



(12) 发明专利申请

(10) 申请公布号 CN 115725533 A

(43) 申请公布日 2023.03.03

(21) 申请号 202211443845.9

A01H 6/46 (2018.01)

(22) 申请日 2022.11.18

(71) 申请人 中国农业科学院作物科学研究所
地址 100081 北京市海淀区中关村南大街
12号

(72) 发明人 夏兰琴 张臣 林勇 张阳军
钟雪

(74) 专利代理机构 北京纪凯知识产权代理有限公司 11245
专利代理师 魏少伟

(51) Int. Cl.

C12N 9/10 (2006.01)

C12N 15/82 (2006.01)

C12N 15/54 (2006.01)

A01H 5/00 (2018.01)

权利要求书2页 说明书17页
序列表(电子公布) 附图3页

(54) 发明名称

一种获得抗草甘膦水稻的方法及其所用双碱基融合编辑系统

(57) 摘要

本发明公开了一种获得抗草甘膦水稻的方法及其所用双碱基融合编辑系统。本发明公开的获得抗草甘膦水稻的方法,包括:利用双碱基融合编辑系统(STCBE)将受体水稻中的EPSPS基因通过基因编辑突变为EPSPS-1422C>T基因,得到抗草甘膦水稻;EPSPS基因编码序列中序列12所示的EPSPS蛋白质;EPSPS-1422C>T基因编码序列中序列13所示的EPSPS-D213N蛋白质。本发明的STCBE系统及创制的无转基因抗草甘膦水稻新种质,为提升稻田杂草的综合治理效果,实现水稻轻简化,保障粮食安全,推动农业可持续发展,提供了重要技术和材料支撑。

1. 培育抗草甘膦水稻的方法,包括:将受体水稻中的EPSPS基因通过基因编辑突变为EPSPS-1422C>T基因,得到抗草甘膦水稻;

所述EPSPS基因编码如下A1) 或A2) 的EPSPS蛋白质:

A1) 序列表中序列12所示的EPSPS蛋白质;

A2) 将序列表中序列12所示的氨基酸序列经过一个或几个氨基酸残基的取代和/或缺失和/或添加、第213位为天冬氨酸残基且具有相同功能的EPSPS蛋白质;

所述EPSPS-1422C>T基因编码如下B1) 或B2) 的EPSPS-D213N蛋白质:

B1) 序列表中序列13所示的EPSPS-D213N蛋白质;

B2) 将序列表中序列13所示的氨基酸序列经过一个或几个氨基酸残基的取代和/或缺失和/或添加、第213位为天冬酰胺残基且具有相同功能的EPSPS-D213N蛋白质。

2. 根据权利要求1所述的方法,其特征在于:所述EPSPS基因为如下a11) 或a12):

a11) 序列表中序列10所示的DNA分子;

a12) 与a11) 具有75%或75%以上同一性,且编码所述EPSPS蛋白质的DNA分子;

所述EPSPS-1422C>T基因为如下b11) 或b12):

b11) 序列表中序列11所示的DNA分子;

b12) 与b11) 具有75%或75%以上同一性,且编码所述EPSPS-D213N蛋白质的DNA分子。

3. 根据权利要求1或2所述的方法,其特征在于:所述基因编辑采用双碱基融合编辑系统进行,所述双碱基融合编辑系统包括evoFERNY、TadA8e、nCas9-NG (D10A) 和2×UGI,所述2×UGI由两个UGI连接得到;

进一步,所述evoFERNY为序列表中序列5的第9-169位所示的蛋白质;

和/或,所述TadA8e为序列表中序列5的第218-383位所示的蛋白质;

和/或,所述nCas9-NG (D10A) 为序列表中序列5的第416-1782位所示的蛋白质;

和/或,所述UGI为序列表中序列5的第1793-1875位所示的蛋白质。

4. 根据权利要求3所述的方法,其特征在于:所述双碱基融合编辑系统由所述evoFERNY、所述TadA8e、所述nCas9-NG (D10A) 和所述2×UGI连接得到。

5. 根据权利要求1或2所述的方法,其特征在于:所述基因编辑通过向所述受体水稻导入含有权利要求3或4中所述双碱基融合编辑系统的编码基因以及转录靶向靶标DNA的gRNA的DNA片段的载体实现,所述靶标DNA含有EPSPS基因的第1422位的C。

6. 根据权利要求5所述的方法,其特征在于:所述双碱基融合编辑系统的编码基因含有所述evoFERNY的编码基因、所述TadA8e的编码基因、所述nCas9-NG (D10A) 的编码基因和所述2×UGI的编码基因;

进一步,所述evoFERNY的编码基因为序列表中序列4的第2023-2505位所示的DNA分子;

和/或,所述TadA8e的编码基因为序列表中序列4的第2650-3147位所示的DNA分子;

和/或,所述nCas9-NG (D10A) 的编码基因为序列表中序列4的第3244-7344位所示的DNA分子;

和/或,所述2×UGI的编码基因为序列表中序列4的第7375-7902位所示的DNA分子。

7. 下述任一产品:

P1) 权利要求3或4中所述双碱基融合编辑系统;

P2) 权利要求5或6中所述双碱基融合编辑系统的编码基因;

P3) 含有权利要求5或6中所述双碱基融合编辑系统的编码基因的表达盒、重组载体、重组微生物或重组细胞系。

8. 权利要求7所述产品在基因编辑或获得基因编辑产品中的应用；

进一步,所述基因编辑为C向T的突变和/或A向G的突变。

9. 权利要求1或2中所述EPSPS-D213N蛋白质或所述EPSPS-1422C>T基因。

10. 权利要求1-6中任一所述的方法,或权利要求7所述产品,或权利要求1或2中所述EPSPS-D213N蛋白质或所述EPSPS-1422C>T基因在培育抗草甘膦水稻中的应用。

一种获得抗草甘膦水稻的方法及其所用双碱基融合编辑系统

技术领域

[0001] 本发明涉及分子育种领域中,一种获得抗草甘膦水稻的方法及其所用双碱基融合编辑系统。

背景技术

[0002] 水稻是我国重要的口粮作物,提高水稻产量和品质将对缓解世界粮食危机和保障粮食安全具有十分重要的意义。然而,在水稻的生产过程中,农田杂草与水稻竞争阳光、水分、养分及生长空间,同时又是作物病原菌及害虫的中间寄主,严重影响水稻的产量与品质。有效控制田间杂草是促进粮食增产,增加农民收入,确保粮食安全的重要措施之一。因此,开发广谱、高效、低毒、易降解、无残留等灭生性除草剂或选择性除草剂耐性作物具有非常广阔的应用价值和市场潜力。

[0003] 单碱基编辑技术(Base editing)是基于CRISPR系统的新型靶基因定点修饰技术,在不产生DNA双链断裂的情况下,利用胞嘧啶脱氨酶或人工进化的腺嘌呤脱氨酶对靶位点进行精准的单碱基编辑,从而实现单核苷酸改变以产生功能获得性突变,可以加速作物的从头驯化或定向进化。单碱基编辑的工作最早来自David Liu实验室,利用活性受损的CRISPR/Cas核酸酶(无诱导DNA双链断裂的能力)与胞嘧啶脱氨基酶(Cytosine deaminase)以及尿嘧啶糖基化酶抑制蛋白(UGI)融合实现C>T的编辑(Komor et al.2016)。现有的单碱基编辑器有三大类:第一类是融合胞嘧啶脱氨酶和UGI的胞嘧啶碱基编辑器(Cytosine base editor,CBE),可催化C>T碱基的转换;第二类则是融合脱氧腺嘌呤脱氨基酶(Adenine deaminase)的腺嘌呤碱基编辑器(Adenine base editor,ABE),可实现A>G碱基的转换;第三类是融合胞嘧啶脱氨酶和尿嘧啶糖基转移酶(UNG)的碱基编辑器(C to G base editor,CGBE),可催化C>G碱基的颠换。CBE系统能催化碱基C脱氨基变成U,而U在DNA复制过程中会被识别成T,经DNA修复和复制后实现C>T的转换(Komor et al.2016)。CBE通常要与UGI融合,UGI能防止细胞内的尿嘧啶糖基化酶将碱基U糖基化引起碱基切除修复,阻止UG逆转成CG,从而提高C向T的转换效率。ABE系统在腺嘌呤脱氨酶的作用下将A脱氨变为I,经DNA修复和复制后实现A向G的转换。在实际应用过程中,常将两个腺嘌呤脱氨酶TadA串联成二聚体的形式进行使用(Gaudelli et al.2017)。为了拓展单碱基编辑系统的编辑范围,研究者使用尿嘧啶糖基转移酶(UNG)代替UGI,UNG可以切除脱氨酶产生的碱基U,形成一个无嘌呤和无嘧啶位点,启动DNA修复过程,实现了C>G的颠换,将这种新型碱基编辑器命名为CGBE系统(Zhao et al.2021;Tian et al.2022)。利用脱氨酶-nCas9或脱氨酶-dCas9的融合,单碱基编辑系统已成功应用于多种重要农作物,如水稻、小麦、玉米、大豆和棉花等作物。

[0004] 为了提高单碱基编辑系统的效率,David Liu实验室突变了TadA的8个氨基酸位点(TadA8e),在哺乳细胞中检测到的脱氨速度比之前的野生型脱氨酶TadA快近千倍,大大提高了碱基编辑效率(Richter et al.2020)。之后将Tad8e应用到水稻中,编辑效率显著提升,有的位点编辑效率甚至能达到100%(Wei et al.2021)。此外,为了在植物中获得更高活性的碱基编辑器,刘耀光团队成功开发了一套高效、广靶向的植物胞嘧啶单碱基编辑器

PevoFERNY-NG,在水稻的NGG-PAM和NG-PAM靶点都有较高的编辑效率(Zeng et al.2020),其中evoFERNY是一种进化脱氨酶,比APOBEC1蛋白更小,并能够在所有测试的序列背景中进行有效编辑。

[0005] 植物定向进化的饱和靶向内源诱变编辑器(Saturated Targeted Endogenous Mutagenesis Editor,STEME)的出现,加快了农作物性状的发展与进化,为快速获得新的优良农艺性状提供了可能(Li et al.2020)。研究者将胞嘧啶脱氨酶hA3A和腺嘌呤脱氨酶TadA二聚体同时融合在nCas9(D10A)的N端,并将UGI以融合或自由表达的形式置于nCas9(D10A)的C端。STEME双碱基编辑器只在一个sgRNA引导下就可以诱导靶位点C>T和A>G的同时突变,显著增加了靶基因碱基突变的饱和度及产生突变类型的多样性。该研究应用STEME系统在水稻中定向进化OsACC(Acetyl-coenzyme A carboxylase)基因获得了除草剂抗性系列突变体。此外,在前期开发的一系列单碱基编辑器基础之上,提出了单碱基编辑技术介导的水稻内源靶标基因的定向进化技术理念,在水稻细胞内人工模拟自然界基因的长期进化过程,短时间内创制大量除草剂抗性基因OsALS(Acetolactatesynthase)的新等位基因材料(Kuang et al.2020)。通过除草剂抗性筛选,研究人员成功鉴定出4个自然界中未曾被发现的、对除草剂具有不同抗性程度的OsALS新等位基因;并通过单碱基编辑技术引入到水稻生产品种南粳46中,将其升级为通过苗期一、两次施药,生育期田间无杂草发生的“洁田稻”。但目前开发的用于定向进化的碱基编辑器,效率较低,且由于PAM限制,靶点选择有限,有时不能在编码区和启动子区做到全覆盖,进而影响了其应用。

[0006] 草甘膦是目前使用最广泛的广谱灭生性除草剂,培育抗草甘膦基因编辑水稻对改善稻田杂草化学防控效果、减少用药量和简化防控措施有重要意义。草甘膦高效且低残留,杂草不易对其产生抗性,成本较低且易于在环境中降解。草甘膦作用于植物体内莽草酸途径的5-烯醇式丙酮酸-3-磷酸莽草酸合成酶(EPSPS),干扰EPSPS正常酶活功能,阻碍酪氨酸、苯丙氨酸和色氨酸等植物生长必需氨基酸的正常合成而影响植物的生长发育,导致死亡(Baerson et al.2002)。目前,培育抗草甘膦农作物品种常用的方法是应用转基因手段将来自外源的草甘膦抗性基因导入农作物中,使其在农作物中表达,从而培育出抗草甘膦作物新品种。但是由于转基因作物的商业化审批严格,目前我国尚未有转基因粮食作物商业化生产的报道。因此,开发高效碱基编辑器,以水稻内源EPSPS基因为靶标基因,通过碱基编辑器介导的定向进化,开发抗草甘膦的新型等位基因资源和新型无转基因抗草甘膦水稻新种质,对提升稻田杂草的综合治理效果,实现水稻轻简化,保障粮食安全具有重要意义。截至目前,相关研究尚未见国内外报道。

发明内容

[0007] 本发明所要解决的一个技术问题是如何培育抗草甘膦水稻。

[0008] 为解决上述技术问题,本发明首先提供了培育(或构建)抗草甘膦水稻的方法,所述方法包括:将受体水稻中的EPSPS基因通过基因编辑突变为EPSPS-1422C>T基因,得到抗草甘膦水稻;

[0009] 所述EPSPS基因编码如下A1)或A2)的EPSPS蛋白质:

[0010] A1) 序列表中序列12所示的EPSPS蛋白质;

[0011] A2) 将序列表中序列12所示的氨基酸序列经过一个或几个氨基酸残基的取代和/

或缺失和/或添加、第213位为天冬氨酸残基且具有相同功能的EPSPS蛋白质；

[0012] 所述EPSPS-1422C>T基因编码如下B1) 或B2) 的EPSPS-D213N蛋白质：

[0013] B1) 序列表中序列13所示的EPSPS-D213N蛋白质；

[0014] B2) 将序列表中序列13所示的氨基酸序列经过一个或几个氨基酸残基的取代和/或缺失和/或添加、第213位为天冬酰胺残基且具有相同功能的EPSPS-D213N蛋白质。

[0015] 其中，所述EPSPS蛋白质与所述EPSPS-D213N蛋白质相比除第213位外其余氨基酸残基均相同，所述EPSPS蛋白质的第213位为天冬氨酸残基，所述EPSPS-D213N蛋白质的第213位为天冬酰胺残基。

[0016] 上述A2) 中的蛋白质，为与序列12所示蛋白质的氨基酸序列具有75%或75%以上同一性且具有相同功能的蛋白质。上述B2) 中的蛋白质，为与序列13所示蛋白质的氨基酸序列具有75%或75%以上同一性且具有相同功能的蛋白质。所述具有75%或75%以上同一性为具有75%、具有80%、具有85%、具有90%、具有95%、具有96%、具有97%、具有98%或具有99%的同一性。

[0017] 上述方法中，所述EPSPS基因可为如下a11) 或a12)：

[0018] a11) 序列表中序列10所示的DNA分子；

[0019] a12) 与a11) 具有75%或75%以上同一性，且编码所述EPSPS蛋白质的DNA分子；

[0020] 所述EPSPS-1422C>T基因可为如下b11) 或b12)：

[0021] b11) 序列表中序列11所示的DNA分子；

[0022] b12) 与b11) 具有75%或75%以上同一性，且编码所述EPSPS-D213N蛋白质的DNA分子。

[0023] 这里使用的术语“同一性”指与天然核酸序列的序列相似性。“同一性”包括与本发明的序列10或11具有75%或更高，或85%或更高，或90%或更高，或95%或更高同一性的核苷酸序列。同一性可以用肉眼或计算机软件进行评价。使用计算机软件，两个或多个序列之间的同一性可以用百分比(%)表示，其可以用来评价相关序列之间的同一性。上述75%或75%以上同一性，可为80%、85%、90%或95%以上的同一性。

[0024] 上述方法中，所述基因编辑具体可为将EPSPS基因第1422位的C突变为T。

[0025] 上述方法中，所述基因编辑可采用双碱基融合编辑系统 (Asurrogate two-component composite base editing system, STCBE) 进行，所述双碱基融合编辑系统可包括胞嘧啶脱氨酶 (evoFERNY)、腺嘌呤脱氨酶 (TadA8e)、nCas9-NG (D10A) 和2×UGI，所述2×UGI由两个尿嘧啶糖基化酶抑制蛋白 (UGI) 连接得到。

[0026] 进一步，所述evoFERNY可为序列表中序列5的第9-169位所示的蛋白质。

[0027] 所述TadA8e可为序列表中序列5的第218-383位所示的蛋白质。

[0028] 所述nCas9-NG (D10A) 可为序列表中序列5的第416-1782位所示的蛋白质。

[0029] 所述UGI可为序列表中序列5的第1793-1875位所示的蛋白质。

[0030] 上述方法中，所述双碱基融合编辑系统可由所述evoFERNY、所述TadA8e、所述nCas9-NG (D10A) 和所述2×UGI依次连接得到。

[0031] 上文中的连接可直接连接，也可通过连接肽连接。所述连接肽可为48aa连接肽、32aa连接肽或10aa连接肽，所述48aa连接肽的序列可如序列表中序列5的第170-217位所示，所述32aa连接肽的序列可如序列表中序列5的第384-415位所示，所述10aa连接肽的序

列可如序列表中序列5的第1783-1792位所示。

[0032] 所述2×UGI可为序列表中序列5的第1793-1968位所示的蛋白质。

[0033] 具体的,所述双碱基融合编辑系统可为序列表中序列5的第9-1968位所示的蛋白质,或为序列5所示的蛋白质。

[0034] 上述方法中,所述基因编辑具体可通过向所述受体水稻导入含有所述双碱基融合编辑系统的编码基因以及转录靶向靶标DNA的gRNA的DNA片段的载体实现,所述靶标DNA含有EPSPS基因的第1422位的C。

[0035] 其中,所述双碱基融合编辑系统的编码基因与所述转录靶向靶标DNA的gRNA的DNA片段可处于同一载体中,也可处于不同载体中。

[0036] 上述方法中,所述双碱基融合编辑系统的编码基因可含有所述evoFERNY的编码基因、所述TadA8e的编码基因、所述nCas9-NG (D10A) 的编码基因和所述2×UGI的编码基因。

[0037] 所述evoFERNY的编码基因可为序列表中序列4的第2023-2505位所示的DNA分子。

[0038] 所述TadA8e的编码基因可为序列表中序列4的第2650-3147位所示的DNA分子。

[0039] 所述nCas9-NG (D10A) 的编码基因可为序列表中序列4的第3244-7344位所示的DNA分子。

[0040] 所述2×UGI的编码基因可为序列表中序列4的第7375-7902位所示的DNA分子。

[0041] 上述方法中,所述双碱基融合编辑系统的编码基因可为由所述evoFERNY的编码基因、所述TadA8e的编码基因、所述nCas9-NG (D10A) 的编码基因和所述2×UGI的编码基因中的部分基因连接或四个基因连接得到的1个、2个或3个DNA分子,也可为由四个独立基因组成的基因组合。所述双碱基融合编辑系统的编码基因由2-4个DNA分子组成时,各DNA分子可处于同一载体中,也可处于不同载体中。

[0042] 具体的,所述双碱基融合编辑系统的编码基因可由所述evoFERNY的编码基因、所述TadA8e的编码基因、所述nCas9-NG (D10A) 的编码基因和所述2×UGI的编码基因依次连接得到。

[0043] 上文中的连接可直接连接,也可通过连接肽的编码基因连接。所述连接肽可为所述48aa连接肽、所述32aa连接肽或所述10aa连接肽。

[0044] 进一步,所述双碱基融合编辑系统的编码基因可为序列表中序列4的第2023-7902位所示的DNA分子,或为序列4的第1999-7935位所示的DNA分子。

[0045] 所述载体具体可为V-STCBE-2,所述V-STCBE-2为将序列1所示pCXUN-Ubi-nCas9-NG (D10A) -HygACG的第1-9272位所示DNA片段替换为序列4所示的DNA片段得到的重组载体。

[0046] 本发明所要解决的另一个技术问题是提供一种双碱基融合编辑系统。

[0047] 本发明首先提供了下述任一产品:

[0048] P1) 所述双碱基融合编辑系统;

[0049] P2) 所述双碱基融合编辑系统的编码基因;

[0050] P3) 含有所述双碱基融合编辑系统的编码基因的表达盒、重组载体、重组微生物或重组细胞系。

[0051] 表达盒,是指能够在宿主细胞中表达目的蛋白质的DNA,该DNA不但可包括启动目的蛋白编码基因转录的启动子,还可包括终止目的蛋白编码基因转录的终止子。进一步,所述表达盒还可包括增强子序列。

- [0052] 在本发明的一个实施例中,所述表达盒为序列4的第1-8788位所示的DNA分子。
- [0053] 可用现有的载体构建含有所述表达盒的重组载体。
- [0054] 所述载体可为质粒、黏粒、噬菌体或病毒载体。所述载体具体可为pCXUN-Ubi-nCas9-NG (D10A) -HygACG。
- [0055] 所述重组载体具体可为所述V-STCBE-2。
- [0056] 所述微生物可为酵母、细菌、藻或真菌。其中,细菌可为农杆菌。
- [0057] 所述重组细胞系不包括繁殖材料。
- [0058] 所述产品在基因编辑或获得(或制备)基因编辑产品中的应用,也属于本发明的保护范围。
- [0059] 上述应用中,所述基因编辑可为C向T的突变和/或A向G的突变。
- [0060] 所述EPSPS-D213N蛋白质或所述EPSPS-1422C>T基因,也属于本发明的保护范围。
- [0061] 所述培育抗草甘膦水稻的方法,或所述产品,或所述EPSPS-D213N蛋白质或所述EPSPS-1422C>T基因在培育抗草甘膦水稻中的应用,也属于本发明的保护范围。
- [0062] 本发明利用本发明构建的双碱基融合编辑系统(STCBE)以水稻为受体材料,对其内源EPSPS基因进行多位点突变,获得了发生了Asp-213-Asn (D213N)突变的EPSPS新等位基因(EPSPS-1422C>T基因),携带该基因的水稻材料表现出明显的草甘膦抗性。本发明的STCBE系统及创制的无转基因抗草甘膦水稻新种质,为提升稻田杂草的综合治理效果,实现水稻轻简化,保障粮食安全,推动农业可持续发展,提供了重要技术和材料支撑。
- [0063] 下面结合具体实施方式对本发明进行进一步的详细描述,给出的实施例仅为了阐明本发明,而不是为了限制本发明的范围。以下提供的实施例可作为本技术领域普通技术人员进行进一步改进的指南,并不以任何方式构成对本发明的限制。

附图说明

- [0064] 图1. 潮霉素筛选的双碱基融合编辑系统及其编辑效率测定。A. 4种潮霉素筛选的双碱基融合编辑系统的载体框架图。B. 4种双碱基融合编辑系统C>T的编辑效率统计,C1-C14分别为所选基因目标靶点的第1位到第14位的碱基C的编辑效率。C. 4种双碱基融合编辑系统A>G的编辑效率统计,A1-A10分别为所选基因目标靶点的第1位到第10位的碱基A的编辑效率。
- [0065] 图2. 水稻EPSPS基因结构和靶标序列示意图。
- [0066] 图3. 抗性愈伤和植株在含有草甘膦的筛选培养基、再生培养基和生根培养基的生长情况。A. NB筛选培养基(10mg/L草甘膦,50mg/L潮霉素)上培养4周后长出的抗性愈伤(左图)和分化培养基(即再生培养基)培养6周后再生出的抗性幼苗(右图)。B. MS生根培养基(10mg/L草甘膦)培养2周后的抗性植株,左图为培养基中的生长情况,右图为野生型植株与抗性植株的比较,Bar=5cm。
- [0067] 图4. 再生抗性植株的Sanger测序鉴定EPSPS等位基因变异类型。

具体实施方式

- [0068] 下述实施例中的实验方法,如无特殊说明,均为常规方法,按照本领域内的文献所描述的技术或条件或者按照产品说明书进行。下述实施例中所用的材料、试剂、仪器等,如

无特殊说明,均可从商业途径得到。以下实施例中的定量试验,均设置三次重复实验,结果取平均值。下述实施例中,如无特殊说明,序列表中各核苷酸序列的第1位均为相应DNA/RNA的5'末端核苷酸,末位均为相应DNA/RNA的3'末端核苷酸。

[0069] MMG溶液:0.4M甘露醇(Mannitol),15mM MgCl₂,4mM MES,余量为水。

[0070] W5溶液:154mM NaCl,125mM CaCl₂,5mM KCl,2mM MES,余量为水。

[0071] WI溶液:0.5M Mannitol,20mM MgCl₂,4mM MES,余量为水。

[0072] 表1.引物列表

引物名称	序列(5' - 3')	用途
HygACG-F	GATAcGAAAAAGCCTGAACTGACGGCGACGTCTGTGCGAGAAGTTTC	构建双碱基编辑载体
HygACG-R	GTCGATCGACAGATCCGGTCGGCATCTACTCTATTTCTTTGCC	
A3A-F	GTGTTACTTCTGCAGCACTAGTATGGCCCAAAGAAGAAGCG	
A3A-R	TCCACTTCGGAAGATCCTCCGCTAGATCCTCCAGAAGACTC	
FER-F	GTTACTTCTGCAGCACTAGTATGAAGAGGACAGCCGACGGC	
FER-R	GAActCGACTTCTGAGGAGCCGCCGGAGCTGCCAC	
TAD-F1	TCAGAAGTCGAGTTCTCCCATG	
TAD-R1	CCGATGGAGTACTTCTTGTGCGAGCCGCCGGAGCTGCCAC	
TAD-F2	AGGATCTCCGAAGTGAATTTAGCCATGAATATTG	
TAD-R2	CAGAAGATCCCCCGAGTTGATGCTACTCTGCGCTTTC	
UGI-F1	CAGCTTGGTGGTATTCTGGTGGATCTGGCGGAAG	
UGI-R1	ACGCGTTCAGACCTTGCCTTCTTC	
UGI-F2	TCTGGTGGTTCTCCAAGAAGAAGAGAAAGTCTCTGGTGGT	
UGI-R2	CAGCCTCGACGACGCTTCAGACTTCTCTTCTTCTTGGGAG	
HiFTL2-F1	ggagtgagtacggtgtgcGCAAACTATAAACCTTCAGGG	PCR 扩增原生质体中所选基因的目 的片段进行 Hi-tom 高通量测序
HiFTL2-R1	gagttggatgctggatggGTCATTGCCGCCGACCTC	
HiFTL2-F2	ggagtgagtacggtgtgcGGTTTTTGTCAAGATCGATGGC	
HiFTL2-R2	gagttggatgctggatggGCACCAACTACGACACATGG	
HiFTL2-F3	gagttggatgctggatggGCGCTAGCTTAACATTAGTG	
HiFTL2-R3	ggagtgagtacggtgtgcGGTACTTAACCAAGATGAC	
HiGRF4-F1	ggagtgagtacggtgtgcTGCCTGTTTTTGTACAGTTGGA	
HiGRF4-R1	gagttggatgctggatggGGTGCATGTGGCGCTCGCAG	
HiGRF4-F2	ggagtgagtacggtgtgcCAAGAAAGCCTGTGAAACG	
HiGRF4-R2	gagttggatgctggatggGCCATTGCTGCCGCAATAGC	

[0073]

[0074]	HiGRF4-F3	gagttggatgctggatggTGCCTCCAATGGCAGCAGCT	
	HiGRF4-R3	ggagtgagtacggtgtgcGGGCAGCATTATCCATGTG	
	HiSD1-F1	ggagtgagtacggtgtgcCTCCAATCTCATGGTGGCCGAGC	
	HiSD1-R1	gagttggatgctggatggGTGCTCGGACACCTGGAAGAACC	
	HiSD1-F2	ggagtgagtacggtgtgcGCACGGTTCTTCCAGGTGTCC	
	HiSD1-R2	gagttggatgctggatggAATGCAATGTCGTCCACCATCG	
	HiSD1-F3	ggagtgagtacggtgtgcCAAACGCGTTTGTGCCGCGCAG	
	HiSD1-R3	gagttggatgctggatggGAAGGTGTCGCCGATGTTGATG	
	HiEPSPS-F1	ggagtgagtacggtgtgcGAGGCCCTGAAAGCCCTCGGGC	
	HiEPSPS-R1	gagttggatgctggatggATTGCAGTCCAGCGTTCCC	
	HiEPSPS-F2	ggagtgagtacggtgtgcTAGTCGTTGGCTGTGGTGGC	
	HiEPSPS-R2	gagttggatgctggatggTTGCATTTCCACCAGCAGCAG	
	EPSPS-F1	CACCAACCGCACCAACCCAA	
EPSPS-R1	ATATGAGTGTGTGTGATTAACC		
EPSPS-F2	TGATCCTTTGACATGTGATTGATCA		
EPSPS-R2	ATGCAGAAGGTTGACAAAGAATTG		
EPSPS-F3	GGTCTTGATGTCTTCTTTGA		
EPSPS-R3	CGTATTTGATTAACAGTGGC		
EPSPS-F4	CCCTGACCTTACATGTTAATAT		
EPSPS-R4	AACAGAACCTAGACCTCACT	检测再生植株的转基因情况	
Hyg-F	CGAGCTTTCGAGATCCGGGG		
Hyg-R	ATTGGAATCCCCGAACATCGC		

[0075] 实施例1、双碱基融合编辑系统及其在获得草甘膦抗性水稻中的应用

[0076] 本实施例采取以下3种策略制备了四种双碱基融合编辑器：

[0077] (1) CRISPR/Cas9系统采用Cas9变体nCas9-NG (D10A) ,其识别的PAM序列为NGN;

[0078] (2) 将人类胞嘧啶脱氨酶APOBEC3A (hA3A) 的第98位和第104位氨基酸进行突变 (W98Y/W104A) ,并与进化脱氨酶evoFERNY进行比较;同时比较腺嘌呤脱氨酶TadA-TadA8e二聚体和Tad8e单体与不同胞嘧啶脱氨酶组合的双碱基融合编辑系统的编辑效率和编辑窗口;

[0079] (3) 将潮霉素抗性基因 (HygR) 的翻译起始密码子ATG突变为ACG,在编辑的阳性植株中,碱基编辑器如果能正常行使功能,突变的HygR起始密码子ACG则会被还原为ATG,从而能使编辑植株发挥潮霉素抗性,达到辅助筛选的作用;

[0080] 本发明构建的4种潮霉素助理筛选的双碱基融合编辑系统STCBE-1,STCBE-2,STCBE-3,STCBE-4如图1所示。

[0081] 本实施例进一步分别将所选基因的靶位点连接到4种编辑系统上转化原生质体 (表2) ,黑暗处培养2天后提取原生质体基因组,PCR扩增所选靶基因的目的片段进行Hi-tom高通量测序。根据水稻原生质体中高通量测序结果,统计和比较了不同编辑系统对靶基因

编码序列的编辑效率和编辑范围。

[0082] 1 材料和方法

[0083] 1.1 实验材料

[0084] 用于水稻转化的水稻材料为中花11, 公众可从申请人处获得该生物材料。

[0085] 基础载体骨架pCXUN-Ubi-nCas9-NG (D10A) 由中国农业科学院作物科学研究所基因编辑新技术和新材料创制实验室保存;

[0086] 基础腺嘌呤碱基编辑载体pCXUN-Ubi-TadA-Tad8e-nCas9 (D10A), 由中国农业科学院作物科学研究所基因编辑新技术和新材料创制实验室保存;

[0087] 基础胞嘧啶碱基编辑载体PevoFERNY-NG由华南农业大学刘耀光课题组提供;

[0088] 基础胞嘧啶碱基编辑载体pCXUN-Ubi-hA3A (W98Y/W104A) -nCas9 (D10A) -UGI, 由中国农业科学院作物科学研究所基因编辑新技术和新材料创制实验室保存。

[0089] 1.2 实验方法

[0090] 1.2.1 双碱基融合编辑系统载体的构建

[0091] 以pCXUN-Ubi-nCas9-NG (D10A) 为模板, 用正向引物HygACG-F和反向引物HygACG-R (表1) 扩增起始密码子ATG突变为ACG的潮霉素抗性基因片段 (将起始密码子ATG突变为ACG的潮霉素抗性基因记为HygRACG基因), 用限制性内切酶AatII和PvuII双酶切载体pCXUN-Ubi-nCas9-NG (D10A), 利用同源重组酶将潮霉素片段融合到酶切后的载体骨架, 将此改造后的载体记为pCXUN-Ubi-nCas9-NG (D10A) -HygACG。pCXUN-Ubi-nCas9-NG (D10A) -HygACG的序列如序列表中序列1所示。pCXUN-Ubi-nCas9-NG (D10A) -HygACG可按照上述方法获得, 也可通过合成获得。

[0092] 序列1中, 第9365-9745位所示为水稻U3启动子, 第9746-9753位所示为PmeI识别序列, 第9754-9829位所示为gRNA scaffold的DNA序列。

[0093] 以pCXUN-Ubi-hA3A (W98Y/W104A) -nCas9 (D10A) -UGI为模板, 用正向引物A3A-F和反向引物A3A-R进行PCR扩增胞嘧啶脱氨酶hA3A (W98Y/W104A) 的DNA片段, 以pCXUN-Ubi-TadA-Tad8e-nCas9 (D10A) 为模板, 用正向引物TAD-F1和反向引物TAD-R1进行PCR扩增腺嘌呤脱氨酶TadA8e的DNA片段 (表1), 利用同源重组酶将两个片段融合到pCXUN-Ubi-nCas9-NG (D10A) -HygACG载体上的Cas9变体nCas9-NG (D10A) (下文简称为nCas9-NG) 的N端 (SpeI酶切), nCas9-NG识别的PAM序列为NGN, 具有较大靶基因编辑范围, 可以完全覆盖基因的编码区; 进一步以PevoFERNY-NG为模板用正向引物UGI-F1和反向引物UGI-R1进行PCR扩增尿嘧啶糖基化酶抑制蛋白 (UGI) 的DNA片段 (表1), 融合到pCXUN-Ubi-nCas9-NG (D10A) -HygACG载体中nCas9-NG的C端 (MluI酶切), 将所得重组载体记为V-STCBE-1, V-STCBE-1中含有双碱基融合编辑系统STCBE-1 (图1)。V-STCBE-1可按照上述方法获得, 也可通过合成获得。

[0094] V-STCBE-1中双碱基融合编辑系统STCBE-1的相关序列如序列表中序列2所示, V-STCBE-1通过将序列1所示pCXUN-Ubi-nCas9-NG (D10A) -HygACG的第1-9272位所示DNA片段替换为序列2所示的DNA片段得到。V-STCBE-1能表达序列3所示的hA3A-TadA8e-nCas9-NG-2×UGI所形成的融合蛋白 (记为hA3A-TadA8e-nCas9-NG-2×UGI) 以及HygRACG抗性蛋白质。

[0095] 序列2中, 第1-1991位所示为Ubi启动子, 第1999-8049位所示为hA3A-TadA8e-nCas9-NG-2×UGI融合蛋白的编码基因, 第8053-8267位所示为P1oyA的DNA序列, 第8268-8902位所示为NOS终止子, 第9230-9907位所示为35S启动子, 第9974-10999位所示为

HygRACG基因,第11039-11213位所示为P1oyA的DNA序列。序列2的第2002-2022位和第8029-8049位所示为核定位信号NLS的编码基因,第2023-2619位所示为hA3A的编码基因,第2620-2763位所示为48aa连接肽的编码序列,第2764-3261位所示为TadA8e的编码基因,第3262-3357位所示为32aa连接肽的编码序列,第3358-7458位所示为nCas9-NG的编码基因,第7459-7488位和第7738-7767位所示为10aa连接肽的编码序列,第7489-7737位和第7768-8016位所示为UGI的编码基因,第7489-8016位所示为2×UGI的编码序列。

[0096] 序列3中,第2-8位和第2011-2017位所示为核定位信号NLS,第9-207位所示为hA3A,第208-255位所示为48aa连接肽,第256-421位所示为TadA8e,第422-453位所示为32aa连接肽,第454-1820位所示为nCas9-NG,第1821-1830位和第1914-1923位所示为10aa连接肽,第1831-1913位和第1924-2006位所示为UGI,第1831-2006位所示为2×UGI。

[0097] 双碱基融合编辑系统STCBE-1包括hA3A、TadA8e、nCas9-NG、2×UGI,各部分由48aa连接肽、32aa连接肽、10aa连接肽连接,该基因编辑系统具体为序列3所示的融合蛋白hA3A-TadA8e-nCas9-NG-2×UGI。

[0098] 以PevofERNY-NG为模板,用正向引物FER-F和反向引物FER-R进行PCR扩增胞嘧啶脱氨酶evoFERNY(下文简称为FERNY)的DNA片段,以pCXUN-Ubi-TadA-Tad8e-nCas9(D10A)为模板,用正向引物TAD-F1和反向引物TAD-R1进行PCR扩增腺嘌呤脱氨酶TadA8e的DNA片段(表1),利用同源重组酶将两个片段融合到pCXUN-Ubi-nCas9-NG(D10A)-HygACG载体上nCas9-NG的N端(SpeI酶切),进一步以PevofERNY-NG为模板用正向引物UGI-F1和反向引物UGI-R1进行PCR扩增UGI的DNA片段(表1),融合到pCXUN-Ubi-nCas9-NG(D10A)-HygACG载体上nCas9-NG的C端(MluI酶切),将所得重组载体记为V-STCBE-2,V-STCBE-2中含有双碱基融合编辑系统STCBE-2(图1)。V-STCBE-2可按照上述方法获得,也可通过合成获得。

[0099] V-STCBE-2中双碱基融合编辑系统STCBE-2的相关序列如序列表中序列4所示,V-STCBE-2通过将序列1所示pCXUN-Ubi-nCas9-NG(D10A)-HygACG的第1-9272位所示DNA片段替换为序列4所示的DNA片段得到。V-STCBE-2能表达序列5所示的FERNY、TadA8e、nCas9-NG、2×UGI所形成的融合蛋白(记为FERNY-TadA8e-nCas9-NG-2×UGI)以及HygRACG抗性蛋白质。

[0100] 序列4中,第1-1991位所示为Ubi启动子,第1999-7935位所示为FERNY-TadA8e-nCas9-NG-2×UGI融合蛋白的编码基因,第7939-8153位所示为P1oyA的DNA序列,第8154-8788位所示为NOS终止子,第9116-9793位所示为35S启动子,第9860-10885位所示为HygRACG基因,第10925-11099位所示为P1oyA的DNA序列。序列4的第2002-2022位和第7915-7935位所示为核定位信号NLS的编码基因,第2023-2505位所示为FERNY的编码基因,第2506-2649位所示为48aa连接肽的编码序列,第2650-3147位所示为TadA8e的编码基因,第3148-3243位所示为32aa连接肽的编码序列,第3244-7344位所示为nCas9-NG的编码基因,第7345-7374位和第7624-7653位所示为10aa连接肽的编码序列,第7375-7623位和第7654-7902位所示为UGI的编码基因,第7375-7902位所示为2×UGI的编码序列。

[0101] 序列5中,第2-8位和第1973-1979位所示为核定位信号NLS,第9-169位所示为FERNY,第170-217位所示为48aa连接肽,第218-383位所示为TadA8e,第384-415位所示为32aa连接肽,第416-1782位所示为nCas9-NG,第1783-1792位和第1876-1885位所示为10aa连接肽,第1793-1875位和第1886-1968位所示为UGI,第1793-1968位所示为2×UGI。

[0102] 双碱基融合编辑系统STCBE-2包括FERNY、TadA8e、nCas9-NG、2×UGI，各部分由48aa连接肽、32aa连接肽、10aa连接肽连接，该基因编辑系统具体为序列5所示的融合蛋白FERNY-TadA8e-nCas9-NG-2×UGI。

[0103] 以pCXUN-Ubi-hA3A (W98Y/W104A) -nCas9 (D10A) -UGI为模板，用正向引物A3A-F和反向引物A3A-R进行PCR扩增胞嘧啶脱氨酶hA3A (W98Y/W104A) 的DNA片段，以pCXUN-Ubi-TadA-Tad8e-nCas9 (D10A) 为模板，用正向引物TAD-F2和反向引物TAD-R2进行PCR扩增腺嘌呤脱氨酶二聚体TadA-TadA8e的DNA片段(表1)，利用同源重组酶将两个片段融合到pCXUN-Ubi-nCas9-NG (D10A) -HygACG载体上nCas9-NG的N端(SpeI酶切)，进一步以pCXUN-Ubi-hA3A (W98Y/W104A) -nCas9 (D10A) -UGI为模板用正向引物UGI-F2和反向引物UGI-R2进行PCR扩增UGI的DNA片段(表1)，融合到pCXUN-Ubi-nCas9-NG (D10A) -HygACG载体上nCas9-NG的C端(MluI酶切)，将所得重组载体记为V-STCBE-3，V-STCBE-3中含有双碱基融合编辑系统STCBE-3(图1)。V-STCBE-3可按照上述方法获得，也可通过合成获得。

[0104] V-STCBE-3中双碱基融合编辑系统STCBE-3的相关序列如序列表中序列6所示，V-STCBE-3通过将序列1所示pCXUN-Ubi-nCas9-NG (D10A) -HygACG的第1-9272位所示DNA片段替换为序列6所示的DNA片段得到。V-STCBE-3能表达序列7所示的hA3A、TadA-TadA8e、nCas9-NG、UGI所形成的融合蛋白(记为hA3A-TadA-TadA8e-nCas9-NG-UGI)以及HygRACG抗性蛋白质。

[0105] 序列6中，第1-1991位所示为Ubi启动子，第1999-8364位所示为hA3A-TadA-TadA8e-nCas9-NG-UGI融合蛋白的编码基因，第8368-8582位所示为P1oyA的DNA序列，第8583-9217位所示为NOS终止子，第9545-10222位所示为35S启动子，第10289-11314位所示为HygRACG基因，第11354-11528位所示为P1oyA的DNA序列。序列6的第2002-2022位和第8344-8364位所示为核定位信号NLS的编码基因，第2023-2619位所示为hA3A的编码基因，第2620-2763位所示为48aa连接肽的编码序列，第2764-3855位所示为TadA-TadA8e的编码基因，第3856-3951位所示为32aa连接肽的编码序列，第3952-8052位所示为nCas9-NG的编码基因，第8053-8082位所示为10aa连接肽的编码序列，第8083-8331位所示为UGI的编码基因。

[0106] 序列7中，第2-8位和第2116-2122位所示为核定位信号NLS，第9-207位所示为hA3A，第208-255位所示为48aa连接肽，第256-619位所示为TadA-TadA8e，第620-651位所示为32aa连接肽，第652-2018位所示为nCas9-NG，第2019-2028位所示为10aa连接肽，第2029-2111位所示为UGI。

[0107] 双碱基融合编辑系统STCBE-3包括hA3A、TadA-TadA8e、nCas9-NG、UGI，各部分由48aa连接肽、32aa连接肽、10aa连接肽连接，该基因编辑系统具体为序列7所示的融合蛋白hA3A-TadA-TadA8e-nCas9-NG-UGI。

[0108] 以PevoFERNY-NG为模板，用正向引物FER-F和反向引物FER-R进行PCR扩增胞嘧啶脱氨酶evoFERNY的DNA片段，以pCXUN-Ubi-TadA-Tad8e-nCas9 (D10A) 为模板，用正向引物TAD-F2和反向引物TAD-R2进行PCR扩增腺嘌呤脱氨酶二聚体TadA-TadA8e的DNA片段(表1)，利用同源重组酶将两个片段融合到pCXUN-Ubi-nCas9-NG (D10A) -HygACG载体上nCas9-NG的N端(SpeI酶切)，进一步以pCXUN-Ubi-hA3A (W98Y/W104A) -nCas9 (D10A) -UGI为模板用正向引物UGI-F2和反向引物UGI-R2进行PCR扩增UGI的DNA片段(表1)，融合到pCXUN-Ubi-nCas9-

NG (D10A) -HygACG载体上nCas9-NG的C端 (MluI酶切),将所得重组载体记为V-STCBE-4,V-STCBE-4中含有双碱基融合编辑系统STCBE-4(图1)。V-STCBE-4可按照上述方法获得,也可通过合成获得。

[0109] V-STCBE-4中双碱基融合编辑系统STCBE-4的相关序列如序列表中序列8所示,V-STCBE-4通过将序列1所示pCXUN-Ubi-nCas9-NG (D10A) -HygACG的第1-9272位所示DNA片段替换为序列8所示的DNA片段得到。V-STCBE-4能表达序列9所示的FERNY、TadA-TadA8e、nCas9-NG、UGI所形成的融合蛋白(记为FERNY-TadA-TadA8e-nCas9-NG-UGI)以及HygRACG抗性蛋白质。

[0110] 序列8中,第1-1991位所示为Ubi启动子,第1999-8250位所示为FERNY-TadA-TadA8e-nCas9-NG-UGI融合蛋白的编码基因,第8254-8468位所示为P1oyA的DNA序列,第8469-9103位所示为NOS终止子,第9431-10108位所示为35S启动子,第10175-11200位所示为HygRACG基因,第11240-11414位所示为P1oyA的DNA序列。序列8的第2002-2022位和第8230-8250位所示为核定位信号NLS的编码基因,第2023-2505位所示为FERNY的编码基因,第2506-2649位所示为48aa连接肽的编码序列,第2650-3741位所示为TadA-TadA8e的编码基因,第3742-3837位所示为32aa连接肽的编码序列,第3838-7938位所示为nCas9-NG的编码基因,第7939-7968位所示为10aa连接肽的编码序列,第7969-8217位所示为UGI的编码基因。

[0111] 序列9中,第2-8位和第2078-2084位所示为核定位信号NLS,第9-169位所示为FERNY,第170-217位所示为48aa连接肽,第218-581位所示为TadA-TadA8e,第582-613位所示为32aa连接肽,第614-1980位所示为nCas9-NG,第1981-1990位所示为10aa连接肽,第1991-2073位所示为UGI。

[0112] 双碱基融合编辑系统STCBE-4包括FERNY、TadA-TadA8e、nCas9-NG、UGI,各部分由48aa连接肽、32aa连接肽、10aa连接肽连接,该基因编辑系统具体为序列9所示的融合蛋白FERNY-TadA-TadA8e-nCas9-NG-UGI。

[0113] 1.2.2水稻原生质体分离、转化和编辑效率的检测

[0114] 原生质体分离:水稻品种为中花11,将中花11在无菌下培养,选取幼苗茎干和叶鞘部分分离原生质体,将所得原生质体利用MMG溶液重悬,得到原生质体悬液,其中原生质体浓度为 2×10^6 /ml,血球计数器计数。

[0115] 基因编辑载体的构建:将表2中各Protospacer所示的DNA片段(即靶标序列)通过PmeI分别插入至V-STCBE-1、V-STCBE-2、V-STCBE-3、V-STCBE-4的gRNA scaffold上游,共得到 $4 \times 14 (=56)$ 个重组载体,所得到的重组载体能够转录靶向各目标基因的目标靶点(各目标靶点均为已报道的在水稻细胞里被成功编辑的靶点)的sgRNA。

[0116] 表2.检测基因编辑载体的编辑效率所选的目标靶点

目标基因	Protospacer (5' - 3')	PAM
<i>FTL2-T1</i>	GGCAGGGACAGGGACCCTCT	TGT

[0118]	<i>FTL2-T2</i>	GCAGCCATTGGACACGGTCT	TGG
	<i>FTL2-T3</i>	CCCTAAGGTTAGGGTCACTT	GGG
	<i>GRF4-T1</i>	TTCGGCAAGAAGCTGGACCC	AGA
	<i>GRF4-T2</i>	ATCTGTTGTTCGGTCTGCGG	CGG
	<i>GRF4-T3</i>	TTGGAAGCTGCTGCCATTGG	AGG
	<i>GRF4-T4</i>	CCTGCTATTGCCGGCAGCAA	TGG
	<i>SD1-T1</i>	ACCGCCGGCTCTGGCATTGC	CGC
	<i>SD1-T2</i>	CCCAAGATCCCGGAGCCATT	CGT
	<i>SD1-T3</i>	CCAGCTCCGCCGCCGACGCC	GGC
	<i>SD1-T4</i>	ACAGCAGCTCAATCATGCGG	TGC
	<i>EPSPS-T1</i>	AAGCAGATAAAGTTGCAAAA	AGA
	<i>EPSPS-T2</i>	TGGAAGTGAATGCGACCAT	TGA
	<i>EPSPS-T3</i>	TCTTCGCATCCTTCTCAAC	AGG

[0119] 转化:将所得56个基因编辑载体分别转化水稻原生质体。利用含有GFP基因的质粒作为对照,观察原生质体细胞的转化效率。具体步骤如下:

[0120] 取20 μ g基因编辑载体于2ml离心管,用去尖的枪头吸取200 μ l原生质体沿管壁缓慢加入离心管中,轻轻吸打混匀,再加220 μ l PEG-4000,轻轻颠倒混匀,避光诱导转化15min;加800 μ l W5溶液(室温)颠倒混匀,23 $^{\circ}$ C,250g水平离心,升3降3,离心3min,弃上清。加1ml WI溶液,颠倒混匀,轻轻转至6孔板中(已预先加入1ml WI)28 $^{\circ}$ C或室温暗处培养。

[0121] 编辑效率的检测:待转化12h后,观察转化GFP质粒原生质体GFP表达情况,统计转化效率;待转化后42h,收集原生质体。用枪吸打悬起原生质体于2ml离心管中,12000rpm,1min,离心后弃上清;提取原生质体基因组,PCR扩增所选靶基因的目的片段进行Hi-tom高通量测序,检测编辑效率。Hi-tom高通量测序所用引物为:HiFTL2-F1与HiFTL2-R1、HiFTL2-F2与HiFTL2-R2、HiFTL2-F3与HiFTL2-R3、HiGRF4-F1与HiGRF4-R1、HiGRF4-F2与HiGRF4-R2、HiGRF4-F3与HiGRF4-R3、HiSD1-F1与HiSD1-R1、HiSD1-F2与HiSD1-R2、HiSD1-F3与HiSD1-R3、HiEPSPS-F1与HiEPSPS-R1、HiEPSPS-F2与HiEPSPS-R2。

[0122] 1.2.3EPSPS基因多位点突变的质粒库构建

[0123] 利用引物PCR扩增所选品种(中花11)中EPSPS基因编码区序列,确定所选水稻品种是否含有核苷酸的变化(表1),野生型EPSPS基因如序列10所示,所用引物为EPSPS-F1与EPSPS-R1、EPSPS-F2与EPSPS-R2、EPSPS-F3与EPSPS-R3、EPSPS-F4与EPSPS-R4。

[0124] 水稻EPSPS基因全长为3255bp,含有8个外显子和7个内含子,CDS长度为1533bp,编码511个氨基酸。根据基因编码区的序列信息,结合碱基编辑活性窗口和窗口内碱基突变能引起氨基酸变异的原则,在EPSPS的保守结构域所覆盖的基因组序列的正反向序列上设计以NGN为PAM的靶标序列,共35个靶标序列(图2,表3);通过Oligo合成和退火产生sgRNAs文库,将sgRNAs通过同源重组的方式分别构建到编辑效率最高的双碱基融合编辑载体(即V-

STCBE-2) 上, 构建打靶质粒库, 使得各sgRNA的DNA片段(即靶标序列)通过PmeI分别插入至gRNA scaffold上游; 将所得打靶质粒库转化农杆菌EHA-105; 然后挑取农杆菌单克隆进行菌落PCR检测, 得到阳性菌株; 再将阳性菌株接种到1mL的含50 μ g/ml卡那霉素和50 μ g/ml利福平的YEP液体培养基中, 得到的菌液用于水稻(中花11)愈伤的转化。

[0125] 表3. 用于水稻EPSPS定向突变的sgRNA序列信息

sgRNA	位置	Protospacer (靶标序列)	PAM	转化分组
0sEPSPS-g1	Exon-1	GGCGGAGGAGATCGTGCTCC	AGC	1
0sEPSPS-g2	Exon-1	CCTCCTCTCCGCCCTCTCCG	AGG	1
0sEPSPS-g3	Exon-1	TCTCCGCCCTCTCCGAGGTG	AGA	1
0sEPSPS-g4	Exon-2	GGACAACTTGCTGAACAGTG	AGG	2
0sEPSPS-g5	Exon-2	CAGTGAGGATGTTCACTACA	TGC	2
0sEPSPS-g6	Exon-2	ATCCTCACTGTTAGCAAGT	TGT	2
0sEPSPS-g7	Exon-2	GTTGCAAAAAGAGCTGTAGT	CGT	3
0sEPSPS-g8	Exon-2	GAAGGATGCGAAAGAGGAAG	TGC	3
0sEPSPS-g9	Exon-2	ATCCTTCTCAACAGGAACT	TGC	3
0sEPSPS-g10	Exon-2	CAACTCTTCTGGGGAACGC	TGG	3
[0126] 0sEPSPS-g11	Exon-2	CAGCGTTCCCAAGAAGAGT	GCA	4
0sEPSPS-g12	Exon-2	GGAACGCTGGAAGTCAATG	CGA	4
0sEPSPS-g13	Exon-2	GCTGTCAATGGTCGCATTGC	ACT	4
0sEPSPS-g14	Exon-2	GCAATGCGACCATTGACAGC	AGC	4
0sEPSPS-g15	Exon-3	GCCACGAATGAGGGAGAGAC	CGA	5
0sEPSPS-g16	Exon-3	TCATTCGTGGCACTCCATCA	AGC	5
0sEPSPS-g17	Exon-3	CCGACAACCAAGTCACCAAT	CGG	5
0sEPSPS-g18	Exon-3	GCGGATGTCGACTGTTTCCT	TGG	6
0sEPSPS-g19	Exon-3	AGTCGACATCCGCACCAAGT	TGT	6
0sEPSPS-g20	Exon-3	GCACTGAATGCCACCTGTT	CGT	6
0sEPSPS-g21	Exon-4	TTGAGTGCCTTGCTGATGGC	TGC	7
0sEPSPS-g22	Exon-4	TCAGCAAGGCACTCAAGTAC	TGA	7

	OsEPSPS-g23	Exon-4	GATGTGGAGATCGAAATCAT	TGA	7
	OsEPSPS-g24	Exon-4	GATTAGTTTGTCATGATTT	CGA	7
	OsEPSPS-g25	Exon-5	CAAGAAATAGCTCGCGCTTG	AGG	8
	OsEPSPS-g26	Exon-5	AGCGCGAGCTATTTCTTGGC	TGG	8
	OsEPSPS-g27	Exon-5	CCACAACCTTGAAGTGTAC	AGT	8
	OsEPSPS-g28	Exon-5	AACTGGTCGTACCACAACCT	TGA	8
[0127]	OsEPSPS-g29	Exon-7	GGAGAGTAAAGGAAACCGAA	AGG	9
	OsEPSPS-g30	Exon-7	AAGGAAACCGAAAGGATGGT	TGC	9
	OsEPSPS-g31	Exon-7	CCGAATTGCAACCATCCTTT	CCG	9
	OsEPSPS-g32	Exon-7	AATTCGGACCGAGCTAACAA	AGG	10
	OsEPSPS-g33	Exon-7	CTAACAAAGGTAAATTCATT	AGG	10
	OsEPSPS-g34	Exon-8	CAGCTGGGAGCATCGGTTGA	AGA	11
	OsEPSPS-g35	Exon-8	CTTCTCAACCGATGCTCCC	AGC	11

[0128] 1.2.4水稻的稳定遗传转化

[0129] 选取饱满的中花11水稻种子,剥去种皮,灭菌洗涤后,均匀的点在带有2,4-D (2mg/L)的灭菌NB固体培养基上,28℃黑暗处培养4-6周诱导愈伤形成;挑取步骤1.2.3制备好的农杆菌的单克隆接种于10mL含相应抗生素的液体培养基中,28℃振荡培养12小时;当培养至对数期时,收集农杆菌菌体,并用侵染液重悬农杆菌,将已培养好的水稻愈伤组织浸泡到此农杆菌重悬液中侵染30分钟;将侵染完的愈伤组织转入共培养基上进行共培养3天;共培养结束后转至NB筛选培养基(含有2mg/L 2,4-D、10mg/L草甘膦和50mg/L潮霉素的NB培养基)上进行筛选培养,28℃培养4-6周;选取生长良好呈淡黄色的愈伤组织,用无菌镊子移至分化培养基(含有1mg/L NAA、5mg/L 6-BA、2mg/L kinetin、10mg/L草甘膦、30mg/L潮霉素的NB培养基),28℃持续光照培养;待分化出来的幼苗长至2至5毫米,转入MS生根培养基(含10mg/L草甘膦的MS培养基)培养2-3周,之后移入土中置于温室生长(温度28-30℃,16小时光照/8小时黑暗)。

[0130] 1.2.5再生植株的检测和突变类型鉴定

[0131] 经草甘膦(10mg/L)和潮霉素(50mg/L)筛选获得抗性T0代植株,对筛选获得的再生植株进行靶位点突变分析,提取T0代植株的基因组DNA,以T0代植株基因组DNA为模板,进行PCR扩增获得EPSPS基因片段,所用引物为表1中的正向引物EPSPS-F2和反向引物EPSPS-R2。

[0132] 对PCR扩增产物进行sanger测序,并将序列与野生型EPSPS基因序列进行比对,鉴定突变类型;评估再生植株中内源EPSPS靶基因编码氨基酸的突变,鉴定蛋白变异类型,并对获得的再生植株进行脱靶分析。

[0133] 2、结果

[0134] 2.1双碱基融合编辑系统的编辑活性检测

[0135] Hi-tom高通量测序发现V-STCBE-2具有较高的编辑效率(图1),C>T的最高编辑效率可达25.5%,A>G的最高效率可达25.3%,其编辑窗口可以覆盖14个核苷酸,可以为目标

基因的定向进化策略提供技术支持。

[0136] 2.2 EPSPS 碱基编辑新材料的获得和基因型鉴定

[0137] 将构建好的 EPSPS 基因的打靶质粒库转入农杆菌, 根据 sgRNA 在基因编码区的位置分组总共分成了 11 组 (表 3), 每组的农杆菌按照等 OD 值进行混池, 混匀后转化水稻愈伤; 经草甘膦 (10mg/L) 和潮霉素 (50mg/L) 筛选和再生培养后获得再生植株 (图 3), 利用正向引物 Hyg-F 和反向引物 Hyg-R 对碱基编辑载体上的潮霉素抗性基因扩增进行 PCR 扩增、测序, 检测 T0 代植株是否含有外源转基因元件。

[0138] 所得 T0 代植株中 24 株 T0 代植株含有外源转基因元件的存在, 其中在 9 株再生植株中检测到了 EPSPS 基因的发生编辑 (图 4, 表 4), 且编辑均发生在第 1422 个碱基处 (位于 EPSPS 基因的第三个外显子上), 碱基编辑类型均为 C>T (反义链为 G>A), 将 EPSPS 基因 (序列 10) 的第 1422 个碱基发生 C>T 编辑得到的基因记为 EPSPS-1422C>T 基因 (序列 11); 蛋白变异类型为 EPSPS 蛋白第 213 位的 Asp 突变为 Asn (D213N) (图 4, 表 4), 将 EPSPS 蛋白质 (序列 12) 的第 213 位的 Asp 突变为 Asn 后的蛋白质记为 EPSPS-D213N 蛋白质 (序列 13), 将所得含有如上突变的植株记为抗性植株。图 3 中 B 图右图为该抗性植株与野生型水稻中花 11 生长状况的比较, 可知, 所得抗性植株的生长状况明显优于野生型植株, 说明 EPSPS 基因突变为 EPSPS-1422C>T 基因后, 可以提高水稻的草甘膦抗性, 可用于培育抗草甘膦水稻。

[0139] 获得如上抗性植株所用的质粒库中的质粒 (记为 V-STCBE-2-D213) 所转录的 sgRNA 为表 3 中的 0sEPSPS-g19, 该质粒的编辑效率统计如表 4 所示, 编辑效率可达 37.5%。

[0140] 表 4. 水稻再生植株中内源 EPSPS 基因的编辑效率统计

靶点	转化的愈伤数 (个)	转基因植株数 (个)	编辑植株数 (个)	碱基编辑 类型	碱基编辑 效率	氨基酸变异 类型
[0141] EPSPS-D213	120	24	9	C>T	37.5% (9/24)	D213N

[0142] 2.3 脱靶分析

[0143] 利用网站 <http://skl.scau.edu.cn/offtarget/> 对 V-STCBE-2-D213 的潜在脱靶位点进行预测, 显示与靶标序列具有碱基错配的潜在脱靶位点有两个 (表 5)。对获得的草甘膦抗性植株的编辑靶点 EPSPS 的脱靶情况进行鉴定, PCR 产物克隆并测序结果表明, V-STCBE-2-D213 的靶点并不存在脱靶情况 (表 5)。

[0144] 表 5. EPSPS 靶点的脱靶分析

靶点	潜在脱靶位点	潜在脱靶位点的基因位置	潜在脱靶位点的基因序列	错配碱基数	测序植株数	潜在脱靶位点的突变的植株数
[0145] EPSPS-D213	EPSPS-OFF1	Chr08:140131002-140131024	AAGTCCACAT AT GCACCAAGTGC	3	24	0
	EPSPS-OFF2	Chr08:141100755-141100777	AATCGCCATCCCCACCAAGTGT	3	24	0

[0146] 注: 表 5 中, PAM 位点用下划线表示, 错配碱基用加粗字体表示。

[0147] 3. 结论

[0148] 本发明公开了通过碱基编辑定向进化水稻内源EPSPS基因获得抗草甘膦水稻的方法及其所用系统。本发明通过改造现有的单碱基编辑系统,提高对水稻内源靶基因的碱基编辑效率,扩展其编辑范围,成功构建了4种潮霉素助理筛选的新型双碱基融合编辑系统(STCBE)。利用原生质体转化和高通量测序,测试了不同编辑系统对水稻内源基因编码序列的编辑效率,获得具有更高编辑效率的新型双碱基融合编辑系统STCBE-2用于目标基因定向进化的研究。进一步,以水稻中花11为受体材料,以水稻内源EPSPS基因为研究对象,设计了35个sgRNA,利用STCBE-2介导的定向进化对基因编码区多个重要功能域进行多位点突变,通过农杆菌转化和草甘膦筛选,鉴定获得了Asp-213-Asn (D213N)突变的EPSPS新型等位基因,携带该基因的水稻表现出草甘膦抗性。本发明通过系统优化,建立了一种高效双碱基融合编辑系统。在此基础上,利用该双碱基融合编辑系统介导的定向进化,率先获得抗草甘膦的新型等位基因和新种质,并通过后代分离,获得了无转基因抗草甘膦新种质。高效双碱基融合编辑系统的创建和无转基因抗草甘膦水稻新种质的获得和利用,将对减轻稻田杂草危害、简化防控措施、降低防控成本,提升稻田杂草的综合治理效果,实现水稻轻简化生产,保障水稻高产优质生产和国家粮食安全,推动农业可持续发展,具有重要的应用潜力和重大生产推广利用价值。

[0149] 4. 参考文献

- [0150] 1. Baerson SR, Rodriguez DJ, Tran M, Feng Y, Biest NA, Dill GM (2002) Identification of a mutation in the target enzyme 5-enolpyruvylshikimate-3-phosphate synthase. *Plant Physiol* 129:1265-1275
- [0151] 2. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, Liu DR (2017) Programmable base editing of A • T to G • C in genomic DNA without DNA cleavage. *Nature* 551:464-471
- [0152] 3. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533:420-424
- [0153] 4. Kuang Y, Li S, Ren B, Yan F, Spetz C, Li X, Zhou X, Zhou H (2020) Base-Editing-Mediated Artificial Evolution of OsALS1 In Planta to Develop Novel Herbicide-Tolerant Rice Germplasms. *Mol Plant* 13:565-572
- [0154] 5. Li C, Zhang R, Meng X, Chen S, Zong Y, Lu C, Qiu JL, Chen YH, Li J, Gao C (2020) Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors. *Nat Biotechnol* 38:875-882
- [0155] 6. Richter MF, Zhao KT, Eton E, Lapinaite A, Newby GA, Thuronyi BW, Wilson C, Koblan LW, Zeng J, Bauer DE, Doudna JA, Liu DR (2020) Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nature Biotechnol* 38:883-891
- [0156] 7. Tian Y, Shen R, Li Z, Yao Q, Zhang X, Zhong D, Tan X, Song M, Han H, Zhu JK, Lu Y (2022) Efficient C-to-G editing in rice using an optimized base editor. *Plant Biotechnol J* 20:1238-1240

[0157] 8. Wei C, Wang C, Jia M, Guo HX, Luo PY, Wang MG, Zhu JK, Zhang H (2021) Efficient generation of homozygous substitutions in rice in one generation utilizing an rABE8e base editor. *J Integr Plant Biol* 63:1595-1599

[0158] 9. Zeng D, Liu T, Tan J, Zhang Y, Zheng Z, Wang B, Zhou D, Xie X, Guo M, Liu YG, Zhu Q (2020) PhieCBEs: Plant High-Efficiency Cytidine Base Editors with Expanded Target Range. *Mol Plant* 13:1666-1669

[0159] 10. Zhao D, Li J, Li S, Xin X, Hu M, Price MA, Rosser SJ, Bi C, Zhang X (2021) Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat Biotechnol* 39:35-40

[0160] 以上对本发明进行了详述。对于本领域技术人员来说,在不脱离本发明的宗旨和范围,以及无需进行不必要的实验情况下,可在等同参数、浓度和条件下,在较宽范围内实施本发明。虽然本发明给出了特殊的实施例,应该理解为,可以对本发明作进一步的改进。总之,按本发明的原理,本申请欲包括任何变更、用途或对本发明的改进,包括脱离了本申请中已公开范围,而用本领域已知的常规技术进行的改变。按以下附带的权利要求的范围,可以进行一些基本特征的应用。

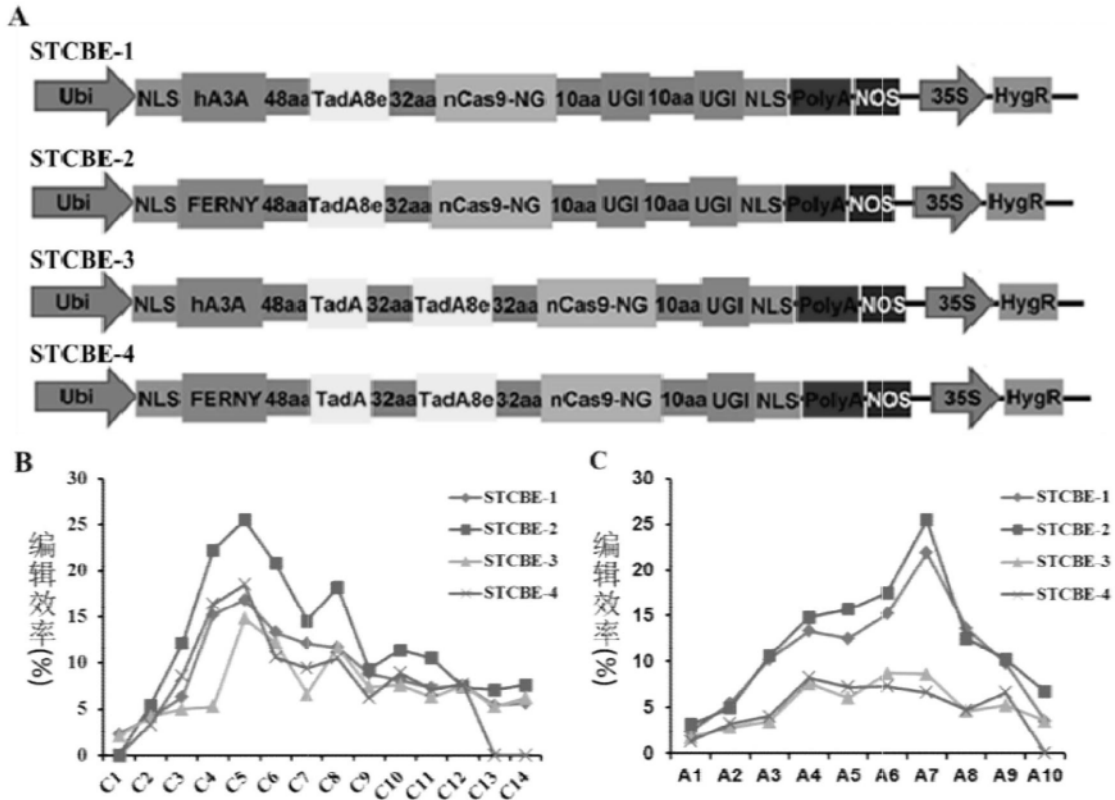


图1

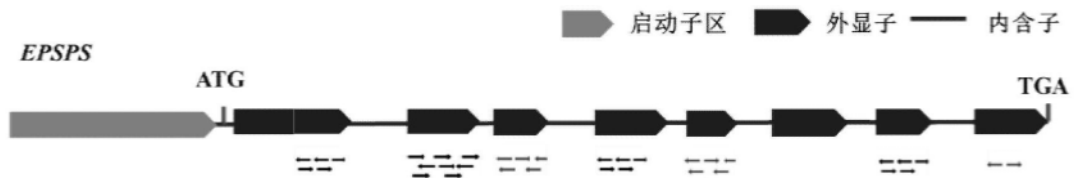


图2

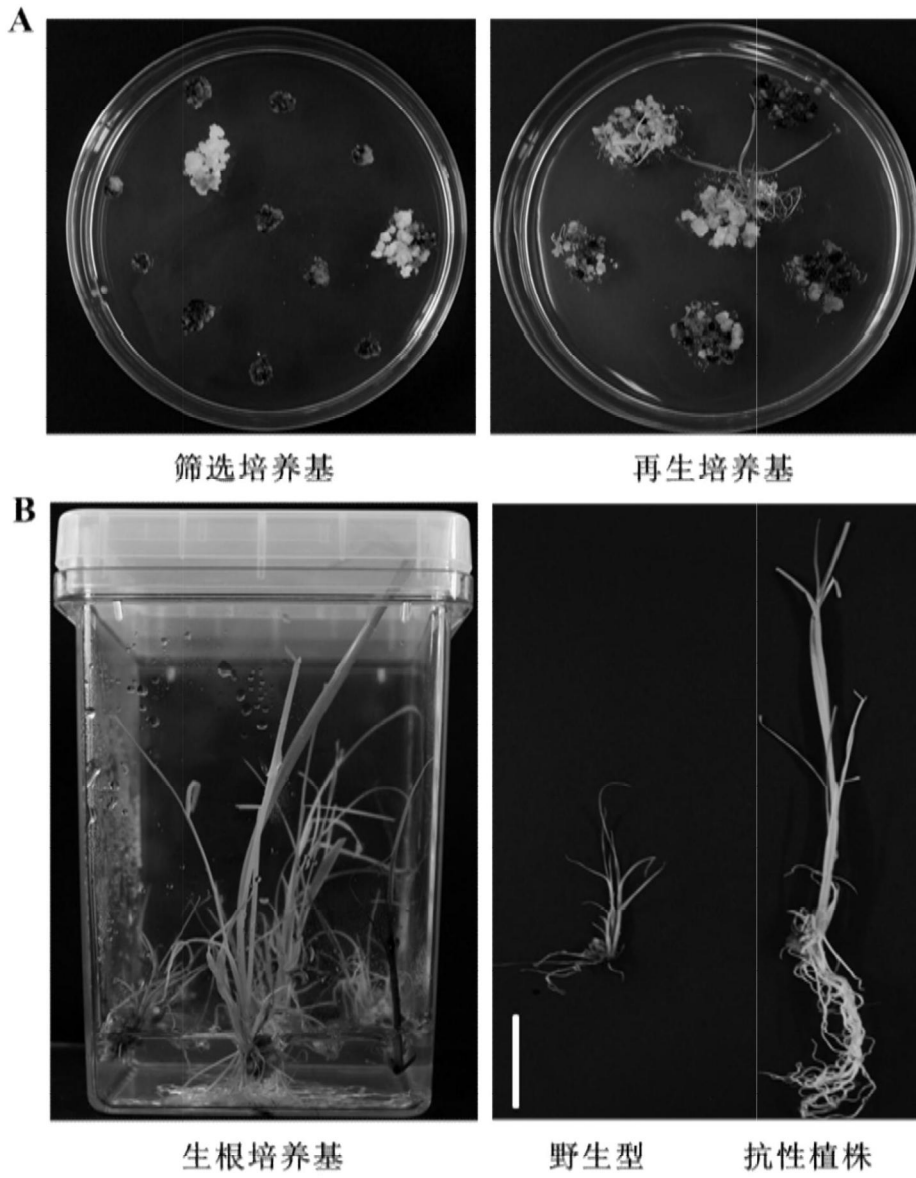


图3

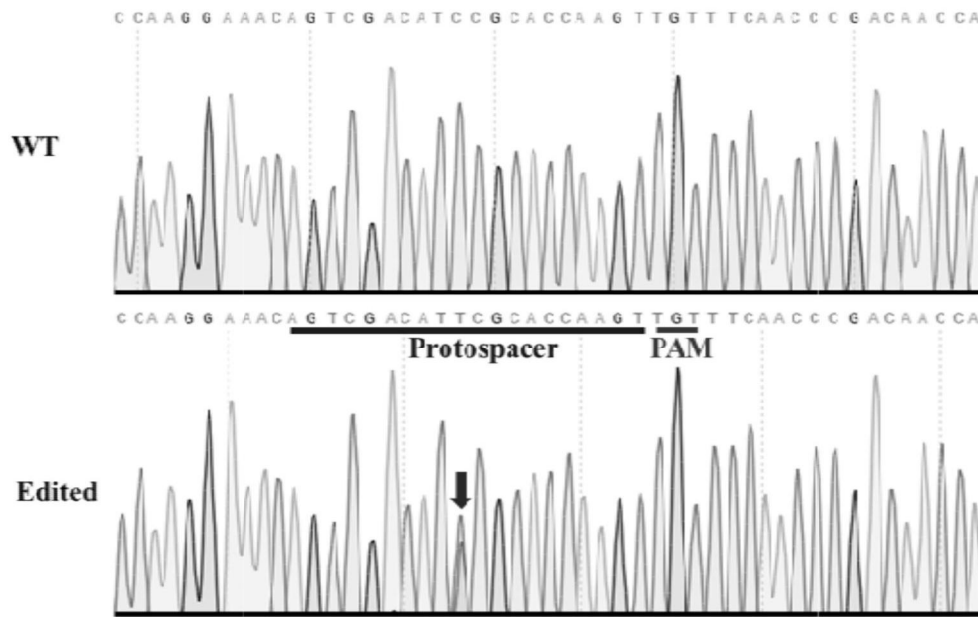


图4