扩增产物。所述扩增产物可用于确定由有性杂交或自交产生的植物是否在靶序列中含有一个或多个修饰并诊断特异性突变体。来自突变植物的样品的扩增产物的长度可以不同于来自具有相同遗传背景的野生型植物的样品的扩增产物的长度。来自突变体样品的扩增产物可以进一步用于在高严格度条件下与包含编码突变蛋白的特定区域的合成的多核苷酸杂交的探针。该方法可以包括进一步检测至少一种探针与靶序列的特异性区域的杂交。

[0098] 制备基因工程植物的方法、提高植物中淀粉水平的方法、农业加工方法、制备动物饲料的方法和用于生产对于编码改变的葡聚糖水合二激酶的工程核酸是纯合的基因工程植物的方法可以包括本文所述的检测方法,作为制备基因工程植物和/或鉴定包含本文中基因工程核酸的植物或植物生物质的部分。

[0099] 以下列表包括本发明的特定实施方式。但是该列表不是限制性的,并且不排除替代实施方式或本文另外描述的实施方式。在下列实施方式列表中描述的同源性百分数是指沿着参考序列的全长所列举序列的同源性。

[0100] 实施方式

[0101] 1、一种合成的核酸启动子,该核酸启动子与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同源性,所述参考序列选自由SEQ ID NO:78 (MzU3.8)、SEQ ID NO:79 (ZmU3)、SEQ ID NO:82 (ZmU3P1)、SEQ ID NO:84 (ZmU3P2) 和SEQ ID NO:86 (MzU3.8P) 所组成的组。

[0102] 2、一种基因构建体,该构建体含有编码Cas9核酸酶的第一工程核酸序列,其中,所述Cas9核酸酶能够切割植物中编码葡聚糖水合二激酶的内源核酸中的靶序列。

[0103] 3、实施方式2中的基因构建体,其中,第一合成的核酸序列与SEQ ID NO:74 (Cas9核酸酶) 或SEQ ID NO:75 (ZmCas9) 的参考序列具有至少70、72、75、80、85、90、91、92、93、94、

95、96、97、98、99或100%同一性。

[0104] 4、实施方式2和实施方式3中的一项或两项的基因构建体,其中,第一核酸与编码至少一个核定位信号(NLS)的多核苷酸序列融合。

[0105] 5、实施方式2-4中任意一项或多项的基因构建体,其中,编码核定位信号的多核苷酸序列选自SEQ ID NO:163-168。

[0106] 6、实施方式2-5中任意一项或多项的基因构建体,其进一步含有编码sgRNA的第二工程核酸序列,并且所述sgRNA能够结合靶序列。

[0107] 7、实施方式6中的基因构建体,其中,所述第二工程核酸含有与参考序列至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性的序列,所述参考序列选自SEQ ID NO:135(ZmU3P1:sgRNA_GWDe24b)、SEQ ID NO:136(ZmU3P2:sgRNA_GWDe24b)、SEQ ID NO:137(ZmU3.8P:sgRNA_GWDe24b)、SEQ ID NO:138(ZmU3P2:sgRNA_GWDe24c)、SEQ ID NO:139(ZmU3P2:sgRNA_GWDe25a)和SEQ ID NO:40(ZmU3P2:sgRNA_GWDe1a)。

[0108] 8、实施方式2-7中任意一项或多项的基因构建体,其中,靶序列与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性,所述参考序列选自SEQ ID NO:91(GWDe1a)、SEQ ID NO:92(GWDe24b)、SEQ ID NO:93(GWDe24c)和SEQ ID NO:94(GWDe25a)所组成的组。

[0109] 9、实施方式2-8中任意一项或多项的基因构建体,其进一步含有与第一工程核酸可操作地连接的第一启动子和与第二工程核酸可操作地连接的第二启动子。

[0110] 10、实施方式9中的基因构建体,其中,所述第一启动子或第二启动子为实施方式1中的合成的核酸启动子。

[0111] 11、实施方式2-10中任意一项或多项的基因构建体,其进一步含有终止子。

[0112] 12、实施方式11中的基因构建体,其中,所述终止子含有与SEQ ID NO:88具有至少90%同一性的核酸序列。

[0113] 13、一种基因构建体,该基因构建体含有编码核酸酶的工程核酸序列,其中,所述核酸酶能够切割包含在编码葡聚糖水合二激酶的内源核酸中的靶序列。

[0114] 14、实施方式13中的基因构建体,其中,所述核酸酶是由与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%同一性的序列编码的大范围核酸酶,所述参考序列选自SEQ ID N:164(4715_大范围核酸酶)和SEQ ID NO:165(4716_大范围核酸酶)所组成的组。

[0115] 15、实施方式13-14中任意一项或多项的基因构建体,其中,所述靶序列含有SEQ ID NO:41(大范围核酸酶GWD-9/10x.272)或SEQ ID NO:42(大范围核酸酶3e GWD-7/8x)的多核苷酸。

[0116] 16、实施方式13-15中任意一项或多项的基因构建体,该基因构建体含有至少一种调节元件,其中,所述调节元件选自启动子、终止子和增强子。

[0117] 17、一种载体,该载体含有实施方式2-16中任意一项或多项的基因构建体。

[0118] 18、一种基因工程植物,该基因工程植物含有编码改变的葡聚糖水合二激酶的工程核酸,并且与具有相同遗传背景的非基因工程植物相比具有升高的淀粉水平。

[0119] 19、实施方式18中的基因工程植物,其中,与具有相同遗传背景的非基因工程植物中的野生型葡聚糖水合二激酶相比,改变的葡聚糖水合二激酶的活性降低。

[0120] 20、实施方式18中的基因工程植物,其中,所述改变的葡聚糖水合二激酶是无活性的。

[0121] 21、实施方式18-20中任意一项或多项所述的基因工程植物,其中,所述工程核酸是编码葡聚糖水合二激酶的基因的一个等位基因的内源核酸的修饰序列。

[0122] 22、实施方式18-21中任意一项或多项的基因工程植物,其中,植物中编码葡聚糖水合二激酶的基因的所有等位基因具有工程核酸的序列。

[0123] 23、实施方式18-22中任意一项或多项的基因工程植物,其中,所述内源核酸含有与SEQ ID NO:1 (Zm GWD编码序列)或SEQ ID NO:2 (Sb GWD编码序列)的参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%,同一性的序列。

[0124] 24、实施方式18-23中任意一项或多项所述的基因工程植物,其中,所述工程核酸含有选自以下至少一种的突变:编码野生型GWD的核酸的内源性序列中一个或多个核苷酸的插入、缺失或取代。

[0125] 25、实施方式24中的基因工程植物,其中,内源核酸中的靶序列中的突变与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性,所述参考序列选自SEQ ID NO:3 (Zm GWD外显子24+内含子)、SEQ ID NO:4 (SbGWD外显子24+内含子)、SEQ ID NO:182 (ZmGWD外显子24)、SEQ ID NO:183 (Sb GWD外显子24)、SEQ ID NO:184 (SbGWD外显子7)和SEQ ID NO:189 (Zm GWD外显子25)。

[0126] 26、实施方式18-25中任意一项或多项的基因工程植物,其中,内源核酸中的靶序列中的突变与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性,所述参考序列选自由SEQ ID NO:91 (GWDe1a)、SEQ ID NO:92 (GWDe24b)、SEQ ID NO:93 (GWDe24c)和SEQ ID NO:94 (GWDe25a)所组成的组。

[0127] 27、实施方式18-26中任意一项或多项的基因工程植物,其中,所述工程核酸含有与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性的多核苷酸,所述参考序列选自由SEQ ID NO:12-40 (Zm GWD突变-外显子24)所组成的组。

[0128] 28、实施方式18-26中任意一项或多项的基因工程植物,其中,所述工程核酸含有与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性的多核苷酸,所述参考序列选自由SEQ ID NO:114-118、188、131-146 (Zm GWD突变-外显子24)和119-120 (Zm GWD突变-外显子25)所组成的组。

[0129] 29、实施方式18-25中任意一项或多项的基因工程植物,其中,所述工程核酸含有与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性的多核苷酸,所述参考序列选自由SEQ ID NO:106 (Sb4715_1 ((WT+ins) _外显子24)和SEQ ID NO:107 (Sb4715_2 (WT+de1) _外显子24)所组成的组。

[0130] 30、实施方式16-28中任意一项或多项基因工程植物,其中,所述改变的葡聚糖水合二激酶含有与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性的氨基酸序列,所述参考序列选自由SEQ ID NO:45-73 (Zm GWD突变蛋白M1-M29)所组成的组。

[0131] 31、实施方式16-25中任意一项或多项基因工程植物,其中,所述改变的葡聚糖水合二激酶含有与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性的氨基酸序列,所述参考序列选自由SEQ ID NO:121-125 (Zm GWD突变蛋白

M32-M36)、126-127 (Zm GWD突变蛋白M38-M39) 和147-162 (Zm GWD突变蛋白M40-M55) 所组成的组。

[0132] 32、实施方式18-26中任意一项或多项的基因工程植物,其中,所述改变的葡聚糖水合二激酶含有与参考序列具有至少70、72、75、80、85、90、91、92、93、94、95、96、97、98、99或100%的同一性的氨基酸序列,所述参考序列选自SEQ ID NO:82 (Sb GWD突变蛋白Sb4715_1WT+ins)或SEQ ID NO:83 (Sb GWD突变蛋白Sb4715_2WT+de1)。

[0133] 33、实施方式18-33中任意一项或多项的基因工程植物,其中,所述植物选自:由单子叶植物、双子叶植物、C4植物、C3植物、番茄、甜菜、甘蔗、桉树,柳树、杨树、玉米、高粱、小麦、苜蓿、大豆、水稻、芒草和柳枝稷所组成的组。

[0134] 34、一种基因工程植物,该基因工程植物含有实施方式2-16中任意一项或多项的基因构建体。

[0135] 35、一种用于产生基因工程植物的方法,该方法包括:

[0136] 使用实施方式17的载体转化植物细胞;

[0137] 选择表达核酸酶并含有编码改变的葡聚糖水合二激酶的工程核酸的转化植物细胞;和

[0138] 从转化的植物细胞再生基因工程植物,其中,与具有相同遗传背景的非基因工程植物相比,基因工程植物或其子代具有升高的淀粉水平。

[0139] 36、实施方式35中的方法,其中,所述核酸酶为大范围核酸酶。

[0140] 37、实施方式35中的方法,其中所述核酸酶为Cas9核酸酶。

[0141] 38、一种用于基因改造含有改变的葡聚糖水合二激酶的植物的方法,该方法包括:

[0142] 将含有编码葡聚糖水合二激酶的内源基因中的靶序列的至少一个植物细胞与含有编码能够诱导靶序列上的单链或双链断裂的核酸酶的第一核酸的载体接触;选择含有所述靶序列改变的植物细胞;

[0143] 再生含有来自植物细胞的改变的基因工程植物。

[0144] 39、实施方式38中的方法,其中,所述基因工程植物对于所述改变是纯合的。

[0145] 40、实施方式38的方法,其中,所述基因工程植物对于所述改变是杂合的。

[0146] 41、实施方式40的方法,该方法还包括使杂合的基因工程植物自交,或杂交至具有相同改变的另一杂合的基因工程植物,并且选择对于所述改变是纯合的第一子代植物。

[0147] 42、实施方式40的方法,该方法还包括将基因工程植物与具有相同遗传背景的野生型植物杂交,并且选择对于所述改变是杂合的第一子代植物。

[0148] 43、实施方式42的方法,该方法还包括使所述第一杂合子代植物自交和选择对于所述改变是纯合的第二子代植物。

[0149] 44、实施方式38-43中任意一项或多项的方法,其中,所述改变是选自靶序列中至少一个核苷酸的插入、缺失或取代中的至少一种的突变。

[0150] 45、实施方式44的方法,其中,所述突变是无效突变。

[0151] 46、实施方式38-44中任意一项或多项的方法,其中,与具有相同遗传背景的非基因工程植物相比,所述基因工程植物或其子代具有升高的淀粉水平。

[0152] 47、实施方式38-46中任意一项或多项的方法,其中,所述核酸酶选自由大范围核酸酶、Cas9核酸酶、锌指核酸酶和转录激活物样效应物核酸酶所组成的组。

[0153] 48、实施方式47的方法,其中,所述核酸酶是大范围核酸酶,并且由与参考序列具有至少90%同一性的序列编码,所述参考序列选自由SEQ ID NO:108(4715_大范围核酸酶)和SEQ ID NO:109(4716_大范围核酸酶)所组成的组。

[0154] 49、实施方式48的方法,其中,所述大范围核酸酶能够切割含有SEQ ID NO:41(4715_WWD-9/10x.272的靶标)或SEQ ID NO:42(4716_3e GWD-7/8x276的靶标)的多核苷酸的靶序列。

[0155] 50、实施方式47的方法,其中,所述核酸酶为Cas9核酸酶。

[0156] 51、实施方式50的方法,其中,所述Cas9核酸酶由与SEQ ID NO:74(Cas9核酸酶)或SEQ ID NO:75(ZmCas9)具有至少90%同一性的核酸编码。

[0157] 52、实施方式51的方法,其中,所述编码Cas9核酸酶的核酸与至少一个核定位信号(NLS)融合,并且所述NLS具有选自SEQ ID NO:163-168的多核苷酸序列。

[0158] 53、实施方式38-47和50-52中任意一项的方法,其中,所述载体还含有编码sgRNA的第二核酸序列。

[0159] 54、实施方式53的方法,其中,所述sgRNA能够结合靶序列,并且所述靶序列选自由SEQ ID NO:91(GWDe1a)、SEQ ID NO:92(GWDe24b)、SEQ ID NO:93(GWDe24c)和SEQ ID NO:94(GWDe25a)所组成的组。

[0160] 55、实施方式53-54中任意一项或多项的方法,其中,所述第二核酸含有与SEQ ID NO:95(ZmU3P1:sgRNA_GWDe24b)、SEQ ID NO:96(ZmU3P2:sgRNA_GWDe24b)、SEQ ID NO:97(ZmU3.8P:sgRNA_GWDe24b)、SEQ ID NO:98(ZmU3P2:sgRNA_GWDe24c)、SEQ ID NO:99(ZmU3P2:sgRNA_GWDe25a)和SEQ ID NO:100(ZmU3P2:sgRNA_GWDe1a)具有至少90%同一性的序列。

[0161] 56、实施方式38-55中任意一项或多项的方法,其中,所述载体还含有与第一核酸或第二核酸可操作地连接的核酸启动子。

[0162] 57、实施方式56的方法,其中,所述核酸启动子含有与参考序列具有至少90%同一性的序列,所述参考序列选自由SEQ ID NO:78(MzU3.8)、SEQ ID NO:79(ZmU3)、SEQ ID NO:82(ZmU3P1)、SEQ ID NO:84(ZmU3P2)和SEQ ID NO:86(MzU3.8)所组成的组。

[0163] 58.由实施方式38-57中任意一项的方法产生的基因工程植物,或其子代或其后代,其中,在所述植物、其子代或其后代中含有所述改变。

[0164] 59、实施方式58的基因工程植物,该基因工程植物与具有相同遗传背景的含有野生型葡聚糖水合二激酶的植物相比具有升高的淀粉水平。

[0165] 60、一种提高植物中淀粉水平的方法,该方法包括在植物中表达核酸,所述核酸编码能够诱导靶序列上的双链断裂的核酸酶,并且选择含有靶序列改变的和具有升高的淀粉水平的纯合植物,其中,所述靶序列被包含在仅编码葡聚糖水合二激酶的内源基因中。

[0166] 61、一种农业加工的方法,该方法包括:

[0167] 在植物中表达编码能够诱导靶序列上的双链断裂的核酸酶的核酸,其中,所述靶序列被包含在编码葡聚糖水合二激酶的内源基因中;

[0168] 选择含有靶序列中的改变并且具有升高的淀粉水平的纯合植物;和

[0169] 加工所述纯合植物,其中,所述加工包括一种或多种选自收获、排水、粉碎、干燥、发酵、使用化学品水解、使用外源酶水解和与植物生物质组合的步骤。该方法还可以包括用

于产生实施方式63-71中任意一项或多项的基因工程植物的方法。

[0170] 62、一种制备动物饲料的方法,该方法包括:

[0171] 在植物中表达编码能够诱导靶序列上的双链断裂的核酸酶的核酸,其中,所述靶序列被包含在编码葡聚糖水合二激酶的内源基因中;

[0172] 选择含有靶序列中的改变并且具有升高的淀粉水平的纯合植物;和

[0173] 进行选自由以下步骤所组成的组:收获、排水、粉碎、干燥、青贮、造粒、与可食用纤维源组合、以及与植物生物质组合。该方法还可以包括用于产生实施方式63-71中任意一项或多项的基因工程植物的方法。

[0174] 63、一种用于产生含有编码改变的葡聚糖水合二激酶的工程核酸的基因工程植物的方法,该方法包括修饰植物中编码葡聚糖水合二激酶的基因的至少一个等位基因的内源核酸序列,其中,修饰的工程核酸是工程核酸,并且修饰的植物是基因工程植物。

[0175] 64、实施方式63的方法,其中,所述基因工程植物对于含有突变的基因是纯合的,并且所有等位基因含有所述工程核酸的序列。

[0176] 65、实施方式63的方法,其中,所述基因工程植物对于含有所述突变的基因是杂合的。

[0177] 66、实施方式63或65中的任意一项或多项的方法,该方法还包括将所述基因工程植物自交并获得子代。

[0178] 67、实施方式63-65中的任意一项或多项的方法,该方法还包括将所述基因工程植物与具有相同遗传背景的非基因工程植物杂交并获得子代。

[0179] 68、实施方式66-67中的任意一项或多项的方法,该方法包括分析子代中改变的葡聚糖水合二激酶的存在并选择含有突变的子代植物。

[0180] 69、实施方式63的方法,该方法包括实施方式18-34和58-59中任意一项或多项的基因工程植物。

[0181] 70、实施方式63的方法,其中,所述修饰的步骤通过使用实施方式35-36中任意一项所述的方法进行。

[0182] 71、实施方式63的方法,其中,所述修饰的步骤通过使用实施方式2-16中任意一项的基因构建体进行。

[0183] 72、一种用于鉴定样品中编码葡聚糖水合二激酶的内源基因的修饰序列的试剂盒,其中,所述试剂盒含有第一引物和第二引物,其中,所述第一引物和第二引物能够扩增包含在编码葡聚糖水合二激酶的内源基因中的靶序列。所述靶序列含有与选自SEQ ID NO: 1-4、75、171-187、189-193的参考序列具有至少90%同一性的核酸序列

[0184] 73、实施方式72的试剂盒,该试剂盒还包含一个或多个用于检测靶序列的扩增区域中的修饰的组件。

[0185] 74、实施方式72-73任意一项或多项的试剂盒,其中,所述第一引物含有选自SEQ ID NO:6、7、9、11、101、103、105、110和111的核酸序列。

[0186] 75、实施方式72-74中任意一项或多项所述的试剂盒,其中,所述第二引物含有选自SEQ ID NO:5、8、10、102和104的核酸序列。

[0187] 76、实施方式72-75中任意一项或多项的试剂盒,其中,第一引物和第二引物能够扩增靶序列以产生含有修饰的靶序列的扩增产物。

[0188] 77、实施方式73-76中任意一项或多项的试剂盒,其中,所述扩增的靶序列含有选自SEQ ID NO:12-40、106-107、114-120、131-146和188的序列。

[0189] 78、实施方式72-76中任意一项或多项的试剂盒,其中,所述修饰的靶序列能够在高严格度条件下与含有选自SEQ ID NO:12-40、106-107、114-120、131-146和188的序列的核酸序列杂交。

[0190] 79、实施方式72-78中任意一项或多项的试剂盒,其中,所述样品含有来源于在编码葡聚糖水合二激酶的内源基因中具有至少一个突变的基因工程植物物质。

[0191] 80、一种鉴定样品中编码葡聚糖水合二激酶的内源基因的修饰序列的方法,该方法包括:

[0192] 将样品与第一引物和第二引物接触;

[0193] 扩增包含在编码葡聚糖水合二激酶的内源基因中的靶序列,并且所述靶序列含有与参考序列具有至少90%同一性的核酸序列,所述参考序列选自SEQ ID NO:1-4、75、170-184、186、187、189-193;和

[0194] 检测靶序列中的修饰。

[0195] 81、实施方式79的方法,其中,所述靶序列中的修饰含有选自SEQ ID NO:12-40、106-107、114-120、131-146和188的序列。该鉴定方法可以被加入实施方式60-71中的任意一项或多项。

[0196] 可以通过由本问的一种或多种其他实施方式的一个或多个要素(element)补充实施方式,和/或使用本文的一种或多种其他实施方式的一个或多个要素代替一种实施方式中的一个或多个要素形成本文进一步地的实施方式。

[0197] 实施例

[0198] 提供以下非限制性的实施例来说明具体的实施方式。全部实施方式均可以由来自以下一个或多个实施例的一处或多处细节补充,和/或一种实施方式中的一个或多个要素可以用一个或多个以下实施例的一处或多处细节代替。

[0199] 实施例1基于大范围核酸酶的玉米和高粱基因组中的GWD基因的修饰

[0200] 设计靶向GWD外显子24的大范围核酸酶构建体,其接近预测的编码的酶的活性位点,目的是引入使GWD失活的突变(无效突变)。在玉米和高粱中鉴定和表征大范围核酸酶诱导的GWD DNA突变体。

[0201] 为了改造具有针对玉米和高粱基因组中的GWD基因的特异性的I-CreI归巢内切核酸酶,从玉米和高粱的之前注释的全长GWD基因中选择两个核苷酸序列。所述序列选择是基于玉米和高粱序列之间的高核苷酸序列同一性(95%核苷酸序列同一性)的存在以及两种作物的外显子#24中存在GWD蛋白活性所需的序列基序。目的是以这样的方式研发两个大范围核酸酶构建体,使得它们中的每一个将特异性用于玉米和高粱中的GWD修饰。使用大范围核酸酶方法在选择GWD序列处的靶向基因组修饰将导致缺乏活性位点的GWD蛋白变体(从含有编码序列的帧错位表达的截短的蛋白质或修饰的蛋白质)的表达,因此是催化失活的。将如下所示的ZmGWD(玉米)和SbGWD(高粱)的选定序列提供给Precision Biosciences公司,用于设计大范围核酸酶GWD9-10x.272和GWD7-8x.226。

[0202] 大范围核酸酶GWD-9/10x.272(pAG4715)的靶序列是:ATCCTTGTGGCAAAGAGTGTC(ASEQ ID NO:41)。

[0203] 大范围核酸酶 GWD-7/8x.226 靶序列 (pAG4716) 的靶序列是：GTAGTTGGTGTAAATTACACCTG (SEQ ID NO:42)。

[0204] 被设计的大范围核酸酶识别的外显子24内的DNA序列加下划线。大写字母的序列显示外显子24,而小写字母的序列表示侧翼内含子。双下划线的“CAT”密码子编码对GWD蛋白活性至关重要的组氨酸残基。

[0205] >ZmGWD_外显子24

[0206]

```
aagtgatactagtgaccctctccacaatTTATgCGaaccacagaaattaataatattctattactctgcacctg
acatctggctcctgctatcagTTGGCAGGTTATAAGCCCGTTGAAGTATCAGGTTATGTGGTTGTGGTTGATGAGT
TACTTGCTGTCCAGAACAATCTTATGATAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGGAGAGGAAGAAATA
CCAGATGGAGTAGTTGGTGTAATTACACCTGATATGCCAGATGTTCTGTCTCAT
GTGTCAGTCCGAGCAAGGAATAGCAAGgtttatcttcacagctatggttgaagatttcttgaatTTTTTctcttgta
ttgatgttgacatactagctTTTTTcctaata (SEQ ID NO:3)
```

[0207] >SbGWD_外显子24

[0208]

```
aagtggtaactagtgacctctccacagTTTTATgtgaaccacagaaattaataatgataatattctattactctgc
acctgacatctggctcctgataacagTTGGCAGGTTATAAGCCAGTTGAAGTATCAGGTTATGTGGTTGTGGTTGA
TGAGTTACTTGCTGTCCAGAACAATCTTATGATAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGGAGAGGAAG
AAATACCAGATGGAGTAGTTGGTGTAATTACACCTGATATGCCAGATGTTCTGTCCCAT
GTGTCAGTCCGAGCAAGGAATAGCAAGgtttatTTTtcacagttatggttgaagcttctcagatTTTTTcttggta
tcgatgttgacataaccagTTTTTcctaata (SEQ ID NO:4)
```

[0209] 使用Clustal软件比对选择的ZmGWD和SbGWD序列 (Larkin MA et al.,2007; Goujon Met al.,2010,二者均通过引用并入本文,如同完整阐述)。

[0210] CLUSTAL 2.1多序列比对

[0211]

SbGWD_外显子 24 aagtggactactagtgacctctccacagttttatgtgaaccacagaaattaatatgataa 59
 ZmGWD_外显子 24 aagtgatactagtgaccctctccacaattttatgccaaccacagaaatta-----ataa 54

SbGWD_外显子 24 tatattctattactctgcacctgacatctggctcctgataacagTTGGCAGGTTATAAGC119
 ZmGWD_外显子 24 tatattctattactctgcacctgacatctggctcctgctatcagTTGGCAGGTTATAAGC114

SbGWD_外显子 24 CCAGTTGAAGTATCAGGTTATGTGGTTGTGGTTGATGAGTTACTTGCTGTCCAGAACAAA179
 ZmGWD_外显子 24 CCGGTTGAAGTATCAGGTTATGTGGTTGTGGTTGATGAGTTACTTGCTGTCCAGAACAAA174
 ** *****

SbGWD_外显子 24 TCTTATGATAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGGAGAGGAAGAAATACCA239
 ZmGWD_外显子 24 TCTTATGATAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGGAGAGGAAGAAATACA234

SbGWD_外显子 24 GATGGAGTAGTTGGTGTAAATTACACCTGATATGCCAGATGTTCTGTCCCATGTGTTCAGTC299
 ZmGWD_外显子 24 GATGGAGTAGTTGGTGTAAATTACACCTGATATGCCAGATGTTCTGTCTCATGTGTTCAGTC294

SbGWD_外显子 24 CGAGCAAGGAATAGCAAGgttttatctttcacagttatggtgcaagettttctcagatttttt359
 ZmGWD_外显子 24 CGAGCAAGGAATAGCAAGgttttatctttcacagctatggtgcaagattttcttgaatttttt354

SbGWD_外显子 24 ttctttgtatcgatggtgacataaccagtttttttcttaaat 397 (SEQ ID NO: 4)

[0212]

ZmGWD_外显子 24 ctctttgtattgatggtgacataactagcttttttcttaaat 392 (SEQ ID NO: 3)

[0213] 用于表达大范围核酸酶的植物转化载体的研发:

[0214] 由Precision Biosciences公司提供的大范围核酸酶序列GWD-9/10x.272[SEQ ID NO:108]和GWD-7/8x.226[SEQ ID NO:109],通过使用PCR方法在5'端添加BamHI限制性位点,在3'端添加AvrII位点进行进一步修饰。随后,将GWD-9/10x.272和GWD-7/8x.226核苷酸序列作为在玉米泛素1基因启动子和Nos转录终止子序列之间的BamHI-AvrII片段克隆到pAG4500载体中,以分别产生植物转化载体pAG4715和pAG4716。图1和图2显示了pAG4715和pAG4716载体的各自的图谱。参考图1和图2,pAG4715和4716含有玉米泛素启动子(ZmUbi1)、玉米泛素内含子(ZmUbi1内含子)和用作转录终止子的聚腺苷酸化信号NosT。两种载体还含有作为选择性标记物的磷酸甘露糖异构酶基因(PMI)、At NLS(核定位序列)、ZmKozak、mUBQmono、T-DNA右和左边界(分别为RB和LB)、链球菌乙酰转移酶(streptothricin acetyltransferase)基因和赋予链霉素抗性的氨基糖苷乙酰转移酶(aadA)基因。pAG4715含有GWD9-10x.272大范围核酸酶序列[SEQ ID NO:108],pAG4716含有GWD7-8x.226大范围核酸酶序列[SEQ ID NO:109]。

[0215] pAG4715和pAG4716被用于在玉米和高粱中产生转基因事件和突变体。

[0216] 本文使用的靶蛋白、基因、突变体和载体的序列列于表1中。

[0217] 表1序列描述

[0218]

SEQ ID NO	描述	类型
1	ZmGWD 编码序列	DNA
2	SbGWD 编码序列	DNA
3	ZmGWD 外显子 24 (包括内含子)	DNA
4	SbGWD 外显子 24 (包括内含子)	DNA
182	ZmGWD 外显子 24 (无内含子)	DNA
183	SbGWD 外显子 24 (无内含子)	DNA
184	SbGWD 外显子 7 (无内含子)	DNA

[0219]

SEQ ID NO	描述	类型
5	Mega-1 (4716) PCR 反向引物	DNA
6	Mega-1 (4716) PCR 正向引物	DNA
7	Mega-2 (4715) PCR 正向引物	DNA
8	Mega-2 (4715) PCR 反向引物	DNA
9	ZmGWD mega-2 PCR 正向引物	DNA
10	ZmGWD mega-2 PCR 反向引物	DNA
11	SbGWD mega-2 PCR 正向引物	DNA
13	SbGWD mega-2 PCR 反向引物	DNA
12	M16 (Zm GWD 外显子 24-无内含子)	DNA
13	M17 (Zm GWD 外显子 24-无内含子)	DNA
14	M18 (Zm GWD 外显子 24-无内含子)	DNA
15	M27 (Zm GWD 外显子 24-无内含子)	DNA
16	M1 (Zm GWD 外显子 24-无内含子)	DNA
17	M11 (Zm GWD 外显子 24-无内含子)	DNA
18	M10 (Zm GWD 外显子 24-无内含子)	DNA
19	M3 (Zm GWD 外显子 24-无内含子)	DNA
20	M8 (Zm GWD 外显子 24-无内含子)	DNA
21	M14 (Zm GWD 外显子 24-无内含子)	DNA
22	M13 (Zm GWD 外显子 24-无内含子)	DNA
23	M12 (Zm GWD 外显子 24-无内含子)	DNA
24	M22 (Zm GWD 外显子 24-无内含子)	DNA
25	M23 (Zm GWD 外显子 24-无内含子)	DNA
26	M24 (Zm GWD 外显子 24-无内含子)	DNA
27	M2 (Zm GWD 外显子 24-无内含子)	DNA
28	M21 (Zm GWD 外显子 24-无内含子)	DNA
29	M4 (Zm GWD 外显子 24-无内含子)	DNA
30	M19 (Zm GWD 外显子 24-无内含子)	DNA
31	M26 (Zm GWD 外显子 24-无内含子)	DNA
32	M25 (Zm GWD 外显子 24-无内含子)	DNA
33	M15 (Zm GWD 外显子 24-无内含子)	DNA
34	M5 (Zm GWD 外显子 24-无内含子)	DNA
35	M2 (Zm GWD 外显子 24-无内含子)	DNA
36	M28 (Zm GWD 外显子 24-无内含子)	DNA
37	M6 (Zm GWD 外显子 24-无内含子)	DNA
38	M9 (Zm GWD 外显子 24-无内含子)	DNA
39	M7 (Zm GWD 外显子 24-无内含子)	DNA
40	M29 (Zm GWD 外显子 24-无内含子)	DNA

[0220]

SEQ ID NO	描述	类型
106	突变体 Sb4715_1 (Wt+ ins)	DNA
107	突变体 Sb4715_2 (WT +del)	DNA
41	大范围核酸酶 GWD-9/10x.272 靶序列 (pAG4715)	DNA
42	大范围核酸酶 GWD-7/8x 靶序列 (pAG4716)	DNA
43	ZmGWD (野生型蛋白)	氨基酸
44	SbGWD (野生型蛋白)	氨基酸
45	ZmGWD M1 (突变的 GWD 蛋白)	氨基酸
46	ZmGWD M2 (突变的 GWD 蛋白)	氨基酸
47	ZmGWD M3 (突变的 GWD 蛋白)	氨基酸
48	ZmGWD M4 (突变的 GWD 蛋白)	氨基酸
49	ZmGWD M5 (突变的 GWD 蛋白)	氨基酸
50	ZmGWD M6 (突变的 GWD 蛋白)	氨基酸
51	ZmGWD M7 (突变的 GWD 蛋白)	氨基酸
52	ZmGWD M8 (突变的 GWD 蛋白)	氨基酸
53	ZmGWD M9 (突变的 GWD 蛋白)	氨基酸
54	ZmGWD M10 (突变的 GWD 蛋白)	氨基酸
55	ZmGWD M11 (突变的 GWD 蛋白)	氨基酸
56	ZmGWD M12 (突变的 GWD 蛋白)	氨基酸
57	ZmGWD M13 (突变的 GWD 蛋白)	氨基酸
58	ZmGWD M14 (突变的 GWD 蛋白)	氨基酸
59	ZmGWD M15 (突变的 GWD 蛋白)	氨基酸
60	ZmGWD M16 (突变的 GWD 蛋白)	氨基酸
61	ZmGWD M17 (突变的 GWD 蛋白)	氨基酸
62	ZmGWD M18 (突变的 GWD 蛋白)	氨基酸
63	ZmGWD M19 (突变的 GWD 蛋白)	氨基酸
64	ZmGWD M20 (突变的 GWD 蛋白)	氨基酸
65	ZmGWD M21 (突变的 GWD 蛋白)	氨基酸
66	ZmGWD M22 (突变的 GWD 蛋白)	氨基酸
67	ZmGWD M23 (突变的 GWD 蛋白)	氨基酸
68	ZmGWD M24 (突变的 GWD 蛋白)	氨基酸
69	ZmGWD M25 (突变的 GWD 蛋白)	氨基酸
70	ZmGWD M26 (突变的 GWD 蛋白)	氨基酸
71	ZmGWD M27 (突变的 GWD 蛋白)	氨基酸
72	ZmGWD M28 (突变的 GWD 蛋白)	氨基酸
73	ZmGWD M29 (突变的 GWD 蛋白)	氨基酸

[0221]

SEQ ID NO	描述	类型
74	突变的蛋白 Sb4715_1 (WT + ins)	氨基酸
75	突变的蛋白 Sb4715_2 (WT + del)	氨基酸

[0222] 实施例2将TALENs应用于靶向高粱基因组中GWD基因的修饰

[0223] 在高粱GWD基因 (SbGWD) 的外显子7和24的每一个中选择两对DNA序列用于研发四个定制化的TAL DNA结合结构域,其将与截短的FokI核酸酶序列融合。选择高粱外显子24是因为其含有GWD活性位点,并且为了与玉米中的其他内源性DNA编辑技术进行比较,例如,大范围核酸酶和CRISP/Cas9技术。在GWD基因序列的上游区域选择高粱外显子7用于产生更短的截短形式的GWD蛋白。使用专有程序在Life Technologies网站上进行DNA结合结构域的序列的选择。融合到截短FokI内切核酸酶的用于靶向GWD基因的外显子7和24中的高粱基因组修饰的两对TAL DNA结合结构域由Life Technologies构建。每对TALEN将在相应的GWD位点识别基因组DNA序列的上链和下链,以靶向作用于FokI核酸酶进行DNA切割。

[0224] 选择用于基于TALEN的GWD修饰的SbGWD核苷酸序列。SbGWD_外显子7序列位于SbGWD编码序列 (SEQ ID NO:2) 的736-969nt内:

[0225] >SbGWD_外显子7

[0226]

GAGGAGTATGAAGCTGCACGAGCTGAGTTAATAGAGGAATTAATAGAGGTGTTTCTTTAGAGAAGCTTCGAGCTAA
ATTGACAAAAACACCTGAAGCACCTGAGTCAGATGAACGTAATCTCCTGCATCTCGAATGCCCGTTGATAAACTTC
CAGAGGACCTTGTACAGGTGCAGGCTTATATAAGGTGGGAGAAAGCGGGCAAGCCAAATTATCCTCCTGAGAAGCAA
CTG (SEQ ID NO:184)

[0227] SbGWD_外显子24序列位于SbGWD编码序列 (SEQ ID NO:2) 的3030-3243nt内:

[0228] >SbGWD_外显子24

[0229]

TTGGCAGGTTATAAGCCCAGTTGAAGTATCAGGTTATGTGGTTGTGGTTGATGAGTTACTTGCTGTCCAGAACAAAT
CTTATGATAAACCAACCATCCTTGTGGCAAAGAGTGTC AAGGGAGAGGAAGAAATACCAGATGGAGTAGTTGGTGTA
ATTACACCTGATATGCCAGATGTTCTGTCC**CAT**GTGTTCAGTCCGAGCAAGGAATAGCAAG (SEQ ID NO:183)

[0230] 每个外显子中下划线的序列代表所选的TAL DNA结合位点,左侧序列对DNA上链是特异性的,右侧序列靶向作用于DNA下链。编码对于GWD蛋白活性具有催化重要性的组氨酸残基的密码子是双下划线的,并且在外显子24内是粗体。对于外显子7或外显子24特异性的TALEN将各自成对被克隆至基于pAG4500的植物转化载体中。

[0231] 实施例3植物转化和分析

[0232] 玉米和高粱转化:使用质粒pSB1操作手册中所述的方案提取来自农杆菌的DNA。使用Qiagen DNeasy Plant Mini试剂盒 (69140) 提取植物DNA。根据Negrotto D et al.Plant Cell Rep 19:798;Ishida Y et al.1996 Nat Biotech 14:74,使用GWD大范围核酸酶靶向构建体pAG4715和/或pAG4716转化玉米和高粱胚,全部通过引用并入本文,如同完整阐述。简言之,将携带合适的转化质粒的LBA4404农杆菌细胞接种于来自野生型AxB玉米的胚发生愈伤组织。农杆菌介导的未成熟玉米胚胎的转化按照Negrotto D等描述的那样进行。使用先前报道的步骤 (Ishida Y et al.1996 Nat Biotech 14:745;Hiei Y et al.1994 Plant

J 6:271;Hiei Y and Komari T 2006 Plant Cell Tissue Organ Cult.85:27;Komari T et al.1996 Plant J 10:165),将用于GWD大范围核酸酶的表达盒克隆到能够与根癌土壤杆菌菌株LBA4404的三亲交配中的pSB1载体重组的中间载体的KpnI-EcoRI位点。将玉米 (*Zea mays*栽培品种HiII,A188或B73)母本植株在温室中在28°C的日光下生长16小时。从籽粒中分离出未成熟的合子胚,并用含有目标基因的土壤杆菌溶液接种。接种后,未成熟胚在组织培养过程中生长10-12周。对具有叶和根的发育良好的幼苗进行取样用于PCR分析,以鉴定含有目标基因的转基因植物。用水冲洗PCR阳性和生根的植物以洗去琼脂培养基,并将其移植到土壤中并在温室中生长以产生种子和秸秆。

[0233] 根据Gao等,2005的方案进行高粱转化。根据Elkonin和Pakhomova,2000进行转基因植物的再生。

[0234] 使用质粒pSB1操作手册中描述的方案提取来自农杆菌的DNA。使用Qiagen DNeasy Plant Mini试剂盒(69140)提取植物DNA。

[0235] 10×TE+十二烷基肌酸钠(Sarkosyl)-用于96孔板的植物DNA分离:简言之,将COSTAR研磨块装填3/4的叶样品,将一个5mm钢珠加入具有样品的每个孔中,并使用存储垫施用器以密封所述块。在研磨前或直至处理时间,将样品在-80°C下储存至少30分钟。为了处理,使用Klecko粉碎机和安全研磨机以最大速度将样品研磨45秒。将密封移除并丢弃。使用多通道移液器和无菌溶液槽向每个样品中加入300微升的10×TE+十二烷基肌酸钠缓冲液(5mL 1M Tris,1mL 0.5M EDTA,0.5g十二烷基肌酸钠,46mL ddH₂O)。将板在振荡器上以300rpm温育10min,并以4000rpm旋转3分钟。除去上清液并弃去,将沉淀重悬于1×TE缓冲液中。将150微升样品等分试样加入96孔PCT板中。所述PCR板用铝箔密封。为了获得最佳结果,在同一天进行DNA分离和PCR。

[0236] 转基因诊断PCR反应设置:“完全”PCR反应混合物如下:15μl 2×GoTaq MM (GoTaq Green Master Mix (PROMEGA#M712)、3μl目标基因特异的正向和反向引物组合(每个以10μM混合)、2μl制备的DNA和水,以将体积调节至30μl。将每孔的28微升“完全”PCR反应混合物等分到PCR板(FISHER,#14230236)中,将2微升植物DNA样品分装到PCR板的每个孔中。在每个PCR反应中使用阳性对照和无模板阴性对照。土壤杆菌DNA对照品在TE缓冲液中以1:100稀释以产生清晰的带,PCR板用密封垫(COSTAR#6555)和滚轮,在BIORAD PTC-100热循环仪上进行PCR。热循环程序如下:1)95°C-3min;30个循环95°C-30sec,55°C-30sec,72°C-45sec;72°C-5min;10°C(保持),和2)90°C30min和10°C(保持)。将每个12微升的PCR反应物加入到Ready Agarose 96 Plus凝胶-3%(BIORAD#161-3062)上并在约100V下电泳20分钟,然后用配备Quantity One软件的BIORAD凝胶系统观察。快速负载的50bp DNA梯状条带(NEB N0473S)用于鉴定PCR片段的大小。使用10×TBE缓冲液(Promega V4251)。

[0237] 使用傅里叶变换近红外(FT-NIR)技术在线监测活玉米叶的淀粉含量:玉米叶组织中的淀粉含量是动物饲料和生物燃料生产的重要因素。常用的GOPOD测定不适用于在活组织中的实时监测,因为该测定是侵入性的,需要物理组织样品,以及密集的劳动。基于天然淀粉和玉米叶淀粉或玉米粉(湿和干)的干混物的FT-NIR光谱的淀粉含量的预测模型使用偏最小二乘回归来进行研发。三个关键因素决定了FT-NIR技术成功应用于快速化学表征:准确和可重复的近红外光谱采集、可靠的校准数据和稳健的化学计量分析。为了分析,使用以下材料:分光光度计(Perkin Elmer Spectrum One NTS沃尔瑟姆,马萨诸塞州)、

Unscrambler[®] (版本10.2., Camo Software公司, 伍德布里奇, 新泽西州)、烘箱、Hi-maize 抗性淀粉 (Honeyville, 布里格姆市, 犹他州) 和淀粉 (产品编号S516-500, Fisher Scientific)。使用去离子水将所有空白和测试样品稀释10倍, 并使用葡萄糖氧化酶-过氧化物酶 (GOPOD) 比色测定法 (Megazyme International, 威克洛, 爱尔兰) 测定未反应的淀粉含量。

[0238] 样品制备: 通过将一定重量比例的抗性淀粉 (Honeyville, 布里格姆市, 犹他州) 与淀粉 (产品编号S516-500, Fisher Scientific) 混合以制备共计56个具有0-33%淀粉含量的干淀粉混合物样品。Honeyville产品含有从通过传统植物育种产生的高直链淀粉玉米杂种中分离的HI-MAIZE 260 (Ingredion, 布里奇沃特, 纽泽西州) 抗性淀粉, 并且含有33%可消化的或血糖淀粉。

[0239] 收集来自不同活的玉米植物的在一定年龄或具有不同淀粉积累的150个绿叶样品。将100个样品烘干并将50个样品未干燥 (“湿”)。将叶片样品研磨至0.5mm并储存在塑料样品袋中用于湿度平衡。使用标准方法测量水分含量。

[0240] 淀粉测定: 使用GOPOD测定分析淀粉共混物、湿和干绿色组织或玉米粉样品的淀粉含量。

[0241] FT-NIR光谱的扫描、处理和分析: 扫描-将大约5g碾磨的样品倒入较小的NIRA杯中, 调平, 并在每次扫描之间手动旋转扫描16次。用单独的子样品重复该程序5次, 并将得到的光谱扫描取平均值。共计56个淀粉共混物样品用于建立淀粉模型。在校准设置中使用36个样本, 在验证设置中使用14个样本, 在测试设置中使用6个样本。使用100个玉米叶或种子样品建立干粉碎的绿色组织或粉末模型。对于校准、验证和测试, 分别使用72、20和8个样品。使用49个玉米叶或玉米种子样品建立湿磨玉米模型。36个样品用于校准, 10个样品用于验证, 并且3样品用于测试设置。

[0242] 使用FT-NIR光谱-Unscrambler[®] (版本10.2., Camo Software公司, 伍德布里奇, 新泽西州) 的处理和分析来处理和分析光谱数据, 建立和验证校准, 并测试回归模型。使用乘法散射校正 (MSC) 和基于二阶导数的平滑技术, 例如, Savitzky-Golay (SG) 技术, 进行数据预处理。使用MSC与SG二阶导数预处理光谱数据组合的部分最小二乘模型被研发用于淀粉共混物和碾磨的绿色组织或磨碎的玉米。表2中显示了校准、验证和测试样品的测量和预测 (MSC+二阶导数模型) 淀粉含量的实施例。

[0243] 表2测量的和预测的 (MSC+二阶导数模型) 校准、验证和测试样品的淀粉含量

[0244]

淀粉含量	淀粉共混物		干玉米粉		湿玉米粉	
	测量(%)	预测(%)	测量(%)	预测(%)	测量(%)	预测(%)
校准	2.50	2.52	5.0	6.5	5.0	5.2
	5.00	4.89	7.5	7.4	10.0	7.9
	10.0	10.1	10.0	10.2	12.5	12.0
	15.0	15.6	15.0	14.5	13.2	13.5
	20.0	20.0				
验证	4.75	4.7	7.5	7.6	7.5	7.5
	10.0	9.8	12.5	12.4	10.0	10.0
	15.7	16.0	17.5	17.5	12.0	12.5
	18.9	18.9				
测试	6.5	6.55	5.0	6.5	7.2	7.3
	12.5	12.45	11.5	12.5	11.2	11.5
	17.5	18.2	14.5	13.2	12.2	13.5

[0245] R^2 : 淀粉共混物: 校准: 0.98, 验证: 0.97, 预测: 0.97

[0246] 干玉米粉: 校准: 0.86, 验证: 0.80, 预测: 0.80

[0247] 湿玉米粉: 校准: 0.94, 验证: 0.80, 预测: 0.75

[0248] 实施例4在玉米和高粱GWD基因中大范围核酸酶诱导的DNA突变的鉴定和表征

[0249] 目标基因 (GOI) 阳性玉米和高粱转化体的鉴定: 从使用pAG4715或pAG4716转化的玉米和高粱植物的叶子取样, 提取DNA, 并筛选存在被包含在pAG4715或pAG4716中的大范围核酸酶转基因。

[0250] 玉米和高粱转化体的筛选: 使用PCR扩增的GWD DNA序列的序列分析, 筛选分别携带pAG4715或pAG4716转基因的被称为4715或4716植物的GWD突变。

[0251] GWD的PCR扩增: 使用表3所示的ZmGWDmega-2和SbGWDmega-2引物扩增外显子24上的大范围核酸酶靶向区周围的DNA序列。

[0252] 表3用于基因分型4715和4716植物的引物

[0253]

引物组	引物名称	正向或反向	序列	产物大小 (bp)
Mega-1 (4716)	Mega-1R	反向	TGATCTTCAGCACGAGGTTG (SEQ ID NO: 5)	265
Mega-1 (4716)	Mega-1F	正向	GGCTCCATCTATGCCTGTATC (SEQ ID NO: 6)	265
Mega-2 (4715)	Mega-2F	正向	GAGCTCAGTTTCGCTGTCTATC (SEQ ID NO: 7)	209
Mega-2 (4715)	Mega-2R	反向	ATGATCTTCAGCACGAGGTTG (SEQ ID NO: 8)	209
ZmGWD mega-2	ZmGWD mega-2F	正向	GGTTATAAGCCCGGTTGAAGTA (SEQ ID NO: 9)	204
ZmGWD mega-2	ZmGWD mega-2R	反向	ctattccttgctcggactgac (SEQ ID NO: 10)	204
SbGWD mega-2	SbGWD mega-2f	正向	GGCAGGTTATAAGCCAGTT (SEQ ID NO: 11)	208
ZmGWD mega-2	SbGWD mega-2r	反向	CTATTCCTTGCTCGGACTGAC (SEQ ID NO: 10)	208

[0254] 将引物组在无核酸酶的水中稀释至终浓度为5 μ M。如上所述进行PCR反应。

[0255] 使用我们的PMI55程序(95 $^{\circ}$ C, 2min; 30个循环[95 $^{\circ}$ C, 30sec; 55 $^{\circ}$ C, 30sec; 72 $^{\circ}$ C, 45sec; 72 $^{\circ}$ C, 8min)在Eppendorf Mastercycler proS(Eppendorf)上运行PCR样品。

[0256] 在Bio-Rad ReadyAgarose 96 Plus Gels, TBE (#161-3062)上分离PCR样品,并用Bio-Rad凝胶成像系统进行可视化。

[0257] 图3示出了显示4715和4716事件的条带或分别成功并入来自pAG4715或pAG4716的转基因的凝胶的实例。参考该图,移位的GWD条带表明在GWD大范围核酸酶靶向位点处存在潜在的插入和缺失(插入缺失,indel),并且用星号标记。

[0258] GWD Indel等位基因的DNA序列表征:

[0259] 初始GWD PCR产物的测序-使用用于扩增的相同引物(ZmGWDmega-2或SbGWDmega-2),在Beckman Coulter Genomics(36Cherry Hill Dr, Danvers, MA 01923)上对PCR产物进行测序。测序允许GWD基因座、野生型、纯合突变体和杂合突变体的三种不同遗传结果之间存在差异。野生型在204或208bp GWD PCR片段内没有突变,纯合子带有插入缺失突变,杂合体带有GWD PCR片段的无法解析的序列区(表明至少存在一处插入缺失)。

[0260] 单个GWD等位基因的克隆和测序。为了开始克隆,使用与来自杂合植物的DNA相同的如上所述的引物组(ZmGWDmega-2或SbGWDmega-2),使用PCR扩增GWD,以表征单个GWD等位基因。如上所述,通过在琼脂糖凝胶上运行8 μ l PCR产物以确认PCR扩增,剩余的22 μ l PCR反应物使用Qiagen PCR纯化试剂盒(28104; Qiagen, 马里兰州, 美国)进行纯化,并用30 μ l洗脱缓冲液(EB)洗脱。

[0261] 使用具有One Shot[®] TOP10感受态大肠杆菌(K4500-01; Life Technologies)的TOPO[®] TA克隆试剂盒根据方案克隆纯化的PCR产物。对于克隆过程,将4 μ l纯化的PCR产物用于连接,将连接物在冰上孵育至少10min,将50 μ l每个转化反应物接种在LB羧苄青霉素

(50 μ g/ml) X-gal板上。

[0262] 使用无菌移液器的枪头从每个反应中挑取8个大肠杆菌菌落,并转移到20 μ l含有羧苄青霉素的无菌液体LB中。然后使用与上述相同的引物组(ZmGWD ω -2或SbGWD ω -2)和2 μ l每种稀释的大肠杆菌克隆培养物通过PCR扩增GWD。确认PCR产物并如上所述进行测序。

[0263] 表4描述了ZmGWD大范围核酸酶事件的接合性、突变类型和位置。在表4中,ZmGWD突变编号为1-28。野生型(WT)植物是指两个GWD野生型等位基因。半合子事件是指一个GWD突变体等位基因和一个GWD野生型等位基因。杂合事件是指两个不同的GWD突变体等位基因。纯合事件是指两个相同的GWD突变体等位基因。

[0264] 表4 ZmGWD大范围核酸酶事件的接合性和突变

[0265]

构建体/事件	接合性	突变位点#(1)	突变类型 位点# (1)	突变位点# (2)	突变类型 位点# (2)
4715_5	半合子	M18	1bp 缺失		
4715_6	野生型		GWD 野生型		
4715_11	半合子	M17	10bp 缺失		
4715_13	半合子	M16	24bp 缺失		
4715_14	半合子	M17	10bp 缺失		
4715_15	半合子	M18	1bp 缺失		
4715_18	野生型		GWD 野生型		
4715_20	杂合子或 纯合子		两个突变 等位基因		
4715_25	半合子	M17	10bp 缺失		
4715_28	半合子	M27	16 bp 缺失		
4716_1	半合子	M1	4bp 取代		
4716_2	纯合子	M15	40bp 缺失		
4716_3	杂合子	M9	15bp 缺失	M6	1bp 缺失
4716_4	杂合子	M11	17bp 缺失	M12	38bp 缺失
4716_5	杂合子	M7	4bp 缺失	M11	17bp 缺失
4716_6	杂合子	M4	9bp 缺失	M14	25bp 缺失
4716_7	半合子	M1	4bp 取代		
4716_8	纯合子	M11	17bp 缺失		
4716_9	野生型		GWD 野生		

[0266]

构建体/事件	接合性	突变位点#(1)	突变类型 位点# (1)	突变位点# (2)	突变类型 位点# (2)
			型		
4716_10	杂合子	M11	17bp 缺失	M10	1bp 插入
4716_11	杂合子	M11	17bp 缺失	M10	1bp 插入
4716_12	杂合子	M3	15bp 缺失	M8	6bp 缺失
4716_13	纯合子	M14	25bp 缺失		
4716_14	纯合子	M13	36bp 缺失		
4716_15	杂合子	M11	17bp 缺失	M12	38bp 缺失
4716_18	半合子	M20	1bp 缺失		
4716_20	野生型		GWD 野生型		
4716_22	杂合子	M5	211bp 插入	M11	17bp 缺失
4716_23	纯合子	M15	40bp 缺失		
4716_24	杂合子	M2	6bp 缺失	M14	25bp 缺失
4716_25	杂合子	M10	1bp 插入	M28	27bp 缺失
4716_26	杂合子	M11	17bp 缺失	M12	38bp 缺失
4716_27	杂合子	M11	17bp 缺失	M10	1bp 插入
4716_151	半合子	M10	1bp 插入		
4716_152	杂合子	M22	10bp 缺失	M23	8bp 缺失 +15bp 插入 (7bp 插入)
4716_153	野生型		GWD 野生型		
4716_154	杂合子	M22	10bp 缺失	M23	8bp 缺失 +15bp 插入 (7bp 插入)
4716_155	纯合子	M4	9bp 缺失		
4716_157	半合子	M24	1bp 缺失		
4716_158	半合子	M20	1bp 缺失		
4716_159	纯合子	M20	1bp 缺失		
4716_160	半合子	M10	1bp 插入		
4716_161	杂合子	M4	9bp 缺失	M21	33bp 缺失 (57bp 缺失+)

[0267]

构建体/事件	接合性	突变位点#(1)	突变类型 位点# (1)	突变位点# (2)	突变类型 位点# (2)
					17bp 插入)
4716_162	杂合子	M20	1bp 缺失	M19	4bp 缺失
4716_163	半合子	M26	211bp 插入		
4716_164	半合子	M20	1bp 缺失		
4716_165	纯合子	M4	9bp 缺失		
4716_166	纯合子	M4	9bp 缺失		
4716_167	杂合子	M13	36bp 缺失	M25	2bp 缺失
4716_201	纯合子	M29	2bp 缺失		

[0268] 使用Vector NTI Advance (版本11.5;Life Technologies)将每个克隆的DNA序列与野生型(WT)GWD进行比较。比较野生型ZmGWD和SbGWD的DNA序列和转基因事件,并显示在以下文件中:ZmGWD大范围核酸酶突变体DNA序列比对;ZmGWD大范围核酸酶突变蛋白序列比对;SbGWD大范围核酸酶突变体DNA序列比对;SbGWD大范围核酸酶突变蛋白序列比对。

[0269] 如下所示,来自三个PCR产物的序列的比对表明了将玉米突变体M5和M26与野生型序列ZmGWD外显子24区分开的插入和缺失。

[0270] 用于ZmGWD的玉米突变体M5和M26的CLUSTAL 0 (1.2.1) 多重序列比对:

[0271] ZmGWD外显子24

[0272]

```

TTGGCAGGTTATAAGCCCGGTTGAAGTATCAGGTTATGTGGTTTGTGGTTGATGAGTTACT 60
M5      TTGGCAGGTTATAAGCCCGGTTGAAGTATCAGGTTATGTGGTTGTGGTTGATGAGTTACT 60
M26     TTGGCAGGTTATAAGCCCGGTTGAAGTATCAGGTTATGTGGTTGTGGTTGATGAGTTACT 60
*****
    
```

[0273] ZmGWD外显子24

[0274]

```

TGCTGTCCAGAACAATCTTATGATAAACCACCATCCTTGTGGCAAAGAGTGTCAAGGG120
M5      TGCTGTCCAGAACAATCTTATGATAAACCACCATCCTTGTGGCAAAGAGTGTCAAGGG120
M26     TGCTGTCCAGAACAATCTTATGATAAACCACCATCCTTGTGGCAAAGAGTGTCAAGGG120
*****
    
```

[0275] ZmGWD外显子24

[0276]

```

AGAGGAAGAAATACCAGATGGAGTA-----145
M5      AGAGGAAGAAATACCAGATGGAGTAGTTGCAGAAATTATTGAATTCCTTCATTAATTGAACT180
M26     AGAGGAAGAAATACCAGATGGAGCAGTGTGCTCGGGTACAGCTTCTTATTTCAATGTCTC180
*****
    
```

[0277] ZmGWD外显子24

[0278]

```

-----145
M5      CTATGATGATGCTTACTT--GATTTGATTATATTTGATGCTCAATCAATATTTGATGATTT238
M26     CAGTGGGCGTCTTACCTCTATGTTTGTGTTTTTTT~TTAAGTGCAGAAATAGAGAAAGTT239
    
```

[0279] ZmGWD外显子24

[0280]

```

-----145
M5      GTTGGAACTTGCTCTCCGATGCAAGGFGATCCAACGGGGTGTGTGCGCAACGTA AAC CAGG298
M26     CTFGCAAATATCTACTCTATGAAAAGGACAGCTAPTTGGAAATA-----TGTGAACAGA293

```

[0281] ZmGWD外显子24

[0282]

```

-----145
M5      GTTTTCG-CACGAGATGGCAATAGCTCTGT-T----AACCTAGCCTCTCACGGGCACGTGTG353
M26     ACTATC CCCCAGTTGCTGEGAAAACCAAGARGAAAGTTCCCTTCAAATACTACTCCATGA353

```

[0283] ZmGWD外显子24

[0284]

```

---GTTGGTGTAAATTACACCTGATATGCCAGATGTTCTGTCTCATGTGTGCAGTCCGAGCA202
M5      CGEGGGTATTTAATTACACCTGATATGCCAGATGTTCTGTCTCATGTGTGCAGTCCGAGCA413
M26     CGACAAGTGTCTATTACACCTGATATGCCAGATGTTCTGTCTCATGTGTGCAGTCCGAGCA413
*****

```

[0285] ZmGWD外显子24

```

AGGAATAGCAAG 214 (SEQ ID NO: 182)
[0286] M5      AGGAATAGCAAG 425 (SEQ ID NO: 34)
M26     AGGAATAGCAAG 425 (SEQ ID NO: 31)
*****

```

[0287] 来自28个PCR产物的序列的以下比对表明了区别玉米突变体M1-M4、M6-M25和M27-M29与野生型序列ZmGWD外显子24 (SEQ ID NO:1 (ZmGWD) 的3030-3243nt) 的修饰 (例如, 缺失和插入):

[0288] 用于玉米突变体M1-M4、M6-M25和M27-M29的CLUSTAL O (1.2.1) 多重序列比对:

[0289] ZmGWD外显子24

[0290]

```

TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M1 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M2 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M3 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M4 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M5 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M6 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M7 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M8 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M9 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M10 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M11 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M12 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M13 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M14 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M15 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M16 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M17 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M18 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M19 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M20 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M21 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M22 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M23 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M24 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M25 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M27 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M28 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
M29 TTGGCAGGTTATAAGCCCCGGTTGAAGTATCAGGTTAATGTTGGTTGTGGTTGATGAGTTACT 60
*****

```

[0291] ZmGWD外显子24

[0292]

	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M1	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M2	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M3	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M4	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M6	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M7	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M8	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M9	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M10	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M11	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M12	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M13	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M14	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M15	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M16	TGCTGTCCAGAACAAATCTTATGATAAAACCAAGGGAGAGGA-----	101
M17	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGGG-----	110
M18	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAA--GAGTGTCAAGGG	119
M19	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M20	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M21	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAAAT	120
M22	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M23	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M24	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M25	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M27	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGGGAGAGAT----	116
M28	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120
M29	TGCTGTCCAGAACAAATCTTATGATAAAACCAACCATCCTTGTGGCAAAGAGTGTCAAGGG	120

[0293] ZmGWD外显子24

[0294]

	AGAGGAAGAAATACCAGATGGAGTAGTTGGTGTGTA-----ATTACACCTGATATGCCAG	173
M1	AGAGGAAGAAATACCAGATGGAGTAGTTGGGAAGA-----AATACACCTGATATGCCAG	173
M2	AGAGGAAGAAATACCAGATGGAGTAGTTGTT-----ACACCTGATATGCCAG	167
M3	AGAGGAAGAAATACCAGATGGAGCACCT-----GATATGCCAG	158
M4	AGAGGAAGAAATACCAGATGGAGTAA-----TTACACCTGATATGCCAG	164
M6	AGAGGAAGAAATACCAGATGGAGTAGTTGGTAA-----ATTACACCTGATATGCCAG	172
M7	AGAGGAAGAAATACCAGATGGAGTAGTTGGT-----TTACACCTGATATGCCAG	169
M8	AGAGGAAGAAATACCAGATGGAGTAGTTGGT-----ATGCCAGATATGCCAG	167
M9	AGAGGAAGAAATACCAGATGGAGTAGTTG-----GATATGCCAG	158
M10	AGAGGAAGAAATACCAGATGGAGTAGTTGGTGTAA-----ATTACACCTGATATGCCAG	174
M11	AGAGGAAGAAATACCAGATGGAGTAGTTGGTGTGAG-----	156
M12	AGAGGAAGAAATACCAGATGGAGTAGTTGGTGTGAGTCCGAGCAAGGAAATAGCAAG----	176
M13	AGAGGAAGAAATACCAGATGTTCTGTCTCATGTGTGTC-----	156
M14	AGAGGAAGAAATACA-----CCTGATATGCCAG	148
M15	CTGATATGC-----CAG	132
M16	-----AGAAATACCAGATGGAGTAGTTGGT-----GTAATTACACCTGATATGCCAG	148
M17	AGAGGAAGAAATACCAGATGGAGTAGTTGGTGT-A-----ATTACACCTGATATGCCAG	163
M18	AGAGGAAGAAATACCAGATGGAGTAGTTGGTGTGTA-----ATTACACCTGATATGCCAG	172
M19	AGAGGAAGAAATACCAGATGGAGTAGTTGGTGT-----ATTACACCTGATATGCCAG	172
M20	AGAGGAAGAAATACCAGATGGAGTAGTTGGTGT-----ATTACACCTGATATGCCAG	172
M21	CTTATGATAAAAC-----ATGCCAG	140
M22	AGAGGAAGAAATACCAGATGGAGTAGTTGGTGTGA-----TATGCCAG	163
M23	AGAGGAAGAAATACCAGATGGAGTAGTTGGCAAAGATAAAACCTTGCACCTGATATGCCAG	180
M24	AGAGGAAGAAATACCAGATGGAGTAGTTGGTGA-----ATTACACCTGATATGCCAG	172
M25	AGAGGAAGAAATACCAGATGGAGTAGTTGG-----TA-----ATTACACCTGATATGCCAG	171
M27	-----ACCAGATGGAGTAGTTGG-----TGTAATTACACCTGATATGCCAG	157
M28	AGAGGAAGAAACACCTGATA-----TGCCAG	146
M29	AGAGGAAGAAATACCAGATGGAGTAGTTGG-----TA-----ATTACACCTGATATGCCAG	171

[0295]

ZmGWD 外显子 24	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	214 (SEQ ID NO: 182)
M1	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	214 (SEQ ID NO: 16)
M2	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	208 (SEQ ID NO: 35)
M3	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	199 (SEQ ID NO: 19)
M4	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	205 (SEQ ID NO: 29)
M6	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	213 (SEQ ID NO: 37)
M7	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	210 (SEQ ID NO: 39)
M8	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	208 (SEQ ID NO: 20)
M9	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	199 (SEQ ID NO: 38)
M10	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	215 (SEQ ID NO: 18)
M11	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	197 (SEQ ID NO: 17)
M12	-----	176 (SEQ ID NO: 23)
M13	-----AGTCCGAGCAAGGAATAGCAAG	178 (SEQ ID NO: 22)
M14	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	189 (SEQ ID NO: 21)
M15	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	173 (SEQ ID NO: 33)
M16	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	189 (SEQ ID NO: 12)
M17	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	204 (SEQ ID NO: 13)
M18	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	213 (SEQ ID NO: 14)
M19	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	213 (SEQ ID NO: 30)
M20	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	213 (SEQ ID NO: 27)
M21	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	181 (SEQ ID NO: 28)
M22	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	204 (SEQ ID NO: 24)
M23	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	221 (SEQ ID NO: 25)
M24	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	213 (SEQ ID NO: 26)
M25	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	212 (SEQ ID NO: 32)
M27	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	198 (SEQ ID NO: 15)
M28	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	187 (SEQ ID NO: 36)
M29	ATGTTCTGTCTCATGTGTGTCAGTCCGAGCAAGGAATAGCAAG	212 (SEQ ID NO: 40)

[0296] 分析来自29个转基因玉米事件和野生型植物的GWD的氨基酸序列,并且显示在野生型ZmGW (SEQ ID NO:185) 蛋白中在氨基酸1040-1120位置的缺失和插入,将该蛋白与以下玉米突变体区分开:ZmGWD_M1 (SEQ ID NO:45)、ZmGWD_M2 (SEQ ID NO:46)、ZmGWD_M3 (SEQ ID NO:47)、ZmGWD_M4 (SEQ ID NO:48)、ZmGWD_M5 (SEQ ID NO:49)、ZmGWD_M6 (SEQ ID NO:50)、ZmGWD_M7 (SEQ ID NO:51)、ZmGWD_M8 (SEQ ID NO:52)、ZmGWD_M9 (SEQ ID NO:53)、ZmGWD_M10 (SEQ ID NO:54)、ZmGWD_M11 (SEQ ID NO:55)、ZmGWD_M12 (SEQ ID NO:56)、ZmGWD_M13 (SEQ ID NO:57)、ZmGWD_M14 (SEQ ID NO:58)、ZmGWD_M15 (SEQ ID NO:59)、ZmGWD_M16 (SEQ ID NO:60)、ZmGWD_M17 (SEQ ID NO:61)、ZmGWD_M18 (SEQ ID NO:62)、ZmGWD_M19 (SEQ ID NO:63)、ZmGWD_M20 (SEQ ID NO:4)、ZmGWD_M21 (SEQ ID NO:65)、ZmGWD_M22 (SEQ ID NO:66)、ZmGWD_M23 (SEQ ID NO:67)、ZmGWD_M24 (SEQ ID NO:68)、ZmGWD_M25 (SEQ ID NO:69)、ZmGWD_M26 (SEQ ID NO:70)、ZmGWD_M27 (SEQ ID NO:71)、ZmGWD_M28 (SEQ ID NO:72) 和ZmGWD_M29 (SEQ ID NO:73)。

[0297] ZmGWD (SEQ ID NO:43) 氨基酸1040-11120的CLUSTAL O (1.2.1) 多重序列比对:

	EmGWD	FTILVAKSVGEESEIPDGVVGVITPDMPD-----VLS---HV-----SVR-
	EmGWD_M1	FTILVAKSVGEESEIPDGVVGRHTPDMPD-----VLS---HV-----SVR-
	EmGWD_M2	FTILVAKSVGEESEIPDGV--VYTPDMPD-----VLS---HV-----SVR-
	EmGWD_M3	FTILVAKSVGEESEIPDGA-----FDMPD-----VLS---HV-----SVR-
	EmGWD_M4	FTILVAKSVGEESEIPDG---VITPDMPD-----VLS---HV-----SVR-
	EmGWD_M5	FTILVAKSVGEESEIPDGVVAVLNSFTIELYDDALLDCIILMLLHILMIVGTCSPNQGD
	EmGWD_M6	FTILVAKSVGEESEIPDGVVGLHALICQM-----FCL---MQSQEQGIARYCLRF
	EmGWD_M7	FTILVAKSVGEESEIPDGVVGLHALICQM-----FCL---MQSQEQGIARYCLRF
	EmGWD_M8	FTILVAKSVGEESEIPDGV--VGMFDMPD-----VLS---HV-----SVR-
	EmGWD_M9	FTILVAKSVGEESEIPDGV-----VGMFD-----VLS---HV-----SVR-
	EmGWD_M10	FTILVAKSVGEESEIPDGVVGVNYT*
	EmGWD_M11	FTILVAKSVGEESEIPDGVVGRVCSVCSVPSKE*
	EmGWD_M12	FTILVAKSVGEESEIPDGVVGVSPK-----E*
	EmGWD_M13	FTILVAKSVGEESEI-----PD-----VLS---HV-----SVR-
	EmGWD_M14	FTILVAKSVGEESE-----LHLLICQM-----FCL---MQSQEQGIARYCLRF
	EmGWD_M15	FTILVAKSNYT*
	EmGWD_M16	F-----RERSKYQME*
	EmGWD_M17	FTIL---VAREKRYQME*
	EmGWD_M18	FTILVARVSRERKRYQME*
	EmGWD_M19	FTILVAKSVGEESEIPDGVVGLVHLICQM-----FCL---MQSQEQGIARYCLRF
	EmGWD_M20	FTILVAKSVGEESEIPDGVVGLVHLICQM-----FCL---MQSQEQGIARYCLRF
	EmGWD_M21	FTILVAKSVVIL*
	EmGWD_M22	FTILVAKSVGEESEIPDGVVGLVHLICQM-----FCL---MQSQEQGIARYCLRF
[0298]	EmGWD_M23	FTILVAKSVGEESEIPDGVVGLKDEPCT*
	EmGWD_M24	FTILVAKSVGEESEIPDGVVGLHALICQM-----FCL---MQSQEQGIARYCLRF
	EmGWD_M25	FTILVAKSVGEESEIPDGVVGNHT*
	EmGWD_M26	FTILVAKSVGEESEIPDGVAVCSGTASYFNVSS-GRLTSMFVFFFK-CRNRESSCKYLLVE
	EmGWD_M27	FTILVA-----RERSKYQME*
	EmGWD_M28	FTILVAKSVGEESE-----TFDMPD-----VLS---HV-----SVR-
	EmGWD_M29	FTILVAKSVGEESEIPDGVVGVNYT*
	EmGWD	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 185)
	EmGWD_M1	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 45)
	EmGWD_M2	-----APHSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 46)
	EmGWD_M3	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 47)
	EmGWD_M4	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 48)
	EmGWD_M5	---FTGVCRHVHVFVARDGHSVHLASRCHCAGVFRYT*----- (SEQ ID NO: 49)
	EmGWD_M6	VL-TTPLYLNLKDMIRNCFPSILLQI*----- (SEQ ID NO: 50)
	EmGWD_M7	VL-TTPLYLNLKDMIRNCFPSILLQI*----- (SEQ ID NO: 51)
	EmGWD_M8	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 52)
	EmGWD_M9	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 53)
	EmGWD_M10	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 54)
	EmGWD_M11	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 55)
	EmGWD_M12	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 56)
	EmGWD_M13	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 57)
	EmGWD_M14	VL-TTPLYLNLKDMIRNCFPSILLQI*----- (SEQ ID NO: 58)
	EmGWD_M15	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 59)
	EmGWD_M16	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 60)
	EmGWD_M17	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 61)
	EmGWD_M18	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 62)
	EmGWD_M19	VL-TTPLYLNLKDMIRNCFPSILLQI*----- (SEQ ID NO: 63)
	EmGWD_M20	VL-TTPLYLNLKDMIRNCFPSILLQI*----- (SEQ ID NO: 64)
	EmGWD_M21	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 65)
	EmGWD_M22	VL-TTPLYLNLKDMIRNCFPSILLQI*----- (SEQ ID NO: 66)
[0299]	EmGWD_M23	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 67)
	EmGWD_M24	VL-TTPLYLNLKDMIRNCFPSILLQI*----- (SEQ ID NO: 68)
	EmGWD_M25	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 69)
	EmGWD_M26	KDEYLEICEQNVYFQLLGRKTK---K-----VPSHIYSNTTSV---YPT*----- (SEQ ID NO: 70)
	EmGWD_M27	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 71)
	EmGWD_M28	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 72)
	EmGWD_M29	-----ARNSKVLKATCFDHTLSELEGYDQKLFSPFPTSADITYREITE (SEQ ID NO: 73)

[0300] 对于高粱,使用两个大范围核酸酶构建体产生GWD突变,4715和4716。第一代(T0)转化的植物可以产生纯合的GWD突变体、半合子(WT+突变)GWD突变体或杂合(两个不同的突变体,例如,等位基因1+等位基因2)GWD突变体。使用以下缩写:del=缺失;ins=插入;sub=取代;SbGWD CDS是野生型序列。例如,序列名“Sb4715_1(WT+ins)”具有以下含义:Sb4715

是高粱中的构建体;1是转基因事件:WT+ins表明T0事件4715_1对于携带WT GWD等位基因和插入(ins)GWD等位基因的GWD突变是半合子的。相同的构建体被用于玉米(Zm)的转化。

[0301] 高粱双色(Sb)GWD序列和高粱双色GWD突变体Sb475_1(WT+ins)和Sb4715_2(WT+del)之间的CLUSTAL核酸比对显示与野生型SbGWD序列相比突变体序列的改变。SbGWD_外显子24序列位于SbGWD编码序列(SEQ ID NO:2)的3030-3243nt内。Sb475_1(WT+ins)的序列含有在SbGWD的位置3139-3149中的13个核苷酸插入和在SbGWD的位置3133-3136的核苷酸取代。

[0302] 如下所示,来自3个PCR产物的序列的比对表明了区分Sb4715_1(WT+ins)和Sb4715_2(WT+del)与SbGWD外显子24区域的插入和缺失。

[0303] CLUSTAL O(1.2.1)多重序列比对:

[0304]

SbGWD_外显子24 DTGGCAGGTTTATAAGCCACATGGAAGTATCAGGTTATGTTGGTTGTGGTTCATGACTTACT 80

[0305]

Sb4715_1 (WT+ins) TTGCGAGGTTTATAAGCCACATGGAAGTATCAGGTTATGTTGGTTGTGGTTCATGACTTACT 80
Sb4715_2 (WT+del) TTGCGAGGTTTATAAGCCACATGGAAGTATCAGGTTATGTTGGTTGTGGTTCATGACTTACT 80
SbGWD_外显子24 TCCCTGTCCAGAAACAATTCCTATGATAAACCACATCCCTGTGTCAGAAACA----- 118
Sb4715_1 (WT+ins) TCCCTGTCCAGAAACAATTCCTATGATAAACCACATCCCTGTGTCAGAAACA----- 120
Sb4715_2 (WT+del) TCCCTGTCCAGAAACAATTCCTATGATAAACCACATCCCTGTGTCAGAAACA----- 118
SbGWD_外显子24 ---GTGTCAGAAACAATTCCTATGATAAACCACATCCCTGTGTCAGAAACA----- 167
Sb4715_1 (WT+ins) GTGTCAGAAACAATTCCTATGATAAACCACATCCCTGTGTCAGAAACA----- 169
Sb4715_2 (WT+del) ---GTGTCAGAAACAATTCCTATGATAAACCACATCCCTGTGTCAGAAACA----- 165
SbGWD_外显子24 TTGCGAGGTTTATAAGCCACATGGAAGTATCAGGTTATGTTGGTTGTGGTTCATGACTTACT 180
Sb4715_1 (WT+ins) TTGCGAGGTTTATAAGCCACATGGAAGTATCAGGTTATGTTGGTTGTGGTTCATGACTTACT 186
Sb4715_2 (WT+del) TTGCGAGGTTTATAAGCCACATGGAAGTATCAGGTTATGTTGGTTGTGGTTCATGACTTACT 182

[0306] 使用序列野生型SbGWD(SEQ ID NO:44)进行大范围核酸酶突变蛋白氨基酸序列的预测。

[0307] 实施例5. 突变植物积累升高水平的绿色组织淀粉

[0308] 在第一代(T0)转化的玉米和高粱GWD大范围核酸酶植物中测定淀粉。收集组织、干燥并研磨成细粉末。通过标准方法(Smith AM and Zeeman SC,Quantification of starch in plant tissues(2006)Nat Protocols 1:11342-1345,其通过引用并入本文,如同完整阐述)测定淀粉含量。通过改编来自Megazyme国际爱尔兰有限公司(Megazyme试剂盒和试剂;货号K-TSTA)的方案测定总淀粉含量。简单地说,建立85°C和50°C的加热块。将5至15mg干燥的碾磨组织置于1.5ml防沸腾微量离心管中。向每个管中加入1毫升70%乙醇,涡旋混合样品并沉淀。向每个样品中加入400微升溶液1。所述溶液1含有1ml的热稳定性α-淀粉酶和29ml的100mM乙酸钠缓冲液,pH5.0。将样品重悬并涡旋。将样品在85°C下孵育12分钟,并在室温下冷却5分钟。将300微升的GOPOD试剂(Megazyme试剂盒,货号K-TST)预先加入至平底96孔测定板的每个孔中。将10微升样品加入到每个孔中,并与也被加入至它们各自孔中的1μL,5μL,10μL和20μL葡萄糖标准品(1mg/ml)相比较。将板在50°C下孵育20min。在510nm测定吸光值。参考图4。图4显示了与野生型植物WT_195、WT_18和WT_6相比,突变体4715_20(两个突变体等位基因)、4716_7(M1)、4716_1(M1)、4716_18(M20)、4716_12(M3/M8)、4716_23(M15)、4716_28(未表征)、4716_3M9/M6)、4716_22(M5/M11)、4716_24(M2/M14)、24716_25

(M10/M28)、4716_15 (M11/M12)、4716_5 (M7/M11)、4716_6 (M4/M14)、4716_27 (M11/M10)、4716_4 (M11/M12)、4716_26 (M11/M12)、4716_2 (M15)、4716_13 (M14)、4716_13 (M14)、4716_11 (M11/M10)、4716_8 (M11)、4716_10 (M11/M10) 和4716_14 (M13) 具有升高的淀粉水平。许多纯合和杂合事件在叶中表现出大于20重量%的淀粉增加量。基于不同组织中淀粉积累的加权平均值,我们估计总植物淀粉(不包括谷物)为约10% (重量/重量)。

[0309] 使用RNA干扰技术,先前在玉米中观察到的水平显著升高,尽管在那些实验中测量到低转录丰度。这是令人惊讶的结果,因为预期基于RNAi的沉默在植物中将式明显的并且具有与基因缺失或敲除策略相同的效果。

[0310] 图5显示了所选择的半合、纯合和杂合事件的绿色组织淀粉。图5显示了转基因玉米事件M17 (4715_14)、M18 (4715_15)、M1 (4716_1)、M20 (4716_18)、M3/M12 (4716_12)、M9 (4716_3)、M7/M11 (4716_5)、M4/M14 (4716_6)、M11/M12 (4716_4)、M15 (4716_2)、M14 (4716_13)、M11 (4716_8)、M11/M10 (4716_10)、M13 (4716_14) 与非转基因对照WT (4716_9) 相比,具有升高的淀粉水平。观察到数个事件 (4716_13 (M14)、4715_15 (M11/M12) 和4716_6 (M4/M14) 具有高于平均的生物物质的量。

[0311] 实施例6.GWD敲除 (GWDko) 的穗轴具有升高的淀粉水平

[0312] TOGWDko和野生型 (wt) 玉米突变体品系自交,培育至成熟,干燥,并进行横切以进行染色。将穗轴切片用Lugol's溶液 (5%KI) 染色4min,并用H₂O过夜脱色。

[0313] 分析突变事件4716_13、4716_26、4716_167、4716_164、4716_9和4716-153的淀粉含量。结果如表5所示。

[0314] 表5突变品系中的淀粉含量

[0315]

构建体_事件	接合性	绿色组织淀粉	秸秆淀粉
4716_13	纯合子	229.7	38.5
4716_26	杂合子	220.7	35.3
4716_167	杂合子	112.0	31.7
4716_164	半合子	30.5	4.0
4716_9	野生型	18.8	8.4
4716_153	野生型	5.7	1.2

[0316] 参考表5和图6,表明与半合子 (4716_164) 和野生型 (4716_9和4716_153) 穗轴相比,纯合的 (4716_13) 和杂合的 (4716_26和4716_167) 穗轴具有加深的淀粉染色。

[0317] 实施例7.CRISPR/Cas玉米转化载体的构建

[0318] 为了构建Cas9表达盒,选择含有N端和C端At核定位序列 (NLS) 的化脓性链球菌Cas9蛋白序列以及紧接在第一个ATG密码子之后的3xFLAG序列 (Jiang et al., 2013) (SEQ ID NO: 74) 用于在玉米中表达。在玉米中表达的含有两个SV40核定位序列 (以粗体字母表示) 和3xFLAG序列在N-末端部分 (下划线序列) 的化脓性链球菌Cas9的序列如下所示:

[0319] MDYKDHDGDYKDHDTDYKDDDDKMA~~XXXXXXXXXX~~VGTHGVPAAADKKYSIGLDIGTNSVGVAVITDEYKVPSK
KFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLV
EEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKL
FIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED

AKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAIILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKAL
 VRQQLEPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKF IKP ILEKMDGTEELLVKNLREDLLRQRTFDNGSIPHQIH
 LGELHA I LRRQEDFYPPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF I
 ERMTNFDKNLPNEKVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF
 KKIECFDSVEISGVEDRFNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIERLKYAHLFDDK
 VMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEH
 IANLAGSPA I KKGILQTVKVVDELVKVMGRHKPENIV IEMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHP
 VENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVK
 KMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVK
 VITLKSCLVSDFRKDFQFYKVRINNYHHAHDAYLNAVVG TALI KKYPKLESEFVYGDYKVYDVRKMIKSEQEIGK
 ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI VWDKGRDFATVRKVL SMPQVNI VKKTEVQTGGFSKE
 SILPKRNSDKLIARKKDWDPKKGFDSPVAVSVLVVAKVEKGKSKKLKSVKELLGITIMERS SFEKNPIDFLEAK
 GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLGKSPEDNEQKQLFVEQH
 KHYLDEIIEQISEFSKRVILADANLDKVL SAYNKHRDKPIREQAENI IHLFTLTNLGAPAAFKYFDTTIDRKRYTST
 KEVL DATLIHQSI TGLYETRIDLSQLGGDRXXXXXXXXVGG (SEQ ID NO:74)

[0320] 将序列反向翻译并且优化玉米密码子以产生ZmCas9 (SEQ ID NO:75)。优化的ZmCas9核苷酸序列由Genscript合成。将ZmCas9作为在玉米泛素1启动子(ZmUbi1P)和胭脂碱合酶转录终止子(NosT)序列之间的BamHI-AvrII片段克隆到pAG4500中以产生pAG4800。

[0321] 关于构建sgRNA盒的工作包括:1) 鉴定和分离玉米RNA聚合酶III启动子以促进sgRNA的表达;2) sgRNA支架的设计与合成;和3) 选择靶基因和该基因内的20bp特异性序列,用于将Cas9核酸内切酶引导至其靶位点。

[0322] 在1994年,Leader等报道了关于编码U3小核RNA(U3snRNA)的玉米序列的首次描述,他从玉米基因组DNA文库中分离出MzU3.8基因(Genebank登录号Z29641)(SEQ ID NO:76),并表明了MzU3.8U3snRNA在玉米原生质体中表达。使用BLASTN算法和Z29641序列搜索玉米遗传学和基因组学数据库(<http://www.maizegdb.org/>),我们鉴定了标记为ZmU3(SEQ ID NO:77)的玉米U3的同源序列。

[0323] 所述ZmU3位于玉米染色体8上并且被包含在具有核苷酸配位163620300-163621800的序列内。MzU3.8(SEQ ID NO:78)和ZmU3(SEQ ID NO:79)的推导启动子区的CLUSTAL 2.1多核苷酸序列比对显示在两个序列之间具有93.8%的同一性。

[0324] CLUSTAL 2.1多重序列比对:

[0325]

```

MzU3.8      GAAATTCATCTAAGTATCTTGGTAAAGCATGGATTAATTTGGATGCTCACTTCAGGTCTA  60
ZmU3        GAAATTCATCTAAGTATCTTGGTAAAGCATGGATTAATTTGGATGCTCACTTCAGGTCTA  60
*****

MzU3.8      TGCAGCTCCGGTGCCTTSTGATTGTGAGTGTGTGACCGATGCTCATGCTAFTTTGCATTTC  120
ZmU3        TGCAGCTCCGGTGCCTTSTGATTGTGAGTGTGTGACCGATGCTCATGCTAFTTTGCATTTC  120
*****

MzU3.8      TGGATGTATGATGCTAGTAGATCTTCAAACCTAACACGGCATGCCATCATATCCACTG  180
ZmU3        TGGATGTATGATGCTAGTAGATCTTCAAACCTAACACGGCATGCCATCATATCCACTG  180
*****

MzU3.8      CTTGATTTTATGCTCTACCGCTGGCCAAAAATGTGATGATGCTCAGAAACCTCAACTACCTT  240
ZmU3        CTTGATTTTATGCTCTACCGCTGGCCAAAAATGTGATGATGCTCAGAAACCTCAACTACCTT  240
*****

MzU3.8      GAATCAACACGGGGCCAGCAGTGTGTATGACGACAGAAACCAAAAAAAAAATGAGCCAAATAG  300
ZmU3        GAATCAACACGGGGCCAGCAGTGTGTATGACGACAGAAAC-AAAAAAAAATGAGCCAAATAG  299
*****

MzU3.8      TTCAGAAGGAGGCACCTATGCTGAAACTACATTTCTGAAGGTGACTAAAAGGTGAGCGTAG  360
ZmU3        TTCAGAAGGAGGCACCTATGCTGAAACTACATTTCTGAAGGTGACTAAAAGGTGAGCGTAG  359
*****

MzU3.8      AGTGTACTTACTAGTAGTTAGCCACCCTTACCCAAATGCTTTGAGCTTGTATTAGAC  420
ZmU3        AGTGTAACTACTAGTAGTTAGCCACCCTTACCCAAATGCTTTGAGCTTGTATTAGAT  419
*****

MzU3.8      TTCCTAAGCTGAGCATCACTACTGATCTGCAGG--AGGGTCGCTTCGCTGCCAAGATCAA  478
ZmU3        TTCCTAAGCTGAGCATCACTACTGATCTGCAGGCGCCACCCTCGCTTCGCTGCCAAGATCAA  479
*****

MzU3.8      CAGCAACCATGTGGCGGCAACATCCAGCAATGCACATGGCTAAAGATGAGCTCTGTGTC  538
ZmU3        CAGCAACCATGTGGCGGCAACATCCAGCAATGCACATGGCTAAAGATGAGCTCTGTGTC  539
*****

MzU3.8      CAAGTGTGAGCTGCAACCATCTAGGGATCAGCTGAGTTTATCAGTCTTTTCTTTTTC  598
ZmU3        C-----TGTCTAGGGATCAGCTGAGTTTATCAGTCTTTTCTTTTTC  584
* * * * *

MzU3.8      TTCTGGTGAAGGCTCAASCTACTACTGCTTCGATCGGTTGGACTTGGACCTGAAGCCAC  656
ZmU3        TCCAGGTGAGGCTCAAGCTACTACTGCTTCGATCGGTTGGACTTGGACCTGAAGCCAC  638
* * * * *

MzU3.8      ATGTAGGATACCGAATGGACCCAGGACG-----TA  693
ZmU3        ATGTAGGATACCGAATGGACCCAGGACCGCAGTATGTTGGCCAGTCCACCCGGTTA  688
*****

MzU3.8      GTGCCACTCTGGTTG-TCACTACTGGTATGAGGCCAGCTTAAAAATTTAGCTTTGGTACT  752
ZmU3        GTGCCACTCTGGTTGCTCACA-TGCGTAGAAGCCAGCTTAAAAATTTAGCTTTGGTACT  757
*****

MzU3.8      CACAGCA 759 (SEQ ID NO: 78)
ZmU3        CACAGCA 764 (SEQ ID NO: 79)
*****

```

[0326] 使用PCR方法,采用正向引物ob2297 (SEQ ID NO:80)和反向引物ob2299 (SEQ ID NO:81),随后从玉米品系AxB的玉米基因组DNA中分离出758bp的ZmU3启动子(ZmU3P1) (SEQ ID NO:82)。正向引物ob2297在其5'端包括AsiSI限制性位点以利于将sgRNA盒克隆至基于pAG4500的载体中。类似地,使用在其5'端含有AsiSI限制性位点的正向引物ob2343 (SEQ ID NO:83),分离出较短的398bp形式的玉米U3启动子(ZmU3P2) (SEQ ID NO:84),用于测试截短的玉米U3启动子的效率。使用在其5'端具有SwaI限制酶位点的正向引物ob2351 (SEQ ID NO:85)扩增ZmU3P2的另外的变体。此外,使用在Leader等(1994)公开的MzU3.8序列上设计

的长引物,显示其在玉米原生质体中表达,合成308bp对照启动子片段ZmU3.8P (SEQ ID NO: 86)。该ZmU3.8P序列还包括5'端的AsiSI限制性位点。将所有扩增的启动子变体克隆到pCR-BluntII-TOPO载体 (Life Technologies) 中,并通过完全测序证实其完整性。

[0327] 本文的sgRNA支架设计基于sgRNA嵌合体的已被公开的组织 (Larson et al., 2013), 并且含有42bp Cas9柄发夹 (SEQ ID NO:87), 其后连有41bp化脓性链球菌终止子 (SEQ ID NO:88)。为了提高玉米中转录终止的效率,从ZmU3 snRNA (SEQ ID NO:77) 中分离37bp的推导的转录终止子序列ZmU3T (SEQ ID NO:89), 并将其融合在化脓性链球菌终止子 (SEQ ID NO:88) 的下游。使用长引物和具有校对读取活性的KOD Xtreme DNA聚合酶通过PCR合成120bp sgRNA支架 (SEQ ID NO:90)。在两个PCR扩增的sgRNA主链DNA片段的3'末端添加SnaBI或AscI限制性位点以利于进一步克隆。将以这种方式合成的sgRNA支架DNA片段克隆到pCR-BluntII-TOPO载体中并验证序列。

[0328] 为了测试玉米中CRISPR/Cas系统的效率,选择编码GWD的玉米基因用于初始靶向修饰。

[0329] 玉米GWD基因早期已经被注释过,并在正义和反义DNA链上筛选AN19NGG靶序列的存在。AN19NGG序列中的5'末端“A”表示在U3RNA聚合酶III启动子的转录起始处的保守的“腺嘌呤”核苷酸,位于3'末端的“NGG”序列对应于CRISPR/Cas系统活性原生质体-相邻基序 (PAM) 序列。在外显子1、24和25以及其侧翼内含子中鉴定的候选靶序列针对玉米GWD进一步筛选,以消除在玉米基因组内具有多重同一性命中的序列。这项工作已经完成,以尽量减少CRISPR/Cas系统脱靶活性的可能性。在该分析中,在Larson等 (2013年) 提出的BLASTN程序中仅使用靶序列的种子序列 (12bp) 加上两个相邻的PAM核苷酸。选择外显子1以产生几乎完全的GWD敲除,而选择外显子24和25以产生缺乏由外显子24编码的活性位点的GWD变体。最终的19bp GWD靶序列 (SEQ ID NO:131-134), 其被鉴定用于sgRNA研发,如表6所示。

[0330] 表6 GWD基因靶序列及其对应的SEQ ID NO

[0331]

SEQ ID NO	序列名称	序列	GWD 链
91	GWDe1a	GGCATGAGGTGCTTACGTC	反义
92	GWDe24b	CATAACCTGATACTTCAAC	反义
93	GWDe24c	TCTGGCTCCTGCTATCAGT	正义
94	GWDe25a	TCTGCAGAAGTAGGCTTGA	反义

[0332] 使用KOD Xtreme DNA聚合酶通过融合PCR的方式将玉米U3启动子的三种变体中的每一种、选择的GWD靶序列和sgRNA骨架组装在一起,以构建六个sgRNA表达盒 (SEQ ID NO: 135-140)。将PCR扩增的片段克隆到pCR-BluntII-TOPO载体中,并通过测序验证合成的sgRNA表达盒的完整性。PCR合成的sgRNA盒的列表如表7所示。

[0333] 表7合成的sgRNA表达盒及其对应的SEQ ID NO

[0334]

SEQ ID NO	sgRNA表达盒	侧翼限制性位点
95	ZmU3P1:sgRNA_GWDe24b	AsiSI-SnaBI
96	ZmU3P2:sgRNA_GWDe24b	AsiSI-SnaBI

97	ZmU3.8P:sgRNA_GWDe24b	AsiSI-SnaBI
98	ZmU3P2:sgRNA_GWDe24c	AsiSI-SnaBI
99	ZmU3P2:sgRNA_GWDe25a	SwaI-AscI
100	ZmU3P2:sgRNA_GWDe1a	AsiSI-SnaBI

[0335] 随后将组装的sgRNA盒作为AsiSI-SnaBI片段克隆到pAG4800中以构建载体pAG4804-4809(表8)。

[0336] 表8用于玉米CRISPR/Cas系统研发的载体及其SEQ ID NO

[0337]

质粒	遗传因子
pAG4800	Ubi1P:Cas9
pAG4804	U3P1:sgRNA_GWDe24b + Ubi1P:ZmCas9
pAG4805	U3P2:sgRNA_GWDe24b + Ubi1P:ZmCas9
pAG4806	U3.8P:sgRNA_GWDe24b + Ubi1P:ZmCas9

[0338]

pAG4807	U3P2:sgRNA_GWDe24c + Ubi1P:ZmCas9
pAG4808	U3P2:sgRNA_GWDe1a + Ubi1P:ZmCas9
pAG4809	U3P2:sgRNA_GWDe25a + Ubi1P:ZmCas9
pAG4817	U3P2:sgRNA_GWDe25a + U3P2:sgRNA_GWDe24c + Ubi1P:ZmCas9

[0339] 通过将作为SwaI-AscI片段的ZmU3P2:sgRNA_GWDe25a盒克隆到pAG4807中构建一个含有两个sgRNA表达盒的另外的载体pAG4817。通过将Cas9核酸内切酶靶向作用于位于玉米GWD基因和侧翼外显子24内相距364bp的两个不同位点,该载体被构建用于完全去除GWD外显子24。

[0340] 为了研发用于玉米的CRISPR/Cas系统而构建的植物转化载体pAG4800和pAG4804的图谱显示在图7-8中。如图7所示,pAG4800载体含有Cas9表达盒和PMI表达盒。所述Cas9表达盒含有核苷酸序列ZmCas9(SEQ ID NO:75)。所述ZmCas9是编码Cas9的化脓性链球菌基因的玉米密码子优化的序列,其被融合到两个At NLS的5'末端和3'末端以及紧接在第一个ATG密码子之后的3xFLAG序列。Zm Cas9编码在用于在玉米中表达的氮末端部分含有两个At核定位序列和3xFLAG序列的化脓性链球菌Cas9蛋白(SEQ ID NO:74)。所述Cas9盒还含有Zm Ubi1启动子、Zm Ubi1内含子、mUBQMono前导序列和NosT终止子。所述PMI盒含有PMI基因、ZmUbi1启动子、mUBQMono、ZmKozak前导序列和NosT终止子。如图8所示,pAG4804载体含有GWDe24b-sgRNA支架盒、Cas9表达盒和PMI表达盒。所述GWDe24b-sgRNA支架盒含有ZmU3P1启动子、GWDe24序列、sgRNA支架和ZmU3T终止子。所述Cas9表达盒含有在5'末端和3'末端与两个At NLS融合的ZmCas9和紧接在第一个ATG密码子之后的3xFLAG序列。所述Cas9盒还含有Zm Ubi1启动子、Zm Ubi1内含子、mUBQMono前导序列和NosT终止子。所述PMI盒含有PMI基因、ZmUbi1启动子,mUBQMono、ZmKozak前导序列和NosT终止子。

[0341] 实施例8. CRISPR/Cas诱导的突变植物的产生

[0342] CRISPR/Cas和玉米NLS大范围核酸酶诱导的玉米GWD基因突变的鉴定和表征

[0343] 根据实施例3中所述的方案进行玉米植物转化。筛选CRISPR/Cas诱导的突变,其与实施例3中描述的筛选大范围核酸酶诱导的突变方法相似,除了用于基因分型的引物和识别突变。

[0344] 表9描述了用于CRISPR/Cas植物基因分型的引物,包括4804-4806、4804-6引物组;将GWD24b-F替换为GWDe24a-F的4809、4817和4804-6引物组和围绕GWD大范围核酸酶靶向区域的用于扩增DNA序列的引物,其包括4804-4807、4804-7mut引物;4817、2856/2858引物;4837-4839,371/429引物。表9用于基因分型4804、4805、4806、4817、4837、4838和4839植物和扩增围绕GWD外显子24靶向区域的DNA序列的引物

[0345]

引物组	引物名称	正向或反向	序列	SEQ ID NO	产物大小(bp)
4804-6	GWDe24b-F	正向	CTCACAGCACATA ACCTGATACT	101	100
4804-6	sgRNA-R	反向	CGACTCGGTGCCA CTTT	102	100
4804-6	ZmCas9-F	正向	AGAATCAGACCAC GCAGAAG	103	186
4804-6	ZmCas9-R	反向	GCTCCTGGTCCACA TACATATC	104	186
4809/4817	GWDe24a-F	正向	TGCAGAAGTAGGC TTGAGTTT	110	89
4804-7mut	GWDe23-F	正向	TGCTCTTCTGAACC GATTTGA	105	560
4804-7mut	ZmGWD mega-2R	反向	CTATTCCTTGCTCG GACTGAC	13	560
4817	2856	正向	GAAGGGGATTGGA GAGGAAG	111	613
4817	2858	反向	CATGACGTTCAAAT AGCCTCA	112	613
4837-4839	371	正向	GGTTATAAGCCCGG TTGAAGTA	12	381
4837-4839	429	反向	GCAGAAGTAGGCT TGAAGGAA	113	381

[0346] 与之前的分析类似,使用Vector NTI Advance (版本11.5;Life Technologies)将每个突变体的DNA序列与WT GWD进行比较。突变体DNA序列描述于表10中。

[0347] 已经产生了携带靶向作用于GWD基因区域的基因编辑构建体的转基因玉米植物,并且正在分析GWD基因的靶区域中的突变。GWD突变和预测的蛋白质列于表11-14中。在这些表中,携带两个不同GWD突变等位基因(杂合子)的事件用-1或-2标记,以表示单个等位基因。内含子序列以小写字母表示,外显子24或25用大写字母表示。因为pAG4817靶向作用于ZmGWD内的两个不同位置,所以为4817_2和4817_52提供了两个突变。第一靶序列位于外显子24的5'末端的上游,插入至该靶位点的额外的“T”显示为在内含子23特异的小写字母内的大写字母“T”(参见M37序列)。M37是4817_2和4817_52中的相同修饰。

[0348] 由CRISPR/Cas9引入的所有修饰被突出显示(粗黑色=插入;灰色=缺失),并且缺失的核苷酸由点显示。删除或插入的核苷酸的相应数目呈现在表11和12的最后一列中。

[0349] 类似地,突出显示M32-M39的推导的蛋白质序列的所有改变。在翻译阅读框架移位和翻译早期终止的情况下,与野生型GWD不同的所有氨基酸也被突出显示,蛋白质的末端用

星号(*)表示。

[0350] 表10 CRISPR/Cas9在单个转基因4804、4806和4817事件中诱导的突变

[0351]

突变	事件和等位基因
M32	4804_2,4804_3-2,4804_4-1,4804_5-2,4806_1
M33	4804_3-1,4804_5-1,4804_7-1
M34	4804_4-2
M35	4804_6
M36	4804_7-2
M37,M38	4817_2
M37,M39	4817_52

[0352] 表11单个4804和4806事件中CRISPR/Cas9诱导的突变的核苷酸序列

[0353]

序列描述	DNA 序列	SEQ ID NO	缺失/插入数
野生型 ZmGWD 外显子 24*	gctcctgctatcagTTGGCAGGTTATAAGCCCGGT TGAAGTATCAGGTTATGTGGTTGTGGTTGATGAGT TACTTG	186	无
M32	gctcctgctatcagTTGGCAGGTTATAAGCCCGGT TTGAAGTATCAGGTTATGTGGTTGTGGTTGATGAG TACTTG	114	+1
M33	gctcctgctatcagTTGGCAGGTTATAAGCCCGGT TGAAGTATCAGGTTATGTGGTTGTGGTTGATGAGT TACTTG	115	-2
M34	gctcctgctatcagTTGGCAGGTTATAAGCCCGGT TGAAGTATCAGGTTATGTGGTTGTGGTTGATGAGT TACTTG	116	-3
M35	gctcctgctatcagTTGGCAGGTTATAAGCCCGGT TGAAGTATCAGGTTATGTGGTTGTGGTTGATGAGT TACTTG	117	-1
M36	gctcctgctatcagTTGGCAGGTTATAAGCCCGGT TGAAGTATCAGGTTATGTGGTTGTGGTTGATGAGT TACTTG	118	-37

[0354] *野生型ZmGWD是外显子24的81-160nt的区域 (SEQ ID NO:3)

[0355] 表12单个4817事件中CRISPR/Cas9诱导突变的核苷酸序列

[0356]

序列描述	DNA 序列	SEQ ID NO	缺失/插入数
野生型 ZmGWD 外显子 24*	gctcctgctatcagttggcaggttataagcccggttgaagtatcaggttatgctggttctggttgatgattacttg	187	无
M37	gctcctgctatcagttggcaggttataagcccggttgaagtatcaggttatgctggttctggttgatgattacttg	188	+1
野生型 ZmGWD 外显子 25	cactctatctgaaacttgaaggatattgatcagaaactgTTTTCTTCAAGCCTACTTCTGCAGATATA	189	无
M38	cactctatctgaaacttgaaggatattgatcagaaactgTTTTCTTCAAGCCTACTTCTGCAGATATA	119	-48
M39	cactctatctgaaacttgaaggatattgatcagaaactgTTTTCTTCAAGCCTACTTCTGCAGATATA	120	+1

[0357] *野生型ZmGWD是外显子24的81-160nt的区域 (SEQ ID NO:3)

[0358] 表13在CRISPR/Cas9突变体4804和4806中外显子24的部分推导的蛋白质序列

[0359]

序列描述	蛋白序列	SEQ ID NO
野生型 ZmGWD 外显子	WQVISPVEVSGYVVVVDPELLAVQNKSYDKPTILVAKSVRGEEREIPDG	190

[0360]

24*		
M32	WQVISPV	121
M33	WQVISPVSIRICGGG*	122
M34	WQVISPVWVSGYVVVVDPELLAVQNKSYDKPTILVARSVEGEEREIPDG	123
M35	WQVISPVKYQVMNLWLMSYLLSRTNLMINQPSLWQVSPERKKYQME*	124
M36	NLWLMSYLLSRTNLMINQPSLWQVSPERKKYQME*	125

[0361] *野生型ZmGWD外显子24是野生型ZmGWD (SEQ ID NO:43) 的1011-1057氨基酸 (aa) 的区域

[0362] 表14在CRISPR/Cas9突变体4817中外显子25的部分推导的蛋白质序列

[0363]

序列描述	蛋白序列	SEQ ID NO
野生型 ZmGWD 外显子 25*	VLFATCFDHTTTLSELEGYDQKLFSEFKPTSADITYR	191
M38	VLFATCFDHTTTLSELEGYDQKLFSEFKPTSADITYR	126
M39	VLFATCFDHTTTLSELEGYDQKLFSEFKPTSADITYRNL*	127

[0364] **野生型ZmGWD外显子25是野生型ZmGWD (SEQ ID NO:43) 的1082-1116aa的区域

[0365] 玉米NLS大范围核酸酶诱导的突变的表征

[0366] 为了构建玉米NLS大范围核酸酶构建体pAG4837-4839,使用来自Opaque2 (Hicks

et al., PNAS, 1995) (表15) 的玉米NLS序列替换pAG4716中的病毒SV40NLS序列。观察到玉米GWD基因的外显子24中诱导的突变的大的变异。这些突变包括1至114个核苷酸的取代、缺失和插入(表16-17)。NLS变体的间接评估的效率被估计为含有GWD基因的靶区域中的任何修饰的事件数除以分析的事件的总数(表15)。每个评估的NLS序列支持诱导突变的产生,其中NLS3和NLS4是最有效的。

[0367] 表15含有植物来源的NLS序列的大范围核酸酶构建体

[0368]

构建体	表达盒	NLS号	蛋白序列	SEQ ID NO	相对效率 (%)
-----	-----	------	------	-----------	----------

[0369]

pAG4837	ZmUbi1P:NLS1: GWD7-8x.226	NLS1	MPTEERVRRKRKES NRESARRSRYRKA AHLKEL	128	59.1
pAG4838	ZmUbi1P:NLS3: GWD7-8x.226	NLS3	MARKRKESNRESA RRSRYRKA AHLKEL	129	75.0
pAG4839	ZmUbi1P:NLS4: GWD7-8x.226	NLS4	MARKRKESNRESA RRSRRSRYRKY	130	71.4

[0370] 表16在4837、4838和4839事件中ZmGWD基因的外显子24中的玉米NLS大范围核酸酶诱导的代表性突变的列表

[0371]

序列描述	突变	DNA 序列	SEQ ID NO	缺失/插入
野生型 ZmGWD*	无	GAAATACCAGATGGAGTAGTTGGTGTAAATTA CACCTGATATGCCAGATGTTCTGTCT	192	无
4837_12	M40	GAAATACCAGATGGAGTAGTTG...TAAATTA CACCTGATATGCCAGATGTTCTGTCT	131	-3
4837_12	M41	GAAATACCAGATGGAGTAGTTGGTATAAATTA ACACCTGATATGCCAGATGTTCTGTCT	132	-2/+3
4837_16	M42	GAAATACCAGATGGAGTAGTTGGTGTAAATTA CACCTGATATGCCAGATGTTCTGTCT	133	-1
4837_19	M43	GAAATACCAGATGGAGTAGTTGGTGTAGAGT ATAACACCTGATATGCCAGATGTTCTGTCT	134	-3/+8
4837_53	M44	GAAATACCAGATGGAGTAGTTGGTGT... ...TCTGTCT	135	-24
4838_1	M45	GAAATACCAGATG... ...TCTGTCT	136	-36
4838_51	M46	GAAATACCAGATGGAGTAGTTGGTGTATGAA CACGTAATTACACCTGATATGCCAGATGTTCT GTCT	137	+10
4838_53	M47	GAAATACCAGATGGAGTAGTTGGTGT... ...CT	138	-29
4839_1	M48	GAAATACCAGATGGAGTAGTTGGTGT...TTA CACCTGATATGCCAGATGTTCTGTCT	139	-3
4839_3	M49	GAAATACCAGATGGAGTAGTTGGTGTAAATTA ACACCTGATATGCCAGATGTTCTGTCT	140	+1
4839_54	M50	GAAATACCAGATGC... ...GATATGCCAGATGTTCTGTCT	141	-22
4839_57	M51	GAAATACCAGATGGAGTAGTTGGTGTCTCAT GCCAGATGTCAAGCAATTACACCTGATATGC CAGATGTTCTGTCT	142	+19
4839_58	M52	GAAATACCAGATGGAGTAGTTGGTGT... ...ATGTTCTGTCT	143	-21
4839_58	M53	GAAATACCAGATGGAGTAGTTGGTGT... ...CAGATATGCCAGATGTTCTGTCT	144	-9/+2
4839_61	M54	GAAATACCAGATGGAGTAGTTGGTGTCAATTA CTCATATTTTCTGTGATTGAATATTTCTTTC CAGATGGAGTGTCAAGGGAGAGGAAGAATA CCAGATGGAGTGTCAAGGGAGAGGAAGAAT ACCAGATGAAGGAAATACACCTGATATGCCA GATGTTCTGTCT	145	-4/+114
4839_61	M55	GAAATACCAGATGGAGT...TA CACCTGATATGCCAGATGTTCTGTCT	146	-12

[0372] *野生型ZmGWD是SEQ ID NO:1的3157-3213nt的区域。

[0373] 表17在CRISPR/Cas9突变体4837、4838和4839中外显子24的部分推导的蛋白质序列

[0374]

序列描述	蛋白序列	SEQ ID NO
野生型 ZmGWD 外显子 24*	IPDGVVGVITPDMPDVL SHVSVRARN SK	193
M40	IPDGVVGVITPDMPDVL SHVSVRARN SK	147
M41	IPDGVVGINYT*	148
M42	IPDGVVGV LHLICQMFCLMCQSEQGIARYCL RPVLTTELYLNLKDMIRNCFPSLLL*	149
M43	IPDGVVGV E*	150
M44	IPDGVVGV.....LSHVSVRARN SK	151
M45	IPD.....VLSHVSVRARN SK	152
M46	IPDGVVGV*	153
M47	IPDGVVGV SCVSPSKE*	154
M48	IPDGVVGVITPDMPDVL SHVSVRARN SK	155
M49	IPDGVVGVNYT*	156
M50	IPDCLCQMFCLMFQSEQGIARYCLRPVLTTE LYLNLKDMIRNCFPSLLLQI*	157
M51	IPDGVVGVSCQN*	158
M52	IPDGVVGV.....DVL SHVSVRARN SK	159
M53	IPDGVVGVRYARCSVSPSKE*	160
M54	IPDGVVGAPTHIFCD*	161
M55	IPDGVV.....ITPDMPDVL SHVSVRARN SK	162

[0375] *野生型ZmGWD外显子24是SEQ ID NO:43的1054-1081aa的区域。

[0376] 绿色组织淀粉的测定

[0377] 测定CRISPR/Cas和玉米NLS大范围核酸酶品系的绿叶组织中的淀粉。根据实施例3中所述的方案测定从40天龄事件收获的叶组织的淀粉。这些数据证实了我们的CRISPR/Cas靶向系统和玉米NLS大范围核酸酶基因编辑构建体的效果。

[0378] 表18 CRISPR/Cas和玉米NLS大范围核酸酶品系中的淀粉含量

[0379]

载体	事件	淀粉 (mg 葡萄糖 /100mg DW) [% DW]	SD
WT	WT	1.0	1.4
4804	58	1.1	0.1
4804	60	1.2	0.1
4804	59	1.3	0.2
4804	61	2.4	0.2
4804	53	9.3	1.9
4804	6	9.6	0.3
4804	4	11.2	1.4
4804	54	11.9	0.5
4804	51	12.4	0.9
4804	57	14.1	1.8
4804	62	16.7	0.4
4804	56	16.7	0.7
4804	7	17.0	0.6
4804	63	17.9	0.4
4804	5	18.0	0.2
4804	52	18.2	0.8
4804	3	18.5	0.9
4804	64	19.0	0.7
4804	2	19.5	1.2
4804	55	20.5	0.9
4804	1	27.8	1.7
4805	1	1.2	0.3
4805	103	11.2	0.7
4805	53	11.5	1.0
4805	56	11.7	0.7
4805	55	13.8	2.3
4805	101	14.2	0.4
4805	104	15.0	1.2
4805	51	15.0	1.2
4805	54	15.6	1.0
4805	2	16.4	0.7

载体	事件	淀粉 (mg 葡萄糖 /100mg DW) [% DW]	SD
4806	2	1.2	0.1
4806	52	1.2	0.6
4806	101	1.4	0.2
4806	53	1.7	0.2
4806	154	2.0	0.3
4806	57	2.2	0.7
4806	152	2.7	0.9
4806	1	3.3	0.3
4806	205	4.0	1.0
4806	207	4.6	0.6
4806	209	5.9	1.0
4806	210	7.1	0.9
4806	208	7.4	0.6
4806	155	7.5	1.0
4806	55	8.6	1.4
4806	56	9.4	1.8
4806	206	10.1	1.4
4806	54	11.0	0.9
4806	151	15.7	1.4
4807	1	0.7	0.0
4807	6	0.7	0.1
4807	5	0.7	0.1
4807	2	0.7	0.0
4807	4	1.0	0.1
4807	3	2.1	0.4
4809	1	10.1	0.8
4809	2	10.6	0.5
4809	4	13.3	0.7
4809	3	17.8	1.3
4817	54	14.2	0.7
4817	55	18.0	1.1
4817	51	19.2	0.8

[0380]

载体	事件	淀粉 (mg 葡萄糖 /100mg DW) [% DW]	SD	载体	事件	淀粉 (mg 葡萄糖 /100mg DW) [% DW]	SD
4806	204	0.9	0.1	4817	1	19.4	0.8
4806	203	0.9	0.1	4817	53	20.7	0.5
4806	202	0.9	0.1	4817	52	20.9	0.8
4806	201	1.0	0.1				

[0381] 表19 CRISPR/Cas和玉米NLS大范围核酸酶4837、4838和4839品系中的淀粉含量

[0382]

载体	事件	淀粉 (mg 葡萄糖/100mg DW) [% DW]	SD
WT	1	1.0	0.2
4837	8	1.3	0.2
4837	51	1.8	0.5
4837	7	1.8	0.3
4837	16	2.2	0.5
4837	4	2.3	0.2
4837	1	2.3	0.2
4837	15	2.5	0.1
4837	53	3.3	0.5
4837	9	3.6	0.5
4837	3	4.7	1.0
4837	12	6.2	1.4
4837	17	16.5	0.5
4837	11	17.0	0.7
4837	14	18.2	0.6
4837	2	18.6	1.1
4837	5	19.3	1.3
4837	19	20.0	1.1
4837	18	20.4	0.4
4837	10	22.7	2.1
4837	6	23.3	0.9
4837	52	24.7	1.4
4838	4	1.2	0.1
4838	3	1.7	0.4
4838	53	1.8	0.7
4838	2	3.3	0.4
4838	54	6.9	1.2

[0383]

载体	事件	淀粉 (mg 葡萄糖/100mg DW) [% DW]	SD
4838	51	19.3	1.4
4838	52	19.9	1.2
4838	1	20.0	1.4
4839	53	1.1	0.1
4839	63	1.1	0.1
4839	62	1.3	0.1
4839	52	1.4	0.2
4839	54	2.1	0.4
4839	56	2.3	0.2
4839	51	2.4	1.1
4839	57	2.7	0.3
4839	58	2.8	0.8
4839	55	22.8	1.7
4839	61	23.5	1.1
4839	60	24.5	1.1
4839	59	28.0	3.3

[0384] 参考表18和19,观察到许多事件表现出高淀粉含量,其范围为约3%-27.8%。

[0385] 实施例10.绿色组织淀粉的测定

[0386] 测定所有CRISPR/Cas细胞系在绿叶组织以及干燥的秸秆叶、茎和穗轴中的淀粉。根据实施例3描述的方案测定从40天龄的CRISPR/Cas事件收获的叶组织的淀粉。图9说明了pAG4804玉米事件中的淀粉积累。如图9所示,所有7个T0玉米4804事件具有高淀粉含量,其范围为9.6-27.8%,这表明所有当前未解析的GWD序列是两个不同GWD突变的结果,而不是一个野生型和一个突变体等位基因(半合子)。图10说明了pAG4806玉米事件中的淀粉积累。如图10所示,T0玉米4806事件都具有低淀粉含量,这表明一个未解析的GWD序列是半合子。这些数据证实了本文的CRISPR/Cas靶向系统的功效,其包括新的GWD引导RNA靶向序列和用于表达引导RNA的新U3启动子。

[0387] 实施例11.用于优良近交基因渗入和测试的隐性突变的育种

[0388] 用于精确DNA基因工程和诱变的新方法的出现提供了产生用于研发新的和具有有益植物性状的靶向隐性突变和显性突变的手段。这些方法中的一些包括用大范围核酸酶、Talens和CRISPR/Cas系统靶向基因的特定区域。跟踪和推进靶向突变对性状研发提出了新的挑战,因为与传统的转基因植物性状不同,它们不携带显性T-DNA表达盒和选择标记物。使用本文所述的TALENs、ZFN、大范围核酸酶和CRISPR/Cas方法产生玉米和高粱中的靶向突变可以导致产生用于筛选和育种这些独特植物性状的新方法。

[0389] 实施例12.跟踪和育种靶向突变

[0390] 转化(T0)生成基因型:使用所述方法鉴定的基因特异性突变导致具有三种不同基因型中的一种的第一代转化(T0)植物:1)一个野生型基因等位基因和一个突变等位基因,2)两个不同的突变等位基因,或3)两个相同的突变等位基因。这些突变型等位基因组合分别被称为半合子、杂合子和纯合子。

[0391] 跟踪靶向突变的分子方法:相对于野生型序列,靶向DNA突变(例如,取代、缺失、插

入或组合)的特定序列的特征是,当将突变育种至其他品系时,或通过育种扩大现有品系时,可以使用本文所述的方法跟踪。为了区分T0和T1+(来源于T0亲本品系及其以后的子代)分离群体中的野生型、半合子、杂合子和纯合子,可以使用至少五种方法。这些方法包括:1)使用凝胶电泳对突变位点进行PCR以进行大小分离,2)使用限制酶消化和凝胶电泳进行PCR以产生突变特异性限制模式,3)使用直接测序进行PCR,4)使用克隆和测序进行PCR,和5)使用结合或不结合突变位点的引物进行PCR。

[0392] 来自野生型或工程的、改变的或优化的内源核酸的纯合DNA序列将容易通过靶向或突变区域中的PCR、大小测定和/或DNA测序来分析。相反,具有不同等位基因的DNA序列可能导致由于两个等位基因序列(例如,野生型和突变体或两个不同的突变体)的差异而难以使用测序、PCR或大小测定来分析的序列的一部分序列。来自这些类型的靶向事件的PCR产物将需要进行克隆以分离和有效地对每个等位基因进行测序。

[0393] 一旦已经确认了靶向突变的序列,就建立了用于跟踪的突变特异性分子策略。如前所述,该策略将取决于突变的特征。在跟踪本文中产生的gwd突变时,针对本文所述的每种突变和工程内源(优化的)核酸研发PCR特异性反应。

[0394] 育种杂交和自交:使用转基因诱导的靶向突变产生的性状的主要目标是通过从转基因中分离所需的突变来分离突变。这可以通过遗传杂交来实现,并且通过异交(更高频率的回收转基因阴性的突变植物)是最有效的,其涉及将T0花粉交叉到非转基因植物的雌性组分。这也可以使用自交以较低的频率完成,其包括自花授粉(使用T0花粉授粉同一T0植株)。近缘授粉(sibbing)是另一种选择,并且涉及两个基因相同的植物之间的杂交,这将预期与自交的结果相同。

[0395] 携带靶向突变的T0植物可以自交以产生纯合植物,并且根据T0接合性将导致不同数量和类型的子代。纯合T0植物将产生100%纯合子代,半合子T0植物对于突变等位基因以1:2:1(纯合:半合子:野生型)分离,杂合T0植物将以1:2:1(纯合靶向等位基因1:杂合:纯合靶向等位基因2)。

[0396] 携带靶向突变的T0植物也可以异交到其他品系中,并且根据T0接合性将导致不同数量和类型的子代。纯合T0植物将产生100%半合子代,半合子T0植物将产生50%半合子代和50%野生型子代,杂合T0植物将产生半合子代,50%具有靶向等位基因和50%具有靶向等位基因2。

[0397] 因为所有T0植物都携带转基因,转基因插入位置是最常见的不同,并且转基因插入的数目可以不同,所以转基因的分离模式在每个T0转化事件/植物之间具有显著变化的可能性。为了鉴定转基因阴性植物,可以将PCR应用于与T0植物进行任意杂交(自交、近缘或异交)的子代。转基因阴性植物将通过不存在转基因特异性PCR产物来鉴定。然后使用在T0植物的初始表征期间定义的分子诊断方法筛选转基因阴性植物的靶向突变。

[0398] 然后可以保持和培育从转基因分离的靶向突变,用于测试和基因渗入(introgressions),继续使用性状特异性分子诊断方案。

[0399] 用于靶向突变的跟踪和育种程序的实例:本文描述了用于靶向突变的分子跟踪和育种过程。本文描述了来自玉米(M20,来自事件4716_164的ZmGWD_M20)的GWD基因的跟踪。该部分内容和其他大范围核酸酶诱导的玉米和高粱中的靶向突变的产生和初始序列表征已经在本文的实施例中描述。T0 4716_164植物最初对M20突变是半合的,并携带未知数量

的T-DNA插入。M20突变是在ZmGWD的外显子24中的隐性单碱基对(bp)缺失,其导致在野生型序列中的MluCI限制酶位点处的突变。这种小尺寸缺失需要使用具有凝胶电泳的MluCI RFLP,因为其不能仅通过凝胶电泳与野生型区分。

[0400] 图11显示了来自于玉米事件4716_164的靶向突变M20的自交和异交的示意图。如图11所示,T04716_164植物自交和异交,分别产生用于功效测试和基因渗入的子代。为了鉴定来自T0自交亲本的纯合M20子代,使用大范围核酸酶靶向基因和GWD基因的靶区域进行PCR。图12显示了来自自交T04716_164 M20植物的T1子代的基因分型。如图12所示,来自大范围核酸酶GOI和ZmGWD靶区域的PCR产物使用MluCI消化,在5%聚丙烯酰胺上分离并使用溴化乙锭染色。这显示了不携带T-DNA(大范围核酸酶)但对于M20是纯合的植物。维持这些植物用于测试。

[0401] 为了鉴定来自T0异型亲本的半合子M20子代,进行选择性标记基因、PMI、大范围核酸酶基因和靶区域的PCR,然后进行3%琼脂糖凝胶电泳和溴化乙锭染色。这也允许鉴定T-DNA阴性植物。图13显示了来自异型T04716_164 M20植物的T1子代的基因分型。然后对来自T-DNA阴性植物的相同PCR产物进行MluCI限制性消化,并在使用溴化乙锭染色的5%聚丙烯酰胺凝胶上将它们分离。图14说明来自异型4716_164 M20植物的T1子代的基因分型。维持这些植物用于进一步基因渗入和将来的测试。

[0402] 参考文献

[0403] An,G et al.,2005.Reverse genetic approaches for functional genomics of rice.Plant molecular biology,59(1),pp.111-23.在<http://www.ncbi.nlm.nih.gov/pubmed/16217606>[Accessed July 25,2012]可以获得。Arnould S,Perez C,Cabaniols JP,Smith J,Gouble A,Grizot S,Epinat JC,Duclert A,Duchateau P,Paques F.(2007)Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells.J.Mol.Biol.371:49-65.

[0404] Arnould S,Delenda C,Grizot S,Desseaux C,Paques F,Silva GH,Smith J.(2011)The I-CreI meganuclease and its engineered derivatives:Applications from cell modification to gene therapy.Protein Eng.Des.Sel.24:27-31.

[0405] Belhaj K,Chaparro-Garcia A,Kamoun S,Nekrasov V.(2013)Plant genome editing made easy:targeted mutagenesis in model and crop plants using the CRISPR/Cas system.Plant Methods 9:39-48.

[0406] Boch J.&Bonas U.(2010)Xanthomonas AvrBs3family-type III effectors:discovery and function.Annu.Rev.Phytopathol.48:419-436.

[0407] Cermak T,Doyle EL,Christian M,Wang L,Zhang Y,Schmidt C,Baller JA,Somia NV,Bogdanove AJ,Voytas DF.(2011)Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting.Nucleic Acids Res.39:e82

[0408] Chi-Ham CL et al.,2010.The intellectual property landscape for gene suppression technologies in plants.Nature Biotechnology,28(1):32-36.

[0409] Christian,M et al.,2010.Targeting DNA double-strand breaks with TAL

effector nucleases. *Genetics*, 186 (2), pp. 757-61. 在 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2942870&tool=pmcentre&rendertype=abstract> [Accessed July 14, 2012] 可以获得。

[0410] Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819-823.

[0411] Djukanovic V, Smith J, Lowe K, Yang M, Gao H, Jones S, Nicholson MG, West A, Lape J, Bidney D, Falco SC, Jantz D, Lyznik LA. (2013) Male-sterile maize plants produced by targeted mutagenesis of the cytochrome P450-like gene (MS26) using a re-designed I-CreI homing endonuclease. *The Plant Journal* 76:888-899.

[0412] Elkonin LA, Pakhaomova NV. (2000) Influence of nitrogen and phosphorus on induction embryogenic callus of sorghum. *Plant Cell Tissue and Organ Culture* 61:115-123.

[0413] Frizzi A & Huang S, 2010. Tapping RNA silencing pathways for plant biotechnology. *Plant biotechnology journal*, 8 (6), pp. 655-77. 在 <http://www.ncbi.nlm.nih.gov/pubmed/20331529> [Accessed July 24, 2012] 可以获得。Gao Z, Xie X, Ling Y, Muthukrishnan S, Liang GH. (2005) *Agrobacterium tumefaciens*-mediated sorghum transformation using a mannose selection. *Plant biotechnology journal*, 3, pp. 591-599.

[0414] Garcia-Bustos J, Heitman J, Hall MN (1991) Nuclear protein localization. *Biochim Biophys Acta* 1071:83-101.

[0415] Gasiunas G, Barrangou R, Horvath P, Siksnys V. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci US A*. 109 (39):2579-2586.

[0416] Goujon M et al., 2010 A new bioinformatics analysis tools framework at EMBL-EBI (2010) *Nucleic acids research* Jul, 38 Suppl:W695-9 doi:10.1093/nar/gkq313 Heath PJ, Stephens KM, Monnat RJ Jr., Stoddard BL. (1997) The structure of I-CreI, a group I intron-encoded homing endonuclease. *Nat. Struct. Biol.* 4:468-76.

[0417] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337:816-821.

[0418] Joung JK & Sander JD. (2013) TALENs: a widely applicable technology for targeted genome editing. *Nature Reviews (Mol Cell Biol)* 14:49-55.

[0419] Kalderon D, Richardson WD, Markham AF, Smith AE (1984) Sequence requirements for nuclear location of simian virus 40 large T antigen. *Nature* 311:33-38.

[0420] Larkin MA et al., 2007 ClustalW and ClustalX version 2 *Bioinformatics*, 23 (21):2947-2948. doi:10.1093/bioinformatics/btm404

[0421] Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS. (2013) CRISPR

interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.* 8(11):2180-2196.

[0422] Leader DJ, Connelly S, Filipowicz W, Brown JWS. (1994) Characterisation and expression of a maize U3 snRNA gene. *Biochimica et Biophysica Acta* 1219:145-147.

[0423] Liang Z, Zhang K, Chen K, Gao C. (2014) Targeted mutagenesis in *Zea mays* using TALENs and the CRISPR/Cas System. *Journal of Genetics and Genomics* 41:63-68. Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, Yang B. (2011) TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.* 39:359-372.

[0424] Maniatis T, Fritsch EF and J. Sambrook J, 1982. *Molecular Cloning* Cold Spring Harbor Laboratory.

[0425] Puchta H, Dujon B and Hohn B, 1993. Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic acids research*, 21(22), pp.5034-40. 在 <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=310614&tool=pmcentrez&rendertype=abstract> 可以获得。

[0426] Raikhel NV (1992) Nuclear targeting in plants. *Plant Physiol* 100:1627-1632.

[0427] Rosen LE et al., (2006) Homing endonuclease I-CreI derivatives with novel DNA target specificities. *Nucleic Acids Res.* 34:4791-4800.

[0428] Shieh MW et al., (1993) Nuclear targeting of the maize R protein requires two nuclear localization sequences. *Plant Physiology* 101:353-361.

[0429] Shukla VK et al., (2009) Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature* 459:437-441.

[0430] Sikora P et al., 2011. Mutagenesis as a tool in plant genetics, functional genomics, and breeding. *International journal of plant genomics*, 2011, p.314829. 在 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3270407&tool=pmcentre_z&rendertype=abstract [Accessed July 25, 2012] 可以获得。

[0431] Smith AM and Zeeman SC, 2006. Quantification of starch in plant tissues. *Nat. Protocols* 1:1342-1345.

[0432] Smith TF, Waterman MS, 1981. Identification of Common Molecular Subsequences. *J Mol Biol* 147:195-197.

[0433] Symington LS, Gautier J. (2011) Double-Strand Break End Resection and Repair Pathway Choice. *Annual Review of Genetics.* 45:247-271.

[0434] Till BJ et al., 2007 Discovery of chemically induced mutations in rice by TILLING. *BMC Plant Biol.* 7:19.

[0435] Upadhyay SK et al., (2014) RNA-guided genome editing for target gene mutations in wheat. *Genes, Genomes, Genetics* 3:2233-2238.

[0436] Vainstein A et al.,2011.Permanent genome modifications in plant cells by transient viral vectors.Trends in biotechnology,29(8),pp.363-9.在<http://www.ncbi.nlm.nih.gov/pubmed/21536337>[Accessed July 30,2012]可以获得。

[0437] Varagona MJ et al.,(1992)Nuclear localization signal(s)required for nuclear targeting of the maize regulatory protein Opaque-2.The Plant Cell 4: 1213-1227.

[0438] Wagner P et al.,1990)Active transport of proteins into the nucleus.FEBS 275:1-5.

[0439] Wehrkamp-Richter S et al.,2009.Characterisation of a new reporter system allowing high throughput in planta screening for recombination events before and after controlled DNA double strand break induction.Plant physiology and biochemistry:PPB/Société *française* de physiologie végétale,47(4),pp.248-55.在<http://www.ncbi.nlm.nih.gov/pubmed/19136269>[Accessed July 17, 2012]可以获得。Weise,S.E.et al.,(2012)Engineering starch accumulation by manipulation of phosphate metabolism of starch.P.Biotech.J.10,545-554.

[0440] Wright D et al.,2005.High-frequency homologous recombination in plants mediated by zinc-finger nucleases.The Plant journal:for cell and molecular biology,44(4),pp.693-705.在<http://www.ncbi.nlm.nih.gov/pubmed/16262717>[Accessed July 19,2012].Yon J,Fried M.(1989)Precise gene fusion by PCR.Nucleic Acids Res.17(12):4895.可以获得。

[0441] Larkin MA et al.,2007ClustalW and ClustalX version 2 Bioinformatics,23(21):2947-2948.doi:10.1093/bioinformatics/btm404

[0442] Goujon M et al.,2010A new bioinformatics analysis tools framework at EMBL-EBI(2010)Nucleic acids research Jul,38 Suppl:W695-9doi:10.1093/nar/gkq313

[0443] 在本申请通篇中引用的参考文献均出于本文显而易见的目的而被纳入本文,如同将它们全文纳入本文一样。出于说明的目的,这些参考文献的特定部分被引用到本文特定的位置。在特定位置的参考文献的引用表明将参考文献中教导的方法纳入本文。然而,在特定位置的参考文献的引用,不限制将所有引用的参考文献教导的方法,出于任何目的而纳入本文。

[0444] 因此,应当理解的是,本发明不限于具体公开的实施方式,旨在涵盖如从属权利要求、如上的说明和/或如附图所示所定义的,在本发明的精神和范围内的所有变型。

SEQUENCE LISTING

<110> 谷万达公司
 <120> 表达修饰的葡聚糖水合二激酶的植物
 <130> AGR-PT024.1W0
 <150> US 62/056,852
 <151> 2014-09-29
 <160> 196
 <170> PatentIn version 3.5
 <210> 1
 <211> 4416
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (4416)
 <223> ZmGWD coding sequence
 <400> 1

```

atgtccggat tcagtgccgc ggccaacgca gcggcggtctg agcgggtgcgc gctcgcgttc 60
cgcgcacggc ccgcggcctc ctgccagcg aagcggcagc agcagccgca gccagcgtcc 120
ctccgacgca gcggggggcca gcgccgcccc acgacgtctt ccgctctag ccgcgcccc 180
gtcgtgccgc gcgccgtcgc cacgtccgcg gaccgcgcgt ccccgacct tatcgaaag 240
ttcacgctgg attccaactc cgagctccag gtcgcagtga acccagcgc gcagggtttg 300
gtgtcagaga ttagcctgga ggtgaccaac acaagcggtt ccttgatttt gcattgggga 360
gcccttcgcc cggacaagag agattggatc ctcccgcca gaaaacctga tggaacgaca 420
gtgtacaaga acagggctct caggacacct ttgttaaagt caggtgataa ctccactcta 480
aggattgaga tagatgatcc tgggggtcac gccattgagt tcctcatctt tgacgagaca 540
cagaacaaat ggtttaaaaa caatggccag aattttcagg ttcagttcca gtcgagccgc 600
catcagggta ctggtgcatc tggtgctcc tctctgcta cttctacctt ggtgccagag 660
gatcttgtgc agatccaagc ttacctcgg tgggaaagaa ggggaaagca gtcatacaca 720
ccagagcaag aaaaggagga gtatgaagct gcacgagctg agttaataga ggaagtaaac 780
agagggtgtt ctttagagaa gcttcgagct aaattgaca aagcacctga agcacctgag 840
tcggatgaaa gtaaactctt tgcattctga atgcccctc gtaaacttcc agaggatctt 900
gtacaggtgc agcttatat aaggtgggag caagcgggca agccaaacta tctctctgag 960
aagcaactgg tagaatttga ggaagcaagg aaggaactgc agctgaggt ggacaaggga 1020
atctctattg atcagttgag gcagaaaatt ttgaaaggaa acattgagag taaagtttcc 1080
aagcagctga agaacaagaa gtacttctct gtagaaagga ttcagcgcga aaagagagat 1140
atcacacaac ttctcagtaa acataagcat acacttgtgg aagataaagt agaggttgta 1200
ccaaaacaac caactgttct tgatctcttc accaagtctt tacatgagaa ggatggctgt 1260

```

gaagttctaa gcagaaagct cttcaagttc ggcgataaag agatactggc aatttctacc 1320
 aaggttcaaa ataaaacaga agttcacttg gcaacaaacc ataccgacce acttattctt 1380
 cactggctctt tggcaaaaaa tgctggagaa tggaaggcac cttctccaaa tatattgcca 1440
 tctggttcca cattgctgga caaggcgtgt gaaactgaat ttactaaatc tgaattggat 1500
 ggtttgcatt accaggttgt tgagatagag cttgatgatg gaggatacaa aggaatgcca 1560
 tttgttcttc ggtctggatg aacatggata aaaaataatg gttctgattt tttcctagat 1620
 ttcagcacc c atgatgtcag aaatattaag gcaattttaa agggcaatgg tgatgctggt 1680
 aaaggtactg ctaaggcatt gctggagaga atagcagatc tggaggaaga tgcccagcga 1740
 tctcttatgc acagattcaa tattgcagca gatctagctg accaagccag agatgctgga 1800
 cttttgggta ttgttgggct ttttgtttgg attagattca tggctaccag gcaactaaca 1860
 tggaataaga actataatgt gaagccacgt gagataagca aagcacagga taggtttaca 1920
 gatgatcttg agaatatgta caaagcttat ccacagtaca gagagatatt aagaatgata 1980
 atggctgctg ttggtcgctg aggtgaaggt gatgtttgctc aacgcattcg tgatgagata 2040
 ttagtaatac agagaaataa tgactgcaaa ggtggaatga tggagaatg gcaccagaaa 2100
 ttgcacaaca atacaagccc agatgatgta gtgatatgcc aggccttaat tgattatate 2160
 aagagtgact ttgatataag cgtttactgg gacacctga acaaaaatgg cataaccaa 2220
 gagcgtctct tgagctatga tcgtgctatt cattcagaac caaatttcag aagtgaacag 2280
 aaggcgggtt tactccgtga cctgggaaat tacatgagaa gcctaaaggc tgtgcattct 2340
 ggtgctgate ttgaatctgc tatagcaagt tgtatgggat acaaatcaga gggatgaaggt 2400
 ttcattggtg gtgttcagat caatccagtg aagggtttac catctggatt tccggagttg 2460
 cttgaatttg tgcttgaaca tgttgaggat aaatcagcgg aaccacttct tgaggggcta 2520
 ttggaagctc gaggttgaact gcgcccttg cttcttgatt cgcgtgaacg catgaaagat 2580
 cttatatttt tggacattgc tcttgattct accttcagga cagcaattga aaggctatat 2640
 gaggagctga atgatgcagc cccagagaaa ataatgtact tcatcagtct tgccttgaa 2700
 aatcttgcgc tttcaattga cgacaatgaa gacatcctgt attgtttaaa gggatggaac 2760
 caagccttgg aaatggctaa gcaaaaagac gaccaatggg cgctctatgc taaagcattt 2820
 cttgacagaa acagacttgc ccttgcgagc aaggggagaac aataccataa tatgatgcag 2880
 cctctgctg agtatcttgg ctcgttactc agcatagacc aatgggcagt caatatcttc 2940
 acagaagaaa ttatacgcgg tggatcagct gctactctgt ctgctcttct gaaccgattt 3000
 gatcctgttt taaggaatgt tgctcacctc ggaagttggc aggttataag cccggttgaa 3060
 gtatcagggt atgtggttgt ggttgatgag ttacttctg tccagaacaa atcttatgat 3120
 aaaccaacca tccttgtggc aaagagtgtc aaggggagagg aagaaatacc agatggagta 3180
 gttggtgtaa ttacacctga tatgccagat gttctgtctc atgtgtcagt ccgagcaagg 3240
 aatagcaagg tactgtttgc gacctgtttt gaccacacca ctctatctga acttgaagga 3300
 tatgatcaga aactgttttc cttcaagcct acctctgcag atataaccta tagggagatc 3360
 acagagagtg aacttcagca atcaagttct ccaaatgcag aagttggcca tgcagtacca 3420
 tctatttcat tggccaagaa gaaatttctt ggaaaatatg caatctcagc cgaagaattc 3480
 tctgaggaaa tggttggggc caagtctcgg aatatagcat acctcaaagg aaaagtacct 3540
 tcatgggtcg gtgtcccaac gtcagttgcg ataccatttg gcacttttga gaaggttttg 3600

tcagatgggc ttaataagga agtagcacag agcatagaga agcttaagat cagacttgcc 3660
 caagaagatt ttagtgctct aggtgaaata agaaaagtcg tccttaatct tactgtctct 3720
 atgcaattgg ttaatgagct gaaggagagg atgctaggct ctggaatgcc ctggcctggt 3780
 gatgaaggag acaagcgttg ggagcaagca tggatggcta ttaaaaaggt ttgggcatca 3840
 aaatggaacg aaagagcata ttttagcaca cgcaaggatga aacttgatca tgagtacctt 3900
 tcgatggctg ttctcgtgca agaagttgtg aatgcagatt atgcttttgt cattcatacc 3960
 acaaacctat cgtctggaga ttcttctgag atatatgctg aagtggatga agggttgcc 4020
 gagacctcgc tgggagccta tcttggtcgt gctatgagct ttgtttgcaa aaaagatgac 4080
 cttgactctc ccaagttact tggttacca agcaagccaa ttggtctctt cataaggcaa 4140
 tcaatcatct tccgttccga ctccaacggt gaggacctgg aaggttatgc tggagcagga 4200
 ttatatgata gtgtaccgat ggatgaggag gatgaggttg tacttgatta tacaactgac 4260
 cctcttatag tagaccgtgg attccgaagc tcaatctct caagcatagc acgggctggc 4320
 catgccatcg aggagctata tggttctct caggacgtcg agggagtagt gaaggatgga 4380
 aaaatctatg tagtccagac aagaccacag atgtag 4416

<210> 2

<211> 4410

<212> DNA

<213> Sorghum bicolor

<220>

<221> misc_feature

<222> (1)..(4410)

<223> SbGWD coding sequence

<400> 2

atgaccggat tcagtgccgc ggcctccgca gcagcggcgg cggagcgggtg cgcgctcgcg 60
 atccgcgcac ggcccgcggc ctctcgcga gcgaagcggc agcagcagtc ggcgtccctc 120
 agacgcagcg ggggccagcg ccgccccacc acgctcgtg cctcccgcgc cagcccagtc 180
 gtcgtgcccc gcgccatcgc cacgtccgcg gaccgcgcgt cccacgacct tgcggaag 240
 ttcacgctgg attccaactc cgagctcctg gttgcagtga acccagcgc gcagggtttg 300
 gtgtcggatga tcggcctgga ggtgaccaac acaagcggtt cctgattct gcattgggga 360
 gtccttcgcc cggacaagag agattggatc ctccatcca gacaacctga tggaaacgac 420
 gtgtacaaga acagggtctt taggacgct tttgtaaagt ctggtgataa ctctactctt 480
 agaattgaga tagatgatec tgcggtgcaa gctattgagt tctctatctt tggcgagaca 540
 cagaacaaat ggtttaaaaa caatggccag aattttcaga ttcagctcca gtcgagccgc 600
 catcagggtg atggtgcatc tgggtgctcc tctctgcta cttctacctt ggtgccagag 660
 gatcttgtgc agatccaagc ttacctcgg tgggaaagaa agggaaagca gtcatacaca 720
 ccagagcaag aaaaggagga gtatgaagc gcacgagctg agttaataga ggaattaaat 780
 agagggtttt cttagagaa gcttcgagct aaattgacaa aacacctga agcacctgag 840
 tcagatgaac gtaaactctc tgcactctga atgcccgttg ataaacttcc agaggacctt 900
 gtacagggtc aggcttatat aaggtgggag aaagcgggca agccaaatta tctctctgag 960

aagcaactgg tagaacttga ggaagcaagg aaggaactgc aggctgaggt ggacaaggga 1020
atctctattg atcaattgag gcagaaaatt ttgaaaggaa acattgagag taaagtttcc 1080
aagcagctga agaacaagaa gtacttctct gtagaaagga ttcagcgcaa aaagagagat 1140
atcatgcaac ttctcagtaa acataagcat acagttatgg aagagaaagt agaggttgca 1200
ccaaaacaac caactgttct tgatctcttc accaagtctt tacatgagaa ggatggctgt 1260
gaagttctaa gcagaaagct cttcaagttc ggtgataaag agatactggc aatttccacc 1320
aaggttcaaa ataaaacaga agttcacttg gcaacaaacc atacggagcc acttattctt 1380
cactggctct tggcaaaaaa ggctggagaa tggaaggcac ctccttcaaa tatattgcca 1440
tctggttcca aattgctaga catggcgtgt gaaactgaat ttactagatc tgaattggat 1500
ggtttgtgtt accaggttgt tgagatagag cttgatgatg gaggatacaa aggaatgcca 1560
tttgttctta ggctcgttga aacatggata aaaaataatg gttccgattt tttcctagat 1620
ttcagcacc cgtgataccag aatatattaag ttaaaggaca atggcgatgc tggtaaaggc 1680
actgctaagg cgttgettga gagaatagca gatctggagg aagatgceca gcgatctctt 1740
atgcataggt tcaatattgc agcagatcta getgacgaag ccagagatgc tggactgttg 1800
ggattgttg gactttttgt ttggattagg ttcatggcta ccaggcaact aacatggaat 1860
aagaactata atgtgaagcc acgtgagata agcaaagcac aagataggtt tacagatgat 1920
cttgagaata tgtacagaac ttatcctcag tacagagaga tactaagaat gataatggct 1980
gctgttggtc gtggagggtga aggtgacgtt ggtcaacgca ttcgtgatga gatattagta 2040
atacagagaa ataatgactg caaaggtgga atgatggaag aatggcacca gaaattgcac 2100
aacaatacaa gccagatga ttagtgata tgccaggcat taattgatta tataaaaaat 2160
gattttgata taagcgttta ctgggacacc ttgaacaaaa atggcataac caaagagcgt 2220
ctcttgagct atgatcgtgc tattcattca gaaccaaatt tcagaagtga acagaaggag 2280
ggtttactcc gtgacctggg aaattacatg agaagcctaa aggctgtgca ttctgggtgt 2340
gatcttgaat ctgctatagc aacttgtatg ggatacaaat cagagggtga aggtttcatg 2400
gttggcgttc agatcaatcc agtgaagggt ttgccatctg gatttctga gttgcttgaa 2460
tttgtgcttg accatgttga ggataaatca gcagaaccac ttcttgaggg gctattggaa 2520
gctcgagttg atctgcgcc tttgcttctt gattcgcctg aacgcatgaa agatcttata 2580
tttttgaca ttgctcttga ttctacctc aggacagcaa ttgaaaggtc atatgaggag 2640
ctcaatgatg cagccccaga gaaaataatg tacttcatca gtcttgcct tgaatatctt 2700
gcgttttcaa ttgacgacaa tgaagacatc ctgtattget taaagggatg gaaccaagcc 2760
ttggaaatgg ctaagcaaaa agacgaccaa tgggctcttt acgctaaagc atttcttgac 2820
agaatcagac ttgcccttgc gagcaaggga gaacagtacc ataatatgat gcagccctca 2880
gctgaatata ttggctcgtt actcagcata gacaaatggg cagtcaatat cttcacagaa 2940
gaaattatac gcggtggatc agctgetact ctgtccgctc ttctgaaccg atttgatcct 3000
gttctaagga acgttgctaa ccttggaaagt tggcaggtta taagcccagt tgaagtatca 3060
ggttatgtgg ttgtggttga tgagttaact getgtccaga acaaatctta tgataaacca 3120
accatccttg tggcaaagag tgcaaggga gaggaagaaa taccagatgg agtagttgg 3180
gtaattacac ctgatatgcc agatgttctg tcccatgtgt cagtccgagc aaggaatagc 3240
aaggtactgt ttgcaacctg ttttgacat accactctgt ctgaacttga aggatatgat 3300

cagaaactgc tttccttcaa gcctacttct gcagatataa cctataggga gatcacagag 3360
 agtgagcttc agcaatcaag ttctccaaat gcagaagttg gccatgcagt accatctatt 3420
 tcattggcca agaagaaatt tcttgaaaaa tatgcaatat cagctgaaga attcaccgag 3480
 gaaatggttg gggccaagtc tcggaatata gcatacctca aaggaaaagt accttcatgg 3540
 gttgggtggtc caacgtcagt tgcgatacca tttggcactt ttgagaaggt tttgtcagat 3600
 ggtcttaata aggaagtagc acaaaccata gagaagctta agatcaggct tgctcaagaa 3660
 gatttttagtg ctctaggtga aataagaaaa gccgttctta atcttactgc tcctatgcaa 3720
 ttggttaatg agctgaagga gaggatgcta ggctctggaa tgccctggcc tggatgatgaa 3780
 ggcaacaggc gctgggagca agcatggatg gctattaaaa aggtttgggc atcaaaatgg 3840
 aatgaaagag catattttag cacacgcaag gtgaaactca atcatgagta cttttcgatg 3900
 gctgttcttg tgcaagaagt tgtgaatgca gattatgctt ttgtcattca tactacaaac 3960
 ccatcgtctg gagattcttc tgagatataat gctgaagtcg tgaaagggct cggagagact 4020
 ctctgtgggag cctatcctgg tcgtgctatg agctttgttt gcaaaaaaga tgaccttgac 4080
 tctccaagt tacttggta cccgagcaag ccaattggtc tcttcataag gcgatcgatc 4140
 atctttcgtt ctgactccaa cggcgaggat ctggaaggtt atgccggagc aggattatat 4200
 gatagtgtac cgatggatga ggaggatgaa gtcgtacttg attacacaac tgacctctt 4260
 atagtagatc gtggattccg aaattcaata ctctcaagca tcgcacgggc tggccatgcc 4320
 attgaagage tatatggttc tcctcaggac gtcgagggtg tagtgaagga tggaaaaatc 4380
 tatgtagtcc agacaagacc acagatgtag 4410

<210> 3

<211> 392

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (392)

<223> ZmGWD Exon 24 _introns

<400> 3

aagtgatact agtgaccctc tccacaattt tatggaacc acagaaatta ataatatatt 60
 ctattactct gcacctgaca tctggtctct gctatcagtt ggcaggttat aagccccggtt 120
 gaagtatcag gttatgtggt tgtggttgat gagttacttg ctgtccagaa caaatcttat 180
 gataaaccaa ccatccttgt ggcaaagagt gtcaagggag aggaagaaat accagatgga 240
 gtagttggtg taattacacc tgatatgcca gatgttctgt ctcatgtgtc agtccgagca 300
 aggaatagca aggtttatct tcacagetat gttgcaagat ttcttgaatt ttttctcttg 360
 tattgatggt gacatactag ctttttccta at 392

<210> 4

<211> 397

<212> DNA

<213> Sorghum bicolor

<220>
<221> misc_feature
<222> (1)..(397)
<223> SbGWD Exon 24-introns
<400> 4
aagtggact agtgacctt ccacagtttt atgtgaacca cagaaattaa atatgataat 60
atattctatt actctgcacc tgacatctgg ctctgataa cagttggcag gttataagcc 120
cagttgaagt atcaggttat gtggttggg ttgatgagtt acttgctgtc cagaacaaat 180
cttatgataa accaaccatc cttgtggcaa agagtgtaa gggagaggaa gaaataccag 240
atggagtagt tggtgtaatt acacctgata tgccagatgt tetgtcccat gtgtcagtcc 300
gagcaaggaa tagcaaggtt tattttcaca gttatgttgc aagctttctc agatTTTTTT 360
tcttgatcag atgttgacat accagTTTTT tecta 397
<210> 5
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, Mega-1(4716) PCR primer reverse
<400> 5
tgatcttcag cacgaggttg 20
<210> 6
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, Mega-1 (4716) PCR primer forward
<400> 6
ggctccatct atgcctgtat c 21
<210> 7
<211> 22
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, Mega-2 (4715) PCR primer forward
<400> 7
gagctcagtt tcgctgtcta tc 22
<210> 8
<211> 21
<212> DNA

<213> Artificial Sequence
 <220>
 <223> Synthetic construct, Mega-2 (4715) PCR primer reverse
 <400> 8
 atgatcttca gcacgaggtt g 21
 <210> 9
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, ZmGWDmega-2 PCR primer forward
 <400> 9
 ggttataagc ccggttgaag ta 22
 <210> 10
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, ZmGWDmega-2 PCR primer reverse
 <400> 10
 ctattccttg ctcggactga c 21
 <210> 11
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, SbGWD mega-2 PCR primer forward
 <400> 11
 ggcaggttat aagcccagtt 20
 <210> 12
 <211> 189
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (189)
 <223> M16
 <400> 12
 ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60

tgctgtccag aacaaatctt atgataaacc aaggagagg aagaaatacc agatggagta 120
gttggtgtaa ttacacctga tatgccagat gttctgtctc atgtgtcagt ccgagcaagg 180
aatagcaag 189

<210> 13

<211> 204

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (204)

<223> M17

<400> 13

ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaggg agaggaagaa 120
ataccagatg gagtagttgg tgtaattaca cctgatatgc cagatgttct gtctcatgtg 180
tcagtccgag caaggaatag caag 204

<210> 14

<211> 213

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (213)

<223> M18

<400> 14

ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaagag tgtcaaggga 120
gaggaagaaa taccagatgg agtagttggt gtaattacac ctgatatgcc agatgttctg 180
tctcatgtgt cagtccgagc aaggaatagc aag 213

<210> 15

<211> 198

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (198)

<223> M27

<400> 15

ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60

tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaggg agagatacca 120
gatggagtag ttggtgtaat tacacctgat atgccagatg tctgtctca tgtgtcagtc 180
cgagcaagga atagcaag 198

<210> 16

<211> 214

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (214)

<223> M1

<400> 16

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaaga gtgtcaaggg 120
agaggaagaa ataccagatg gagtagttgg aagaaataca cctgatatgc cagatgttct 180
gtctcatgtg tcagtccgag caaggaatag caag 214

<210> 17

<211> 197

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (197)

<223> M11

<400> 17

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaaga gtgtcaaggg 120
agaggaagaa ataccagatg gagtagttgg tgtcagatgt tctgtctcat gtgtcagtcc 180
gagcaaggaa tagcaag 197

<210> 18

<211> 215

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (215)

<223> M10

<400> 18

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60

tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaaga gtgtcaaggg 120
agaggaagaa ataccagatg gagtagttgg tgtaaattac acctgatatg ccagatgttc 180
tgtctcatgt gtcagtccga gcaaggaata gcaag 215

<210> 19

<211> 199

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (199)

<223> M3

<400> 19

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaaga gtgtcaaggg 120
agaggaagaa ataccagatg gagcacctga tatgccagat gttctgtctc atgtgtcagt 180
ccgagcaagg aatagcaag 199

<210> 20

<211> 208

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (208)

<223> M8

<400> 20

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaaga gtgtcaaggg 120
agaggaagaa ataccagatg gagtagttgg tatgccagat atgccagatg ttctgtctca 180
tgtgtcagtc cgagcaagga atagcaag 208

<210> 21

<211> 189

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (189)

<223> M14

<400> 21

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60

tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaaga gtgtcaagg 120
agaggaagaa ttacacctga tatgccagat gttctgtctc atgtgtcagt ccgagcaagg 180
aatagcaag 189

<210> 22

<211> 178

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (178)

<223> M13

<400> 22

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaaga gtgtcaagg 120
agaggaagaa ataccagatg ttctgtctca tgtgtcagtc cgagcaagga atagcaag 178

<210> 23

<211> 176

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (176)

<223> M12

<400> 23

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaaga gtgtcaagg 120
agaggaagaa ataccagatg gagtagttgg tgtcagtcag agcaaggaat agcaag 176

<210> 24

<211> 204

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (204)

<223> M22

<400> 24

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaaga gtgtcaagg 120
agaggaagaa ataccagatg gagtagttgg tgtgatatgc cagatgttct gtctcatgtg 180

tcagtccgag caaggaatag caag 204
 <210> 25
 <211> 221
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (221)
 <223> M23
 <400> 25
 ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
 agaggaagaa ataccagatg gagtagttgg caaagataaa cttgcacct gatatgccag 180
 atgttctgtc tcatgtgtca gtccgagcaa ggaatagcaa g 221
 <210> 26
 <211> 213
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (213)
 <223> M24
 <400> 26
 ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
 agaggaagaa ataccagatg gagtagttgg tgaattacac ctgatatgcc agatgttctg 180
 tctcatgtgt cagtccgagc aaggaatagc aag 213
 <210> 27
 <211> 213
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (213)
 <223> M20
 <400> 27
 ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
 agaggaagaa ataccagatg gagtagttgg tgtattacac ctgatatgcc agatgttctg 180

tctcatgtgt cagtccgagc aaggaatagc aag 213

<210> 28

<211> 181

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (181)

<223> M21

<400> 28

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaaat 120
 cttatgataa accatgccag atgttctgtc tcatgtgtca gtccgagcaa ggaatagcaa 180
 g 181

<210> 29

<211> 205

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (205)

<223> M4

<400> 29

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
 agaggaagaa ataccagatg gagtaattac acctgatatg ccagatgttc tgtctcatgt 180
 gtcagtccga gcaaggaata gcaag 205

<210> 30

<211> 210

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (210)

<223> M19

<400> 30

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
 agaggaagaa ataccagatg gagtagttgg tgtacacctg atatgccaga tgttctgtct 180

catgtgtcag tccgagcaag gaatagcaag 210
 <210> 31
 <211> 425
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (425)
 <223> M26
 <400> 31
 ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
 agaggaagaa ataccagatg gagcagtgtg ctgggttaca gttcttatt tcaatgtctc 180
 cagtgggctt cttacctta tgtttgtgtt ttttttaag tgcagaaata gagaaagttc 240
 ttgcaaatat ctactctatg aaaaggacag ctatttggaa atatgtgaac agaactatcc 300
 ccagttgctg ggaaaaacca agaagaaagt tccttcaaat atctactcca tgacgacaag 360
 tgtctattac acctgatatg ccagatgttc tgtctcatgt gtcagtccga gcaaggaata 420
 gcaag 425
 <210> 32
 <211> 212
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (212)
 <223> M25
 <400> 32
 ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
 agaggaagaa ataccagatg gagtagttgg taattacacc tgatatgcca gatgttctgt 180
 ctcatgtgtc agtccgagca aggaatagca ag 212
 <210> 33
 <211> 173
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (173)
 <223> M15

<400> 33

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaaga gtaattacac 120
ctgatatgcc agatgttctg tctcatgtgt cagtccgagc aaggaatagc aag 173

<210> 34

<211> 425

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (425)

<223> M5

<400> 34

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaaga gtgtcaaggg 120
agaggaagaa ataccagatg gagtagttgc agaattattg aattctttca taattgaact 180
ctatgatgat gctttacttg attgtattat attgatgctc aatcatatat tgatgattgt 240
tggaaacttgc tctccgatgc aaggatgatcc aacgggggtg tgtcgcaacg taaacagggt 300
tttcgcacga gatggcaata gctctgttaa cctagcctct cacgggcaact gtgcgggggt 360
atttaattac acctgatatg ccagatgttc tgtctcatgt gtcagtccga gcaaggaata 420
gcaag 425

<210> 35

<211> 208

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (208)

<223> M2

<400> 35

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcett gtggcaaaga gtgtcaaggg 120
agaggaagaa ataccagatg gagtagttgt tacacctgat atgccagatg ttctgtctca 180
tgtgtcagtc cgagcaagga atagcaag 208

<210> 36

<211> 187

<212> DNA

<213> Zea mays

<220>

<221> misc_feature
<222> (1)..(187)
<223> M28
<400> 36
ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
agaggaagaa acacctgata tgccagatgt tctgtctcat gtgtcagtcc gagcaaggaa 180
tagcaag 187
<210> 37
<211> 213
<212> DNA
<213> Zea mays
<220>
<221> misc_feature
<222> (1)..(213)
<223> M6
<400> 37
ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
agaggaagaa ataccagatg gagtagttgg taaattacac ctgatatgcc agatgttctg 180
tctcatgtgt cagtccgagc aaggaatagc aag 213
<210> 38
<211> 199
<212> DNA
<213> Zea mays
<220>
<221> misc_feature
<222> (1)..(199)
<223> M9
<400> 38
ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
agaggaagaa ataccagatg gagtagttgg tatgccagat gttctgtctc atgtgtcagt 180
ccgagcaagg aatagcaag 199
<210> 39
<211> 210
<212> DNA
<213> Zea mays
<220>

<221> misc_feature
 <222> (1)..(210)
 <223> M7
 <400> 39
 ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
 agaggaagaa ataccagatg gagtagttgg tttacacctg atatgccaga tgttctgtct 180
 catgtgtcag tccgagcaag gaatagcaag 210
 <210> 40
 <211> 212
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1)..(212)
 <223> M29
 <400> 40
 ttggcagggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
 agaggaagaa ataccagatg gagtagttgg taattacacc tgatatgcca gatgttctgt 180
 ctcatgtgtc agtccgagca aggaatagca ag 212
 <210> 41
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, Meganuclease GWD target sequence pAG4715
 <400> 41
 atccttgtgg caaagagtgt ca 22
 <210> 42
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, Meganuclease target sequence pAG4716
 <400> 42
 gtagttgggtg taattacacc tg 22
 <210> 43
 <211> 1469

<212> PRT
 <213> Zea mays
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (1469)
 <223> ZmGWD
 <400> 43
 Met Ser Gly Phe Ser Ala Ala Ala Asn Ala Ala Ala Ala Glu Arg Cys
 1 5 10 15
 Ala Leu Ala Phe Arg Ala Arg Pro Ala Ala Ser Ser Pro Ala Lys Arg
 20 25 30
 Gln Gln Gln Pro Gln Pro Ala Ser Leu Arg Arg Ser Gly Gly Gln Arg
 35 40 45
 Arg Pro Thr Thr Leu Ser Ala Ser Ser Arg Gly Pro Val Val Pro Arg
 50 55 60
 Ala Val Ala Thr Ser Ala Asp Arg Ala Ser Pro Asp Leu Ile Gly Lys
 65 70 75 80
 Phe Thr Leu Asp Ser Asn Ser Glu Leu Gln Val Ala Val Asn Pro Ala
 85 90 95
 Pro Gln Gly Leu Val Ser Glu Ile Ser Leu Glu Val Thr Asn Thr Ser
 100 105 110
 Gly Ser Leu Ile Leu His Trp Gly Ala Leu Arg Pro Asp Lys Arg Asp
 115 120 125
 Trp Ile Leu Pro Ser Arg Lys Pro Asp Gly Thr Thr Val Tyr Lys Asn
 130 135 140
 Arg Ala Leu Arg Thr Pro Phe Val Lys Ser Gly Asp Asn Ser Thr Leu
 145 150 155 160
 Arg Ile Glu Ile Asp Asp Pro Gly Val His Ala Ile Glu Phe Leu Ile
 165 170 175
 Phe Asp Glu Thr Gln Asn Lys Trp Phe Lys Asn Asn Gly Gln Asn Phe
 180 185 190
 Gln Val Gln Phe Gln Ser Ser Arg His Gln Gly Thr Gly Ala Ser Gly
 195 200 205
 Ala Ser Ser Ser Ala Thr Ser Thr Leu Val Pro Glu Asp Leu Val Gln
 210 215 220
 Ile Gln Ala Tyr Leu Arg Trp Glu Arg Arg Gly Lys Gln Ser Tyr Thr
 225 230 235 240
 Pro Glu Gln Glu Lys Glu Glu Tyr Glu Ala Ala Arg Ala Glu Leu Ile
 245 250 255

Glu Glu Val Asn Arg Gly Val Ser Leu Glu Lys Leu Arg Ala Lys Leu
 260 265 270
 Thr Lys Ala Pro Glu Ala Pro Glu Ser Asp Glu Ser Lys Ser Ser Ala
 275 280 285
 Ser Arg Met Pro Ile Gly Lys Leu Pro Glu Asp Leu Val Gln Val Gln
 290 295 300
 Ala Tyr Ile Arg Trp Glu Gln Ala Gly Lys Pro Asn Tyr Pro Pro Glu
 305 310 315 320
 Lys Gln Leu Val Glu Phe Glu Glu Ala Arg Lys Glu Leu Gln Ala Glu
 325 330 335
 Val Asp Lys Gly Ile Ser Ile Asp Gln Leu Arg Gln Lys Ile Leu Lys
 340 345 350
 Gly Asn Ile Glu Ser Lys Val Ser Lys Gln Leu Lys Asn Lys Lys Tyr
 355 360 365
 Phe Ser Val Glu Arg Ile Gln Arg Lys Lys Arg Asp Ile Thr Gln Leu
 370 375 380
 Leu Ser Lys His Lys His Thr Leu Val Glu Asp Lys Val Glu Val Val
 385 390 395 400
 Pro Lys Gln Pro Thr Val Leu Asp Leu Phe Thr Lys Ser Leu His Glu
 405 410 415
 Lys Asp Gly Cys Glu Val Leu Ser Arg Lys Leu Phe Lys Phe Gly Asp
 420 425 430
 Lys Glu Ile Leu Ala Ile Ser Thr Lys Val Gln Asn Lys Thr Glu Val
 435 440 445
 His Leu Ala Thr Asn His Thr Asp Pro Leu Ile Leu His Trp Ser Leu
 450 455 460
 Ala Lys Asn Ala Gly Glu Trp Lys Ala Pro Ser Pro Asn Ile Leu Pro
 465 470 475 480
 Ser Gly Ser Thr Leu Leu Asp Lys Ala Cys Glu Thr Glu Phe Thr Lys
 485 490 495
 Ser Glu Leu Asp Gly Leu His Tyr Gln Val Val Glu Ile Glu Leu Asp
 500 505 510
 Asp Gly Gly Tyr Lys Gly Met Pro Phe Val Leu Arg Ser Gly Glu Thr
 515 520 525
 Trp Ile Lys Asn Asn Gly Ser Asp Phe Phe Leu Asp Phe Ser Thr His
 530 535 540
 Asp Val Arg Asn Ile Lys Leu Lys Gly Asn Gly Asp Ala Gly Lys Gly
 545 550 555 560
 Thr Ala Lys Ala Leu Leu Glu Arg Ile Ala Asp Leu Glu Glu Asp Ala

	565	570	575
Gln Arg Ser Leu Met His Arg Phe Asn Ile Ala Ala Asp Leu Ala Asp			
	580	585	590
Gln Ala Arg Asp Ala Gly Leu Leu Gly Ile Val Gly Leu Phe Val Trp			
	595	600	605
Ile Arg Phe Met Ala Thr Arg Gln Leu Thr Trp Asn Lys Asn Tyr Asn			
	610	615	620
Val Lys Pro Arg Glu Ile Ser Lys Ala Gln Asp Arg Phe Thr Asp Asp			
625	630	635	640
Leu Glu Asn Met Tyr Lys Ala Tyr Pro Gln Tyr Arg Glu Ile Leu Arg			
	645	650	655
Met Ile Met Ala Ala Val Gly Arg Gly Gly Glu Gly Asp Val Gly Gln			
	660	665	670
Arg Ile Arg Asp Glu Ile Leu Val Ile Gln Arg Asn Asn Asp Cys Lys			
	675	680	685
Gly Gly Met Met Glu Glu Trp His Gln Lys Leu His Asn Asn Thr Ser			
	690	695	700
Pro Asp Asp Val Val Ile Cys Gln Ala Leu Ile Asp Tyr Ile Lys Ser			
705	710	715	720
Asp Phe Asp Ile Ser Val Tyr Trp Asp Thr Leu Asn Lys Asn Gly Ile			
	725	730	735
Thr Lys Glu Arg Leu Leu Ser Tyr Asp Arg Ala Ile His Ser Glu Pro			
	740	745	750
Asn Phe Arg Ser Glu Gln Lys Ala Gly Leu Leu Arg Asp Leu Gly Asn			
	755	760	765
Tyr Met Arg Ser Leu Lys Ala Val His Ser Gly Ala Asp Leu Glu Ser			
	770	775	780
Ala Ile Ala Ser Cys Met Gly Tyr Lys Ser Glu Gly Glu Gly Phe Met			
785	790	795	800
Val Gly Val Gln Ile Asn Pro Val Lys Gly Leu Pro Ser Gly Phe Pro			
	805	810	815
Glu Leu Leu Glu Phe Val Leu Glu His Val Glu Asp Lys Ser Ala Glu			
	820	825	830
Pro Leu Leu Glu Gly Leu Leu Glu Ala Arg Val Glu Leu Arg Pro Leu			
	835	840	845
Leu Leu Asp Ser Arg Glu Arg Met Lys Asp Leu Ile Phe Leu Asp Ile			
	850	855	860
Ala Leu Asp Ser Thr Phe Arg Thr Ala Ile Glu Arg Ser Tyr Glu Glu			
865	870	875	880

Leu Asn Asp Ala Ala Pro Glu Lys Ile Met Tyr Phe Ile Ser Leu Val
 885 890 895
 Leu Glu Asn Leu Ala Leu Ser Ile Asp Asp Asn Glu Asp Ile Leu Tyr
 900 905 910
 Cys Leu Lys Gly Trp Asn Gln Ala Leu Glu Met Ala Lys Gln Lys Asp
 915 920 925
 Asp Gln Trp Ala Leu Tyr Ala Lys Ala Phe Leu Asp Arg Asn Arg Leu
 930 935 940
 Ala Leu Ala Ser Lys Gly Glu Gln Tyr His Asn Met Met Gln Pro Ser
 945 950 955 960
 Ala Glu Tyr Leu Gly Ser Leu Leu Ser Ile Asp Gln Trp Ala Val Asn
 965 970 975
 Ile Phe Thr Glu Glu Ile Ile Arg Gly Gly Ser Ala Ala Thr Leu Ser
 980 985 990
 Ala Leu Leu Asn Arg Phe Asp Pro Val Leu Arg Asn Val Ala His Leu
 995 1000 1005
 Gly Ser Trp Gln Val Ile Ser Pro Val Glu Val Ser Gly Tyr Val
 1010 1015 1020
 Val Val Val Asp Glu Leu Leu Ala Val Gln Asn Lys Ser Tyr Asp
 1025 1030 1035
 Lys Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu
 1040 1045 1050
 Ile Pro Asp Gly Val Val Gly Val Ile Thr Pro Asp Met Pro Asp
 1055 1060 1065
 Val Leu Ser His Val Ser Val Arg Ala Arg Asn Ser Lys Val Leu
 1070 1075 1080
 Phe Ala Thr Cys Phe Asp His Thr Thr Leu Ser Glu Leu Glu Gly
 1085 1090 1095
 Tyr Asp Gln Lys Leu Phe Ser Phe Lys Pro Thr Ser Ala Asp Ile
 1100 1105 1110
 Thr Tyr Arg Glu Ile Thr Glu Ser Glu Leu Gln Gln Ser Ser Ser
 1115 1120 1125
 Pro Asn Ala Glu Val Gly His Ala Val Pro Ser Ile Ser Leu Ala
 1130 1135 1140
 Lys Lys Lys Phe Leu Gly Lys Tyr Ala Ile Ser Ala Glu Glu Phe
 1145 1150 1155
 Ser Glu Glu Met Val Gly Ala Lys Ser Arg Asn Ile Ala Tyr Leu
 1160 1165 1170
 Lys Gly Lys Val Pro Ser Trp Val Gly Val Pro Thr Ser Val Ala

1175	1180	1185
Ile Pro Phe Gly Thr Phe	Glu Lys Val Leu Ser	Asp Gly Leu Asn
1190	1195	1200
Lys Glu Val Ala Gln Ser	Ile Glu Lys Leu Lys	Ile Arg Leu Ala
1205	1210	1215
Gln Glu Asp Phe Ser Ala	Leu Gly Glu Ile Arg	Lys Val Val Leu
1220	1225	1230
Asn Leu Thr Ala Pro Met	Gln Leu Val Asn Glu	Leu Lys Glu Arg
1235	1240	1245
Met Leu Gly Ser Gly Met	Pro Trp Pro Gly Asp	Glu Gly Asp Lys
1250	1255	1260
Arg Trp Glu Gln Ala Trp	Met Ala Ile Lys Lys	Val Trp Ala Ser
1265	1270	1275
Lys Trp Asn Glu Arg Ala	Tyr Phe Ser Thr Arg	Lys Val Lys Leu
1280	1285	1290
Asp His Glu Tyr Leu Ser	Met Ala Val Leu Val	Gln Glu Val Val
1295	1300	1305
Asn Ala Asp Tyr Ala Phe	Val Ile His Thr Thr	Asn Pro Ser Ser
1310	1315	1320
Gly Asp Ser Ser Glu Ile	Tyr Ala Glu Val Val	Lys Gly Leu Gly
1325	1330	1335
Glu Thr Leu Val Gly Ala	Tyr Pro Gly Arg Ala	Met Ser Phe Val
1340	1345	1350
Cys Lys Lys Asp Asp Leu	Asp Ser Pro Lys Leu	Leu Gly Tyr Pro
1355	1360	1365
Ser Lys Pro Ile Gly Leu	Phe Ile Arg Gln Ser	Ile Ile Phe Arg
1370	1375	1380
Ser Asp Ser Asn Gly Glu	Asp Leu Glu Gly Tyr	Ala Gly Ala Gly
1385	1390	1395
Leu Tyr Asp Ser Val Pro	Met Asp Glu Glu Asp	Glu Val Val Leu
1400	1405	1410
Asp Tyr Thr Thr Asp Pro	Leu Ile Val Asp Arg	Gly Phe Arg Ser
1415	1420	1425
Ser Ile Leu Ser Ser Ile	Ala Arg Ala Gly His	Ala Ile Glu Glu
1430	1435	1440
Leu Tyr Gly Ser Pro Gln	Asp Val Glu Gly Val	Val Lys Asp Gly
1445	1450	1455
Lys Ile Tyr Val Val Gln	Thr Arg Pro Gln Met	
1460	1465	

<210> 44
 <211> 1469
 <212> PRT
 <213> Sorghum bicolor
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (1469)
 <223> SbGWD
 <400> 44
 Met Thr Gly Phe Ser Ala Ala Ala Ser Ala Ala Ala Ala Ala Glu Arg
 1 5 10 15
 Cys Ala Leu Ala Ile Arg Ala Arg Pro Ala Ala Ser Ser Pro Ala Lys
 20 25 30
 Arg Gln Gln Gln Ser Ala Ser Leu Arg Arg Ser Gly Gly Gln Arg Arg
 35 40 45
 Pro Thr Thr Leu Ala Ala Ser Arg Arg Ser Pro Val Val Val Pro Arg
 50 55 60
 Ala Ile Ala Thr Ser Ala Asp Arg Ala Ser His Asp Leu Val Gly Lys
 65 70 75 80
 Phe Thr Leu Asp Ser Asn Ser Glu Leu Leu Val Ala Val Asn Pro Ala
 85 90 95
 Pro Gln Gly Leu Val Ser Val Ile Gly Leu Glu Val Thr Asn Thr Ser
 100 105 110
 Gly Ser Leu Ile Leu His Trp Gly Val Leu Arg Pro Asp Lys Arg Asp
 115 120 125
 Trp Ile Leu Pro Ser Arg Gln Pro Asp Gly Thr Thr Val Tyr Lys Asn
 130 135 140
 Arg Ala Leu Arg Thr Pro Phe Val Lys Ser Gly Asp Asn Ser Thr Leu
 145 150 155 160
 Arg Ile Glu Ile Asp Asp Pro Ala Val Gln Ala Ile Glu Phe Leu Ile
 165 170 175
 Phe Gly Glu Thr Gln Asn Lys Trp Phe Lys Asn Asn Gly Gln Asn Phe
 180 185 190
 Gln Ile Gln Leu Gln Ser Ser Arg His Gln Gly Asn Gly Ala Ser Gly
 195 200 205
 Ala Ser Ser Ser Ala Thr Ser Thr Leu Val Pro Glu Asp Leu Val Gln
 210 215 220
 Ile Gln Ala Tyr Leu Arg Trp Glu Arg Lys Gly Lys Gln Ser Tyr Thr
 225 230 235 240

Pro Glu Gln Glu Lys Glu Glu Tyr Glu Ala Ala Arg Ala Glu Leu Ile
 245 250 255
 Glu Glu Leu Asn Arg Gly Val Ser Leu Glu Lys Leu Arg Ala Lys Leu
 260 265 270
 Thr Lys Thr Pro Glu Ala Pro Glu Ser Asp Glu Arg Lys Ser Pro Ala
 275 280 285
 Ser Arg Met Pro Val Asp Lys Leu Pro Glu Asp Leu Val Gln Val Gln
 290 295 300
 Ala Tyr Ile Arg Trp Glu Lys Ala Gly Lys Pro Asn Tyr Pro Pro Glu
 305 310 315 320
 Lys Gln Leu Val Glu Leu Glu Glu Ala Arg Lys Glu Leu Gln Ala Glu
 325 330 335
 Val Asp Lys Gly Ile Ser Ile Asp Gln Leu Arg Gln Lys Ile Leu Lys
 340 345 350
 Gly Asn Ile Glu Ser Lys Val Ser Lys Gln Leu Lys Asn Lys Lys Tyr
 355 360 365
 Phe Ser Val Glu Arg Ile Gln Arg Lys Lys Arg Asp Ile Met Gln Leu
 370 375 380
 Leu Ser Lys His Lys His Thr Val Met Glu Glu Lys Val Glu Val Ala
 385 390 395 400
 Pro Lys Gln Pro Thr Val Leu Asp Leu Phe Thr Lys Ser Leu His Glu
 405 410 415
 Lys Asp Gly Cys Glu Val Leu Ser Arg Lys Leu Phe Lys Phe Gly Asp
 420 425 430
 Lys Glu Ile Leu Ala Ile Ser Thr Lys Val Gln Asn Lys Thr Glu Val
 435 440 445
 His Leu Ala Thr Asn His Thr Glu Pro Leu Ile Leu His Trp Ser Leu
 450 455 460
 Ala Lys Lys Ala Gly Glu Trp Lys Ala Pro Pro Ser Asn Ile Leu Pro
 465 470 475 480
 Ser Gly Ser Lys Leu Leu Asp Met Ala Cys Glu Thr Glu Phe Thr Arg
 485 490 495
 Ser Glu Leu Asp Gly Leu Cys Tyr Gln Val Val Glu Ile Glu Leu Asp
 500 505 510
 Asp Gly Gly Tyr Lys Gly Met Pro Phe Val Leu Arg Ser Gly Glu Thr
 515 520 525
 Trp Ile Lys Asn Asn Gly Ser Asp Phe Phe Leu Asp Phe Ser Thr Arg
 530 535 540
 Asp Thr Arg Asn Ile Lys Leu Lys Asp Asn Gly Asp Ala Gly Lys Gly

545	550	555	560
Thr Ala Lys Ala Leu Leu Glu Arg Ile Ala Asp Leu Glu Glu Asp Ala			
	565	570	575
Gln Arg Ser Leu Met His Arg Phe Asn Ile Ala Ala Asp Leu Ala Asp			
	580	585	590
Glu Ala Arg Asp Ala Gly Leu Leu Gly Ile Val Gly Leu Phe Val Trp			
	595	600	605
Ile Arg Phe Met Ala Thr Arg Gln Leu Thr Trp Asn Lys Asn Tyr Asn			
	610	615	620
Val Lys Pro Arg Glu Ile Ser Lys Ala Gln Asp Arg Phe Thr Asp Asp			
625	630	635	640
Leu Glu Asn Met Tyr Arg Thr Tyr Pro Gln Tyr Arg Glu Ile Leu Arg			
	645	650	655
Met Ile Met Ala Ala Val Gly Arg Gly Gly Glu Gly Asp Val Gly Gln			
	660	665	670
Arg Ile Arg Asp Glu Ile Leu Val Ile Gln Arg Asn Asn Asp Cys Lys			
	675	680	685
Gly Gly Met Met Glu Glu Trp His Gln Lys Leu His Asn Asn Thr Ser			
	690	695	700
Pro Asp Asp Val Val Ile Cys Gln Ala Leu Ile Asp Tyr Ile Lys Asn			
705	710	715	720
Asp Phe Asp Ile Ser Val Tyr Trp Asp Thr Leu Asn Lys Asn Gly Ile			
	725	730	735
Thr Lys Glu Arg Leu Leu Ser Tyr Asp Arg Ala Ile His Ser Glu Pro			
	740	745	750
Asn Phe Arg Ser Glu Gln Lys Glu Gly Leu Leu Arg Asp Leu Gly Asn			
	755	760	765
Tyr Met Arg Ser Leu Lys Ala Val His Ser Gly Ala Asp Leu Glu Ser			
	770	775	780
Ala Ile Ala Thr Cys Met Gly Tyr Lys Ser Glu Gly Glu Gly Phe Met			
785	790	795	800
Val Gly Val Gln Ile Asn Pro Val Lys Gly Leu Pro Ser Gly Phe Pro			
	805	810	815
Glu Leu Leu Glu Phe Val Leu Asp His Val Glu Asp Lys Ser Ala Glu			
	820	825	830
Pro Leu Leu Glu Gly Leu Leu Glu Ala Arg Val Asp Leu Arg Pro Leu			
	835	840	845
Leu Leu Asp Ser Pro Glu Arg Met Lys Asp Leu Ile Phe Leu Asp Ile			
	850	855	860

Ala Leu Asp Ser Thr Phe Arg Thr Ala Ile Glu Arg Ser Tyr Glu Glu			
865	870	875	880
Leu Asn Asp Ala Ala Pro Glu Lys Ile Met Tyr Phe Ile Ser Leu Val			
	885	890	895
Leu Glu Asn Leu Ala Phe Ser Ile Asp Asp Asn Glu Asp Ile Leu Tyr			
	900	905	910
Cys Leu Lys Gly Trp Asn Gln Ala Leu Glu Met Ala Lys Gln Lys Asp			
	915	920	925
Asp Gln Trp Ala Leu Tyr Ala Lys Ala Phe Leu Asp Arg Ile Arg Leu			
	930	935	940
Ala Leu Ala Ser Lys Gly Glu Gln Tyr His Asn Met Met Gln Pro Ser			
945	950	955	960
Ala Glu Tyr Leu Gly Ser Leu Leu Ser Ile Asp Lys Trp Ala Val Asn			
	965	970	975
Ile Phe Thr Glu Glu Ile Ile Arg Gly Gly Ser Ala Ala Thr Leu Ser			
	980	985	990
Ala Leu Leu Asn Arg Phe Asp Pro Val Leu Arg Asn Val Ala Asn Leu			
	995	1000	1005
Gly Ser Trp Gln Val Ile Ser Pro Val Glu Val Ser Gly Tyr Val			
	1010	1015	1020
Val Val Val Asp Glu Leu Leu Ala Val Gln Asn Lys Ser Tyr Asp			
	1025	1030	1035
Lys Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu			
	1040	1045	1050
Ile Pro Asp Gly Val Val Gly Val Ile Thr Pro Asp Met Pro Asp			
	1055	1060	1065
Val Leu Ser His Val Ser Val Arg Ala Arg Asn Ser Lys Val Leu			
	1070	1075	1080
Phe Ala Thr Cys Phe Asp His Thr Thr Leu Ser Glu Leu Glu Gly			
	1085	1090	1095
Tyr Asp Gln Lys Leu Leu Ser Phe Lys Pro Thr Ser Ala Asp Ile			
	1100	1105	1110
Thr Tyr Arg Glu Ile Thr Glu Ser Glu Leu Gln Gln Ser Ser Ser			
	1115	1120	1125
Pro Asn Ala Glu Val Gly His Ala Val Pro Ser Ile Ser Leu Ala			
	1130	1135	1140
Lys Lys Lys Phe Leu Gly Lys Tyr Ala Ile Ser Ala Glu Glu Phe			
	1145	1150	1155
Thr Glu Glu Met Val Gly Ala Lys Ser Arg Asn Ile Ala Tyr Leu			

1160	1165	1170
Lys Gly Lys Val Pro Ser Trp Val Gly Val Pro Thr Ser Val Ala		
1175	1180	1185
Ile Pro Phe Gly Thr Phe Glu Lys Val Leu Ser Asp Gly Leu Asn		
1190	1195	1200
Lys Glu Val Ala Gln Thr Ile Glu Lys Leu Lys Ile Arg Leu Ala		
1205	1210	1215
Gln Glu Asp Phe Ser Ala Leu Gly Glu Ile Arg Lys Ala Val Leu		
1220	1225	1230
Asn Leu Thr Ala Pro Met Gln Leu Val Asn Glu Leu Lys Glu Arg		
1235	1240	1245
Met Leu Gly Ser Gly Met Pro Trp Pro Gly Asp Glu Gly Asn Arg		
1250	1255	1260
Arg Trp Glu Gln Ala Trp Met Ala Ile Lys Lys Val Trp Ala Ser		
1265	1270	1275
Lys Trp Asn Glu Arg Ala Tyr Phe Ser Thr Arg Lys Val Lys Leu		
1280	1285	1290
Asn His Glu Tyr Leu Ser Met Ala Val Leu Val Gln Glu Val Val		
1295	1300	1305
Asn Ala Asp Tyr Ala Phe Val Ile His Thr Thr Asn Pro Ser Ser		
1310	1315	1320
Gly Asp Ser Ser Glu Ile Tyr Ala Glu Val Val Lys Gly Leu Gly		
1325	1330	1335
Glu Thr Leu Val Gly Ala Tyr Pro Gly Arg Ala Met Ser Phe Val		
1340	1345	1350
Cys Lys Lys Asp Asp Leu Asp Ser Pro Lys Leu Leu Gly Tyr Pro		
1355	1360	1365
Ser Lys Pro Ile Gly Leu Phe Ile Arg Arg Ser Ile Ile Phe Arg		
1370	1375	1380
Ser Asp Ser Asn Gly Glu Asp Leu Glu Gly Tyr Ala Gly Ala Gly		
1385	1390	1395
Leu Tyr Asp Ser Val Pro Met Asp Glu Glu Asp Glu Val Val Leu		
1400	1405	1410
Asp Tyr Thr Thr Asp Pro Leu Ile Val Asp Arg Gly Phe Arg Asn		
1415	1420	1425
Ser Ile Leu Ser Ser Ile Ala Arg Ala Gly His Ala Ile Glu Glu		
1430	1435	1440
Leu Tyr Gly Ser Pro Gln Asp Val Glu Gly Val Val Lys Asp Gly		
1445	1450	1455

Lys Ile Tyr Val Val Gln Thr Arg Pro Gln Met
 1460 1465

<210> 45

<211> 74

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (74)

<223> ZmGWD_M1_aa_1040-1120

<400> 45

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
 1 5 10 15

Asp Gly Val Val Gly Arg Asn Thr Pro Asp Met Pro Asp Val Leu Ser
 20 25 30

His Val Ser Val Arg Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly
 35 40 45

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

Pro Asp Val Leu Ser His Val Ser Val Arg
 65 70

<210> 46

<211> 79

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (79)

<223> ZmGWD_M2 _1040-1120)

<400> 46

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
 1 5 10 15

Asp Gly Val Val Val Thr Pro Asp Met Pro Asp Val Leu Ser His Val
 20 25 30

Ser Val Arg Ala Arg Asn Ser Lys Val Leu Phe Ala Thr Cys Phe Asp
 35 40 45

His Thr Thr Leu Ser Glu Leu Glu Gly Tyr Asp Gln Lys Leu Phe Ser
 50 55 60

Phe Lys Pro Thr Ser Ala Asp Ile Thr Tyr Arg Glu Ile Thr Glu

<210> 49

<211> 95

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (95)

<223> ZmGWD_M5_aa_1040-1120

<400> 49

```

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
1           5           10           15
Asp Gly Val Val Ala Glu Leu Leu Asn Ser Phe Ile Ile Glu Leu Tyr
           20           25           30
Asp Asp Ala Leu Leu Asp Cys Ile Ile Leu Met Leu Asn His Ile Leu
           35           40           45
Met Ile Val Gly Thr Cys Ser Pro Met Gln Gly Asp Pro Thr Gly Val
           50           55           60
Cys Arg Asn Val Asn Arg Val Phe Ala Arg Asp Gly Asn Ser Ser Val
65           70           75           80
Asn Leu Ala Ser His Gly His Cys Ala Gly Val Phe Asn Tyr Thr
           85           90           95

```

<210> 50

<211> 73

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (73)

<223> ZmGWD_M6_aa_1040-1120

<400> 50

```

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
1           5           10           15
Asp Gly Val Val Gly Lys Leu His Leu Ile Cys Gln Met Phe Cys Leu
           20           25           30
Met Cys Gln Ser Glu Gln Gly Ile Ala Arg Tyr Cys Leu Arg Pro Val
           35           40           45
Leu Thr Thr Pro Leu Tyr Leu Asn Leu Lys Asp Met Ile Arg Asn Cys
           50           55           60
Phe Pro Ser Ser Leu Leu Leu Gln Ile

```

65 70
 <210> 51
 <211> 72
 <212> PRT
 <213> Zea mays
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (72)
 <223> ZmGWD_M7_aa_1040-1120
 <400> 51
 Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
 1 5 10 15
 Asp Gly Val Val Gly Leu His Leu Ile Cys Gln Met Phe Cys Leu Met
 20 25 30
 Cys Gln Ser Glu Gln Gly Ile Ala Arg Tyr Cys Leu Arg Pro Val Leu
 35 40 45
 Thr Thr Pro Leu Tyr Leu Asn Leu Lys Asp Met Ile Arg Asn Cys Phe
 50 55 60
 Pro Ser Ser Leu Leu Leu Gln Ile
 65 70
 <210> 52
 <211> 79
 <212> PRT
 <213> Zea mays
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (79)
 <223> ZmGWD_M8_aa_1040-1120
 <400> 52
 Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
 1 5 10 15
 Asp Gly Val Val Gly Met Pro Asp Met Pro Asp Val Leu Ser His Val
 20 25 30
 Ser Val Arg Ala Arg Asn Ser Lys Val Leu Phe Ala Thr Cys Phe Asp
 35 40 45
 His Thr Thr Leu Ser Glu Leu Glu Gly Tyr Asp Gln Lys Leu Phe Ser
 50 55 60
 Phe Lys Pro Thr Ser Ala Asp Ile Thr Tyr Arg Glu Ile Thr Glu
 65 70 75

<210> 53
 <211> 76
 <212> PRT
 <213> Zea mays
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (76)
 <223> ZmGWD_M9_aa_1040-1120
 <400> 53
 Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
 1 5 10 15
 Asp Gly Val Val Gly Met Pro Asp Val Leu Ser His Val Ser Val Arg
 20 25 30
 Ala Arg Asn Ser Lys Val Leu Phe Ala Thr Cys Phe Asp His Thr Thr
 35 40 45
 Leu Ser Glu Leu Glu Gly Tyr Asp Gln Lys Leu Phe Ser Phe Lys Pro
 50 55 60
 Thr Ser Ala Asp Ile Thr Tyr Arg Glu Ile Thr Glu
 65 70 75

<210> 54
 <211> 25
 <212> PRT
 <213> Zea mays
 <220>
 <221> MISC_FEATURE
 <223> ZmGWD_M10_aa_1040-1120
 <400> 54
 Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
 1 5 10 15
 Asp Gly Val Val Gly Val Asn Tyr Thr
 20 25

<210> 55
 <211> 34
 <212> PRT
 <213> Zea mays
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (34)
 <223> ZmGWD_M11_aa_1040-1120

<400> 55

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
 1 5 10 15
 Asp Gly Val Val Gly Val Arg Cys Ser Val Ser Cys Val Ser Pro Ser
 20 25 30

Lys Glu

<210> 56

<211> 27

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (27)

<223> ZmGWD_M12 _1040-1120

<400> 56

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
 1 5 10 15
 Asp Gly Val Val Gly Val Ser Pro Ser Lys Glu
 20 25

<210> 57

<211> 60

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (60)

<223> ZmGWD_M13 _aa_1040-1120

<400> 57

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
 1 5 10 15
 Asp Val Leu Ser His Val Ser Val Arg Ala Arg Asn Ser Lys Val Leu
 20 25 30

Phe Ala Thr Cys Phe Asp His Thr Thr Leu Ser Glu Leu Glu Gly Tyr
 35 40 45

Asp Gln Lys Leu Phe Ser Phe Lys Pro Thr Ser Ala
 50 55 60

<210> 58

<211> 65

<212> PRT

<213> Zea mays
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (65)
 <223> ZmGWD_M14_aa_1040-1120
 <400> 58
 Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Leu His
 1 5 10 15
 Leu Ile Cys Gln Met Phe Cys Leu Met Cys Gln Ser Glu Gln Gly Ile
 20 25 30
 Ala Arg Tyr Cys Leu Arg Pro Val Leu Thr Thr Pro Leu Tyr Leu Asn
 35 40 45
 Leu Lys Asp Met Ile Arg Asn Cys Phe Pro Ser Ser Leu Leu Leu Gln
 50 55 60
 Ile
 65
 <210> 59
 <211> 11
 <212> PRT
 <213> Zea mays
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (11)
 <223> ZmGWD_M15_aa_1040-1120
 <400> 59
 Pro Thr Ile Leu Val Ala Lys Ser Asn Tyr Thr
 1 5 10
 <210> 60
 <211> 10
 <212> PRT
 <213> Zea mays
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (10)
 <223> ZmGWD_M16_aa_1040-1120
 <400> 60
 Pro Arg Glu Arg Lys Lys Tyr Gln Met Glu
 1 5 10
 <210> 61

<211> 15

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (15)

<223> ZmGWD_M17_aa_1040-1120

<400> 61

Pro Thr Ile Leu Val Ala Arg Glu Arg Lys Lys Tyr Gln Met Glu

1 5 10 15

<210> 62

<211> 18

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (18)

<223> ZmGWD_M18_aa_1040-1120

<400> 62

Pro Thr Ile Leu Val Ala Arg Val Ser Arg Glu Arg Lys Lys Tyr Gln

1 5 10 15

Met Glu

<210> 63

<211> 71

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (71)

<223> ZmGWD_M19_aa_1040-1120

<400> 63

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro

1 5 10 15

Asp Gly Val Val Gly Val His Leu Ile Cys Gln Met Phe Cys Leu Met

20 25 30

Cys Gln Ser Glu Gln Gly Ile Ala Arg Tyr Cys Leu Arg Pro Val Leu

35 40 45

Thr Thr Pro Leu Tyr Leu Asn Leu Lys Asp Met Ile Arg Asn Cys Phe

50 55 60

<223> ZmGWD_M22_aa_1040-1120

<400> 66

```

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
1           5           10           15
Asp Gly Val Val Gly Val Ile Cys Gln Met Phe Cys Leu Met Cys Gln
           20           25           30
Ser Glu Gln Gly Ile Ala Arg Tyr Cys Leu Arg Pro Val Leu Thr Thr
           35           40           45
Pro Leu Tyr Leu Asn Leu Lys Asp Met Ile Arg Asn Cys Phe Pro Ser
           50           55           60
Ser Leu Leu Leu Gln Ile
65           70

```

<210> 67

<211> 27

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (27)

<223> ZmGWD_M23_aa_1040-1120

<400> 67

```

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
1           5           10           15
Asp Gly Val Val Gly Lys Asp Lys Pro Cys Thr
           20           25

```

<210> 68

<211> 73

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (73)

<223> ZmGWD_M24_aa_1040-1120

<400> 68

```

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
1           5           10           15
Asp Gly Val Val Gly Glu Leu His Leu Ile Cys Gln Met Phe Cys Leu
           20           25           30
Met Cys Gln Ser Glu Gln Gly Ile Ala Arg Tyr Cys Leu Arg Pro Val

```

35 40 45
 Leu Thr Thr Pro Leu Tyr Leu Asn Leu Lys Asp Met Ile Arg Asn Cys
 50 55 60

Phe Pro Ser Ser Leu Leu Leu Gln Ile
 65 70

<210> 69

<211> 24

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (24)

<223> ZmGWD_M25_aa_1040-1120

<400> 69

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
 1 5 10 15

Asp Gly Val Val Gly Asn Tyr Thr
 20

<210> 70

<211> 95

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (95)

<223> ZmGWD_M26_aa_1040-1120

<400> 70

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
 1 5 10 15

Asp Gly Ala Val Cys Ser Gly Thr Ala Ser Tyr Phe Asn Val Ser Ser
 20 25 30

Gly Arg Leu Thr Ser Met Phe Val Phe Phe Phe Lys Cys Arg Asn Arg
 35 40 45

Glu Ser Ser Cys Lys Tyr Leu Leu Tyr Glu Lys Asp Ser Tyr Leu Glu
 50 55 60

Ile Cys Glu Gln Asn Tyr Pro Gln Leu Leu Gly Lys Thr Lys Lys Lys
 65 70 75 80

Val Pro Ser Asn Ile Tyr Ser Met Thr Thr Ser Val Tyr Tyr Thr
 85 90 95

<210> 71
 <211> 13
 <212> PRT
 <213> Zea mays
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (13)
 <223> ZmGWD_M27_aa_1040-1120
 <400> 71
 Pro Thr Ile Leu Val Ala Arg Glu Arg Tyr Gln Met Glu
 1 5 10
 <210> 72
 <211> 72
 <212> PRT
 <213> Zea mays
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (72)
 <223> ZmGWD_M28_aa_1040-1120
 <400> 72
 Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Thr Pro
 1 5 10 15
 Asp Met Pro Asp Val Leu Ser His Val Ser Val Arg Ala Arg Asn Ser
 20 25 30
 Lys Val Leu Phe Ala Thr Cys Phe Asp His Thr Thr Leu Ser Glu Leu
 35 40 45
 Glu Gly Tyr Asp Gln Lys Leu Phe Ser Phe Lys Pro Thr Ser Ala Asp
 50 55 60
 Ile Thr Tyr Arg Glu Ile Thr Glu
 65 70
 <210> 73
 <211> 24
 <212> PRT
 <213> Zea mays
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (24)
 <223> ZmGWD_M29_aa_1040-1120
 <400> 73

Pro Thr Ile Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro
 1 5 10 15
 Asp Gly Val Val Gly Asn Tyr Thr
 20
 <210> 74
 <211> 1417
 <212> PRT
 <213> Streptococcus pyogenes
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (1147)
 <223> Cas9 protein
 <400> 74
 Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp
 1 5 10 15
 Tyr Lys Asp Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val
 20 25 30
 Gly Ile His Gly Val Pro Ala Ala Asp Lys Lys Tyr Ser Ile Gly Leu
 35 40 45
 Asp Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr
 50 55 60
 Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His
 65 70 75 80
 Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu
 85 90 95
 Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr
 100 105 110
 Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu
 115 120 125
 Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe
 130 135 140
 Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn
 145 150 155 160
 Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His
 165 170 175
 Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu
 180 185 190
 Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu
 195 200 205

Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe			
210	215	220	
Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile			
225	230	235	240
Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser			
	245	250	255
Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys			
	260	265	270
Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr			
	275	280	285
Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln			
	290	295	300
Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln			
305	310	315	320
Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser			
	325	330	335
Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr			
	340	345	350
Lys Ala Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His			
	355	360	365
Gln Asp Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu			
	370	375	380
Lys Tyr Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly			
385	390	395	400
Tyr Ile Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys			
	405	410	415
Pro Ile Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu			
	420	425	430
Asn Arg Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser			
	435	440	445
Ile Pro His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg			
	450	455	460
Gln Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu			
465	470	475	480
Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg			
	485	490	495
Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile			
	500	505	510
Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln			

515	520	525
Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu		
530	535	540
Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr		
545	550	555
Asn Glu Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro		
565	570	575
Ala Phe Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe		
580	585	590
Lys Thr Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe		
595	600	605
Lys Lys Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp		
610	615	620
Arg Phe Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile		
625	630	635
Lys Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu		
645	650	655
Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu		
660	665	670
Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys		
675	680	685
Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys		
690	695	700
Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp		
705	710	715
Phe Leu Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile		
725	730	735
His Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val		
740	745	750
Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly		
755	760	765
Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp		
770	775	780
Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile		
785	790	795
Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser		
805	810	815
Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser		
820	825	830

Gln Ile Leu Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu		
835	840	845
Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp		
850	855	860
Gln Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp His Ile		
865	870	875
Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu		
	885	890
Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu		
	900	910
Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala		
	915	920
Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg		
	930	940
Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu		
945	950	955
Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser		
	965	970
Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val		
	980	985
Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys Asp		
	995	1000
Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His Ala		
	1010	1015
His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile Lys		
	1025	1030
Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr Lys		
	1040	1045
Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln Glu Ile		
	1055	1060
Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met Asn		
	1070	1075
Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg Lys		
	1085	1090
Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val Trp		
	1100	1105
Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser Met		
	1115	1120
Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly Gly		

1130	1135	1140
Phe Ser Lys Glu Ser Ile	Leu Pro Lys Arg Asn	Ser Asp Lys Leu
1145	1150	1155
Ile Ala Arg Lys Lys Asp	Trp Asp Pro Lys Lys	Tyr Gly Gly Phe
1160	1165	1170
Asp Ser Pro Thr Val Ala	Tyr Ser Val Leu Val	Val Ala Lys Val
1175	1180	1185
Glu Lys Gly Lys Ser Lys	Lys Leu Lys Ser Val	Lys Glu Leu Leu
1190	1195	1200
Gly Ile Thr Ile Met Glu	Arg Ser Ser Phe Glu	Lys Asn Pro Ile
1205	1210	1215
Asp Phe Leu Glu Ala Lys	Gly Tyr Lys Glu Val	Lys Lys Asp Leu
1220	1225	1230
Ile Ile Lys Leu Pro Lys	Tyr Ser Leu Phe Glu	Leu Glu Asn Gly
1235	1240	1245
Arg Lys Arg Met Leu Ala	Ser Ala Gly Glu Leu	Gln Lys Gly Asn
1250	1255	1260
Glu Leu Ala Leu Pro Ser	Lys Tyr Val Asn Phe	Leu Tyr Leu Ala
1265	1270	1275
Ser His Tyr Glu Lys Leu	Lys Gly Ser Pro Glu	Asp Asn Glu Gln
1280	1285	1290
Lys Gln Leu Phe Val Glu	Gln His Lys His Tyr	Leu Asp Glu Ile
1295	1300	1305
Ile Glu Gln Ile Ser Glu	Phe Ser Lys Arg Val	Ile Leu Ala Asp
1310	1315	1320
Ala Asn Leu Asp Lys Val	Leu Ser Ala Tyr Asn	Lys His Arg Asp
1325	1330	1335
Lys Pro Ile Arg Glu Gln	Ala Glu Asn Ile Ile	His Leu Phe Thr
1340	1345	1350
Leu Thr Asn Leu Gly Ala	Pro Ala Ala Phe Lys	Tyr Phe Asp Thr
1355	1360	1365
Thr Ile Asp Arg Lys Arg	Tyr Thr Ser Thr Lys	Glu Val Leu Asp
1370	1375	1380
Ala Thr Leu Ile His Gln	Ser Ile Thr Gly Leu	Tyr Glu Thr Arg
1385	1390	1395
Ile Asp Leu Ser Gln Leu	Gly Gly Asp Arg Pro	Lys Lys Lys Arg
1400	1405	1410
Lys Val Gly Gly		
1415		

<210> 75
 <211> 4272
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, ZmCas9
 <400> 75
 ggatcctaaa ccatggatta caaggaccac gacggcgatt acaaggacca cgacattgat 60
 tacaaggacg acgacgataa gatggctccc aagaagaaga ggaaggttgg catccacggg 120
 gtgccggctg ctgacaagaa gtactcgatc ggctcagata ttgggactaa ctctgttggc 180
 tgggccgtga tcaccgacga gtacaaggtg cctcgaaga agttcaaggt cctgggcaac 240
 accgatcggc attctatcaa gaagaatctc attggcgctc tctgttctga ctcaggggag 300
 accgctgagg ctacgagget caagaggacc gcccgcagge ggtacacgcg caggaagaat 360
 cgcattctgt acctgcagga gattttctcc aacgagatgg cgaaggttga cgattctttc 420
 ttccacagge tggaggagtc attcctcgtg gaggaggata agaagcacga gcggcatcca 480
 atcttcggca acattgtcga cgaggttgcc taccacgaga agtaccctac gatctacat 540
 ctgcggaaga agctcgtgga ctccacagat aaggcggacc tccgctgat ctacctcgt 600
 ctggcccaca tgattaagtt caggggcat ttctgatcg aggggatct caaccggac 660
 aatagcgatg ttgacaagct gttcatccag ctctgtcaga cgtacaacca gctcttcgag 720
 gagaacccca ttaatgcgtc aggcgtcgac gcgaaggcta tctgtcctgc tcgctctcgc 780
 aagtctagga ggctggagaa cctgatcgcc cagctgccgg gcgagaagaa gaacggcctg 840
 ttccgggaatc tcatcgtctc cagcctgggg ctacgcca acttcaagtc gaatttcgat 900
 ctgctgagg acgccaagct gcagctctcc aaggacacat acgacgatga cctggataac 960
 ctctggccc agatcggcga tcagtacgcg gacctgttcc tcgctgcca gaatctgtcg 1020
 gacgcatcc tctgtctga tattctcagg gtgaacaccg agattacgaa ggctccgctc 1080
 tcagcctcca tgatcaagcg ctacgacgag caccatcagg atctgacct cctgaaggcg 1140
 ctggctcagc agcagctccc cgagaagtac aaggagattt tcttcgatca gtccaagaac 1200
 ggctacgctg ggtacattga cggcggggcc agccaggagg agttctacaa gttcatcaag 1260
 ccgattctgg agaagatgga cggcacggag gagctcctgg tgaagctcaa tcgagggac 1320
 ctctgagga agcagcggac attcgataac ggcagatcc cacaccagat tcattctcggg 1380
 gagctgcacg ccattcctgag gcggcaggag gaattctacc ctttctcaa ggataaccgc 1440
 gagaagatcg agaagattct gaccttccgc atcccgtact acgtcggccc actcgcctgc 1500
 ggcaactccc gcttcgcttg gatgaccgc aagtcagagg agaccatcac gccgtggaac 1560
 ttcgaggagg tggctgacaa gggcgttagc gctcagctgt tcattcagag gatgacgaat 1620
 ttcgacaaga acctgcaaaa tgagaaggtg ctccctaage actcgtctct gtacgagtac 1680
 ttcacagtct acaacgagct cactaaggtg aagtatgtga ccgagggcat gaggaagccg 1740
 gctttcctgt ctggggagca gaagaaggcc atcgtggacc tctgttcaa gaccaaccgg 1800
 aaggtcacgg ttaagcagct caaggaggac tacttcaaga agattgagtg cttcgattcg 1860
 gtcgagatca gcggcgttga ggacaggttc aacgcctccc tggggaccta ccacgatctc 1920

ctgaagatca ttaaggataa ggacttcctg gacaacgagg agaatgagga tatectggag 1980
 gacattgtgc tgacactcac tctgttcgag gaccgggaga tgatcgagga gcgcctgaag 2040
 acttacgccc atctcttcga tgacaaggtc atgaagcagc tcaagaggag gaggtacacc 2100
 ggctggggga ggctgagcag gaagctcatc aacggcattc gggacaagca gtccgggaag 2160
 acgatcctcg acttcctgaa gagcgatggc ttgcgcaacc gcaatttcat gcagctgatt 2220
 cacgatgaca gcctcacatt caaggaggat atccagaagg ctgaggtgag cggccagggg 2280
 gactcgctgc acgagcatat cgcgaaacctc gctggctcgc cagctatcaa gaaggggatt 2340
 ctgcagaccg tgaaggttgt ggacgagctc gtgaaggtca tgggcaggca caagcctgag 2400
 aacatcgtca ttgagatggc ccgcgagaat cagaccacgc agaagggccca gaagaactca 2460
 cgcgagagga tgaagaggat cgaggagggc attaaggagc tgggggtcca gatcctcaag 2520
 gagcaccggg tggagaacac gcagctgcag aatgagaagc tctacctgta ctacctccag 2580
 aatggccgcg atatgtatgt ggaccaggag ctggatatta acaggctcag cgattacgac 2640
 gtcgatcata tcgttccaca gtcattcctg aaggatgact ccattgacaa caaggtcctc 2700
 accaggtcgg acaagaaccg gggcaagtct gataatgttc cttcagagga ggtcgttaag 2760
 aagatgaaga actactggcg ccagctcctg aatgccaagc tgatcacgca gcggaagtcc 2820
 gataacctca caaaggctga gaggggcggg ctctctgagc tggacaaggc gggcttcctc 2880
 aagaggcagc tggctcgagac acggcagatc actaagcacg ttgcgcagat tctcgactca 2940
 cggatgaaca ctaagtacga tgagaatgac aagctgatcc gcgaggtgaa ggtcatcacc 3000
 ctgaagtcaa agctcgtctc cgacttcagg aaggatttcc agttctacaa ggttcgggag 3060
 atcaacaatt accaccatgc ccatgacgcg tactgaacg cggtggtcgg cacagctctg 3120
 atcaagaagt acccaaagct ggagtccgag ttctgttacg gggactacaa ggtttacgat 3180
 gtgcgcaaga tgatcgccaa gtcggagcag gagattggca aggctaccgc caagtacttc 3240
 ttctactcta acattatgaa tttcttcaag acagagatca ctctggccaa tggcgagatc 3300
 cggaagcgcc ccctcattga gaccaacggc gagacggggg agatcgtgtg ggacaagggc 3360
 agggatttctg cgaccgtcag gaaggttctc tccatgccac aagtgaatat cgtcaagaag 3420
 acagaggctc agactggcgg gttctctaag gactcaattc tgcctaagcg gaacagcgac 3480
 aagctcatcg cccgcaagaa ggactgggac ccgaagaagt acggcggggtt cgacagcccc 3540
 actgtggcct actcggctct gttgtggcg aaggttgaga agggcaagtc caagaagctc 3600
 aagagcgtga aggagctcct ggggatcacg attatggaga ggtccagctt cgagaagaac 3660
 ccgatcgatt tcctggagge gaagggtac aaggaggatga agaaggacct gatcattaag 3720
 ctccccaaagt actcactctt cgagctggag aacggcagga agcggatgct ggcttccgct 3780
 ggcgagctcc agaaggggaa tgagctcgtc ctgccgtcca agtatgtgaa ctctctctac 3840
 ctggcctccc actacgagaa gctcaagggc agccccgagg acaacgagca gaagcagctg 3900
 ttctcgcgagc agcacaagca ttacctcgac gagatcattg agcagatttc cgagttcagc 3960
 aagcgcgtga tcctggccga cgcgaaatctg gataaggtcc tcagcgcgta caacaagcac 4020
 cgcgacaagc caatcaggga gcaggetgag aatateatc atctcttcc cctgacgaac 4080
 ctggcgccc ctgctgcttt caagtacttc gacacaacta tcgatcgcaa gaggtacaca 4140
 tcgactaagg aggtcctgga cgcgacctc atccaccagt ctattacagg cctgtacgag 4200
 actcggattg atctgtcgca gctcggcggg gataggccca agaagaagag gaaggtcggc 4260

ggctgaccta gg 4272
 <210> 76
 <211> 1007
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (1007)
 <223> MzU3.8
 <400> 76
 gaattccatc taagtatctt ggtaaagcat ggattaattt ggatgctcac ttcaggtcta 60
 tgcagctccg gtgccttgtg attgtgagtt gtgaccgatg ctcatgctat tttgcatttc 120
 tgcgatgtat gatgctagta gatcttcaaa actaacageg catgceatca tcatccactg 180
 cttgatttta gtctcaccgc tggccaaaaa tgtgatgatg ccagaaacct caactacctt 240
 gaatcaacac gggcccagca gtgtgatgac gacagaaacc aaaaaaaaaat gagccaatag 300
 ttcagaagga ggcactatgc agaaactaca tttctgaagg tgactaaaag gtgagcgtag 360
 agtgtactta ctagtagttt agccaccatt acccaaatgc tttcgagctt gtattaagac 420
 ttccaaagct gagcatcatc actgatctgc aggagggtcg cttcgctgcc aagatcaaca 480
 gcaaccatgt ggcggcaaca tccagcattg cacatgggct aaagattgag ctctgtgcca 540
 agtgtgagct gcaaccatct agggatcagc tgagtttate agtctttcct ttttttcatt 600
 ctgggtgaggc atcaagctac tactgcctcg atcggttgga cttggacctg aagcccat 660
 gtaggatacc agaatggacc gaccaggac gtagtgccac ctcggttgte aactgcgta 720
 gaagccagct taaaaattta gctttggtga ctacagcac gaccttactt gaacaggatc 780
 tgttctatag gatcgtactg ttgcatcttt gattaataag aaggcaagta cttaacctg 840
 gttgatgaga atttgacctg tgggccagag cgtgatttaa cggccaggac tttgccttgg 900
 tgcattgtct ggagctgcag atgatcgtt ttggccaggc ttaatgtctg gctagggtgg 960
 cctacaggct gtttgacagg tttctcaatt ttttgctct gctgcag 1007
 <210> 77
 <211> 1501
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (1501)
 <223> ZmU3
 <400> 77
 tagttgtcta ttaataaatt ttatcatgtg tagctgactt aaaagacatg taatctagtg 60
 cgcacatcaat ctcagcatgc aaacatata atttttgaac ttgtgatatt tttatacagt 120
 atatcataat agataaaaatt agacaacaca gaactaaaat tataatatta atactaattt 180

ggaccatacc attaccaa atgttgaact aaatcattct tgaagtcaat atgcttttat 240
 agtttgat atccatgatt tctgaattcc atctaagat gttggtaaag catggattaa 300
 tttggatgcc cacttcaggt ctatgcagct ccggtgcett gtgattgtga gttgtgaccg 360
 atgctcatgc tattctgcat ttctgcgatg tatgtagcta gtagatcttc aaaactaaca 420
 ccgcatgcca tcatcatcca ctgcttgatt ttagtctcac cgctggccaa aaatgtgatg 480
 atgccagaaa cctcaactac cttgaatcaa cacgggceca acagtgtgat gacgacagaa 540
 acaaaaaaaaa atgagccaat agttcagaag gaggcactat gcagaaacta catttctgaa 600
 ggtgactaaa aggtgagcgt agagtgtaat tactagtagt ttagccacca ttacccaaat 660
 gctttcgagc ttgtattaag atttctaag ctgagcatca tcaactgatc gcaggccacc 720
 ctgccttcgc tgccaagatc aacagcaacc atgtggcggc aacatccagc attgcacatg 780
 ggctaaagat tgagctttgt gctcgtcta gggatcagct gaggttatca gtctttcctt 840
 tttttcatcc aggtgaggca tcaagctact actgectega ttggctggac ccgaagccca 900
 catgtaggat accagaatgg gccgaccag gacgcagat gttggccagt cccaccggtt 960
 agtgccatct cggttgctca catgcgtaga agccagetta aaaatttagc tttggtaact 1020
 cacagcacga ccttacttga acaggatctg ttctatagga tcgtactgtt gcacttttga 1080
 ttaataagaa ggcaagtact taaacctggt tgatgagaat ttgacctgtg ggccagagcg 1140
 tgattaacgg ccaggactct ttgccttggg gcattgtctg gagctgcaga tgatcgttct 1200
 tggccaggct taatgtctgg ctagggtggc ctacaggctg tttgacaggc ctctcaattt 1260
 ttttgctctg ctgcaggatg tcatttgaact caacgccatt aatgattgac tttttgatct 1320
 gtgctgcgtt tgaagaaacc tactccagct agcttttctt cagcatttgc actcaaatta 1380
 agagggccag atatcttgcg cgtttttgcc atcagtaata aagttttctt taggtgtgat 1440
 gcatttgaag gggatttaag gaggttattt ctgtcaccag ctgtttttgc ttagtgttgc 1500
 t 1501

<210> 78

<211> 761

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (761)

<223> MzU3.8 promoter region

<400> 78

gaattccatc taagtatctt ggtaaagcat ggattaattt ggatgctcac ttcaggctca 60
 mtgcagctcc ggtgccttgt gattgtgagt tgtgaccgat gctcatgcta ttttgcattt 120
 cmtgcgatgt atgatgctag tagatcttca aaactaacag cgcattgcat catcatccac 180
 tgccttgatt tagtctcacc gctggccaaa aatgtgatga tgccagaaac ctcaactacc 240
 ttgaatcaac acgggcccag cagtgtgatg acgacagaaa ccaaaaaaaaa atgagccaat 300
 agttcagaag gaggcactat gcagaaacta catttctgaa ggtgactaaa aggtgagcgt 360
 agagtgtact tactagtagt ttagccacca ttacccaaat gctttcgagc ttgtattaag 420

acttcctaag ctgagcatca tcaactgatct gcaggagggt cgettegetg ccaagatcaa 480
 cagcaaccat gtggcggcaa catccagcat tgcacatggg ctaaagattg agctctgtgc 540
 caagtgtgag ctgcaaccat ctagggatca gctgagttta tcagtctttc ctttttttca 600
 ttctgggtgag gcatcaagct actactgcct cgatcggttg gacttggacc tgaagccac 660
 atgtaggata ccagaatgga ccgaccagg acgtagtgcc acctcggttg tcacactgcg 720
 tagaagccag cttaaaaatt tagctttggt gactcacagc a 761

<210> 79

<211> 764

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (764)

<223> ZmU3 promoter region

<400> 79

gaattccatc taagtatggt ggtaaagcat ggattaattt ggatgccac ttcaggctca 60
 tgcagctccg gtgccttggt attgtgagtt gtgaccgatg ctcagtctat tctgcatttc 120
 tgcgatgtat gtagctagta gatcttcaaa actaacaccg catgccaatca tcactccactg 180
 cttgatttta gtctcaccgc tggccaaaaa tgtgatgatg ccagaaacct caactacctt 240
 gaatcaacac gggcccaaca gtgtgatgac gacagaaaca aaaaaaatg agccaatagt 300
 tcagaaggag gcactatgca gaaactacat ttctgaaggt gactaaaagg tgagcgtaga 360
 gtgtaattac tagtagttta gccaccatta cccaaatget ttcgagcttg tattaagatt 420
 tctaagctg agcatcatca ctgatctgca ggccaccctc gcttcgctgc caagatcaac 480
 agcaaccatg tggcggcaac atccagcatt gcacatgggc taaagattga gctttgtgcc 540
 tcgtctaggg atcagctgag gttatcagtc tttccttttt ttcattccagg tgaggcatca 600
 agctactact gcctcgattg gctggaccgc aagccacat gtaggatacc agaatgggcc 660
 gaccaggac gcagtatggt ggccagtccc accggtagt gccatctcgg ttgctcatat 720
 gcgtagaagc cagcttaaaa atttagcttt ggtaactcac agca 764

<210> 80

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, ob 2297 forward primer

<400> 80

gcgatcgcca tctaagtatg ttggtaaage atgg 34

<210> 81

<211> 36

<212> DNA

<213> Artificial Sequence
 <220>
 <223> Synthetic construct, ob2299 reverse primer
 <400> 81
 tgctgtgagt taccaaagct aaatttttaa gctggc 36
 <210> 82
 <211> 758
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (758)
 <223> ZmU3P1
 <400> 82
 catctaagta tgttggtaaa gcatggatta atttggatgc ccaattcagg tctatgcagc 60
 tccggtgcct tgtgattgtg agttgtgacc gatgctcatg ctattctgca tttctgcgat 120
 gtatgtagct agtagatctt caaaactaac accgcatgcc atcatcatcc actgcttgat 180
 tttagtctca ccgctggcca aaaatgtgat gatgccagaa acctcaacta cettgaatca 240
 acacgggccc aacagtgtga tgacgacaga aacaaaaaaaa aatgagccaa tagttcagaa 300
 ggaggcacta tgcagaaact acatttctga aggtgactaa aaggtgagcg tagagtgtaa 360
 ttactagtag tttagccacc attacceaaa tgettctgag cttgtattaa gatttcttaa 420
 gctgagcate atcactgac tgcaggccac cctcgetteg ctgccaagat caacagcaac 480
 catgtggcgg caacatccag cattgcacat gggctaaaga ttgagctttg tgccctcgtct 540
 agggatcagc tgaggttacc agtctttcct tttttcacc caggtgaggc atcaagctac 600
 tactgcctcg attggctgga cccgaagccc acatgtagga taccagaatg ggccgaccca 660
 ggacgcagta tgttggccag tcccaccggt tagtgccatc tcggttgctc acatgcgtag 720
 aagccagctt aaaaatttag ctttggtaac tcacagca 758
 <210> 83
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, ob2343 forward primer
 <400> 83
 gcgatgcag tttagccacc attacceaaa tgc 33
 <210> 84
 <211> 398
 <212> DNA
 <213> Zea mays

<220>

<221> misc_feature

<222> (1)..(398)

<223> ZmU3P2

<400> 84

gcgatcgcag tttagccacc attacccaaa tgctttcgag cttgtattaa gatttcctaa 60
gctgagcadc atcaactgac tgcaggccac cctcgcttcg ctgccaagat caacagcaac 120
catgtggcgg caacatccag cattgcacat gggctaaaga ttgagctttg tgccctgtct 180
agggatcagc tgaggttatc agtctttcct tttttcattc caggtgaggc atcaagctac 240
tactgcctcg attggctgga cccgaagccc acatgtagga taccagaatg ggccgaccca 300
ggacgcagta tgttgccag tcccaccggt tagtgccatc tcggttctc acatgcgtag 360
aagccagctt aaaaatttag ctttggtaac tcacagca 398

<210> 85

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, ob2351 forward primer

<400> 85

cgatttaaag agtttagcca ccattacca aatgc 35

<210> 86

<211> 308

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1)..(308)

<223> ZmU3.8P

<400> 86

gcgatcgcgc ttcgctgcca agatcaacag caacatgtg gcggaacat ccagcattgc 60
acatgggcta aagattgagc tctgtgcca gtgtgagctg caacatcta gggatcagct 120
gagttatca gtctttcctt tttttcattc tggtgagcca teaagctact actgcctcga 180
tcggttgac ttggacctga agcccacatg taggatacca gaatggaccg acccaggacg 240
tagtgccacc tcggttgca cactgcgtag aagccagctt aaaaatttag ctttggtgac 300
tcacagca 308

<210> 87

<211> 42

<212> DNA

<213> Streptococcus pyogenes

<220>
 <221> misc_feature
 <222> (1)..(42)
 <223> Cas9 handle hairpin
 <400> 87
 gttttagagc tagaaatagc aagttaaaat aaggctagtc cg 42
 <210> 88
 <211> 41
 <212> DNA
 <213> Streptococcus pyogenes
 <220>
 <221> misc_feature
 <222> (1)..(41)
 <223> S. pyogenes terminator
 <400> 88
 ttatcaactt gaaaaagtgg caccgagtcg gtgctttttt t 41
 <210> 89
 <211> 37
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1)..(37)
 <223> ZmU3T
 <400> 89
 gctctgctgc aggtgatcat ttgactcaac gccatta 37
 <210> 90
 <211> 120
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, sgRNA scaffold
 <400> 90
 gttttagagc tagaaatagc aagttaaaat aaggctagtc cgttatcaac ttgaaaaagt 60
 ggcaaccgagt cgggtgctttt tttgctctgc tgcaggtgat catttgactc aacgccatta 120
 <210> 91
 <211> 19
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Synthetic construct, GWDe1a antisense
 <400> 91
 ggcatgaggt gcttacgtc 19
 <210> 92
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, GWDe24b antisense
 <400> 92
 cataacctga tacttcaac 19
 <210> 93
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, GWDe24c sense
 <400> 93
 tctggctcct gctatcagt 19
 <210> 94
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, GWDe25a antisense
 <400> 94
 tctgcagaag taggcttga 19
 <210> 95
 <211> 911
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, ZmU3P1:sgRNA_GWDe24b
 <400> 95
 gcgatgccca tctaagtatg ttggtaaage atggattaat ttggatgecc acttcaggtc 60
 tatgcagctc cggtgecttg tgattgtgag ttgtgaccga tgctcatgct attctgcatt 120
 tctgcgatgt atgtagctag tagatettca aaactaacac cgcgatgceat catcatccac 180
 tgettgatatt tagtctcacc gctggccaaa aatgtgatga tgccagaaac ctcaactacc 240

ttgaatcaac acgggcccac cagtgtgatg acgacagaaa caaaaaaaaa tgagccaata 300
 gttcagaagg aggcactatg cagaaactac atttctgaag gtgactaaaa ggtgagcgta 360
 gagtgtaatt actagtagtt tagccacat tacccaaatg ctttcgagct tgtattaaga 420
 tttcctaage tgagcatcat cactgatctg caggccacce tcgcttcgct gccaagatca 480
 acagcaacca tgtggcggca acatccagca ttgcatatgg gctaaagatt gagctttgtg 540
 cctcgtctag ggatcagctg aggttatcag tctttccttt ttttcatcca ggtgaggcat 600
 caagctacta ctgcctcgat tggctggacc cgaagcccac atgtaggata ccagaatggg 660
 ccgacccagg acgcagtatg ttggccagtc ccaccggta gtgccatctc ggttgctcac 720
 atgcgtagaa gccagcttaa aaatttagct ttgtaactc acagcacata acctgatact 780
 tcaacgtttt agagctagaa atagcaagtt aaaataaggc tagtccgta tcaacttgaa 840
 aaagtggcac cgagtcggtg ctttttttgc tctgctgcag gtgatcattt gactcaacgc 900
 cattatacgt a 911

<210> 96

<211> 543

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, ZmU3P2:sgRNA_GWDe24b

<400> 96

gcgatcgcag tttagccacc attacccaaa tgctttcgag cttgtattaa gatttcctaa 60
 gctgagcate atcactgac tgcaggccac cctcgettcg ctgccaagat caacagcaac 120
 catgtggcgg caacatccag cattgcacat gggctaaaga ttgagctttg tgctcgtct 180
 agggatcagc tgaggttacc agtctttcct tttttcacc caggtgaggc atcaagctac 240
 tactgcctcg attggctgga cccgaagccc acatgtagga taccagaatg ggccgaccca 300
 ggacgcagta tgttgccag tcccaccggt tagtgccatc tcggttgctc acatgcgtag 360
 aagccagctt aaaaatttag ctttggtaac tcacagcaca taacctgata cttcaacggt 420
 ttagagctag aatagcaag ttaaaataag gctagtccgt tatcaacttg aaaaagtggc 480
 accgagtcgg tgcttttttt gctctgctgc aggtgatcat ttgactcaac gccattatac 540
 gta 543

<210> 97

<211> 453

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, ZmU3.8P:sgRNA_GWDe24b

<400> 97

gcgatcgcgc ttcgctgcca agatcaacag caacctgtg gcggcaacat ccagcattgc 60
 acatgggcta aagattgagc tctgtgccaa gtgtgagctg caacctcta gggatcagct 120
 gagtttatca gtctttcctt tttttcattc tggtaggca tcaagctact actgcctcga 180

tcggttgac ttggacctga agcccacatg taggatacca gaatggaccg acccaggacg 240
 tagtgccacc tcggttgca cactgcgtag aagccagctt aaaaatttag ctttggtgac 300
 tcacagcaca taacctgata cttcaacgtt ttagagctag aaatagcaag ttaaaataag 360
 gctagtccgt tatcaacttg aaaaagtggc accgagtcgg tgcttttttt gctctgctgc 420
 aggtgatcat ttgactcaac gccattatac gta 453

<210> 98

<211> 543

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, ZmU3P2:sgRNA_GWDe24c

<400> 98

gcgatcgcag tttagccacc attacceaaa tgetttegag cttgtattaa gatttcctaa 60
 gctgagcate atcactgate tgcaggccac cctcgetteg ctgccaagat caacagcaac 120
 catgtggcgg caacatccag cattgcacat gggctaaaga ttgagctttg tgctctgtct 180
 agggatcagc tgaggttate agtetttctt ttttttate caggtgagge atcaagctac 240
 tactgcctcg attggetgga cccgaagccc acatgtagga taccagaatg ggccgaccca 300
 ggacgcagta tgttgccag tcccaccggt tagtgccate tcggttgctc acatgcgtag 360
 aagccagctt aaaaatttag ctttggtaac tcacagcate tggetctgc tatcagtgtt 420
 ttagagctag aaatagcaag ttaaaataag gctagtccgt tatcaacttg aaaaagtggc 480
 accgagtcgg tgcttttttt gctctgctgc aggtgatcat ttgactcaac gccattatac 540
 gta 543

<210> 99

<211> 545

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, ZmU3P2:sgRNA_GWDe24a

<400> 99

atttaaatag tttagccacc attacceaaa tgetttegag cttgtattaa gatttcctaa 60
 gctgagcate atcactgate tgcaggccac cctcgetteg ctgccaagat caacagcaac 120
 catgtggcgg caacatccag cattgcacat gggctaaaga ttgagctttg tgctctgtct 180
 agggatcagc tgaggttate agtetttctt ttttttate caggtgagge atcaagctac 240
 tactgcctcg attggetgga cccgaagccc acatgtagga taccagaatg ggccgaccca 300
 ggacgcagta tgttgccag tcccaccggt tagtgccate tcggttgctc acatgcgtag 360
 aagccagctt aaaaatttag ctttggtaac tcacagcate tgcagaagta ggcttgagtt 420
 ttagagctag aaatagcaag ttaaaataag gctagtccgt tatcaacttg aaaaagtggc 480
 accgagtcgg tgcttttttt gctctgctgc aggtgatcat ttgactcaac gccattagge 540
 gcgcc 545

<210> 100
 <211> 543
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, ZmU3P2:sgRNA_GWDe1a
 <400> 100
 gcgatcgag tttagccacc attacceaaa tgettccgag cttgtattaa gatttcctaa 60
 gctgagcadc atcactgac tgcaggccac cctcgettcg ctgccaagat caacagcaac 120
 catgtggcgg caacatccag cattgcacat gggctaaaga ttgagctttg tgccctcgtct 180
 agggatcagc tgaggttate agtctttcct ttttttcate caggtgagge atcaagctac 240
 tactgcctcg attggctgga cccgaagccc acatgtagga taccagaatg ggccgaccca 300
 ggacgcagta tgttgccag tcccaccggt tagtgccate tcggttgctc acatgcgtag 360
 aagccagctt aaaaatttag ctttggtaac tcacagcagg catgaggtgc ttacgtcgtt 420
 ttagagctag aaatagcaag ttaaaataag getagtcctg tatcaacttg aaaaagtggc 480
 accgagtcgg tgettttttt getctgetgc aggtgatcat ttgactcaac gccattatac 540
 gta 543
 <210> 101
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, GWDe24b-F primer
 <400> 101
 ctcacagcac ataacctgat act 23
 <210> 102
 <211> 17
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, sgRNA-R primer
 <400> 102
 cgactcggcg ccaacttt 17
 <210> 103
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, ZmCas9-F primer

<400> 103
 agaatcagac cacgcagaag 20
 <210> 104
 <211> 22
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, ZmCas9-R primer
 <400> 104
 gctcctggtc cacatacata tc 22
 <210> 105
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, GWDex23-F primer
 <400> 105
 tgctcttctg aaccgatttg a 21
 <210> 106
 <211> 227
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, Sb4715_1 (WT+INS)_Exon24
 <400> 106
 ttggcagggtt ataagcccag ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggagtagt tgggtagatt 120
 ggtgtatcaa gggagaggaa gaaataccag atggagtagt tgggtgaatt acacctgata 180
 tgccagatgt tctgtcccat gtgtcagtc gagcaaggaa tagcaag 227
 <210> 107
 <211> 202
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, Sb4715_2 (WT+del)_Exon24
 <400> 107
 ttggcagggtt ataagcccag ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
 agaggaagaa ataccagatg gagttacacc tgatatgcca gatgttctgt cccatgtgtc 180

agtccgagca aggaatagca ag 202

<210> 108

<211> 1095

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, meganuclease 4715

<400> 108

atggcaccga agaagaagcg caaggtgcat atgaatacaa aatataataa agagttctta 60
 ctctacttag cagggtttgt agacggtgac ggttccatct ttgccaggat caggccttct 120
 caatctcgga agttcaagca ccagctgacg ctcgagttea aggtcactca gaagacacag 180
 cgccgttggc tctctgacaa gctggtggac gagatcggcg tgggttacgt gacggacgat 240
 ggcagcgtct ccttttactc tctgtcccag atcaagcett tgcataatth tttaacacaa 300
 ctacaacctt ttctaaaact aaaacaaaaa caagcaagtt tagtttttaa aattattgag 360
 caacttccgt cagcaaaaaga atccccggac aaattcttag aagtttgtac atgggtggat 420
 caaattgcag ctctgaatga ttcgaagacg cgtaaaacaa cttctgaaac cgttcgtgct 480
 gtgctagaca gtttaccagg atccgtggga ggtctatcgc catctcaggc atccagcgc 540
 gcatcctcgg ctctctcaag cccgggttca gggatctccg aagcactcag agctggagca 600
 ggttccggca ctggatacaa caaggaattc ctgctctacc tggcgggctt cgtcagcggg 660
 gacggctcca tctatgccac tatcaggccg aggcatcgg tgaagttcaa gcactttctg 720
 gagctcagtt tcgctgtcta tcagaagaca cagcgcctt ggttctctga caagctggtg 780
 gacgagatcg gtgtgggtta cgtgatgac agtggcagta ctcccgta cctgctgtcc 840
 gagatcaagc ctctgcacaa ctctctgacc cagctccagc ccttctgaa gctcaagcag 900
 aagcaggcca acctcgtgct gaagatcatc gagcagctgc cctccgctaa ggaatccccg 960
 gacaagttcc tggagggtg cacctgggtg gaccagatcg ccgctctgaa cgactccaag 1020
 acccgcaaga cacttccga aaccgtccgc gccgttctag acagtctctc cgagaagaag 1080
 aagtcgtccc cctaa 1095

<210> 109

<211> 1095

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, meganuclease 4716

<400> 109

atggcaccga agaagaagcg caaggtgcat atgaatacaa aatataataa agagttctta 60
 ctctacttag cagggtttgt agacggtgac ggttccatct atgctctgat cagcctagt 120
 caacatctga agttcaagca ccagctgagg ctctggttcg atgtcgtca gaagacacag 180
 cgccgttggc tctctgacaa gctggtggac gagatcggcg tgggttacgt gtatgaccag 240
 ggcagcgtct cctattaccg tctgtccgag atcaagcett tgcataatth tttaacacaa 300

ctacaacctt ttctaaaact aaaacaaaa caagcaaatt tagttttaa aattattgaa 360
 caacttccgt cagcaaaaga atccccggac aaattcttag aagtttgtac atgggtggat 420
 caaattgcag ctctgaatga ttcgaagacg cgtaaaaca cttctgaaac cgttcgtgct 480
 gtgctagaca gtttaccagg atccgtggga ggtctatcgc catctcaggc atccagcgcc 540
 gcatcctcgg cttcctcaag cccgggttca gggatctccg aagcactcag agctggagca 600
 ggttccggca ctggatacaa caaggaattc ctgctctacc tggcgggctt cgtcgacggg 660
 gacggctcca tctatgcctg tatccatcct gatcaagcta ataagttcaa gcaccggctg 720
 cggctctatt tcattgtcag tcagaagaca cagegcgctt ggttcctcga caagctggtg 780
 gacgagatcg gtgtgggtta cgtgtatgac aggggcggcg tctcccatta ccagctgtcc 840
 cagatcaagc ctctgcacaa cttctgacc cagctccage ccttctgaa gctcaagcag 900
 aagcaggcca acctcgtgct gaagatcacc gagcagctgc cctccgcaa ggaatccccg 960
 gacaagttcc tggagggtg cacctgggtg gaccagatcg ccgctctgaa cgactccaag 1020
 acccgcaaga ccaattccga aaccgtccgc gccgttctag acagtctctc cgagaagaag 1080
 aagtcgtccc cctaa 1095

<210> 110

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, GWDe24a-F

<400> 110

tgcagaagta ggcttgagtt t 21

<210> 111

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, 2856 forward primer

<400> 111

gaaggggatt ggagaggaag 20

<210> 112

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, 2858 reverse primer

<400> 112

catgacgttc aaatagctc a 21

<210> 113

<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, 429 reverse primer
<400> 113
gcagaagtag gcttgaagga a 21
<210> 114
<211> 77
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, M32
<400> 114
gctcctgcta tcagttggca ggttataage ccggtttgaa gtatcaggtt atgtggttgt 60
ggttgatgag ttacttg 77
<210> 115
<211> 74
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, M33
<400> 115
gctcctgcta tcagttggca ggttataage ccggttagta tcaggttatg tggttgtggt 60
tgatgagtta cttg 74
<210> 116
<211> 73
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, M34
<400> 116
gctcctgcta tcagttggca ggttataage ccggttgat caggttatgt ggttgtggtt 60
gatgagttac ttg 73
<210> 117
<211> 75
<212> DNA
<213> Artificial Sequence
<220>

<223> Synthetic construct, M35
<400> 117
gctcctgcta tcagttggca ggtataagc ccggtgaagt atcaggttat gtggttgg 60
ttgatgagtt acttg 75
<210> 118
<211> 39
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, M36
<400> 118
gctcctgcta tcagttggtt gtggttgatg agttacttg 39
<210> 119
<211> 19
<212> DNA
<213> Artificial Sequence
<220>
<223> Syntehtic construct, e25a - 48
<400> 119
cactctatct gcagatata 19
<210> 120
<211> 68
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, e25a+1
<400> 120
cactctatct gaacttgaag gatatgatca gaaactgttt tccttcacag cctacttctg 60
cagatata 68
<210> 121
<211> 7
<212> PRT
<213> Artificial Sequence
<220>
<223> Synthetic construct, M32 peptide
<400> 121
Trp Gln Val Ile Ser Pro Val
1 5
<210> 122

<211> 15
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, M33 peptide
 <400> 122
 Trp Gln Val Ile Ser Pro Val Ser Ile Arg Leu Cys Gly Cys Gly
 1 5 10 15
 <210> 123
 <211> 46
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, M34 peptide
 <400> 123
 Trp Gln Val Ile Ser Pro Val Val Ser Gly Tyr Val Val Val Val Asp
 1 5 10 15
 Glu Leu Leu Ala Val Gln Asn Lys Ser Tyr Asp Lys Pro Thr Ile Leu
 20 25 30
 Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro Asp Gly
 35 40 45
 <210> 124
 <211> 47
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, M35 peptide
 <400> 124
 Trp Gln Val Ile Ser Pro Val Lys Tyr Gln Val Met Trp Leu Trp Leu
 1 5 10 15
 Met Ser Tyr Leu Leu Ser Arg Thr Asn Leu Met Ile Asn Gln Pro Ser
 20 25 30
 Leu Trp Gln Arg Val Ser Arg Glu Arg Lys Lys Tyr Gln Met Glu
 35 40 45
 <210> 125
 <211> 35
 <212> PRT
 <213> Artificial Sequence
 <220>

<223> Synthetic construct, M36 peptide

<400> 125

Trp Leu Trp Leu Met Ser Tyr Leu Leu Ser Arg Thr Asn Leu Met Ile

1 5 10 15

Asn Gln Pro Ser Leu Trp Gln Arg Val Ser Arg Glu Arg Lys Lys Tyr

 20 25 30

Gln Met Glu

 35

<210> 126

<211> 19

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M38 peptide

<400> 126

Val Leu Phe Ala Thr Cys Phe Asp His Thr Thr Leu Ser Ala Asp Ile

1 5 10 15

Thr Tyr Arg

<210> 127

<211> 34

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M39 peptide

<400> 127

Val Leu Phe Ala Thr Cys Phe Asp His Thr Thr Leu Ser Glu Leu Glu

1 5 10 15

Gly Tyr Asp Gln Lys Leu Phe Ser Phe Thr Ala Tyr Phe Cys Arg Tyr

 20 25 30

Asn Leu

<210> 128

<211> 32

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, NLS 1 protein

<400> 128

Met Pro Thr Glu Glu Arg Val Arg Lys Arg Lys Glu Ser Asn Arg Glu

1 5 10 15

Ser Ala Arg Arg Ser Arg Tyr Arg Lys Ala Ala His Leu Lys Glu Leu
 20 25 30

<210> 129

<211> 27

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, NLS3 protein

<400> 129

Met Ala Arg Lys Arg Lys Glu Ser Asn Arg Glu Ser Ala Arg Arg Ser
 1 5 10 15

Arg Tyr Arg Lys Ala Ala His Leu Lys Glu Leu
 20 25

<210> 130

<211> 24

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, NLS4 protein

<400> 130

Met Ala Arg Lys Arg Lys Glu Ser Asn Arg Glu Ser Ala Arg Arg Ser
 1 5 10 15

Arg Arg Ser Arg Tyr Arg Lys Val
 20

<210> 131

<211> 54

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, M40

<400> 131

gaaataccag atggagtagt tgtaattaca cctgatatgc cagatgttct gtct 54

<210> 132

<211> 58

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, M41

<400> 132

gaaataccag atggagtagt tggataaat tacacctgat atgccagatg ttctgtct 58

<210> 133

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, M42

<400> 133

gaaataccag atggagtagt tgggtgatta cacctgatat gccagatggt ctgtct 56

<210> 134

<211> 62

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, M43

<400> 134

gaaataccag atggagtagt tgggtgtagag taataacacc tgatatgcca gatgttctgt 60
ct 62

<210> 135

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, M44

<400> 135

gaaataccag atggagtagt tgggtgttctg tct 33

<210> 136

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, M45

<400> 136

gaaataccag atgttctgtc t 21

<210> 137

<211> 67

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, M46
<400> 137
gaaataccag atggagtagt tgggtgatga acacgtaatt acacctgata tgccagatgt 60
tctgtct 67
<210> 138
<211> 28
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, M47
<400> 138
gaaataccag atggagtagt tgggtgtct 28
<210> 139
<211> 54
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, M48
<400> 139
gaaataccag atggagtagt tgggtgttaca cctgatatgc cagatgttct gtct 54
<210> 140
<211> 58
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, M49
<400> 140
gaaataccag atggagtagt tgggtgtaaatac tacacctgat atgccagatg ttctgtct 58
<210> 141
<211> 35
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, M50
<400> 141
gaaataccag atgggatatg ccagatgttc tgtct 35
<210> 142
<211> 76
<212> DNA

- <213> Artificial Sequence
<220>
<223> Synthetic construct, M51
<400> 142
gaaataccag atggagtagt tgggtgtctca tgccagatgt gaagaaatta cacctgatat 60
gccagatggt ctgtct 76
<210> 143
<211> 36
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, M52
<400> 143
gaaataccag atggagtagt tggatgatggt ctgtct 36
<210> 144
<211> 49
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, M53
<400> 144
gaaataccag atggagtagt tgggtgcaga tatgccagat gttctgtct 49
<210> 145
<211> 167
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, M54
<400> 145
gaaataccag atggagtagt tgggtgcattt actcatattt tctgtgattg aatattcttt 60
tccagatgga gtgtcaaggg agaggaagaa ataccagatg gagggtcaag ggagaggaag 120
aaataccaga tgaaggaaat acacctgata tgccagatgt tctgtct 167
<210> 146
<211> 38
<212> DNA
<213> Artificial Sequence
<220>
<223> Synthetic construct, M55
<400> 146

<213> Artificial Sequence

<220>

<223> Synthetic construct, M43 peptide

<400> 150

Ile Pro Asp Gly Val Val Gly Val Glu

1 5

<210> 151

<211> 20

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M44 peptide

<400> 151

Ile Pro Asp Gly Val Val Gly Val Leu Ser His Val Ser Val Arg Ala

1 5 10 15

Arg Asn Ser Lys

20

<210> 152

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M45 peptide

<400> 152

Ile Pro Asp Val Leu Ser His Val Ser Val Arg Ala Arg Asn Ser Lys

1 5 10 15

<210> 153

<211> 8

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M46 peptide

<400> 153

Ile Pro Asp Gly Val Val Gly Val

1 5

<210> 154

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M47 peptide

<400> 154

Ile Pro Asp Gly Val Val Gly Val Ser Cys Val Ser Pro Ser Lys Glu
1 5 10 15

<210> 155

<211> 27

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M48 peptide

<400> 155

Ile Pro Asp Gly Val Val Gly Val Thr Pro Asp Met Pro Asp Val Leu
1 5 10 15
Ser His Val Ser Val Arg Ala Arg Asn Ser Lys
 20 25

<210> 156

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M49 peptide

<400> 156

Ile Pro Asp Gly Val Val Gly Val Asn Tyr Thr
1 5 10

<210> 157

<211> 52

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M50 peptide

<400> 157

Ile Pro Asp Gly Ile Cys Gln Met Phe Cys Leu Met Phe Gln Ser Glu
1 5 10 15
Gln Gly Ile Ala Arg Tyr Cys Leu Arg Pro Val Leu Thr Thr Pro Leu
 20 25 30
Tyr Leu Asn Leu Lys Asp Met Ile Arg Asn Cys Phe Pro Ser Ser Leu
 35 40 45
Leu Leu Gln Ile

50

<210> 158

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M51 peptide

<400> 158

Ile Pro Asp Gly Val Val Gly Val Ser Cys Gln Met

1 5 10

<210> 159

<211> 21

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M52 peptide

<400> 159

Ile Pro Asp Gly Val Val Gly Asp Val Leu Ser His Val Ser Val Arg

1 5 10 15

Ala Arg Asn Ser Lys

 20

<210> 160

<211> 23

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M53 peptide

<400> 160

Ile Pro Asp Gly Val Val Gly Val Arg Tyr Ala Arg Cys Ser Val Ser

1 5 10 15

Cys Val Ser Pro Ser Lys Glu

 20

<210> 161

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M54 peptide

<400> 161

Ile Pro Asp Gly Val Val Gly Ala Phe Thr His Ile Phe Cys Asp

1 5 10 15

<210> 162

<211> 24

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, M55 peptide

<400> 162

Ile Pro Asp Gly Val Thr Pro Asp Met Pro Asp Val Leu Ser His Val

1 5 10 15

Ser Val Arg Ala Arg Asn Ser Lys

20

<210> 163

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, SV40 NLS coding sequence

<400> 163

ccgaagaaga agcgcaaggt g 21

<210> 164

<211> 96

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, NLS1 coding sequence

<400> 164

atgcctaccg aggaaagagt gaggaaaaga aaggaatcca atagagaatc agccagacgc 60

tccagataca ggaaagccgc tcacctgaaa gaactg 96

<210> 165

<211> 81

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, NLS3 coding sequence

<400> 165

atggccagga aaagaaagga atccaataga gaatcagcca gacgctccag atacaggaaa 60

gccgctcacc tgaagaact g 81

<210> 166
 <211> 72
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, NLS4 coding sequence
 <400> 166
 atggccagga aaagaaagga atccaataga gaatcagcca gacgctccag acgctccaga 60
 tacaggaagg tg 72
 <210> 167
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, NLS5 coding sequence
 <400> 167
 atgtcggagc gaaagcgacg agagaagctc 30
 <210> 168
 <211> 39
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, NLS6 coding sequence
 <400> 168
 atgatcagcg aggctcttcg caaagctata gggaagcgg 39
 <210> 169
 <211> 10
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, NLS5 peptide
 <400> 169
 Met Ser Glu Arg Lys Arg Arg Glu Lys Leu
 1 5 10
 <210> 170
 <211> 13
 <212> PRT
 <213> Artificial Sequence
 <220>

<223> Synthetic construct, NLS6 peptide
 <400> 170
 Met Ile Ser Glu Ala Leu Arg Lys Ala Ile Gly Lys Arg
 1 5 10
 <210> 171
 <211> 204
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1)..(204)
 <223> T1_ZmGWDmega-2F-2R
 <400> 171
 ggttataagc ccggttgaag tatcaggta tgtggttgtg gttgatgagt tacttgctgt 60
 ccagaacaaa tcttatgata aaccaacct cettgtggca aagagtgtca agggagagga 120
 agaaatacca gatggagtag ttggtgtaat tacacctgat atgccagatg ttctgtctca 180
 tgtgtcagtc cgagcaagga atag 204
 <210> 172
 <211> 560
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1)..(560)
 <223> T2_GWDex23-F + ZmGWDmega-2R
 <400> 172
 tgctcttctg aaccgatttg atcctgtttt aaggaatggt gctcacctcg gaaggtaaaa 60
 atgtaaaatc tatgactgct gttgaacttc ttttactttg tatccccagt atatgaacac 120
 ataattctaa ggactacttt gggaaactcaa atccccttcg ggattgaagg ggattggaga 180
 ggaagttagt ttattttcac ctcaatctc tectateccg aaggggattt gaggttccca 240
 aagtagccct aaaagtgata ctagtgacce tetccacaat tttatgcgaa ccacagaaat 300
 taataatata ttctattact ctgcacctga catctggetc ctgctatcag ttggcaggtt 360
 ataagcccgg ttgaagtate aggttatgtg gttgtggttg atgagttact tgctgtccag 420
 aacaaatctt atgataaacc aaccatcett gtggcaaaga gtgtcaaggg agaggaagaa 480
 ataccagatg gagtagttgg tgtaattaca cctgatatgc cagatgttct gtctcatgtg 540
 tcagtccgag caaggaatag 560
 <210> 173
 <211> 613
 <212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1)..(613)

<223> T3_2856 + 2858

<400> 173

```
gaaggggatt ggagaggaag ttagtttatt ttcacctcaa tcctctccta tcccgaagg 60
gatttgaggt tcccaaagta gccctaaaag tgatactagt gaccctctcc acaattttat 120
gcgaaccaca gaaattaata atatattcta ttactctgea cctgacatct ggctcctgct 180
atcagttggc aggttataag cccggttgaa gtatcagggt atgtggttgt ggttgatgag 240
ttacttgctg tccagaacaa atcttatgat aaaccaacca tccttggtgc aaagagtgtc 300
aagggagagg aagaaatacc agatggagta gttggtgtaa ttacacctga tatgccagat 360
gttctgtctc atgtgtcagt ccgagcaagg aatagcaagg tttatcttca cagctatggt 420
gcaagatttc ttgaattttt tctcttgat tgatgttgac atactagctt tttcctaata 480
aaggtactgt ttgcgacctg ttttgaccac accactctat ctgaacttga aggatatgat 540
cagaaactgt tttccttcaa gcctacttct gcagatataa cctataggta cttgaggcta 600
tttgaacgtc atg 613
```

<210> 174

<211> 381

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1)..(381)

<223> T4_ZmGWDmega-2F + 429

<400> 174

```
ggttataagc ccggttgaag tadcaggta tgtggttgtg gttgatgagt tacttgctgt 60
ccagaacaaa tcttatgata aaccaacct ccttggtgca aagagtgtca aggagagga 120
agaaatacca gatggagtag ttggtgtaat tacacctgat atgccagatg ttctgtctca 180
tgtgtcagtc cgagcaagga atagcaaggt ttatcttcaac agctatgttg caagatttct 240
tgaatttttt ctcttgatt gatgttgaca tactagcttt ttcctaataga aggtactggt 300
tgcgacctgt tttgaccaca ccactctate tgaacttgaa ggatgatgatc agaaactggt 360
ttccttcaag cctacttctg c 381
```

<210> 175

<211> 737

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (737)

<223> T5_GWDex23-F + 429

<400> 175

```

tgctcttctg aaccgatttg atcctgtttt aaggaatggt gctcacctcg gaaggtaaaa 60
atgtaaaatc tatgactgct gttgaacttc ttttactttg tatccccagt atatgaacac 120
ataattctaa ggactacttt gggaaactcaa atccccctcg ggattgaagg ggattggaga 180
ggaagttagt ttattttcac ctcaatcctc tctatccccg aaggggattt gaggttccca 240
aagtagccct aaaagtgata ctagtgacc tctccacaat tttatgcaa ccacagaaat 300
taataatata ttctattact ctgcacctga catctggctc ctgctatcag ttggcagggt 360
ataagcccggttgaagtate aggttatgtg gttgtgggtg atgagttact tgctgtccag 420
aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg agaggaagaa 480
ataccagatg gagtagttgg tgtaattaca cctgatatgc cagatgttct gtctcatgtg 540
tcagtccgag caaggaatag caaggtttat cttcacagct atgttgcaag atttcttgaa 600
ttttttctct tgtattgatg ttgacatact agctttttcc taatgaaggt actgtttgcg 660
acctgttttg accacaccac tctatctgaa cttgaaggat atgatcagaa actgttttcc 720
ttcaagccta cttctgc 737

```

<210> 176

<211> 778

<212> DNA

<213> *Zea mays*

<220>

<221> misc_feature

<222> (1) .. (778)

<223> T6_GWDex23-F + 2858

<400> 176

```

tgctcttctg aaccgatttg atcctgtttt aaggaatggt gctcacctcg gaaggtaaaa 60
atgtaaaatc tatgactgct gttgaacttc ttttactttg tatccccagt atatgaacac 120
ataattctaa ggactacttt gggaaactcaa atccccctcg ggattgaagg ggattggaga 180
ggaagttagt ttattttcac ctcaatcctc tctatccccg aaggggattt gaggttccca 240
aagtagccct aaaagtgata ctagtgacc tctccacaat tttatgcaa ccacagaaat 300
taataatata ttctattact ctgcacctga catctggctc ctgctatcag ttggcagggt 360
ataagcccggttgaagtate aggttatgtg gttgtgggtg atgagttact tgctgtccag 420
aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg agaggaagaa 480
ataccagatg gagtagttgg tgtaattaca cctgatatgc cagatgttct gtctcatgtg 540
tcagtccgag caaggaatag caaggtttat cttcacagct atgttgcaag atttcttgaa 600
ttttttctct tgtattgatg ttgacatact agctttttcc taatgaaggt actgtttgcg 660
acctgttttg accacaccac tctatctgaa cttgaaggat atgatcagaa actgttttcc 720
ttcaagccta cttctgcaga tataacctat aggtacttga ggctatttga acgtcatg 778

```

<210> 177

<211> 395

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (395)

<223> T7_2856 + ZmGWDmega-2R

<400> 177

```
gaaggggatt ggagaggaag ttagtttatt ttcacctca tctctccta tcccgaagg 60
gatttgaggt tcccaaagta gccctaaaag tgatactagt gaccctctcc acaattttat 120
gcgaaccaca gaaattaata atatattcta ttactctgca cctgacatct ggctcctgct 180
atcagttggc aggttataag cccggttgaa gtatcaggtt atgtggttgt ggttgatgag 240
ttacttgctg tccagaacaa atcttatgat aaaccaacca tcttgtggc aaagagtgtc 300
aagggagagg aagaaatacc agatggagta gttggtgtaa ttacacctga tatgccagat 360
gttctgtctc atgtgtcagt ccgagcaagg aatag 395
```

<210> 178

<211> 572

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (572)

<223> T8_2856 + 429

<400> 178

```
gaaggggatt ggagaggaag ttagtttatt ttcacctca tctctccta tcccgaagg 60
gatttgaggt tcccaaagta gccctaaaag tgatactagt gaccctctcc acaattttat 120
gcgaaccaca gaaattaata atatattcta ttactctgca cctgacatct ggctcctgct 180
atcagttggc aggttataag cccggttgaa gtatcaggtt atgtggttgt ggttgatgag 240
ttacttgctg tccagaacaa atcttatgat aaaccaacca tcttgtggc aaagagtgtc 300
aagggagagg aagaaatacc agatggagta gttggtgtaa ttacacctga tatgccagat 360
gttctgtctc atgtgtcagt ccgagcaagg aatagcaagg tttatcttca cagctatggt 420
gcaagatttc ttgaattttt tctcttgtat tgatgttgac ataactagctt tttcctaag 480
aaggtactgt ttgcgacctg ttttgaccac accactetat ctgaacttga aggatatgat 540
cagaaactgt tttccttcaa gcctaacttct gc 572
```

<210> 179

<211> 381

<212> DNA

<213> Zea mays

<220>

<221> misc_feature
 <222> (1) .. (381)
 <223> T9_ZmGWDmega-2F+429
 <400> 179
 ggttataagc ccggttgaag tadcaggta tgtggttggtg gttgatgagt tacttgctgt 60
 ccagaacaaa tcttatgata aaccaacat ccttgtggca aagagtgtca aggagagga 120
 agaaatacca gatggagtag ttggtgtaat tacacctgat atgccagatg ttctgtctca 180
 tgtgtcagtc cgagcaagga atagcaaggt ttatcttcac agctatggtg caagatttct 240
 tgaatttttt ctcttgatt gatgttgaca tactagcttt ttctaataga aggtactgtt 300
 tgcgacctgt tttgaccaca cactctatc tgaacttgaa ggatatgatc agaaactgtt 360
 ttcttcaag cctacttctg c 381
 <210> 180
 <211> 422
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (422)
 <223> T10_ZmGWDmega-2F+ 2858
 <400> 180
 ggttataagc ccggttgaag tadcaggta tgtggttggtg gttgatgagt tacttgctgt 60
 ccagaacaaa tcttatgata aaccaacat ccttgtggca aagagtgtca aggagagga 120
 agaaatacca gatggagtag ttggtgtaat tacacctgat atgccagatg ttctgtctca 180
 tgtgtcagtc cgagcaagga atagcaaggt ttatcttcac agctatggtg caagatttct 240
 tgaatttttt ctcttgatt gatgttgaca tactagcttt ttctaataga aggtactgtt 300
 tgcgacctgt tttgaccaca cactctatc tgaacttgaa ggatatgatc agaaactgtt 360
 ttcttcaag cctacttctg cagatataac ctataggtac ttgaggctat ttgaacgtca 420
 tg 422
 <210> 181
 <211> 208
 <212> DNA
 <213> Sorghum bicolor
 <220>
 <221> misc_feature
 <222> (1) .. (208)
 <223> T11_SbGWDmega-2F + ZmGWDmega-2R
 <400> 181
 ggcaggttat aagcccagtt gaagtatcag gttatgtggt tgtggttgat gagttacttg 60
 ctgtccagaa caaatcttat gataaaccaa ccatccttgt ggcaaagagt gtcaaggag 120

aggaagaaat accagatgga gtagttggtg taattacacc tgatatgcca gatgttctgt 180
 cccatgtgtc agtccgagca aggaatag 208

<210> 182

<211> 214

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (214)

<223> Zm GWD Exon 24

<400> 182

ttggcaggtt ataagcccgg ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
 agaggaagaa ataccagatg gagtagttgg tgtaattaca cctgatatgc cagatgttct 180
 gtctcatgtg tcagtccgag caaggaatag caag 214

<210> 183

<211> 214

<212> DNA

<213> Sorghum bicolor

<220>

<221> misc_feature

<222> (1) .. (214)

<223> SbGWD Exon 24

<400> 183

ttggcaggtt ataagcccag ttgaagtatc aggttatgtg gttgtggttg atgagttact 60
 tgctgtccag aacaaatctt atgataaacc aaccatcctt gtggcaaaga gtgtcaaggg 120
 agaggaagaa ataccagatg gagtagttgg tgtaattaca cctgatatgc cagatgttct 180
 gtcccatgtg tcagtccgag caaggaatag caag 214

<210> 184

<211> 234

<212> DNA

<213> Sorghum bicolor

<220>

<221> misc_feature

<222> (1) .. (234)

<223> SbGWD Exon 7

<400> 184

gaggagtatg aagctgcacg agctgagtta atagaggaat taaatagagg tgtttcttta 60
 gagaagcttc gagctaaatt gacaaaaaca cctgaagcac ctgagtcaga tgaacgtaa 120

tctcctgcat ctggaatgcc cgttgataaa cttccagagg accttgtaga ggtgcaggct 180
tatataaggt gggagaaagc gggcaagcca aattatcctc ctgagaagca actg 234

<210> 185

<211> 81

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (81)

<223> ZmGWD aa1040-1120

<400> 185

Pro	Thr	Ile	Leu	Val	Ala	Lys	Ser	Val	Lys	Gly	Glu	Glu	Glu	Ile	Pro
1				5					10					15	
Asp	Gly	Val	Val	Gly	Val	Ile	Thr	Pro	Asp	Met	Pro	Asp	Val	Leu	Ser
			20					25					30		
His	Val	Ser	Val	Arg	Ala	Arg	Asn	Ser	Lys	Val	Leu	Phe	Ala	Thr	Cys
			35				40						45		
Phe	Asp	His	Thr	Thr	Leu	Ser	Glu	Leu	Glu	Gly	Tyr	Asp	Gln	Lys	Leu
			50			55					60				
Phe	Ser	Phe	Lys	Pro	Thr	Ser	Ala	Asp	Ile	Thr	Tyr	Arg	Glu	Ile	Thr
65					70					75					80
Glu															

<210> 186

<211> 76

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (76)

<223> WT ZmGWD_nt 81-160 Exon 24

<400> 186

gctcctgcta tcagttggca gggtataagc ccggttgaag taccagggtta tgggtgtgtg 60
ggtgatgagt tacttg 76

<210> 187

<211> 76

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (76)
 <223> Wt ZmGWD Exon 24
 <220>
 <221> misc_feature
 <222> (1) .. (76)
 <223> Wt ZmGWD Exon 24_nt 81-160
 <400> 187
 gctcctgcta tcagttggca ggttataage ccggttgaag tadcaggta tgtggttgtg 60
 gttgatgagt tacttg 76
 <210> 188
 <211> 77
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, M37
 <400> 188
 gctcctgcta tctagttgga aggttataag cccggttga gttatcaggtt atgtggttgt 60
 gggtgatgag ttacttg 77
 <210> 189
 <211> 67
 <212> DNA
 <213> Zea mays
 <220>
 <221> misc_feature
 <222> (1) .. (67)
 <223> Wt ZmGWD Exon 25
 <400> 189
 cactctatct gaacttgaag gatatgatca gaaactgttt tcttcaagc ctacttctgc 60
 agatata 67
 <210> 190
 <211> 47
 <212> PRT
 <213> Zea mays
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (47)
 <223> Wt ZmGWD aa 1011-1057
 <400> 190
 Trp Gln Val Ile Ser Pro Val Glu Val Ser Gly Tyr Val Val Val Val

1 5 10 15
 Asp Glu Leu Leu Ala Val Gln Asn Lys Ser Tyr Asp Lys Pro Thr Ile
 20 25 30
 Leu Val Ala Lys Ser Val Lys Gly Glu Glu Glu Ile Pro Asp Gly
 35 40 45

<210> 191

<211> 35

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (35)

<223> Wt ZmGWD aa 1082-1116

<400> 191

Val Leu Phe Ala Thr Cys Phe Asp His Thr Thr Leu Ser Glu Leu Glu
 1 5 10 15
 Gly Tyr Asp Gln Lys Leu Phe Ser Phe Lys Pro Thr Ser Ala Asp Ile
 20 25 30

Thr Tyr Arg

35

<210> 192

<211> 57

<212> DNA

<213> Zea mays

<220>

<221> misc_feature

<222> (1) .. (57)

<223> Wt ZmGWD_nt 3157-3213

<400> 192

gaaataccag atggagtagt tgggtgtaatt acacctgata tgccagatgt tctgtct 57

<210> 193

<211> 28

<212> PRT

<213> Zea mays

<220>

<221> MISC_FEATURE

<222> (1) .. (28)

<223> Wt ZmGDW_aa1054-1081

<400> 193

Ile Pro Asp Gly Val Val Gly Val Ile Thr Pro Asp Met Pro Asp Val
 1 5 10 15
 Leu Ser His Val Ser Val Arg Ala Arg Asn Ser Lys
 20 25
 <210> 194
 <211> 1045
 <212> PRT
 <213> Sorghum bicolor
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (1045)
 <223> Sb4715_2 WT + del
 <400> 194
 Met Thr Gly Phe Ser Ala Ala Ala Ser Ala Ala Ala Ala Ala Glu Arg
 1 5 10 15
 Cys Ala Leu Ala Ile Arg Ala Arg Pro Ala Ala Ser Ser Pro Ala Lys
 20 25 30
 Arg Gln Gln Gln Ser Ala Ser Leu Arg Arg Ser Gly Gly Gln Arg Arg
 35 40 45
 Pro Thr Thr Leu Ala Ala Ser Arg Arg Ser Pro Val Val Val Pro Arg
 50 55 60
 Ala Ile Ala Thr Ser Ala Asp Arg Ala Ser His Asp Leu Val Gly Lys
 65 70 75 80
 Phe Thr Leu Asp Ser Asn Ser Glu Leu Leu Val Ala Val Asn Pro Ala
 85 90 95
 Pro Gln Gly Leu Val Ser Val Ile Gly Leu Glu Val Thr Asn Thr Ser
 100 105 110
 Gly Ser Leu Ile Leu His Trp Gly Val Leu Arg Pro Asp Lys Arg Asp
 115 120 125
 Trp Ile Leu Pro Ser Arg Gln Pro Asp Gly Thr Thr Val Tyr Lys Asn
 130 135 140
 Arg Ala Leu Arg Thr Pro Phe Val Lys Ser Gly Asp Asn Ser Thr Leu
 145 150 155 160
 Arg Ile Glu Ile Asp Asp Pro Ala Val Gln Ala Ile Glu Phe Leu Ile
 165 170 175
 Phe Gly Glu Thr Gln Asn Lys Trp Phe Lys Asn Asn Gly Gln Asn Phe
 180 185 190
 Gln Ile Gln Leu Gln Ser Ser Arg His Gln Gly Asn Gly Ala Ser Gly
 195 200 205

Ala Ser Ser Ser Ala Thr Ser Thr Leu Val Pro Glu Asp Leu Val Gln
 210 215 220
 Ile Gln Ala Tyr Leu Arg Trp Glu Arg Lys Gly Lys Gln Ser Tyr Thr
 225 230 235 240
 Pro Glu Gln Glu Lys Glu Glu Tyr Glu Ala Ala Arg Ala Glu Leu Ile
 245 250 255
 Glu Glu Leu Asn Arg Gly Val Ser Leu Glu Lys Leu Arg Ala Lys Leu
 260 265 270
 Thr Lys Thr Pro Glu Ala Pro Glu Ser Asp Glu Arg Lys Ser Pro Ala
 275 280 285
 Ser Arg Met Pro Val Asp Lys Leu Pro Glu Asp Leu Val Gln Val Gln
 290 295 300
 Ala Tyr Ile Arg Trp Glu Lys Ala Gly Lys Pro Asn Tyr Pro Pro Glu
 305 310 315 320
 Lys Gln Leu Val Glu Leu Glu Glu Ala Arg Lys Glu Leu Gln Ala Glu
 325 330 335
 Val Asp Lys Gly Ile Ser Ile Asp Gln Leu Arg Gln Lys Ile Leu Lys
 340 345 350
 Gly Asn Ile Glu Ser Lys Val Ser Lys Gln Leu Lys Asn Lys Lys Tyr
 355 360 365
 Phe Ser Val Glu Arg Ile Gln Arg Lys Lys Arg Asp Ile Met Gln Leu
 370 375 380
 Leu Ser Lys His Lys His Thr Val Met Glu Glu Lys Val Glu Val Ala
 385 390 395 400
 Pro Lys Gln Pro Thr Val Leu Asp Leu Phe Thr Lys Ser Leu His Glu
 405 410 415
 Lys Asp Gly Cys Glu Val Leu Ser Arg Lys Leu Phe Lys Phe Gly Asp
 420 425 430
 Lys Glu Ile Leu Ala Ile Ser Thr Lys Val Gln Asn Lys Thr Glu Val
 435 440 445
 His Leu Ala Thr Asn His Thr Glu Pro Leu Ile Leu His Trp Ser Leu
 450 455 460
 Ala Lys Lys Ala Gly Glu Trp Lys Ala Pro Pro Ser Asn Ile Leu Pro
 465 470 475 480
 Ser Gly Ser Lys Leu Leu Asp Met Ala Cys Glu Thr Glu Phe Thr Arg
 485 490 495
 Ser Glu Leu Asp Gly Leu Cys Tyr Gln Val Val Glu Ile Glu Leu Asp
 500 505 510
 Asp Gly Gly Tyr Lys Gly Met Pro Phe Val Leu Arg Ser Gly Glu Thr

515	520	525
Trp Ile Lys Asn Asn Gly Ser Asp Phe Phe Leu Asp Phe Ser Thr Arg		
530	535	540
Asp Thr Arg Asn Ile Lys Leu Lys Asp Asn Gly Asp Ala Gly Lys Gly		
545	550	555
Thr Ala Lys Ala Leu Leu Glu Arg Ile Ala Asp Leu Glu Glu Asp Ala		
565	570	575
Gln Arg Ser Leu Met His Arg Phe Asn Ile Ala Ala Asp Leu Ala Asp		
580	585	590
Glu Ala Arg Asp Ala Gly Leu Leu Gly Ile Val Gly Leu Phe Val Trp		
595	600	605
Ile Arg Phe Met Ala Thr Arg Gln Leu Thr Trp Asn Lys Asn Tyr Asn		
610	615	620
Val Lys Pro Arg Glu Ile Ser Lys Ala Gln Asp Arg Phe Thr Asp Asp		
625	630	635
Leu Glu Asn Met Tyr Arg Thr Tyr Pro Gln Tyr Arg Glu Ile Leu Arg		
645	650	655
Met Ile Met Ala Ala Val Gly Arg Gly Gly Glu Gly Asp Val Gly Gln		
660	665	670
Arg Ile Arg Asp Glu Ile Leu Val Ile Gln Arg Asn Asn Asp Cys Lys		
675	680	685
Gly Gly Met Met Glu Glu Trp His Gln Lys Leu His Asn Asn Thr Ser		
690	695	700
Pro Asp Asp Val Val Ile Cys Gln Ala Leu Ile Asp Tyr Ile Lys Asn		
705	710	715
Asp Phe Asp Ile Ser Val Tyr Trp Asp Thr Leu Asn Lys Asn Gly Ile		
725	730	735
Thr Lys Glu Arg Leu Leu Ser Tyr Asp Arg Ala Ile His Ser Glu Pro		
740	745	750
Asn Phe Arg Ser Glu Gln Lys Glu Gly Leu Leu Arg Asp Leu Gly Asn		
755	760	765
Tyr Met Arg Ser Leu Lys Ala Val His Ser Gly Ala Asp Leu Glu Ser		
770	775	780
Ala Ile Ala Thr Cys Met Gly Tyr Lys Ser Glu Gly Glu Gly Phe Met		
785	790	795
Val Gly Val Gln Ile Asn Pro Val Lys Gly Leu Pro Ser Gly Phe Pro		
805	810	815
Glu Leu Leu Glu Phe Val Leu Asp His Val Glu Asp Lys Ser Ala Glu		
820	825	830

Pro Leu Leu Glu Gly Leu Leu Glu Ala Arg Val Asp Leu Arg Pro Leu
 835 840 845
 Leu Leu Asp Ser Pro Glu Arg Met Lys Asp Leu Ile Phe Leu Asp Ile
 850 855 860
 Ala Leu Asp Ser Thr Phe Arg Thr Ala Ile Glu Arg Ser Tyr Glu Glu
 865 870 875 880
 Leu Asn Asp Ala Ala Pro Glu Lys Ile Met Tyr Phe Ile Ser Leu Val
 885 890 895
 Leu Glu Asn Leu Ala Phe Ser Ile Asp Asp Asn Glu Asp Ile Leu Tyr
 900 905 910
 Cys Leu Lys Gly Trp Asn Gln Ala Leu Glu Met Ala Lys Gln Lys Asp
 915 920 925
 Asp Gln Trp Ala Leu Tyr Ala Lys Ala Phe Leu Asp Arg Ile Arg Leu
 930 935 940
 Ala Leu Ala Ser Lys Gly Glu Gln Tyr His Asn Met Met Gln Pro Ser
 945 950 955 960
 Ala Glu Tyr Leu Gly Ser Leu Leu Ser Ile Asp Lys Trp Ala Val Asn
 965 970 975
 Ile Phe Thr Glu Glu Ile Ile Arg Gly Gly Ser Ala Ala Thr Leu Ser
 980 985 990
 Ala Leu Leu Asn Arg Phe Asp Pro Val Leu Arg Asn Val Ala Asn Leu
 995 1000 1005
 Gly Ser Trp Gln Val Ile Ser Pro Val Glu Val Ser Gly Tyr Val
 1010 1015 1020
 Val Val Val Asp Glu Leu Leu Ala Val Gln Asn Lys Ser Tyr Asp
 1025 1030 1035
 Lys Pro Thr Ile Leu Val Glu
 1040 1045
 <210> 195
 <211> 840
 <212> PRT
 <213> Sorghum bicolor
 <220>
 <221> MISC_FEATURE
 <222> (1) .. (840)
 <223> Sb4715_2 WT + del
 <400> 195
 Val Val Pro Arg Ala Ile Ala Thr Ser Ala Asp Arg Ala Ser His Asp
 1 5 10 15

Leu Val Gly Lys Phe Thr Leu Asp Ser Asn Ser Glu Leu Leu Val Ala
 20 25 30
 Val Asn Pro Ala Pro Gln Gly Leu Val Ser Val Ile Gly Leu Glu Val
 35 40 45
 Thr Asn Thr Ser Gly Ser Leu Ile Leu His Trp Gly Val Leu Arg Pro
 50 55 60
 Asp Lys Arg Asp Trp Ile Leu Pro Ser Arg Gln Pro Asp Gly Thr Thr
 65 70 75 80
 Val Tyr Lys Asn Arg Ala Leu Arg Thr Pro Phe Val Lys Ser Gly Asp
 85 90 95
 Asn Ser Thr Leu Arg Ile Glu Ile Asp Asp Pro Ala Val Gln Ala Ile
 100 105 110
 Glu Phe Leu Ile Phe Gly Glu Thr Gln Asn Lys Trp Phe Lys Asn Asn
 115 120 125
 Gly Gln Asn Phe Gln Ile Gln Leu Gln Ser Ser Arg His Gln Gly Asn
 130 135 140
 Gly Ala Ser Gly Ala Ser Ser Ser Ala Thr Ser Thr Leu Val Pro Glu
 145 150 155 160
 Asp Leu Val Gln Ile Gln Ala Tyr Leu Arg Trp Glu Arg Lys Gly Lys
 165 170 175
 Gln Ser Tyr Thr Pro Glu Gln Glu Lys Glu Glu Tyr Glu Ala Ala Arg
 180 185 190
 Ala Glu Leu Ile Glu Glu Leu Asn Arg Gly Val Ser Leu Glu Lys Leu
 195 200 205
 Arg Ala Lys Leu Thr Lys Thr Pro Glu Ala Pro Glu Ser Asp Glu Arg
 210 215 220
 Lys Ser Pro Ala Ser Arg Met Pro Val Asp Lys Leu Pro Glu Asp Leu
 225 230 235 240
 Val Gln Val Gln Ala Tyr Ile Arg Trp Glu Lys Ala Gly Lys Pro Asn
 245 250 255
 Tyr Pro Pro Glu Lys Gln Leu Val Glu Leu Glu Glu Ala Arg Lys Glu
 260 265 270
 Leu Gln Ala Glu Val Asp Lys Gly Ile Ser Ile Asp Gln Leu Arg Gln
 275 280 285
 Lys Ile Leu Lys Gly Asn Ile Glu Ser Lys Val Ser Lys Gln Leu Lys
 290 295 300
 Asn Lys Lys Tyr Phe Ser Val Glu Arg Ile Gln Arg Lys Lys Arg Asp
 305 310 315 320
 Ile Met Gln Leu Leu Ser Lys His Lys His Thr Val Met Glu Glu Lys

	325		330		335
Val Glu Val Ala Pro Lys Gln Pro Thr Val Leu Asp Leu Phe Thr Lys					
	340		345		350
Ser Leu His Glu Lys Asp Gly Cys Glu Val Leu Ser Arg Lys Leu Phe					
	355		360		365
Lys Phe Gly Asp Lys Glu Ile Leu Ala Ile Ser Thr Lys Val Gln Asn					
	370		375		380
Lys Thr Glu Val His Leu Ala Thr Asn His Thr Glu Pro Leu Ile Leu					
385		390		395	400
His Trp Ser Leu Ala Lys Lys Ala Gly Glu Trp Lys Ala Pro Pro Ser					
	405		410		415
Asn Ile Leu Pro Ser Gly Ser Lys Leu Leu Asp Met Ala Cys Glu Thr					
	420		425		430
Glu Phe Thr Arg Ser Glu Leu Asp Gly Leu Cys Tyr Gln Val Val Glu					
	435		440		445
Ile Glu Leu Asp Asp Gly Gly Tyr Lys Gly Met Pro Phe Val Leu Arg					
	450		455		460
Ser Gly Glu Thr Trp Ile Lys Asn Asn Gly Ser Asp Phe Phe Leu Asp					
465		470		475	480
Phe Ser Thr Arg Asp Thr Arg Asn Ile Lys Leu Lys Asp Asn Gly Asp					
	485		490		495
Ala Gly Lys Gly Thr Ala Lys Ala Leu Leu Glu Arg Ile Ala Asp Leu					
	500		505		510
Glu Glu Asp Ala Gln Arg Ser Leu Met His Arg Phe Asn Ile Ala Ala					
	515		520		525
Asp Leu Ala Asp Glu Ala Arg Asp Ala Gly Leu Leu Gly Ile Val Gly					
	530		535		540
Leu Phe Val Trp Ile Arg Phe Met Ala Thr Arg Gln Leu Thr Trp Asn					
545		550		555	560
Lys Asn Tyr Asn Val Lys Pro Arg Glu Ile Ser Lys Ala Gln Asp Arg					
	565		570		575
Phe Thr Asp Asp Leu Glu Asn Met Tyr Arg Thr Tyr Pro Gln Tyr Arg					
	580		585		590
Glu Ile Leu Arg Met Ile Met Ala Ala Val Gly Arg Gly Gly Glu Gly					
	595		600		605
Asp Val Gly Gln Arg Ile Arg Asp Glu Ile Leu Val Ile Gln Arg Asn					
	610		615		620
Asn Asp Cys Lys Gly Gly Met Met Glu Glu Trp His Gln Lys Leu His					
625		630		635	640

Asn Asn Thr Ser Pro Asp Asp Val Val Ile Cys Gln Ala Leu Ile Asp
 645 650 655
 Tyr Ile Lys Asn Asp Phe Asp Ile Ser Val Tyr Trp Asp Thr Leu Asn
 660 665 670
 Lys Asn Gly Ile Thr Lys Glu Arg Leu Leu Ser Tyr Asp Arg Ala Ile
 675 680 685
 His Ser Glu Pro Asn Phe Arg Ser Glu Gln Lys Glu Gly Leu Leu Arg
 690 695 700
 Asp Leu Gly Asn Tyr Met Arg Ser Leu Lys Ala Val His Ser Gly Ala
 705 710 715 720
 Asp Leu Glu Ser Ala Ile Ala Thr Cys Met Gly Tyr Lys Ser Glu Gly
 725 730 735
 Glu Gly Phe Met Val Gly Val Gln Ile Asn Pro Val Lys Gly Leu Pro
 740 745 750
 Ser Gly Phe Pro Glu Leu Leu Glu Phe Val Leu Asp His Val Glu Asp
 755 760 765
 Lys Ser Ala Glu Pro Leu Leu Glu Gly Leu Leu Glu Ala Arg Val Asp
 770 775 780
 Leu Arg Pro Leu Leu Leu Asp Ser Pro Glu Arg Met Lys Asp Leu Ile
 785 790 795 800
 Phe Leu Asp Ile Ala Leu Asp Ser Thr Phe Arg Thr Ala Ile Glu Arg
 805 810 815
 Ser Tyr Glu Glu Leu Asn Asp Ala Ala Pro Glu Lys Ile Met Tyr Phe
 820 825 830
 Ile Ser Leu Val Leu Glu Asn Leu
 835 840

<210> 196

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, SV NLS peptide

<400> 196

Pro Lys Lys Lys Arg Lys Val

1

5

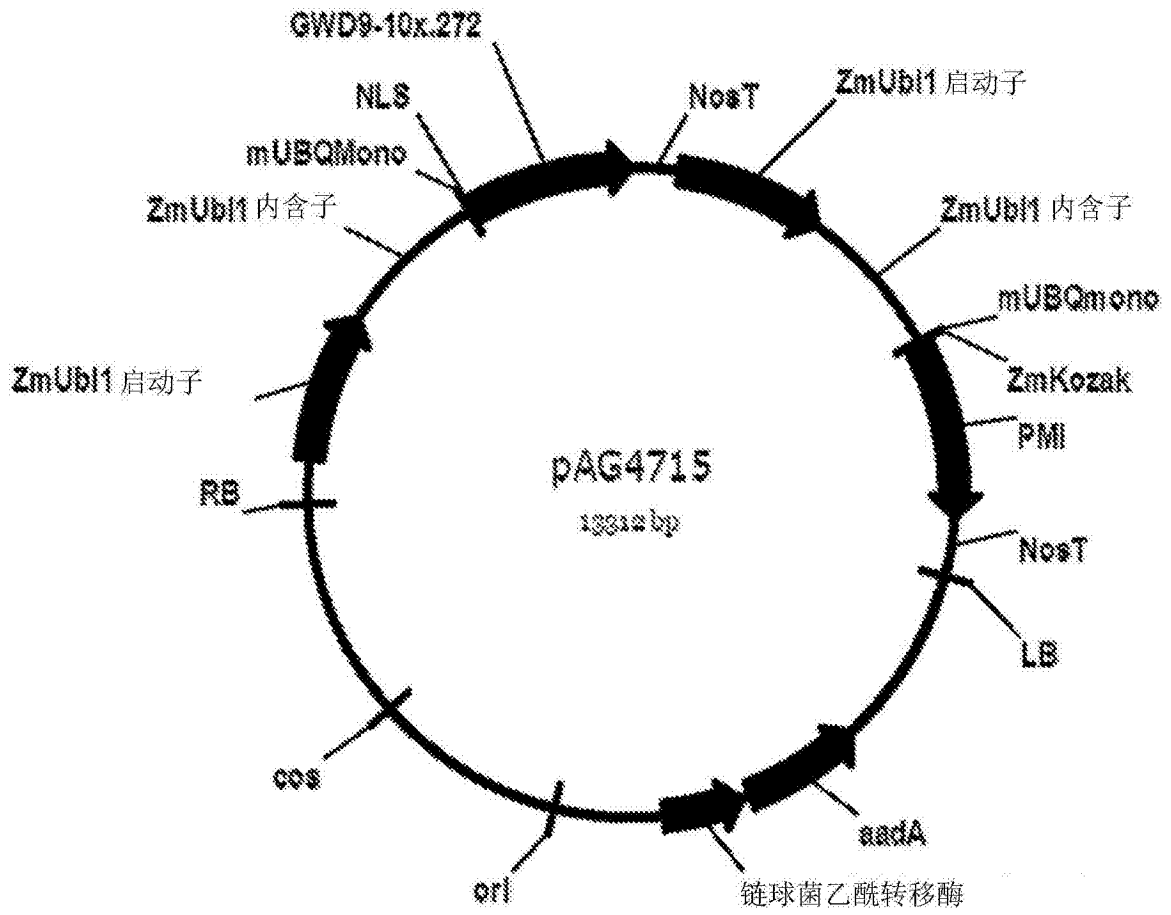


图1

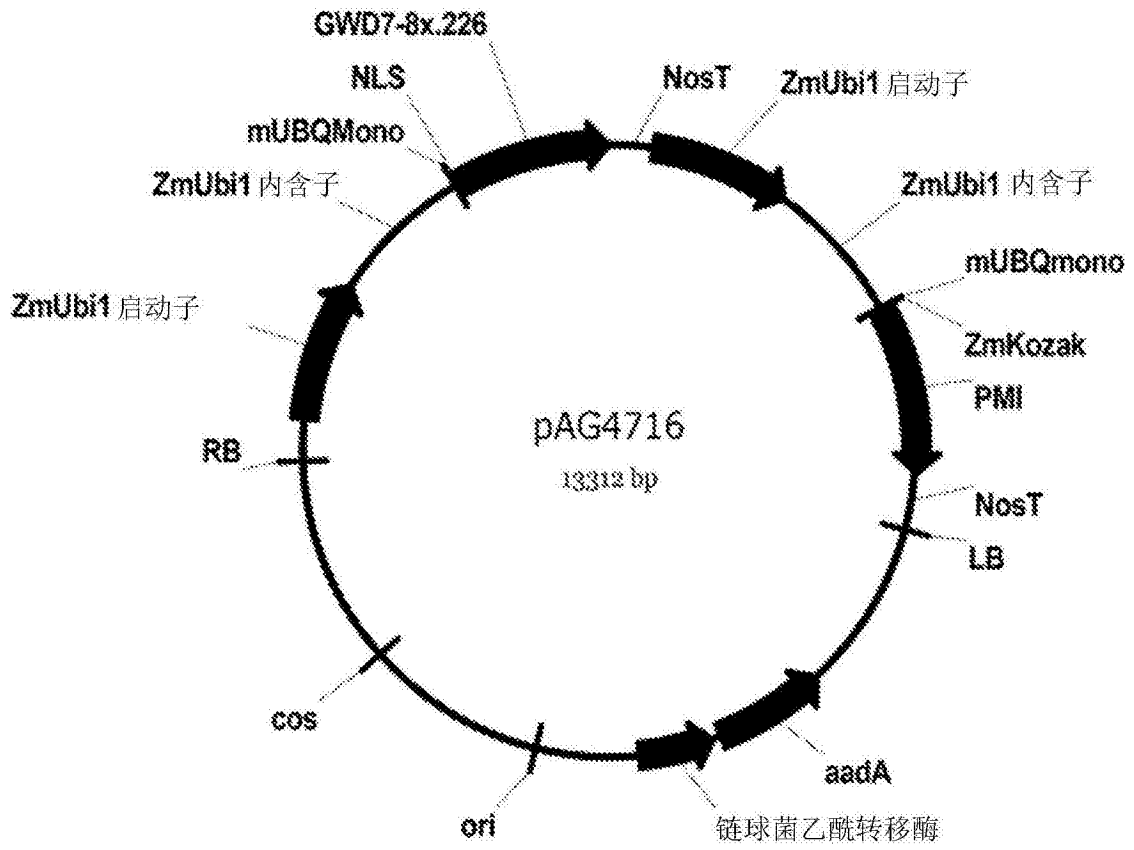


图2

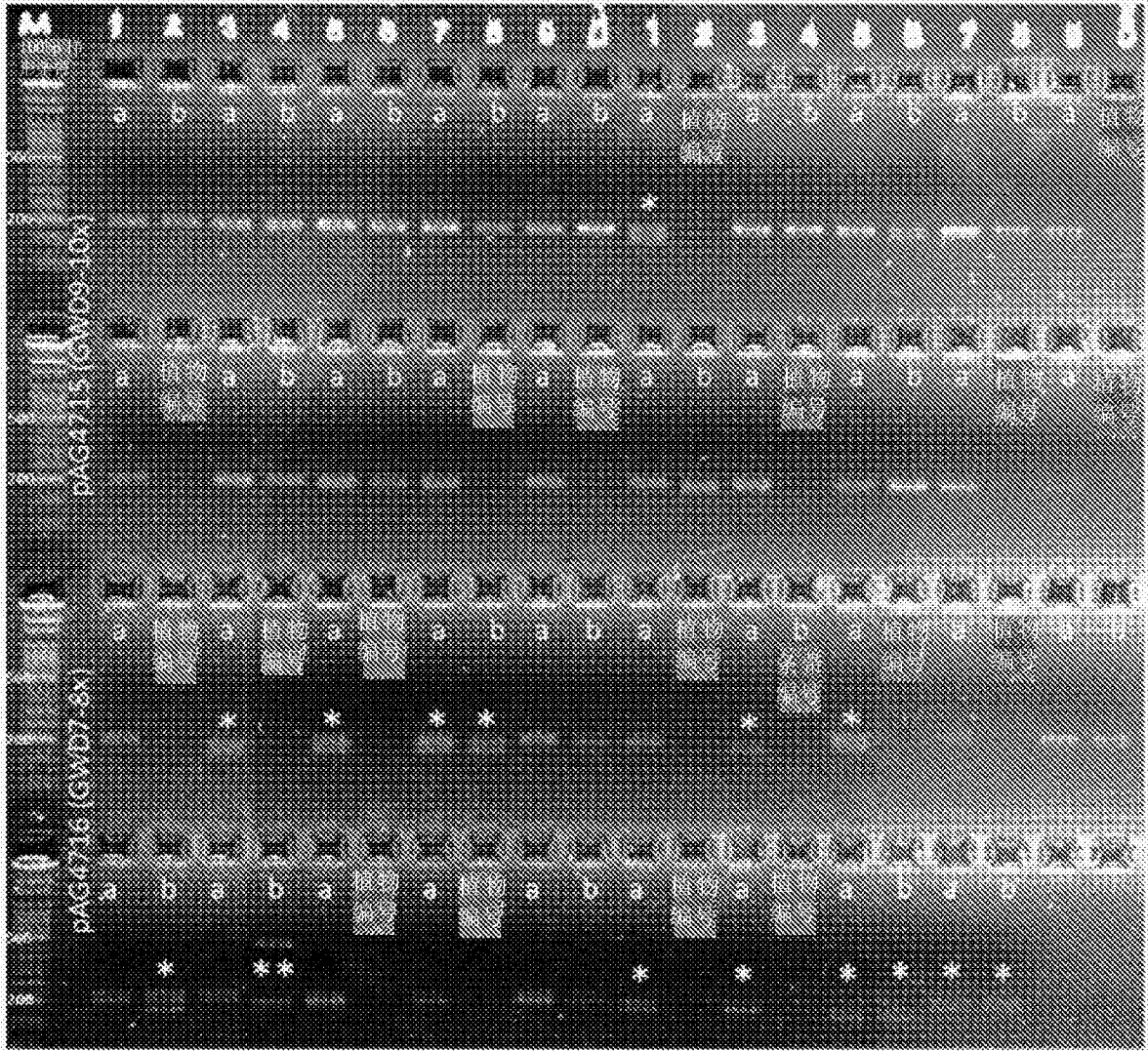


图3

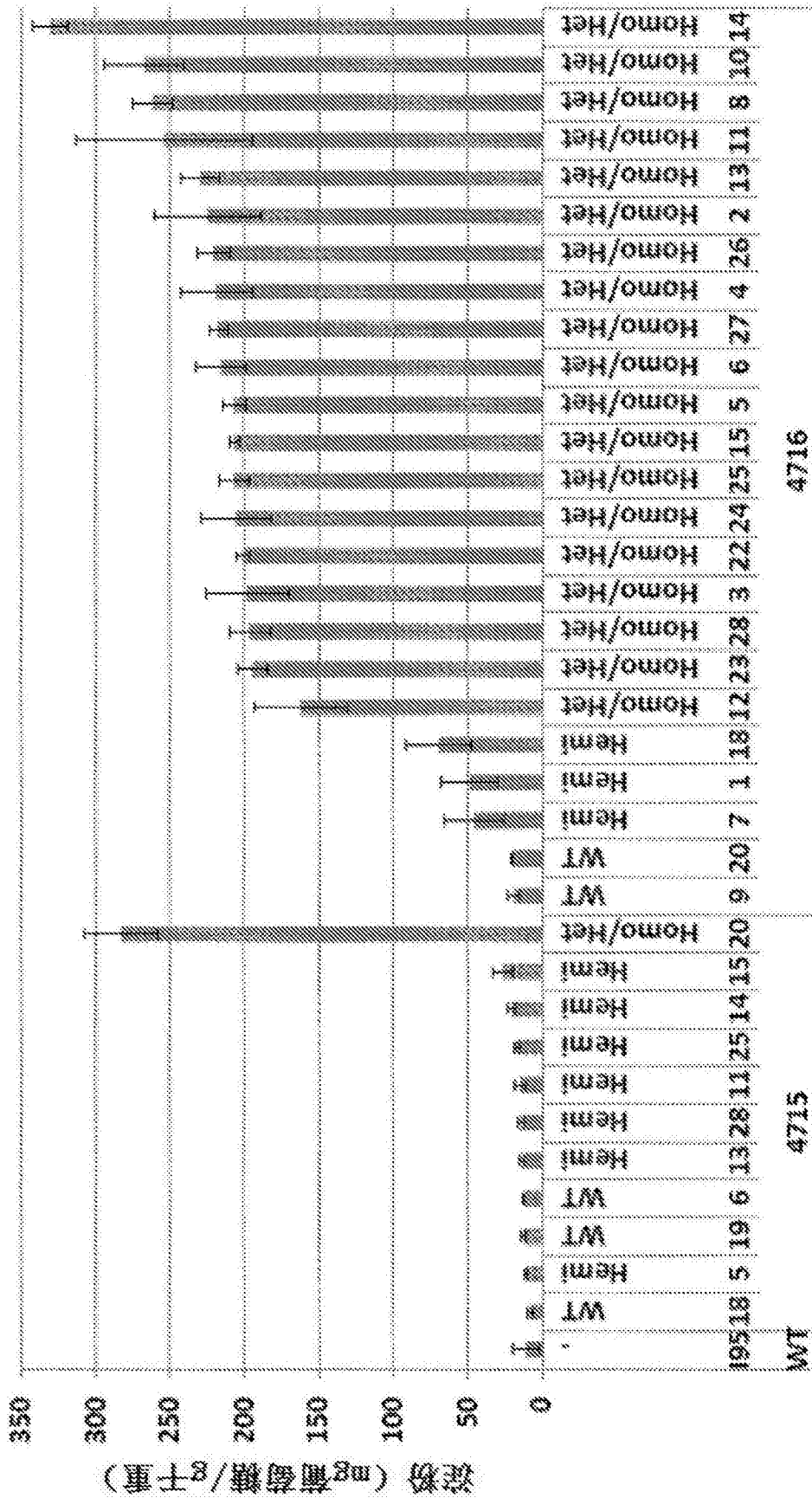


图4

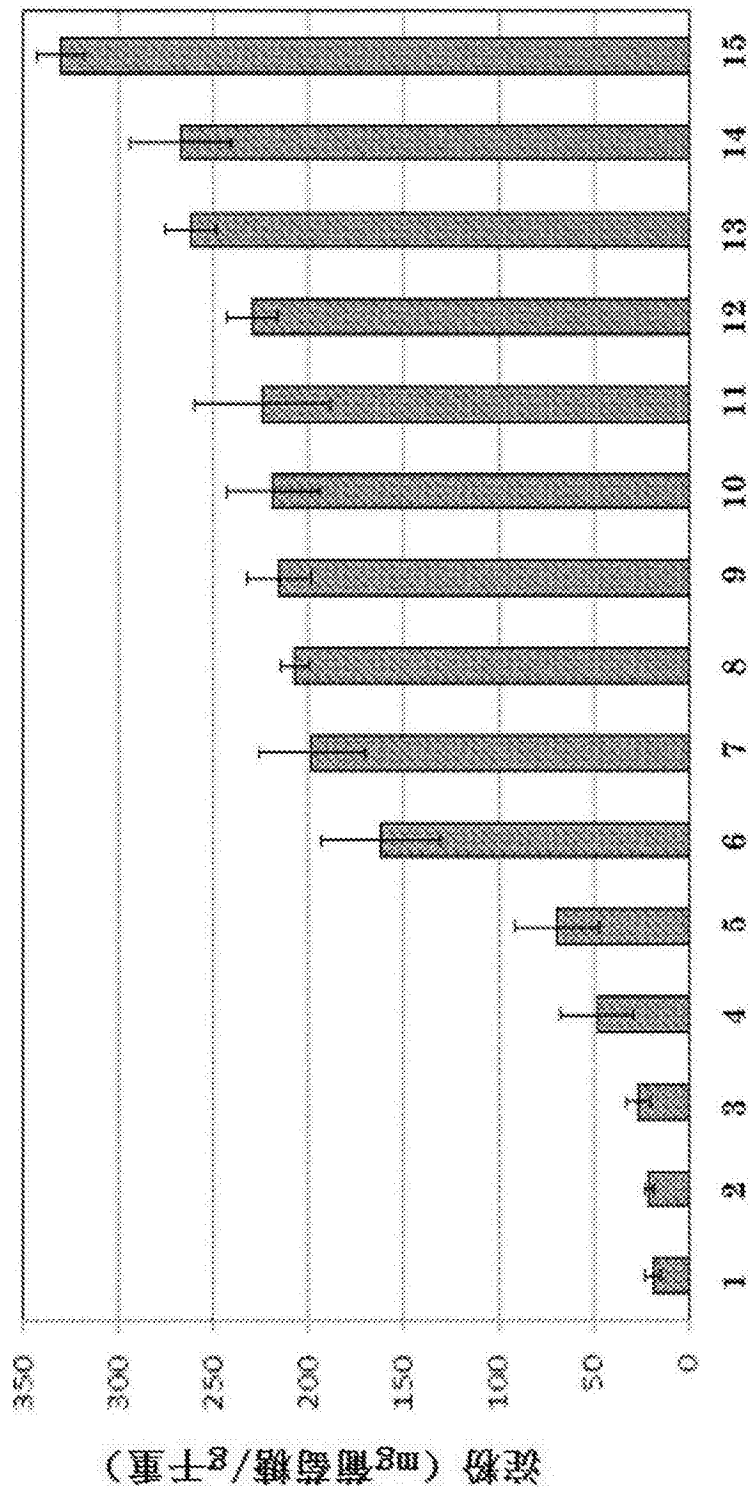


图5

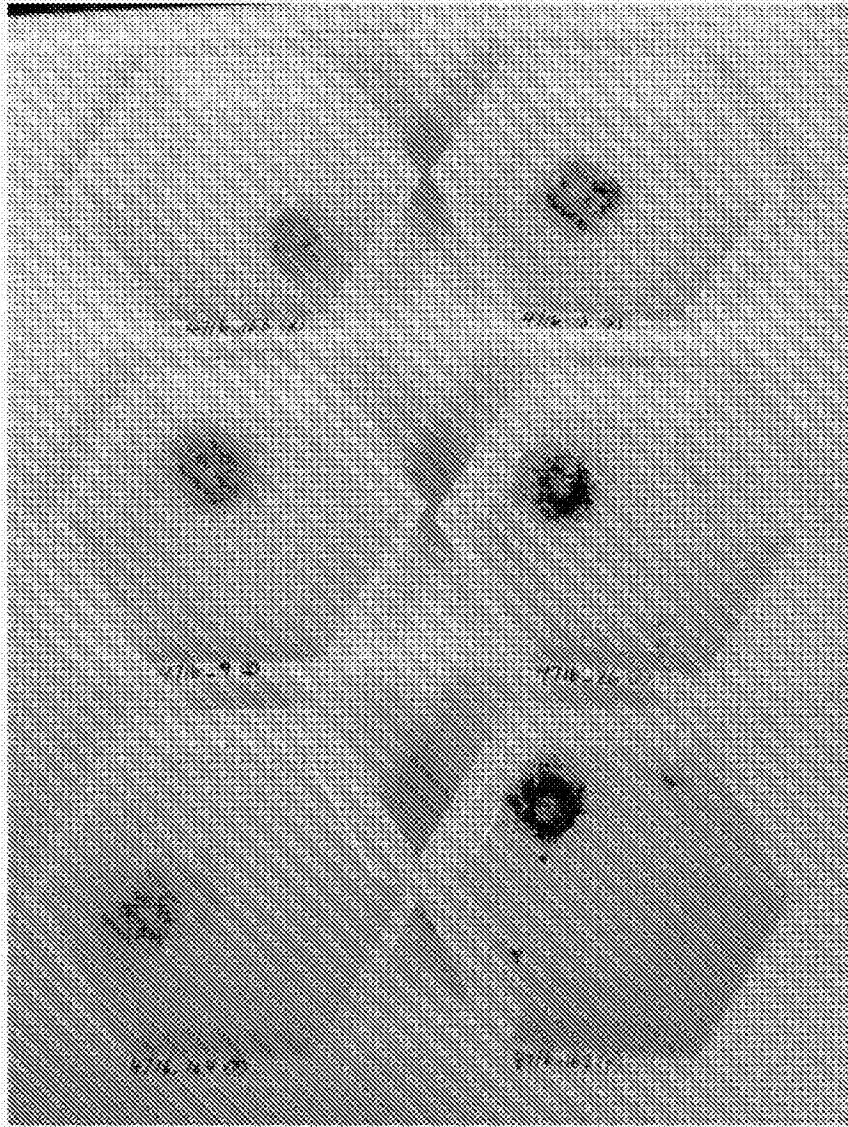


图6

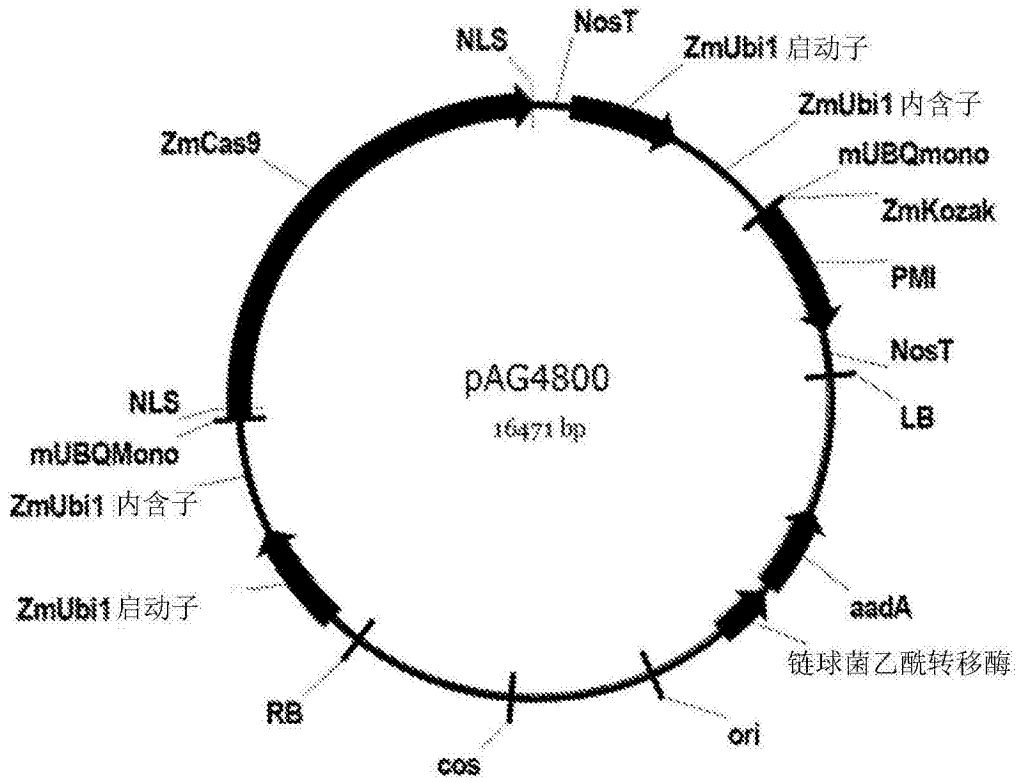


图7

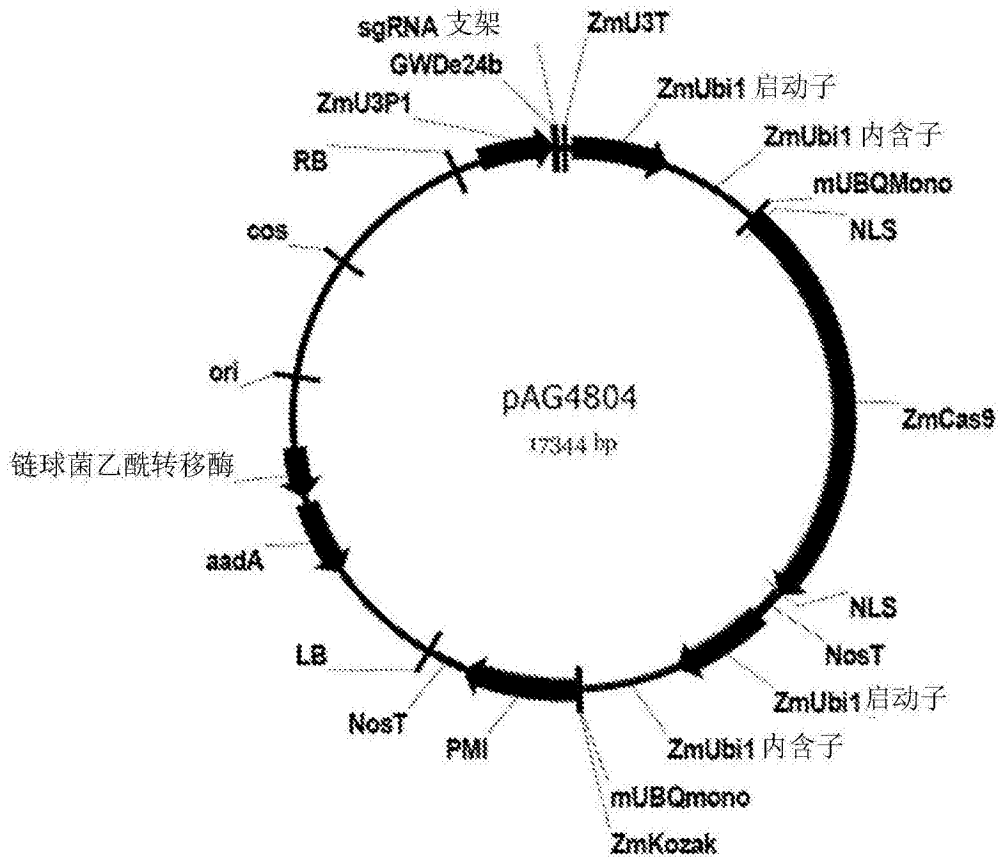


图8

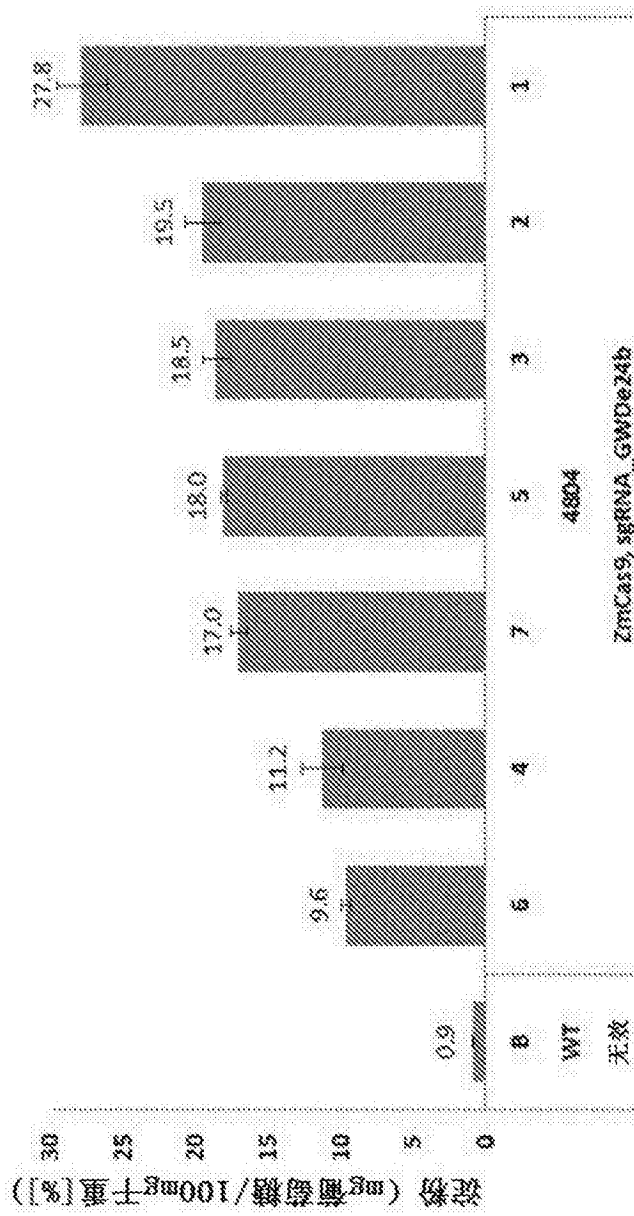


图9

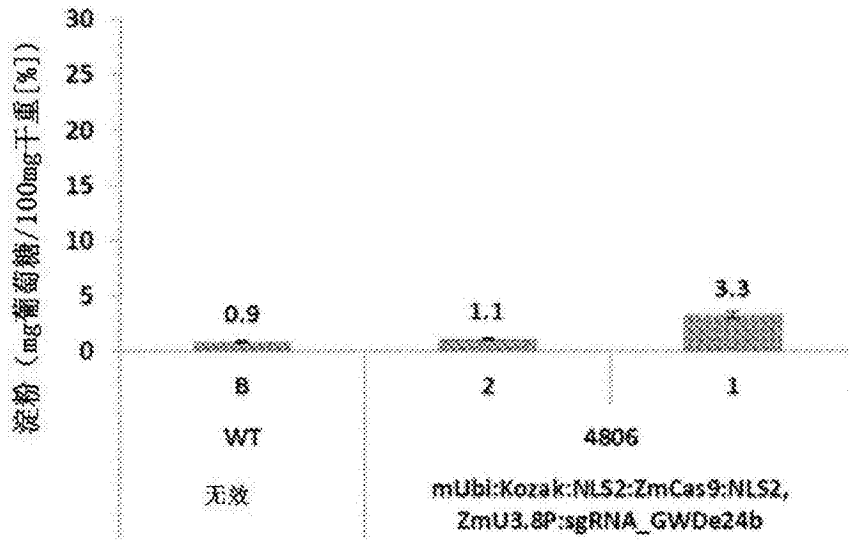


图10

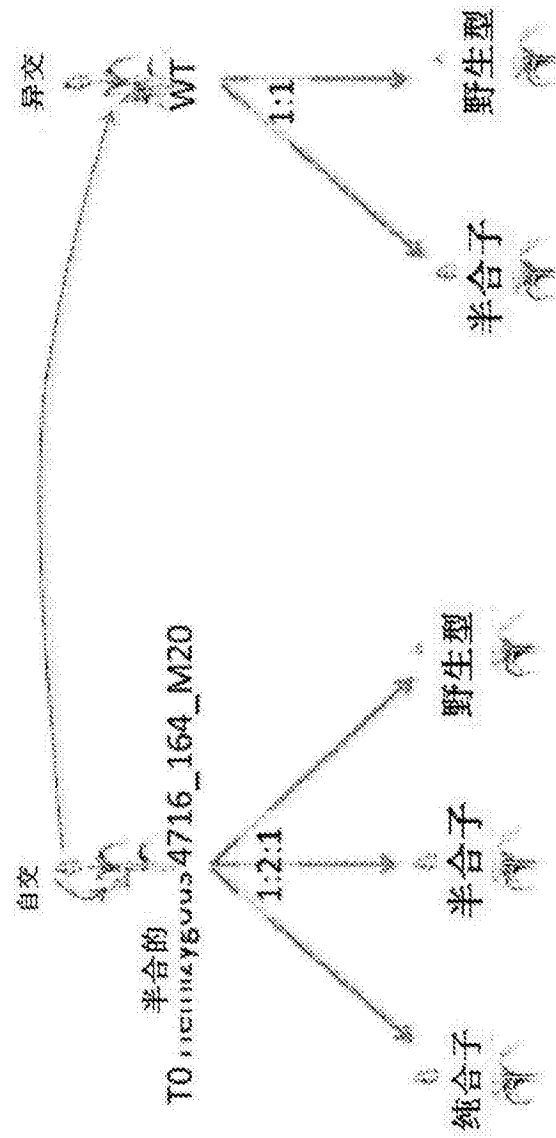


图11

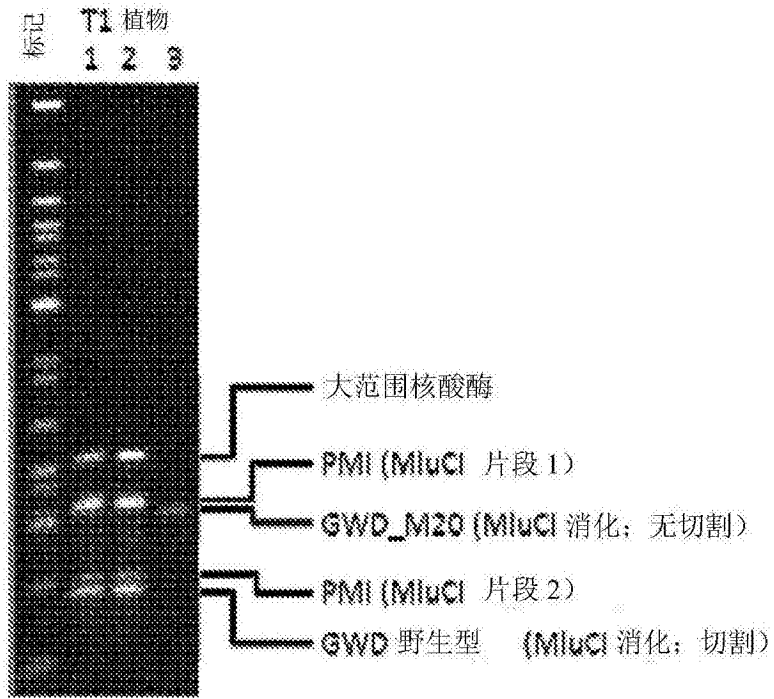


图12

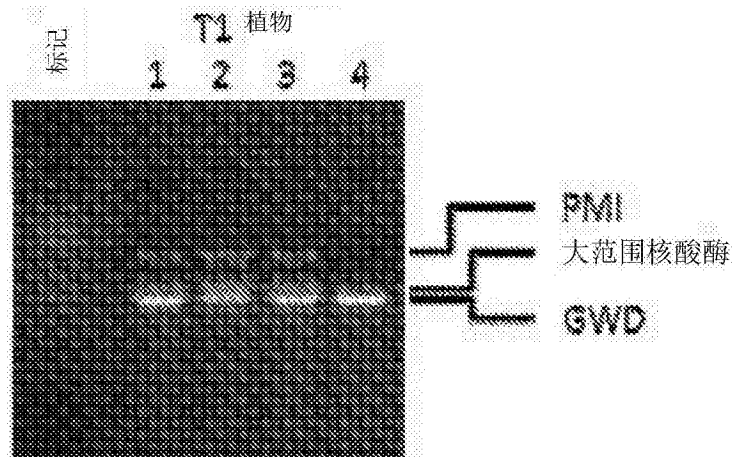


图13

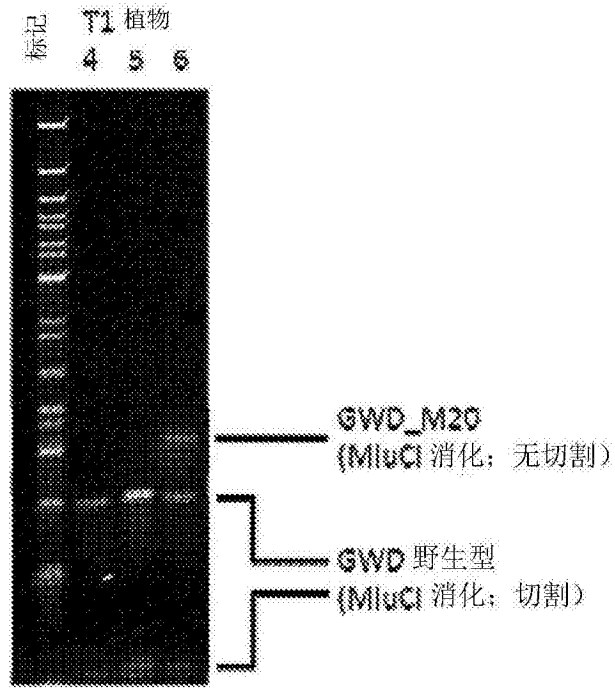


图14