



(12) 发明专利申请

(10) 申请公布号 CN 119031835 A

(43) 申请公布日 2024. 11. 26

(21) 申请号 202380030161.2

(22) 申请日 2023.01.25

(30) 优先权数据

10-2022-0009633 2022.01.24 KR

(85) PCT国际申请进入国家阶段日

2024.09.24

(86) PCT国际申请的申请数据

PCT/KR2023/001125 2023.01.25

(87) PCT国际申请的公布数据

W02023/140722 KO 2023.07.27

(71) 申请人 吉弗拉斯生命科学株式会社

地址 韩国忠清北道

(72) 发明人 崔圣和 崔善美 金珍华 金铤模

刘敏景 全伦我

(74) 专利代理机构 华进联合专利商标代理有限公司 44224

专利代理师 郑元博

(51) Int.Cl.

A01H 6/82 (2018.01)

A01H 5/10 (2018.01)

A01H 1/06 (2006.01)

C12N 15/82 (2006.01)

C12N 15/113 (2010.01)

C12N 9/02 (2006.01)

C12N 9/22 (2006.01)

权利要求书2页 说明书23页

序列表(电子公布) 附图28页

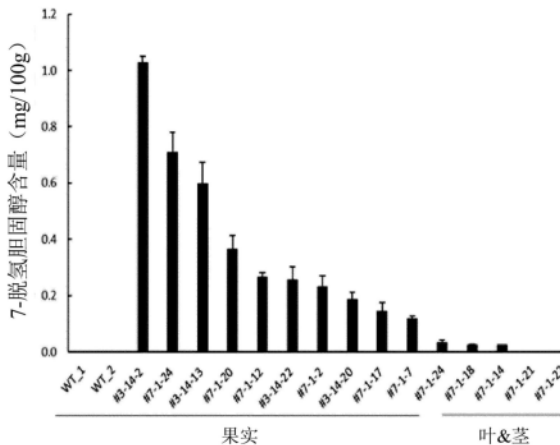
(54) 发明名称

含高浓度7-脱氢胆固醇的番茄及其生产方

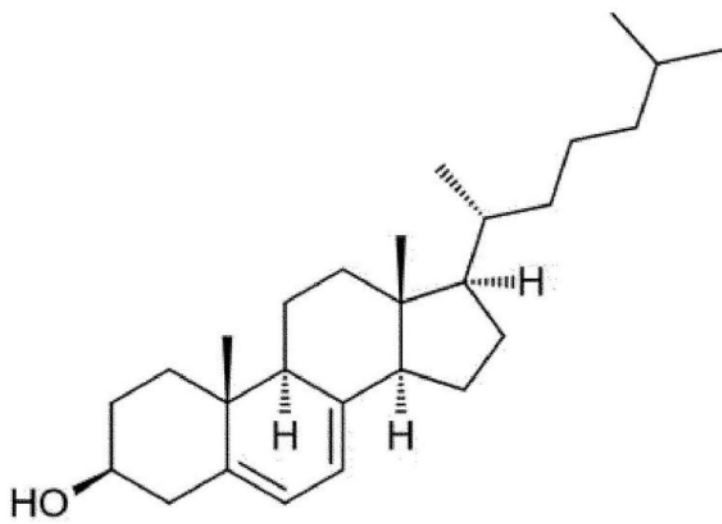
法

(57) 摘要

本发明涉及一种利用基因编辑技术生产的含高浓度维生素D3前体的番茄及其生产方法。具体地,已确认,缺失了DWF5-1基因的番茄会积累高浓度7-脱氢胆固醇,并且可由其产生种子。因此,通过这种方式生产的番茄可以含有高浓度的维生素D3前体,并且具有高商业适用性。



1. 一种含有高浓度下式1的7-脱氢胆固醇的转化番茄：
[式1]



2. 根据权利要求1所述的转化番茄,其中所述番茄的果实和根含有高浓度的7-脱氢胆固醇。
3. 根据权利要求1所述的转化番茄,其中7-脱氢胆固醇的浓度为每100g重量0.01mg至1.5mg。
4. 根据权利要求1所述的转化番茄,其中所述转化番茄经过基因改造以与野生型番茄相比降低DWF5-1基因或DWF5-1蛋白的表达或活性。
5. 根据权利要求4所述的转化番茄,其中所述DWF5-1基因包含核酸序列SEQ ID NO:63。
6. 根据权利要求4所述的转化番茄,其中所述基因改造是通过所述DWF5-1基因的核酸序列的修饰来诱导的。
7. 根据权利要求6所述的转化番茄,其中所述核酸序列的修饰包括通过替换、缺失或插入而与所述野生型番茄不同的序列。
8. 根据权利要求4所述的转化番茄,其中所述番茄进一步经过基因改造以降低CPD基因或CPD蛋白的表达或活性。
9. 根据权利要求8所述的转化番茄,其中所述CPD包含核酸序列SEQ ID NO:5。
10. 根据权利要求8所述的转化番茄,其中所述基因改造是通过所述CPD基因的核酸序列的修饰来诱导的。
11. 根据权利要求8所述的转化番茄,其中所述番茄进一步经过基因改造以降低SMT1基因或SMT1蛋白的表达或活性。
12. 根据权利要求11所述的转化番茄,其中所述SMT1包含核酸序列SEQ ID NO:7。
13. 根据权利要求11所述的转化番茄,其中所述基因改造是通过所述SMT1基因的核酸序列的修饰来诱导的。
14. 根据权利要求4所述的转化番茄,其中所述转化番茄是纯合子。
15. 根据权利要求4所述的转化番茄,其中所述转化番茄形成种子。
16. 一种载体,包含至少一个与DWF5-1基因的核苷酸序列互补结合的sgRNA(单链向导RNA)和编码CRISPR(成簇的规律间隔的回文重复序列)相关蛋白的核苷酸序列。

17. 根据权利要求16所述的载体, 其中所述sgRNA包含选自由SEQ ID NO:20、SEQ ID NO:21及其组合组成的组中的任一个核苷酸序列。

18. 根据权利要求16所述的载体, 其中所述载体还包含一个或多个与CPD基因的核苷酸序列互补结合的sgRNA。

19. 根据权利要求18所述的载体, 其中所述sgRNA包含选自由SEQ ID NO:28、SEQ ID NO:29及其组合组成的组中的任一个核苷酸序列。

20. 根据权利要求16所述的载体, 其中所述载体还包含一个或多个与SMT1基因的核苷酸序列互补结合的sgRNA。

21. 根据权利要求20所述的载体, 其中所述sgRNA包含选自由SEQ ID NO:32、SEQ ID NO:33及其组合组成的组中的任一个核苷酸序列。

22. 根据权利要求16所述的载体, 其中所述CRISPR相关蛋白为选自由以下组成的组中的任意一种: Cas9、Cpf1、c2c1、C2c2、Cas13、c2c3、Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas5e (CasD)、Cas6、Cas6e、Cas6f、Cas7、Cas8a、Cas8a1、Cas8a2、Cas8b、Cas8c、Csn1、Csx12、Cas10、Cas10d、Cas10、Cas10d、CasF、CasG、CasH、Csy1、Csy2、Csy3、Cse1 (CasA)、Cse2 (CasB)、Cse3 (CasE)、Cse4 (CasC)、Csc1、Csc2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx15、Csf1、Csf2、Csf3、Csf4和Cul966。

23. 根据权利要求22所述的载体, 其中所述CRISPR相关蛋白是Cas9。

24. 一种生产含有高浓度7-脱氢胆固醇的转化番茄的方法, 包括利用农杆菌将根据权利要求16至23中任一项所述的载体引入番茄中。

含高浓度7-脱氢胆固醇的番茄及其生产方法

技术领域

[0001] 本发明涉及一种含有高浓度7-脱氢胆固醇(维生素D3的前体)的番茄及其生产方法。

背景技术

[0002] 维生素D是维持生命所必需的营养素。据报道,维生素D不足不仅会增加佝偻病、软骨病、骨质疏松症的发病率,还会增加乳腺癌、结直肠癌、前列腺癌等恶性肿瘤;包括高血压在内的心血管疾病;糖尿病;多发性硬化症;银屑病;类风湿性关节炎;肺结核等等的发病率。维生素D是一种可以通过食物摄入的维生素,但也可以通过充足的阳光在体内合成。因此,如果你缺乏足够的阳光照射,应该摄入富含维生素D的食物或补充维生素D营养补充剂。

[0003] 尤其是,现代人一天中的大部分时间都在室内,即使在户外活动时也会用衣服和帽子遮挡阳光,因此需要注意维生素D的缺乏。在日光照射不足的情况下,必须通过食物摄入足够的维生素D,但存在的问题是维生素D含量高的食物并不多。即使已知食物富含维生素D,但每100克食物中所含的维生素D的量为:0.29 μ g/约7片(约100g)奶酪片、2.9 μ g/100g金枪鱼、1.36 μ g/2个(约100g)鸡蛋、4.4 μ g/约1杯(约100g)葡萄干。因此,从这些食物中不可能满足每日建议的维生素D摄入量。

[0004] 此外,作为维生素D的营养补充剂,合成维生素D被报道存在副作用和不良反应,而天然维生素D存在非常昂贵且难以获取的问题。另一方面,韩国专利公布号:10-2017-0138657公开了当使DWF5(δ 5,7-固醇 δ 7还原酶,DWARF5)、CPD(组成型光形态发生DWARF, DWF3)和SMT1(固醇甲基转移酶1)基因缺失时,可以生产出富含维生素D的前体7-脱氢菜油甾醇的莴苣。然而,没有对富含维生素D3的前体7-脱氢胆固醇的番茄的生产的研究。

发明内容

[0005] 技术问题

[0006] 因此,本发明人通过CRISPR技术缺失DWF5-1、CPD及SMT1基因来进行实验。结果证实,当DWF5-1基因缺失时,生产出含有高浓度7-脱氢胆固醇的番茄。基于上述内容,本发明人完成了本发明。具体而言,本发明旨在确定针对开发含有高浓度7-脱氢胆固醇(维生素D3产生的前体)的番茄而优化的DWF5-1、CPD和SMT1缺失位点,并开发包含它们的富含维生素D3前体的番茄。

[0007] 解决方案

[0008] 为了实现上述目的,在本发明的一个方面,提供了一种含有高浓度的下述式1的7-脱氢胆固醇的转化番茄。

[0009] 在本发明的另一个方面,提供了一种载体,其包含一个或多个与DWF5-1基因的核苷酸序列互补结合的sgRNA(单链向导RNA)和编码CRISPR(成簇的规律间隔的回文重复序列)相关蛋白的核苷酸序列。

[0010] 在本发明的另一个方面,提供了一种生产含有高浓度7-脱氢胆固醇的转化番茄的

方法,包括利用农杆菌 (*Agrobacterium*) 将载体引入番茄。

[0011] 发明效果

[0012] 当根据本发明生产的转化番茄是包含DWF5-1基因突变的纯合子时,番茄的果实和根会积累大量的7-脱氢胆固醇。此外,纯合子 (T2) 能够产生种子,从而保留所需的番茄特性。因此,本发明的转化番茄具有高商业适用性。

附图说明

[0013] 图1是显示DWF5-1 sgRNA所靶向的DWF5-1位点的图。

[0014] 图2是显示CPD sgRNA所靶向的CPD位点的图。

[0015] 图3是显示SMT1 sgRNA所靶向的SMT1位点的图。

[0016] 图4是显示使用RGEN确认DWF5-1 sgRNA的基因编辑效率所获得的结果的图。

[0017] 图5是显示使用RGEN确认CPD sgRNA的CPD基因编辑效率所获得的结果的图。

[0018] 图6是显示使用RGEN确认SMT1 sgRNA的SMT1基因编辑效率所获得的结果的图。

[0019] 图7显示了负载有DWF5-1 sgRNA的Cas9-sgRNA载体的示意图。

[0020] 图8显示了负载有DWF5-1 sgRNA和CPD sgRNA的Cas9-sgRNA载体的示意图。

[0021] 图9显示了负载有DWF5-1 sgRNA、CPD sgRNA和SMT1 sgRNA的Cas9-sgRNA载体的示意图。

[0022] 图10是显示利用农杆菌构建转化番茄的过程的流程图。

[0023] 图11显示了通过基因编辑构建的番茄的发芽过程。

[0024] 图12是显示对DWF5-1基因被编辑的番茄植株体进行观察所获得的结果的图。

[0025] 图13是显示对DWF5-1和CPD基因被编辑的番茄植株体进行观察所获得的结果的图。

[0026] 图14是显示对DWF5-1、CPD和SMT1基因被编辑的番茄植株体进行观察所获得的结果的图。

[0027] 图15是显示使用T7E1检测确认DWF5-1、CPD或SMT1基因是否被编辑所获得的结果的图。

[0028] 图16是显示利用Sanger测序确认DWF5-1基因是否缺失所获得的结果的图。

[0029] 图17是显示通过下一代测序对DWF5-1基因缺失的T1代转化番茄植株体中的DWF5-1基因表型进行确认所获得的结果的图。

[0030] 图18a是显示DWF5-1基因缺失的T1代纯合转化番茄植株体的基因编辑效率和收获种子量(上)、及核苷酸序列(下)的图。

[0031] 图18b是显示DWF5-1基因缺失的T1代纯合转化番茄植株体的核苷酸序列的图。

[0032] 图19是显示用于测量DWF5-1基因缺失的纯合子 (T1) 的T2代转化番茄植株中所含的7-脱氢胆固醇的样本制备过程的流程图。

[0033] 图20是显示通过LC/MS测量DWF5-1基因缺失的纯合子 (T1) 的T2代转化番茄植株中所含的7-脱氢胆固醇的测量条件的图。

[0034] 图21是显示以市售的7-脱氢胆固醇为标准试剂,通过LC/MS测量野生型 (WT) 番茄植株体中所含的7-脱氢胆固醇、以及DWF5-1基因缺失的纯合子 (T1) 的T2代转化番茄植株体中所含的7-脱氢胆固醇所获得的结果的图。

[0035] 图22是显示通过LC/MS测量DWF5-1基因缺失的纯合子(T1)的T2代转化番茄植株体的果实中所含的7-脱氢胆固醇所获得的结果的图。

[0036] 图23是显示通过RT-PCR确认DWF5-1基因缺失的纯合子(T1)的T2代转化番茄植株体中的潮霉素及草丁膦(PPT)抗性基因的表达的结果的图。

[0037] 图24是显示对抗生素抗性基因未表达的T2代转化番茄植株体(D100_3-14-29和D100_7-1-15)的DWF5-1的外显子6域的核苷酸序列与野生型(WT)番茄植株体的核苷酸序列进行比较分析所获得的结果的图。

[0038] 图25是显示对抗生素抗性基因未表达的T2代转化番茄植株体(D100_3-14-29和D100_7-1-15)的DWF5-1的外显子6域的氨基酸序列与野生型(WT)番茄植株体的氨基酸序列进行比较分析所获得的结果的图。

[0039] 图26是显示DWF5-1基因缺失的纯合(T2)转化番茄植株体(T2)的表型的图。

[0040] 图27是显示通过q-PCR确认番茄植株体的根、茎、叶及果实中的DWF5-1或DWF5-2基因的表达量所获得的结果的图。**** $p < 0.0001$

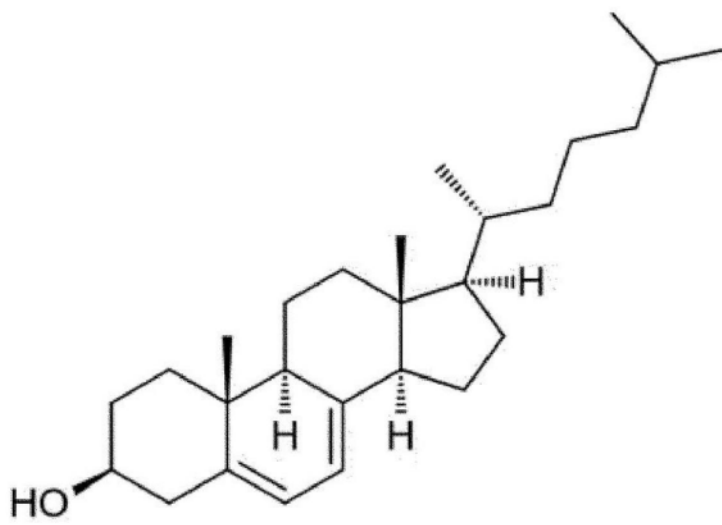
[0041] 图28是显示通过GC/MS测量DWF5-1基因缺失的纯合子(T1)的T3代转化番茄植株体的根(培养根)中所含的7-脱氢胆固醇所获得的结果的图。

[0042] 本发明的最佳实施方式

[0043] 在本发明的一个方面中,提供了一种含有高浓度的下述式1的7-脱氢胆固醇的转化番茄,

[0044] [式1]

[0045]



[0046] 如本文所用,术语“番茄(tomato)”的学名为Solanum lycopersicum,是指属于茄科、茄目植物的一种植物或其果实。其是一年生植物,原产拉丁美洲,高1至3米,花黄色。这种水果因含有番茄红素而呈红色,用于食用。本说明书中,番茄可以为植株体,可以是植物器官(例如,根、茎、叶、花瓣(花)、种子、果实等)、植物组织(表皮、筛管、软组织、木质部、维管组织)等。具体来说,它可以是果实和/或根。在这种情况下,番茄植株体可以包括含有根培养物等的细胞培养物。

[0047] 本文使用的术语“7-脱氢胆固醇”是转化为维生素D3的维生素D3原。在本发明中,转化番茄可以含有每100g重量0.01mg至1.5mg的7-脱氢胆固醇。具体地,转化番茄可以含有每100g重量0.01mg、0.10mg、0.20mg、0.30mg、0.40mg、0.50mg、0.60mg、0.70mg、0.80mg、

0.90mg、1.00mg、1.10mg、1.20mg、1.30mg、1.40mg或1.50mg的7-脱氢胆固醇。

[0048] 在本发明中,可以对转化番茄进行基因改造,以与野生型番茄相比降低番茄DWF5-1 (DWF5-1) 基因或DWF5-1蛋白的表达或活性。在这种情况下,可以通过核酸序列的修饰来诱导基因改造。

[0049] 本文中使用的术语“DWF5”是指7-脱氢胆固醇还原酶。这种酶也被称为 $\delta 5,7$ -固醇 $\delta 7$ 还原酶。该酶具有还原植物固醇代谢途径的中间代谢物5-脱氢表固醇并将其转化为24-亚甲基胆固醇的活性。

[0050] 已知在拟南芥 (*Arabidopsis thaliana*) 中,DWF5以单一基因的形式存在于植物中,当DWF5基因缺失时,植物固醇代谢会急剧下降,从而导致侏儒症,无法生长至成年期。在番茄中,DWF5以两种形式存在:DWF5-1和DWF5-2。在这种情况下,DWF5-1和DWF5-2可分别包含氨基酸序列SEQ ID NO:2和SEQ ID NO:4。据报道,DWF5-1基因在植物固醇代谢途径中发挥作用,DWF5-2基因会通过还原7-脱氢胆固醇 (7-脱氢胆固醇还原酶) 并将其转化为胆固醇而在胆固醇生成途径中发挥作用。因此,当DWF5-1基因缺失时,可以预期由于植物固醇代谢很可能会降低,因此导致生长异常。但由于有报道称植物中的胆固醇途径是从植物固醇代谢途径进化而来的,因此不能完全排除DWF5-1和DWF5-2彼此具有功能同源性的可能性。本发明对DWF5-1和DWF5-2基因表达模式进行比较分析的结果显示,在番茄的叶、茎、根、花、未成熟果实和成熟果实中,DWF5-1的表达量均是DWF5-2基因的5-10倍以上,而DWF5-2基因在所有组织中也均以低表达水平表达。因此,假设DWF5-1和DWF5-2具有功能同源性,推测DWF5-1基因可以更有效地提高番茄果实中的7-脱氢胆固醇含量。

[0051] 本发明中,仅缺失DWF5-1基因的突变植株体在未成熟和成熟的番茄果实和根组织中均产生了高含量的维生素D3原,其生长和种子产生与野生型番茄相似 (图18、图26和表13)。因此,在本发明中,通过至少缺失DWF5-1基因,开发出了在成熟果实中产生高浓度维生素D3原的番茄,并且确认了单独缺失DWF5-1基因是能够培育并生产与野生型相似的果实的工业实用性的技术。

[0052] 在本发明中,DWF5可以为DWF5-1。具体地,在本发明中,DWF5可以为包含氨基酸SEQ ID NO:2的DWF5-1。DWF5-1基因可以是包含编码DWF5-1的核苷酸序列的基因组DNA、cDNA或RNA。更具体地,在本发明中,DWF5-1基因可以是基因组DNA。在本发明中,DWF5、DWF5-1和S1DWF5-1可以互换使用。

[0053] 本文中使用的术语“基因组DNA”是指染色体DNA,是指真核细胞中遗传信息的编码形式。真核生物基因组DNA (以下简称DNA) 包含外显子和内含子。外显子是包含编码蛋白质的核酸序列的部分,内含子是不参与蛋白质合成的部分。为了产生生物体功能所必需的各种蛋白质,DNA被转录成RNA,此时除内含子以外的外显子被连接在一起。内含子包含有助于转录的信息,例如诱导转录起始的启动子,并用于在转录过程中制造前mRNA,但是会被截去而不用来制造成熟的mRNA。

[0054] 在本发明中,DWF5-1基因可以包含12个外显子和11个内含子。具体地,DWF5-1可以包括外显子1、外显子2、外显子3、外显子4、外显子5、外显子6、外显子7、外显子8、外显子9、外显子10、外显子11或外显子12。此外,DWF5-1可包含内含子1、内含子2、内含子3、内含子4、内含子5、内含子6、内含子7、内含子8、内含子9、内含子10或内含子11。

[0055] 更具体地,DWF5-1基因的外显子1至外显子12可以分别包含核酸序列SEQ ID NO:

40至SEQ ID NO:51,或由核酸序列SEQ ID NO:40至SEQ ID NO:51组成。DWF5-1基因的内含子1至内含子11可以分别包含核酸序列SEQ ID NO:52至SEQ ID NO:62或由核酸序列SEQ ID NO:52至SEQ ID NO:62组成。

[0056] 在一个实施方式中,在本发明中,DWF5-1基因可以包括核酸序列SEQ ID NO:63或由核酸序列SEQ ID NO:63组成。

[0057] 在本发明中,DWF5-1基因可以包含调节DWF5-1转录的侧翼区。

[0058] 如本文所用,术语“侧翼区”是指在特定基因两侧延伸的DNA序列,并且是不转录成RNA的位点。与基因5'端相邻的DNA区域称为5'侧翼区,与基因3'端相邻的DNA区域称为3'侧翼区。侧翼区包含调控序列,通过上述序列与参与转录的蛋白质结合,从而调控基因转录。特别地,在真核生物的情况下,5'侧翼区包含增强子、沉默子、启动子等,并通过与转录因子和RNA聚合酶等蛋白质结合来调节转录。

[0059] 在本发明的一个实施方式中,DWF5-1基因的5'侧翼区可以包含核酸序列SEQ ID NO:64。在一个实施方式中,DWF5-1基因的3'侧翼区可以包含核酸序列SEQ ID NO:65。

[0060] 更具体地,在本发明中,DWF5-1基因可以是cDNA。在一个实施方式中,它可以包含核酸序列SEQ ID NO:1或由核酸序列SEQ ID NO:1组成。

[0061] 本文使用的术语“基因改造”或“基因改造的”是指向细胞或由其产生的细胞引入一个或多个基因修饰的行为。

[0062] 如本文所用,术语“基因或蛋白的表达或活性的降低”是指与可以比较靶基因或蛋白质的相同物种的野生型相比,靶基因或蛋白的表达或活性更低。另外,“失活”是指与野生型靶基因或蛋白相比,产生不表达的蛋白质或即使表达也没有活性的蛋白质。

[0063] 也就是说,DWF5-1基因或蛋白的表达或活性的降低或失活,意味着野生型番茄的DWF5-1基因正常发挥的生物学功能或作用全部或部分丧失。例如,DWF5-1基因表达的蛋白质可能过早终止或失去其作为蛋白质的正常功能。具体地,基因编辑可以通过在靶区域生成终止密码子或者生成编码与野生型不同的氨基酸的密码子来敲除DWF5-1基因。此外,还可以向不产生蛋白质的非编码DNA序列中引入突变,但不限于此。

[0064] 具体地,与野生型番茄的DWF5-1基因或蛋白相比,经基因改造的番茄可以具有减少约20%或更多、约30%或更多、约40%或更多、约50%或更多、约60%或更多、约70%或更多、约80%或更多、约85%或更多、约90%或更多、约95%或更多、或约100%的DWF5-1基因或蛋白。可以使用本领域已知的任何方法测定DWF5-1基因或DWF5-1蛋白的表达或活性。

[0065] 可以通过核酸序列的修饰来诱导基因改造。具体而言,可以通过替换(转换)、缺失或插入来修饰核酸序列,从而实现基因改造。通过上述基因改造,本发明的转化番茄的DWF5-1基因中可以包含与野生型番茄的DWF5-1基因不同的基因序列,并且上述基因的生物学功能可能丧失。

[0066] 更具体地说,DWF5-1基因的工程改造可以通过选自由以下组成的组中的核酸序列的任一种修饰来诱导:

[0067] i) 缺失DWF5-1基因的全部或1至50个连续核苷酸序列;

[0068] ii) 将DWF5-1基因中的1至50个核苷酸替换为与野生型DWF5-1基因不同的核苷酸;

[0069] iii) 将1至50个各自独立地选自A、T、C和G的核苷酸插入DWF5-1基因中;以及

[0070] iv) 上述各项的组合。

[0071] 更具体地说,可以通过选自由以下组成的组中的核酸序列的任一种修饰来诱导基因改造:

[0072] i) 全部或部分缺失DWF5-1基因,例如缺失DWF5-1基因的一个或多个核苷酸,例如缺失1至50、1至45、1至40、1至35、1至30、1至25、1至20、1至15、1至10、1至5、1至3或1个核苷酸;

[0073] ii) 将DWF5-1基因中的一个或多个核苷酸,例如1至50、1至45、1至40、1至35、1至30、1至25、1至20、1至15、1至10、1至5、1至3或1个核苷酸,替换为与野生型DWF5-1基因不同的核苷酸;

[0074] iii) 将一个或多个选自A、T、C和G的核苷酸,例如1至50、1至45、1至40、1至35、1至30、1至25、1至20、1至15、1至10、1至5、1至3或1个核苷酸,插入DWF5-1基因的任意位点;以及

[0075] iv) i) 至iii) 的组合。

[0076] DWF5-1基因的核酸序列被修饰的部分(“靶区域”)可以是上述基因中的1个或更多个、3个或更多个、5个或更多个、7个或更多个、10个或更多个、12个或更多个、15个或更多个、17个或更多个、20个或更多个、25个或更多个、27个或更多个、30个或更多个、33个或更多个、37个或更多个、40个或更多个、43个或更多个、47个或更多个或50个或更多个的连续核苷酸序列的区域。

[0077] 在本发明的一个实施方式中,DWF5-1基因的核酸序列的修饰可以包括在选自由外显子6 (SEQ ID NO:9)、外显子7 (SEQ ID NO:10)、外显子9 (SEQ ID NO:11)、外显子10 (SEQ ID NO:12)、外显子11 (SEQ ID NO:13) 及其组合组成的组中的任一个域的核酸序列中。在本发明的一个实施方式中,DWF5-1基因的核酸序列的修饰可以包括在外显子6域的核酸序列中。具体而言,包含DWF5-1的外显子6域的核酸序列的SEQ ID NO:9的核苷酸序列可以在核酸序列中通过替换、缺失或插入进行修饰。更具体地,可以通过在包含DWF5-1基因的外显子6域的核酸序列的SEQ ID NO:9的核苷酸序列中进行替换、缺失或插入来诱导序列的修饰,以降低DWF5-1基因或DWF5-1蛋白的表达或活性,或导致存在终止密码子。

[0078] 本发明的转化番茄可以进一步经过基因改造以降低CPD基因或CPD蛋白的表达或活性。基因改造与上面描述的相同。

[0079] 如本文所用,术语“CPD(组成型光形态发生DWARF)”是细胞色素P450单加氧酶的CP90A家族的一种酶,也称为DWF3。在芸苔素内酯(brassinolide)的生物合成过程中,其作用于C6氧化途径,将芸苔素甾酮(cathasterone)转化为茶甾酮(teasterone),将6-脱氧芸苔素甾酮转化为6-脱氧茶甾酮。在本发明中,上述酶的作用是将7-脱氢菜油甾醇(7-dehydrocampesterol)还原为7-脱氢菜油甾烷醇(7-dehydrocampestanol)。

[0080] 在这种情况下,CPD可包含氨基酸序列SEQ ID NO:6。此外,编码其的核酸可以包含核苷酸序列SEQ ID NO:5。

[0081] 在本发明中,CPD基因或CPD蛋白的表达或活性的降低或失活,意味着野生型番茄的CPD基因正常发挥的生物学功能或作用全部或部分丧失。此时,基因或蛋白的表达或活性的降低或失活与上述相同。例如,CPD基因表达的蛋白质可能过早终止或失去其作为蛋白质的正常功能。具体地,基因编辑可以通过在靶区域生成终止密码子或者生成编码与野生型不同的氨基酸的密码子来敲除CPD基因。此外,还可以向不产生蛋白质的非编码DNA序列中引入突变,但不限于此。

[0082] 具体地,与野生型番茄的CPD基因或蛋白相比,经基因改造的番茄可以具有减少约20%或更多、约30%或更多、约40%或更多、约50%或更多、约60%或更多、约70%或更多、约80%或更多、约85%或更多、约90%或更多、约95%或更多、或约100%的CPD基因或蛋白。可以使用本领域已知的任何方法测定CPD基因或CPD蛋白的表达或活性。

[0083] 可以通过核酸序列的修饰来诱导基因改造。具体而言,可以通过替换、缺失或插入来修饰核酸序列,从而实现基因改造。通过上述基因改造,本发明的转化番茄的CPD基因中可以包含与野生型番茄的CPD基因不同的基因序列,并且上述基因的生物功能可能丧失。

[0084] 更具体地说,CPD基因的工程改造可以通过选自由以下组成的组中的核酸序列的任一种修饰来诱导:

[0085] i) 缺失CPD基因的全部或1至50个连续核苷酸序列;

[0086] ii) 将CPD基因中的1至50个核苷酸替换为与野生型CPD基因不同的核苷酸;

[0087] iii) 将1至50个各自独立地选自A、T、C和G的核苷酸插入CPD基因中;以及

[0088] iv) 上述各项的组合。

[0089] 更具体地说,可以通过选自由以下组成的组中的核酸序列的任一种修饰来诱导基因改造:

[0090] i) 全部或部分缺失CPD基因,例如缺失CPD基因的一个或多个核苷酸,例如缺失1至50、1至45、1至40、1至35、1至30、1至25、1至20、1至15、1至10、1至5、1至3或1个核苷酸;

[0091] ii) 将CPD基因中的一个或多个核苷酸,例如1至50、1至45、1至40、1至35、1至30、1至25、1至20、1至15、1至10、1至5、1至3或1个核苷酸,替换为与野生型CPD基因不同的核苷酸;

[0092] iii) 将一个或多个选自A、T、C和G的核苷酸,例如1至50、1至45、1至40、1至35、1至30、1至25、1至20、1至15、1至10、1至5、1至3或1个核苷酸,插入CPD基因的任意位点;以及

[0093] iv) i) 至iii) 的组合。

[0094] CPD基因的核酸序列被修饰的部分(“靶区域”)可以是上述基因中的1个或更多个、3个或更多个、5个或更多个、7个或更多个、10个或更多个、12个或更多个、15个或更多个、17个或更多个、20个或更多个、25个或更多个、27个或更多个、30个或更多个、33个或更多个、37个或更多个、40个或更多个、43个或更多个、47个或更多个或50个或更多个的连续核苷酸序列的区域。

[0095] 在本发明的一个实施方式中,CPD基因的核酸序列的修饰可以包括在选自由外显子7(SEQ ID NO:14)、外显子8(SEQ ID NO:15)及其组合组成的组中的任一个域的核酸序列中。在本发明的一个实施方式中,CPD基因的核酸序列的修饰可以是外显子8域的核酸序列的修饰。具体地,包含CPD基因的外显子8域的核酸序列的SEQ ID NO:15的核苷酸序列可以在核酸序列中通过替换、缺失或插入进行修饰。更具体地说,可以通过在包含CDP基因的外显子8域的核酸序列的SEQ ID NO:15的核苷酸序列中进行替换、缺失或插入来诱导序列的修饰,以降低CDP的活性或导致存在终止密码子。

[0096] 本发明的包含DWF5-1和CPD基因突变的转化番茄可以进一步经过基因改造以降低SMT1基因或SMT1蛋白的表达或活性。基因改造与上面描述的相同。

[0097] 如本文所用,术语“SMT1(固醇甲基转移酶1)”是一种将甲基基团添加到类固醇侧链的酶。具体来说,SMT1的作用是将5-脱氢表甾醇转化为7-脱氢菜油甾醇。

[0098] 在这种情况下,SMT1可包含氨基酸序列SEQ ID NO:8。此外,编码其的核酸可以包含核苷酸序列SEQ ID NO:7。

[0099] 在本发明中,SMT1基因或SMT1蛋白的表达或活性的降低或失活,意味着野生型番茄的SMT1基因正常发挥的生物学功能或作用全部或部分丧失。此时,基因或蛋白的表达或活性的降低或失活与上述相同。例如,SMT1基因表达的蛋白质可能过早终止或失去其作为蛋白质的正常功能。具体地,基因编辑可以通过在靶区域生成终止密码子或者生成编码与野生型不同的氨基酸的密码子来敲除SMT1基因。此外,还可以向不产生蛋白质的非编码DNA序列中引入突变,但不限于此。

[0100] 具体地,与野生型番茄的SMT1基因或SMT1蛋白相比,经基因改造的番茄可以具有减少约20%或更多、约30%或更多、约40%或更多、约50%或更多、约60%或更多、约70%或更多、约80%或更多、约85%或更多、约90%或更多、约95%或更多、或约100%的SMT1基因或SMT1蛋白。可以使用本领域已知的任何方法测定SMT1基因或SMT1蛋白的表达或活性。

[0101] 可以通过核酸序列的修饰来诱导基因改造。具体而言,可以通过替换、缺失或插入来修饰核酸序列,从而实现基因改造。通过上述基因改造,本发明的转化番茄的SMT1基因中可以包含与野生型番茄的SMT1基因不同的基因序列,并且上述基因的生物学功能可能丧失。

[0102] 更具体地说,SMT1基因的工程改造可以通过选自由以下组成的组中的核酸序列的任一种修饰来诱导:

[0103] i) 缺失SMT1基因的全部或1至50个连续核苷酸序列区域;

[0104] ii) 将SMT1基因中的1至50个核苷酸替换为与野生型SMT1基因不同的核苷酸;

[0105] iii) 将1至50个各自独立地选自A、T、C和G的核苷酸插入SMT1基因中;以及

[0106] iv) 上述各项的组合。

[0107] 更具体地说,可以通过选自由以下组成的组中的核酸序列的任一种修饰来诱导基因改造:

[0108] i) 全部或部分缺失SMT1基因,例如缺失SMT1基因的一个或多个核苷酸,例如缺失1至50、1至45、1至40、1至35、1至30、1至25、1至20、1至15、1至10、1至5、1至3或1个核苷酸;

[0109] ii) 将SMT1基因中的一个或多个核苷酸,例如1至50、1至45、1至40、1至35、1至30、1至25、1至20、1至15、1至10、1至5、1至3或1个核苷酸,替换为与野生型SMT1基因不同的核苷酸;

[0110] iii) 将一个或多个选自A、T、C和G的核苷酸,例如1至50、1至45、1至40、1至35、1至30、1至25、1至20、1至15、1至10、1至5、1至3或1个核苷酸,插入SMT1基因的任意位点;以及

[0111] iv) i) 至iii) 的组合。

[0112] SMT1基因的核酸序列被修饰的部分(“靶区域”)可以是上述基因中的1个或更多个、3个或更多个、5个或更多个、7个或更多个、10个或更多个、12个或更多个、15个或更多个、17个或更多个、20个或更多个、25个或更多个、27个或更多个、30个或更多个、33个或更多个、37个或更多个、40个或更多个、43个或更多个、47个或更多个或50个或更多个的连续核苷酸序列的区域。

[0113] 在本发明的一个实施方式中,SMT1基因的核酸序列的修饰可以包括在选自由外显子7(SEQ ID NO:16)、外显子9(SEQ ID NO:17)、外显子11(SEQ ID NO:18)、外显子12(SEQ

ID NO:19) 及其组合组成的组中的任一个域的核酸序列中。在本发明的一个实施方式中, SMT1基因的突变可以包括外显子7和外显子9域的核酸序列的修饰。具体地, SMT1基因的突变可以是对包含外显子7和外显子9域的核酸序列的SEQ ID NO:16和SEQ ID NO:17的核苷酸序列进行替换、缺失或插入的修饰。更具体地说, 可以通过在包含SMT1基因的外显子7和外显子9域的核酸序列的SEQ ID NO:16和SEQ ID NO:17的核苷酸序列中进行替换、缺失或插入来诱导序列的修饰, 以降低SMT1的活性或导致存在终止密码子。

[0114] 转化番茄可以是纯合子。也就是说, 核酸序列的修饰可以包括在编码DWF5-1的两个基因中。在一个实施方式中, 核酸序列的修饰可以包括在编码选自DWF5-1基因的外显子6、外显子7、外显子9、外显子10、外显子11和其组合组成的组中的任一个域的两个基因中。在本发明的一个实施方式中, 转化番茄可以在编码DWF5-1基因的外显子6域的两个基因中具有修饰的序列, 该修饰的序列与野生型DWF5-1基因的核酸序列不同。

[0115] 在转化番茄中, 核酸序列的修饰可以包括在编码DWF5-1、CPD和SMT1的两个基因中。在这种情况下, 转化番茄可以在编码选自CPD基因的外显子7、外显子8和其组合组成的组中的任一个域的两个基因中具有修饰的序列, 该修饰的序列与野生型CPD基因的核酸序列不同。此外, 核酸序列的修饰可以包括在编码选自SMT1基因的外显子7、外显子9、外显子11、外显子12及其组合组成的组中的任一个域的两个基因中。在这种情况下, 核酸序列的修饰与上面描述的不同。

[0116] 在本发明的一个实施方式中, 在转化番茄中, 核酸序列的修饰可以包括在各自编码DWF5-1的外显子6域、CDP的外显子7域以及SMT1的外显子7和外显子9域的两个基因中。

[0117] 当本发明的转化番茄为纯合子, 且DWF5-1基因或蛋白的表达或活性降低或失活时, 番茄可以形成种子。当转化番茄为纯合子, 且DWF5-1、CPD和SMT1基因或蛋白的表达或活性降低或失活时, 番茄可能不会结果, 也不会形成种子。

[0118] 在本发明的另一个方面, 提供了一种载体, 其包含一个或多个与DWF5-1基因的核苷酸序列互补结合的sgRNA和编码CRISPR相关蛋白的核苷酸序列。此处, DWF5-1与上面描述的不同。此外, DWF5-1基因可以是基因组DNA、cDNA或RNA。

[0119] 本文中使用的术语“向导RNA(gRNA)”是指通过基因组编辑识别细胞中的靶核酸并切割、插入或连接靶核酸的多核苷酸。gRNA可以包含与靶核酸中的靶序列互补的序列。gRNA可以是与靶核酸中PAM的5'或3'方向的2至24个连续核苷酸(以下称为'nt')的核苷酸序列互补的多核苷酸。gRNA的长度可以是17nt、18nt、19nt、20nt、21nt、22nt、23nt或24nt。

[0120] gRNA可以是单链向导RNA(sgRNA)。sgRNA可以包括针对靶核酸序列的CRISPR RNA(crRNA)和/或与CRISPR相关蛋白形成复合物的反式激活crRNA(tracrRNA)。sgRNA可以包含具有与靶基因中的靶核酸序列(靶区域)互补的核苷酸序列(靶向序列)的部分(也称为间隔区、靶DNA识别序列、碱基配对区等)和用于结合CRISPR相关蛋白的发夹结构。更具体地说, 它可以包含与靶基因中的靶核酸序列互补的核苷酸序列(靶向序列)的部分、用于结合CRISPR相关蛋白的发夹结构、以及终止子序列。上述结构可以从5'至3'依次存在, 但不限于此。可以使用任何类型的sgRNA, 只要sgRNA包含crRNA和tracrRNA的主要区域以及与靶基因互补的核苷酸序列即可。

[0121] 在本发明中, sgRNA可以与DWF5-1的基因组DNA的核酸序列互补结合。具体来说, sgRNA可以与SEQ ID NO:63互补结合。更具体地, sgRNA可以与选自DWF5-1基因的外显子1

至外显子12、内含子1至内含子11、5'侧翼区和3'侧翼区组成的组中的一个或多个于一个域的核酸序列互补结合。更具体地说,sgRNA可以与选自SEQ ID NO:40至SEQ ID NO:62、SEQ ID NO:64和SEQ ID NO:65组成的组中的一个或多个于一个域的核酸序列互补结合。

[0122] 在本发明中,sgRNA可以与DWF5-1的cDNA的核酸序列互补结合。具体地,sgRNA可以与SEQ ID NO:1的核酸序列互补结合。

[0123] 在本发明的一个实施方式中,sgRNA可以与选自DWF5-1基因的外显子6、外显子7、外显子9、外显子10和外显子11组成的组中的任一个域的核酸序列互补结合。在一个实施方式中,与DWF5-1基因的外显子6域的核酸序列结合的sgRNA可以包括核苷酸序列SEQ ID NO:20或SEQ ID NO:21。在一个实施方式中,与外显子7域的核酸序列结合的sgRNA可以包括核苷酸序列SEQ ID NO:22。在一个实施方式中,与外显子9域的核酸序列结合的sgRNA可以包括核苷酸序列SEQ ID NO:23或SEQ ID NO:24。在一个实施方式中,与外显子10域结合的sgRNA可以包含核苷酸序列SEQ ID NO:25。在一个实施方式中,与外显子11域的核酸序列结合的sgRNA可以包括核苷酸序列SEQ ID NO:26。优选地,在本发明的一个实施方式中,sgRNA可以包括与外显子6域的核酸序列互补结合的核苷酸序列SEQ ID NO:20和SEQ ID NO:21或由其组成。

[0124] 在本发明中,载体还可以包含与CPD基因的核苷酸序列互补结合的sgRNA。与CPD基因互补结合的sgRNA可以与CPD基因的外显子7或外显子8域的核苷酸序列互补结合。与CPD基因外显子7域结合的sgRNA可以包含核苷酸序列SEQ ID NO:27。与CPD基因外显子8域结合的sgRNA可以包含核苷酸序列SEQ ID NO:28、SEQ ID NO:29、SEQ ID NO:30或SEQ ID NO:31。在本发明的一个实施方式中,sgRNA可以包含与外显子8域的核苷酸序列互补结合的核苷酸序列SEQ ID NO:28和SEQ ID NO:29。

[0125] 在本发明中,包含与DWF5基因和CPD基因的核苷酸序列互补结合的sgRNA的载体还可以包含与SMT1基因的核苷酸序列互补结合的sgRNA。在这种情况下,与SMT1基因的核苷酸序列互补结合的sgRNA可以与选自SMT1基因的外显子7、外显子9、外显子11和外显子12组成的组中的一个或多个于一个域的核苷酸序列互补结合。与SMT1基因的外显子7域结合的sgRNA可以包含核苷酸序列SEQ ID NO:32。与SMT1基因的外显子9域结合的sgRNA可以包含核苷酸序列SEQ ID NO:33。与SMT1基因的外显子11域结合的sgRNA可以包含核苷酸序列SEQ ID NO:34。与SMT1基因的外显子12域结合的sgRNA可以包含核苷酸序列SEQ ID NO:35。在本发明的一个实施方式中,sgRNA可以包含与外显子7和外显子9域的核苷酸序列互补结合的核苷酸序列SEQ ID NO:32和SEQ ID NO:33。此时,与DWF5基因和CPD基因的核苷酸序列互补结合的sgRNA与上述相同。

[0126] 本文中使用的术语“CRISPR (成簇的规律间隔的回文重复序列) 相关蛋白”(CRISPR 相关蛋白)是指当核酸例如DNA或RNA具有双链或单链(dsDNA/RNA和ssDNA/RNA)时能够识别并切割核酸的酶。具体来说,它可以识别与sgRNA结合的双链或单链核酸并将其切割。所述的CRISPR相关蛋白包括CRISPR相关蛋白以及具有其功能的突变体。

[0127] CRISPR相关蛋白可以衍生自:酿脓链球菌(*Streptococcus pyogenes*)、嗜热链球菌(*Streptococcus thermophilus*)、链球菌属(*Streptococcus* sp.)、金黄色葡萄球菌(*Staphylococcus aureus*)、达松维尔诺卡氏菌(*Nocardiosis dassonvillei*)、始旋链霉菌(*Streptomyces pristinae spiralis*)、绿色色素链霉菌(*Streptomyces*

viridochromogenes)、玫瑰链孢囊菌(*Streptosporangium roseum*)、脂环酸芽孢杆菌(*Alicyclobacillus acidocaldarius*)、假真菌样芽孢杆菌(*Bacillus pseudomycoides*)、硒还原芽孢杆菌(*Bacillus selenitireducens*)、西伯利亚微小杆菌(*Exiguobacterium sibiricum*)、德氏乳杆菌(*Lactobacillus delbrueckii*)、唾液乳杆菌(*Lactobacillus salivarius*)、海洋微颤菌(*Microscilla marina*)、伯克霍尔德氏菌(*Burkholderiales bacterium*)、食茶极地单胞菌(*Polaromonas naphthalenivorans*)、极地单胞菌属(*Polaromonas* sp.)、瓦氏鳄球藻(*Crocospaera watsonii*)、蓝丝菌属(*Cyanothece* sp.)、铜绿微囊藻(*Microcystis aeruginosa*)、铜绿假单胞菌(*Pseudomonas aeruginosa*)、聚球藻属(*Synechococcus* sp.)、阿拉伯糖醋盐杆菌(*Acetohalobium arabaticum*)、丹氏制氮菌(*Ammonifex degensii*)、贝西嗜热纤维分解菌(*Caldicelulosiruptor beccarii*)、候选金矿菌(*Candidatus Desulforudis*)、肉毒梭菌(*Clostridium botulinum*)、艰难梭菌(*Clostridium difficile*)、大芬戈尔德菌(*Finegoldia magna*)、嗜热盐碱厌氧菌(*Natronaerobius thermophilus*)、嗜热丙酸互营细菌(*Pelotomaculum thermopropionicum*)、嗜温嗜酸硫杆菌(*Acidithiobacillus caldus*)、嗜酸氧化亚铁硫杆菌(*Acidithiobacillus ferrooxidans*)、酒色异着色菌(*Allochromatium vinosum*)、海洋杆菌属(*Marinobacter* sp.)、嗜盐亚硝化球菌(*Nitrosococcus halophilus*)、瓦氏亚硝化球菌(*Nitrosococcus watsonii*)、游海假交替单胞菌(*Pseudoalteromonas haloplanktis*)、消旋纤线杆菌(*Ktedonobacter racemifer*)、伊夫氏甲烷盐菌(*Methanohalobium evestigatum*)、多变鱼腥藻(*Anabaena variabilis*)、泡沫节球藻(*Nodularia spumigena*)、念珠藻属(*Nostoc* sp.)、极大节旋藻(*Arthrospira maxima*)、钝顶节旋藻(*Arthrospira platensis*)、节旋藻属(*Arthrospira* sp.)、鞘丝藻属(*Lyngbya* sp.)、原型微鞘藻(*Microcoleus chthonoplastes*)、颤藻属(*Oscillatoria* sp.)、运动石袍菌(*Petrogobacter mobilis*)、非洲栖热腔菌(*Thermosiphon africanus*)、深海单细胞蓝藻(*Acaryochloris marina*)、沙氏纤毛菌(*Leptotrichia shahii*)、普氏菌(*Prevotella*)或新凶手弗朗西丝菌(*Francisella novicida*)。

[0128] 在本发明中,CRISPR相关蛋白可以是选自由以下组成的组中的任意一种:Cas9、Cpf1、c2c1、C2c2、Cas13、c2c3、Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas5e (CasD)、Cas6、Cas6e、Cas6f、Cas7、Cas8a、Cas8a1、Cas8a2、Cas8b、Cas8c、Csn1、Csn12、Cas10、Cas10d、Cas10、Cas10d、CasF、CasG、CasH、Csy1、Csy2、Csy3、Cse1 (CasA)、Cse2 (CasB)、Cse3 (CasE)、Cse4 (CasC)、Csc1、Csc2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx15、Csf1、Csf2、Csf3、Csf4和Cu1966。具体地,CRISPR相关蛋白可以是Cas9。

[0129] Cas9蛋白可与向导sgRNA形成复合物,充当核糖核酸蛋白(RNP)的形式。此处,“核糖核蛋白(RNP)”为RNA与蛋白质的复合物,在本发明中可以为Cas9-sgRNA复合物。

[0130] 如本文所用,术语“载体”能够被引入宿主细胞,然后重组并插入宿主细胞的基因组中。或者,载体应理解为包含能够作为附加体自主复制的多核苷酸序列的核酸运载工具。此类载体包括线性核酸、质粒、噬菌粒、粘粒、RNA载体、病毒载体、微型染色体及其类似物。病毒载体的例子包括但不限于逆转录病毒、腺病毒和腺相关病毒。

[0131] 在本发明的一个实施方式中,载体可以包含核苷酸序列SEQ ID NO:36至SEQ ID

N0:38中的任一个。此外,在本发明中,载体与SEQ ID N0:36至SEQ ID N0:38中任一个具有约70%、约75%、约80%、约85%、约90%、约91%、约92%、约93%、约94%、约95%、约96%、约97%、约98%、约99%或约100%的序列同源性。此外,载体可以部分或完全针对在靶生物或细胞中表达而进行密码子优化。

[0132] 该载体可以可操作地连接到适当的启动子,使得多核苷酸可以在宿主细胞中表达。当将本发明的载体应用于植物细胞时,本发明的启动子可以包括用于将基因引入植株体的启动子。启动子的例子可以包括但不限于SP6启动子、T7启动子、T3启动子、PM启动子、玉米泛素启动子(Ubi)、花椰菜花叶病毒(CaMV) 35S启动子、胭脂碱合酶(nos)启动子、玄参花叶病毒35S启动子、甘蔗杆状病毒启动子、鸭跖草黄斑病毒启动子、核酮糖-1,5-二磷酸羧化酶小亚基(ssRUBISCO)光诱导型启动子、水稻胞浆磷酸三糖异构酶(TPI)启动子、拟南芥腺嘌呤磷酸核糖基转移酶(APRT)启动子和章鱼碱合酶启动子。

[0133] 具体地,载体可以是质粒DNA、噬菌体DNA等。另外,载体可以为商业上开发的质粒(pUC18、pBAD、pIDTSAMRT-AMP等)、大肠杆菌来源的质粒(pYG601BR322、pBR325、pUC118、pUC119等)、枯草芽孢杆菌来源的质粒(pUB110、pTP5等)、酵母来源的质粒(YEp13、YEp24、YCp50等)、噬菌体DNA(Charon4A、Charon21A、EMBL3、EMBL4、λgt10、λgt11、λZAP等)、动物病毒载体(逆转录病毒、腺病毒、痘苗病毒等)、昆虫病毒载体(杆状病毒等)等。由于载体的蛋白质的表达量和修饰根据宿主细胞而不同,因此优选选择并使用最适合目的的宿主细胞。

[0134] 在本发明中,载体可以包括选择性标记。选择性标记是具有可通过常规化学方法选择的特征的核酸序列,包括所有能够区分转化细胞与非转化细胞的基因。标记的例子可以包括但不限于除草剂抗性基因,例如草甘膦、草铵膦或草丁膦,以及抗生素抗性基因,例如卡那霉素、G418、博来霉素、潮霉素和氯霉素。在本发明的一个实施方式中,选择性标记可以是潮霉素和/或草丁膦。

[0135] 在本发明的另一个方面,提供了一种生产含有高浓度7-脱氢胆固醇的转化番茄的方法,包括利用农杆菌将包含编码sgRNA的基因和编码CRISPR相关蛋白的基因的载体引入番茄中。

[0136] 可以通过本领域已知的植物转化方法来构建转化植物。本领域的普通技术人员可以根据所选宿主植物的特性,选择并实施适合于特定植物的已知转化方法。具体而言,利用农杆菌的转化方法可以用作植物转化方法。

[0137] “利用农杆菌的转化方法”是利用农杆菌将外部基因转移到植物细胞中的方法,农杆菌是一种革兰氏阴性土壤细菌,会导致植物的根和茎产生肿瘤。这是利用农杆菌(例如根癌农杆菌(*Agrobacterium tumefaciens*)和发根农杆菌(*Agrobacterium rhizogenes*))中发现的肿瘤诱导质粒(Ti质粒)的T-DNA(转移DNA)插入植物基因组的现象的方法。

[0138] 在本发明的一个实施方式中,当在DWF5-1基因的外显子6域中包含核酸序列的修饰时,纯合番茄(T1)可以产生种子(图18)。此外,DWF5-1基因的外显子6域中核酸序列发生修饰的纯合T2代番茄能够正常生长并产生种子(图26),并且7-脱氢胆固醇在果实和根中以高浓度积累(表14、图21和图22)。然而,当转化番茄的DWF5-1基因的外显子6域和CPD基因外显子8域中的核酸序列包括了修饰时,番茄无法长成完整的植株体。另外,当转化番茄为DWF5-1基因的外显子6域、CPD基因的外显子8域、以及SMT1基因的外显子7和外显子9域的核酸序列具有修饰的纯合子时,不会产生果实。

具体实施方式

[0139] 通过以下实施例对本发明进行更详细的说明。但下述实施例仅用于说明本发明，本发明的范围不仅限于这些实施例。

[0140] 实施例1. 用于基因编辑DWF5、CPD和SMT1的Cas9-sgRNA载体的构建

[0141] 通过Cas-Designer网站 (<http://www.rgenome.net/cas-designer/>) 设计和选择用于基因编辑的DWF5-1 (以下简称“DWF5”, 图1)、CPD (图2) 和SMT1 (图3) 的sgRNA。选定靶基因的名称或序列、动物或植物的物种, 进而选定所要使用的Cas9, 根据基因中所需外显子位点、sgRNA的GC含量、错配处的脱靶等结果设计sgRNA。Cas9的PAM序列为NGG-3', 选择番茄 (*S. lycopersicum*) 作为植物 (图1至3)。此外, 还筛选了sgRNA以确认基因编辑效率 (表2和图1至3)。

[0142] 具体地, 使用RGEN (RNA导向的核酸内切酶) 分析来进行sgRNA的靶向适当性。使用PCR纯化试剂盒从上述扩增的PCR产物 (PCR产物) 中纯化DNA模板, 然后测量DNA模板的量, 并按照表1所示的组合进行RGEN。37°C孵育1小时后, 各加入3μL 6xDNA-Purple染料, 冷冻, 然后上样至2.5%凝胶。通过分析DNA模板中未切割的条带的程度来分析基因编辑效率。

[0143] [表1]

[0144]	RGEN 终浓度条件	
	10X 3.1 NEB 缓冲液	2 μL
	sgRNA (120 nM)	5 μL
	spyCas9 (100 nM)	5 μL
	DNA 模板 (PCR 产物)	200 ng
	DW 补足至 20 μL, 37°C 1h	

[0145] 结果显示, 在构建的靶向DWF5基因的sgRNA中, 确认了除D5-4和D5-6以外的所有sgRNA均能够靶向DWF5基因 (图4)。除C1之外的所有CPD sgRNA都靶向CPD基因 (图5), 并且所有构建的SMT1 sgRNA都靶向CPD基因 (图6)。

[0146] [表2]

[0147]	靶基因 (Cas9)	sgRNA (引物名称)	DNA 切割试验	sgRNA (子名称)	sgRNA 序列 (PAM 位点 3')	基因中位点 (外显子)
--------	------------	-----------------	-------------	----------------	-------------------------	----------------

[0148]

SolycDWF5 (SEQ ID NO: 1)	SolycDWF5-1	O	D5-1	TGGAGCTTTATCCTCGTAT CGGG (SEQ ID NO: 20)	6 (SEQ ID NO: 9)
SolycDWF5 (SEQ ID NO: 1)	SolycDWF5-2	O	D5-2	GCAGATTGGTATGATGTC TTGG (SEQ ID NO: 21)	6 (SEQ ID NO: 9)
SolycDWF5 (SEQ ID NO: 1)	SolycDWF5-3	O	D5-3	TGTAAATACGATATTGATG TTGG (SEQ ID NO: 22)	7 (SEQ ID NO: 10)
SolycDWF5 (SEQ ID NO: 1)	SolycDWF5-4	△	D5-4	GACCAGCCACTAGTATGTA AAGG (SEQ ID NO: 23)	9 (SEQ ID NO: 11)
SolycDWF5 (SEQ ID NO: 1)	SolycDWF5-5	O	D5-5	TGCATTTTCCATTGTTCTC CGG (SEQ ID NO: 24)	9 (SEQ ID NO: 11)
SolycDWF5 (SEQ ID NO: 1)	SolycDWF5-6	X	D5-6	CCTCTGGTGAGAAAAAGA CCAGC (SEQ ID NO: 25)	10 (SEQ ID NO: 12)
SolycDWF5 (SEQ ID NO: 1)	SolycDWF5-7	O	D5-7	AGAAAGATGACGTAGAAG TAGGG (SEQ ID NO: 26)	11 (SEQ ID NO: 13)
SolycCPD (SEQ ID NO: 5)	SolycCPD gRNA1	△	C1	AACGAGTCACCAAGTGGT GAAGG (SEQ ID NO: 27)	7 (SEQ ID NO: 14)
SolycCPD (SEQ ID NO: 5)	SolycCPD gRNA2	O	C2	CGCTTCAGCATCCTTGTCG TTGG (SEQ ID NO: 28)	8 (SEQ ID NO: 15)
SolycCPD (SEQ ID NO: 5)	SolycCPD gRNA3	O	C3	TGAGCAGAAAGAAGAAG AGAAGG (SEQ ID NO: 29)	8 (SEQ ID NO: 15)
SolycCPD (SEQ ID NO: 5)	SolycCPD gRNA4	O	C4	TGAGATTGTAGAATGCAG AGTGG (SEQ ID NO: 30)	8 (SEQ ID NO: 15)
SolycCPD (SEQ ID NO: 5)	SolycCPD gRNA5	O	C5	GTAGAATGCAGAGTGGAG GGAGG (SEQ ID NO: 31)	8 (SEQ ID NO: 15)
SolycSMT1 (SEQ ID NO: 7)	SolycSMT1 gRNA1	O	S1	GTTGAACCGCAAAGTGGG ATTGG (SEQ ID NO: 32)	7 (SEQ ID NO: 16)
SolycSMT1 (SEQ ID NO: 7)	SolycSMT1 gRNA2	O	S2	GATGCTATAGAGAGATCTA TAGG (SEQ ID NO: 33)	9 (SEQ ID NO: 17)
SolycSMT1 (SEQ ID NO: 7)	SolycSMT1 gRNA4	O	S4	GTTTCGGCACTTGAATATG TGGG (SEQ ID NO: 34)	11 (SEQ ID NO: 18)
SolycSMT1 (SEQ ID NO: 7)	SolycSMT1 gRNA5	O	S5	GCCGCAGAAGGTCTTGTT GGTGG (SEQ ID NO: 35)	12 (SEQ ID NO: 19)

[0149] 根据sgRNA筛选的结果,最终构建了包含构建为靶向DWF5、CPD或SMT1的sgRNA的Cas9-sgRNA载体。pNGPJ0014载体(SEQ ID NO:39)是基于具有卡那霉素和潮霉素抗性基因

的pCAMBIA1300质粒构建的。

[0150] pNGPJ0014载体具有用于卡那霉素和潮霉素抗生素选择的抗生素盒,以及用Cas9合成的多顺反子tRNA-gRNA盒。Cas9构建为由拟南芥泛素10 (Ubi 10) 启动子表达,tRNA-gRNA构建为由拟南芥泛素6启动子表达。

[0151] 为了将Cas9蛋白递送到细胞核内,在Cas9开放阅读框的N端添加了SV40 (PKKKRKV) 核定位信号序列,在C端添加了Bipartite (KEPAATKKAGQAKKKK) 核定位信号序列。此时,Cas9-sgRNA载体被构建为靶向一种至三种类型的基因。此外,它被构建为每个靶标负载两种类型的sgRNA。包含靶向DWF5基因的D5-1 sgRNA (SEQ ID NO:20) 和D5-2 sgRNA (SEQ ID NO:21) 的核苷酸序列的Cas-sgRNA载体被命名为“D100” (SEQ ID NO:36) (图7)。

[0152] 靶向CPD和DWF5基因的Cas9-sgRNA载体被命名为“D120” (SEQ ID NO:37)。此时,CPD sgRNA包含C2 (SEQ ID NO:28) 和C3 (SEQ ID NO:29) 的核苷酸序列,DWF5 sgRNA包含D5-1 (SEQ ID NO:20) 和D5-2 (SEQ ID NO:21) 的核苷酸序列 (图8)。

[0153] 靶向SMT1、CPD和DWF5基因的Cas9-sgRNA载体被命名为“D121” (SEQ ID NO:38)。此时,SMT1 sgRNA包含S1 (SEQ ID NO:32) 和S2 (SEQ ID NO:33) 的核苷酸序列,CPD sgRNA包含C2 (SEQ ID NO:28) 和C3 (SEQ ID NO:29) 的核苷酸序列,DWF5 sgRNA包含D5-1 (SEQ ID NO:20) 和D5-2 (SEQ ID NO:21) 的核苷酸序列 (图9)。

[0154] 实施例2.通过DWF5、CPD和SMT1基因编辑转化番茄

[0155] 番茄的转化按照图10的示意图进行。各步骤所用的培养基和激素的制备和使用方式与表3至表8中所述相同。

[0156] [表3]

[0157]	TMS (V) 培养基	MS 盐	4.3 g/L	制备 100 倍原液后添加四种维生素。 使用 100 倍 (100X) 维生素原液, 浓度各为 10mL/L。
		肌醇	100 mg/L	
		烟酸	0.5 mg/L	
		盐酸吡哆醇	0.5 mg/L	
		盐酸硫胺	0.1 mg/L	
		蔗糖	30 g/L	
		植物琼脂	0.6-0.7%/L	

[0158]	制备培养基, 然后使用 1N NaOH 将 pH 值调节至 5.7, 然后在 121°C 的高压灭菌器中灭菌 15 分钟, 并将约 30mL 分装到 90x15mm 的培养皿中。培养基灭菌后, 当培养基温度为 55°C 至 60°C 时添加维生素原液。
--------	--

[0159] [表4]

[0160]

TPC 培养基 （番茄预培养培养基）	MS 盐	4.3 g/lL	制备 100 倍原液后添加四种维生素。 使用 100 倍（100X）维生素原液，浓 度各为 10mL/L。
	肌醇	100 mg/L	
	烟酸	0.5 mg/L	
	盐酸吡哆醇	0.5 mg/L	
	盐酸硫胺	0.1 mg/L	
	蔗糖	30 g/L	
	结冷胶	0.25%/L	
	BA	1 mg/L	
	NAA	0.1 mg/L	
制备培养基，然后使用 1N NaOH 将 pH 值调节至 5.7，然后在 121℃的高压灭菌器中灭菌 15 分钟，并将约 30mL 分装到 90x15mm 的培养皿中。培养基灭菌后，当培养基温度为 55℃至 60℃时添加维生素原液。			

[0161] [表5]

[0162]

TCC 培养基 (番茄共培养培养基)	MS 盐	4.3 g/L	制备 100 倍原液后添加四种维生素。 使用 100 倍 (100X) 维生素原液, 浓度各为 10mL/L。
	肌醇	100 mg/L	
	烟酸	0.5 mg/L	
	盐酸吡哆醇	0.5 mg/L	
	盐酸硫胺	0.1 mg/L	
	蔗糖	30 g/L	
	结冷胶	0.25%/L	
	BA	1 mg/L	
	NAA	0.1 mg/L	
	乙酰丁香酮 (AS)	100 mM/L	
制备培养基, 然后使用 1NNaOH 将 pH 值调节至 5.2, 然后在 121℃的高压灭菌器中灭菌 15 分钟, 并将约 30mL 分装到 90x15mm 的培养皿中。培养基灭菌后, 当培养基温度为 55℃至 60℃时添加维生素原液、BA、NA 和 AS。			

[0163] [表6]

[0164]

细菌稀释培养基	MS 盐	4.3 g/L	使用 1N NaOH 和 1N HCl 将 pH 调节至 5.7, 然后将其在 121°C 的高压灭菌器中灭菌 15 分钟, 并储存在冰箱中。
	蔗糖	30 g/L	

[0165] [表7]

[0166]

TSI 培养基 (番茄芽诱导培养基)	MS 盐	4.3 g/lL	制备 100 倍原液后添加四种维生素。 使用 100 倍（100X）维生素原液，浓度各为 10mL/L。
	肌醇	100 mg/L	
	烟酸	0.5 mg/L	
	盐酸吡哆醇	0.5 mg/L	
	盐酸硫胺	0.1 mg/L	
	蔗糖	30 g/L	
	结冷胶	0.25%	
	IAA	0.2 mg/L	
	玉米素	2 mg/L	
	PPT	5 mg/L	
	羧苄青霉素	5 mg/L	
制备培养基，然后使用 1N NaOH 将 pH 值调节至 5.7，然后在 121℃的高压灭菌器中灭菌 15 分钟，并将约 30mL 分装到 90x15mm 的培养皿中。培养基灭菌后，在培养基温度为 55℃至 60℃时添加维生素原液、BA、NA 和 AS。			

[0167] [表8]

生根培养基	1/2 MS 盐 (含维生素)	2.2 g/L
	蔗糖	30 g/L
	植物琼脂	0.6-0.7%/L
	IAA	0.2 mg/L
	羧苄青霉素	300 mg/L
制备培养基, 然后使用 1N NaOH 将 pH 值调节至 5.7, 然后在 121°C 的高压灭菌器中灭菌 15 分钟, 并将其分装到洋红色盒中。培养基灭菌后, 当培养基温度为 55 至 60°C 时添加 IAA。		

[0169] [表9]

[0170] 激素和维生素原液的制备

	浓度	产品名称或制备方法
BA	1 mg/mL	购买并使用了 Caisson (BSL01-100mL) 1mg/mL 溶液。
NAA	1 mg/mL	购买并使用了 Caisson (NSL01-100mL) 1mg/mL 溶液。
玉米素	2 mg/mL	购买并使用了 phytotechlab (Z860) 1mg/mL 溶液。
IAA (吲哚乙酸)	1 mg/mL	将 50mL Duchefa, I0901 试剂溶解于 50mL D.W 中, 过滤混合物然后分装, 放入冰箱保存。
乙酰丁香酮 (AS)	100 mM/mL	将 0.5887g 乙酰丁香酮溶解于 20mL DMSO 中, 加入 10mL D.W 以使总量为 30mL。将配制的溶液以 1mL 为单位分装并保存在冰箱中 (不冷冻)。
[0171] 潮霉素	10 mg/mL	购买并使用 Biomedic (CAT.No.BM463H) 10mg/mL 溶液。
羧苄青霉素	300 mg/mL	将 1.5g Duchefa (C0109) 试剂溶解于 50mL D.W 中, 过滤混合物然后分装, 放入冰箱保存。
改良维生素溶液	100x	将其制备成 100 倍原液并在制备培养基时使用。 将肌醇 (MB cell, MB-I4715) 10g/L、烟酸 (Showa, 1414-0130) 50mg/L、盐酸吡哆醇 (Sigma, P-8666) 50mg/L、盐酸硫胺 (Duchefa, T0614) 10mg/L 溶解于 1L D.W 中, 在 121°C 高压灭菌器中灭菌 15 分钟, 然后放入冰箱中保存。

[0172] 具体来说, 番茄种子用70%乙醇消毒5秒, 用2%漂白剂消毒20分钟。将消毒好的种子用无菌水清洗4至5次, 然后放在无菌滤纸上, 除去水分。将消毒后的番茄种子转移至播种培养基TMS (表3) 中, 25°C暗培养条件下培养3天, 转入光照培养条件下培养。播种后, 当子叶长出1cm至1.5cm (9至12天) 时, 将番茄下胚轴顶端剪至0.7cm至1.0cm, 约18至20个 (各自有4、5、5和4个, 或5、5、5和5个, 共18至20个) 种植于TPC培养基平板 (表4) 上, 在光照培养条件 (16L/8D) 下培养1天。

[0173] 将农杆菌 (GV3101菌株, DSMZ, CAT.No.DSM12364) 在含有100mM AS (乙酰丁香酮) 和50mg/L利福平、50mg/L卡那霉素 (Km) 的50mL细菌培养基 (YEP) 中, 在28°C下暗培养条件下培养1至2天, 以获得0.8至1.2的OD600值。

[0174] 将培养的细菌在4°C、6,000rpm下离心15分钟, 以除去上清液, 并将沉淀重悬于等量的番茄共培养 (TCC) 培养基 (表5) 中并稀释至OD值为0.5。

[0175] 将按上述方法制备的40mL细菌培养液 (OD600值为0.5) 和光照培养条件下预处理的植株体置于50mL试管中 (基于播种200粒种子), 在25°C室温下培养20至30分钟。将共培养的外植体转移到无菌滤纸上, 尽量晾干以除去水分, 然后将约18至20个 (各自有4、5、5和4个, 或5、5、5和5个, 共18至20个) 接种到TCC培养基平板上 (表5), 在光照培养条件 (16L/8D) 下于25°C培养2天。共培养2天后, 将外植体转移至芽选择培养基 (TSI, 表7) 中, 并在光照条

件(16L/8D)下于25℃的植物培养箱中培养。此时,当农杆菌正在生长时,将300mg/L羧苄青霉素添加到共培养液体培养基中,洗涤,然后在滤纸上晾干,然后重新种植在培养基上。

[0176] 种植后3至4周,愈伤组织从切割部分分化出来。当分化的愈伤组织诱导出芽并且长至大于1cm时,将其转移到生根培养基(表9)中以诱导根。此时,尽可能去除愈伤组织,并通过移植芽部分来诱导生根。待根完全生根后,用自来水冲洗根,除去根上残留的农杆菌,移栽到土壤中,在保持培养室环境的同时,适应外界环境(湿度变化和非无菌条件)进行培养。

[0177] 图11显示了愈伤组织中随着植株体分化阶段的变化。

[0178] 结果如图12至图14所示,用D100(图12)或D121(图14)转化的番茄芽正常生长为植株体,但用D120转化的番茄外植体生长不正常(图13)。

[0179] 实施例3.DWF5、CPD和SMT1基因编辑生成的番茄植株体的测序

[0180] 为了确认实施例2中获得的番茄植株体的转化,对用D100或D121转化的番茄对象进行了T7E1测定(图15)和Sanger测序(图16和表12)。

[0181] 具体地,在T7E1测定中,首先通过PCR扩增靶区域,然后利用PCR纯化试剂盒纯化扩增的PCR产物。使用纯化的PCR产物作为DNA模板,按照表11的条件和表10所示的组合进行异源双链形成。然后添加1μL T7内切酶I(T7E1),37℃下反应30分钟。最后用1μL蛋白酶K处理,37℃下反应5分钟,使T7E1酶活性失活。通过PCR产物中的片段浓度来评估基因编辑效率。

[0182] [表10]

[0183]	PCR产物(250ng)	5μL
	10xNEB缓冲液2	2μL
	无核酸酶水	补足至14μL

[0184] [表11]

	温度(℃)	升温速率
[0185]	95	4.0/sec
	95-85	-2.0/sec

	85-25	-0.1/sec
[0186]	25-4	4.0/sec

[0187] Sanger测序如下进行。使用Q5高保真DNA聚合酶(NewEngland Biolabs)在20μL反应体积中由基因组DNA提取物扩增sgRNA靶区域。随后,使用All in one Cloning Kit(Biofact,South Korea)将PCR产物克隆到TA载体中,并对每个样本的成功克隆的15到20个克隆进行单独测序。结果,鉴定了通过D100转化产生的14个具有DWF5基因缺失的番茄纯合子(图15和16)。此外,还鉴定出通过D121转化的一个具有DWF5、CPD和SMT1基因缺失的番茄纯合子(图15和表12)。以D121转化植株体为例,由于基因缺失造成植物激素失衡,从而出现无法形成果实的现象。

[0188] [表12]

[0189]

样本		D121-4	D121-5	D121-6	D121-7	D121-8	D121-9	D121-10
DWF5-1	突变	0/-158; 2 WT; 18	-157; 3 -5/-5; 1 -5/-3; 1 -5/-0; 4 WT; 11	-1; 4 WT; 5	+1/-4; 6 -13/1, 0/-1, -8/-1, -3/-1; 1	-158; 5 -7, -4/-7; 1 -6; 1 WT; 5	-659; 1 -8/0, -6/0, -4/0; 1 WT; 2	-4; 1 -6; 3 +1; 1 WT; 7
<i>DWF5-1</i>	接合性	异	异	异	同	异	异	异
CPD	突变	-136	-136 WT; 2	-2; 7 -5; 4	无	-2; 2 -3; 1 WT; 6	-2; 2 -3; 6 WT; 2	-2; 14 -3; 2 -4; 3
<i>CPD</i>	接合性	同	异	同	WT	异	异	同
SMT1	突变	-4; 1 -5/-4; 1 WT; 9	-8; 3 -4; 2 WT; 7	-4; 1 -2; 3 WT; 2	无	-4; 4 -2; 1 WT; 2	-1; 5 -2; 6 -7; 9 -91, WT; 1	-2; 5 -4; 5 WT; 1
<i>SMT1</i>	接合性	异	异	异	WT	异	异	异
敲除		CPD		CPD	DWF5			CPD
样本		D121-13	D121-14	D121-15	D121-16	D121-17	D121-18	D121-19
DWF5-1	突变	-8; 1 WT; 9	-8; 2 WT; 8	-4; 4 +1; 1 WT; 3	-5; 5 -5/-15; 1 -5/-1; 1	-7; 2 WT; 5	-4; 5 -6; 2 -8/-7; 1 +1/-6; 1	-1; 1 -4; 3 -8; 1 WT; 4
<i>DWF5-1</i>	接合性	异	异	异	同	异	同	异

[0190]

CPD	突变	-8; 1 -4; 2 WT; 7	无	无	-136; 8 -6; 2	无	无	无
<i>CPD</i>	接合性	异	WT	WT	同	WT	WT	WT
SMT1	突变	-45; 2 WT; 4	-1; 2 -2; 6 -3; 1	-4; 7 -2; 3	-4; 3 -8; 5	-1; 3 -2; 6 -4; 1	-1; 2 -2; 5 -4; 2	-1; 2 -2; 6 -4; 2
<i>SMT1</i>	接合性	异	同	同	同	同	同	同
敲除			SMT1	SMT1	DWF5 CPD SMT1	SMT1	DWF5 SMT1	SMT1

[0191] 实施例4. 获取DWF5基因编辑产生的番茄种子及番茄基因表型分析

[0192] 将从通过基因剪刀使得DWF5基因缺失的番茄(T0代)中的7个番茄获得的种子(DWF5纯合子)播种,进行下一代测序(NGS),以确定是否获得了下一代(T1代)DWF5基因缺失的番茄。

[0193] 结果证实,从7个T0对象中获得的23粒种子(T1)中,有12颗是纯合的(图17)。另外,从上述12粒种子中,选出基因编辑效率(indel效率)为100%的6粒种子,并确认种子(T2)的发生情况以及获得的种子量。结果,6个对象中有5个产生了种子(图18和表13)。此外,观察到DWF5基因缺失的番茄的果实、叶和茎的表型与野生型番茄相比没有显著差异。这里,野生型番茄(野生型,WT)是对照组。

[0194] [表13]

[0195]	标签	总读段	sgRNA 1			sgRNA 2			T1 接合子	T2 种子量 (g)	
	Label	Total reads	插入	缺失	Indel 频率 (%)	插入	缺失	Indel 频率 (%)	T1 Zygote	T2 the amount of seeds(g)	
	D100-1-4	54192	53220	51	100%	0	52983	100%	单等位 基因	0.37	1.06
	D100-7-1	89757	89183	54	100%	0	88705	100%	单等位 基因	0.35	1.06
	D100-12-8	7532	7377	13	100%	0	7267	100%	单等位 基因	0.34	1.06
	D100-3-11	54964	0	54465	100%	0	53656	100%	单等位 基因	0.13	0.45
[0196]	D100-3-14	44027	0	42936	100%	0	42304	100%	单等位 基因	0.32	0.45
	D100-4-8	83297	0	82686	100%	0	81960	100%	单等位 基因	-	0.45

[0197] 实施例5.T2代的表型分析

[0198] 实施例5.1.维生素D3前体的分析

[0199] 播种实施例4得到的种子(T2),用LC/MS测定T2代番茄中所含的维生素D3原的含量。此时,以野生型番茄(WT)、#3-14株系(D100-3-14)的T2代(4个对象)和#7-1株系(D100-7-1)的T2代(11个对象)作为样本。使用动物维生素D3原(维生素D3原,sigma,Cat30800, Lot.BCBS6021)的前体7-脱氢胆固醇作为标准试剂。另外,作为对照组,使用野生型(WT)番茄作为样本。

[0200] 具体地,称取一定量的各样本,添加30mL乙醇和1mL 10%乙醇焦碳酸(ethanolic pyrogallo)。然后添加3mL 90%氢氧化钾(KOH),在80℃水浴中反应30分钟,每10分钟涡旋一次。反应完成后,将混合物冷却至室温,然后添加30mL蒸馏水和30mL正己烷,震荡混合1小时。然后,将上清液转移至分液漏斗中,向水层中加入30mL正己烷,涡旋混合。将上清液再次转移至分液漏斗中,向己烷层中加入100mL蒸馏水,以用水进行洗涤。水洗涤过程重复进行

两次。随后,用无水硫酸钠对己烷层进行脱水,然后浓缩,溶解在5mL甲醇(MeOH)中,并通过聚四氟乙烯(PTFE)过滤器过滤,以制备测试溶液(图19)。

[0201] 在图20的条件下,通过LC/MS测量按上述方法制备的测试溶液中所含的7-脱氢胆固醇。

[0202] 结果,在WT番茄的果实部分中未检测到7-脱氢胆固醇,而在T2代的#3-14株系中的4个对象中和#7-1株系中的11个对象(共计15个对象)中检测到了7-脱氢胆固醇(表14、图21和图22)。

[0203] [表14]

[0204]

No	样本名称	样本状态	终体积 (L)	标准形式纯度 (%)	结果 (mg/100 g)
1	WT 番茄	成熟果实	0.005	95.3	未检出
2	WT 番茄	成熟果实	0.005	95.3	未检出
3	#3-14-2 番茄	成熟果实	0.005	95.3	1.03
4	#3-14-13 番茄	成熟果实	0.005	95.3	0.60
5	#3-14-20 番茄	成熟果实	0.005	95.3	0.19
6	#3-14-22 番茄	成熟果实	0.005	95.3	0.26
7	#7-1-2 番茄	成熟果实	0.005	95.3	0.23
8	#7-1-7 番茄	成熟果实	0.005	95.3	0.12
9	#7-1-12 番茄	包括未成熟果实	0.005	95.3	0.27

[0205]

10	#7-1-14 番茄	叶茎	0.005	95.3	0.02
11	#7-1-17 番茄	成熟果实	0.005	95.3	0.14
12	#7-1-18 番茄	叶茎	0.005	95.3	0.03
13	#7-1-20 番茄	未成熟果实	0.005	95.3	0.36
14	#7-1-21 番茄	叶茎	0.005	95.3	未检出
15	#7-1-22 番茄	叶茎	0.005	95.3	未检出
16	#7-1-24 番茄	未成熟果实	0.005	95.3	0.71
17	#7-1-27 番茄	叶茎	0.005	95.3	0.03

[0206] 实施例5.2. 番茄根中维生素D3前体的分析

[0207] 将实施例4得到的种子(T2)用70%乙醇消毒5秒,用2%漂白剂消毒20分钟。将消毒好的种子用无菌水清洗4至5次,然后放在无菌滤纸上,除去水分。将种子播种在MS培养基(表15)中,在25℃、16小时光照/8小时黑暗循环条件下培养。待2至3周后,当根形成后,从培养皿中取出根,然后收集根(毛状根),用GC-MS法分析7-DHC含量。

[0208] [表15]

[0209]

培养基组成	1L
MS 粉末	4.4g
蔗糖	10g
pH 5.8	
植物琼脂	7g

[0210] 通过将10mg根与3mL 0.1%抗坏血酸-乙醇和0.05mL 5 α -胆甾烷混合以获得根提

取物,进行GC-MS代谢分析。将提取物与80%氢氧化钾混合进行皂化,然后与己烷混合以分离亲脂性物质。将分离出的亲脂性物质用吡啶和N-甲基-N-三甲基硅基三氟乙酰胺进行衍生化,然后进行GC-MS(气相色谱-四极杆质谱)分析。对于定量和定性分析,使用由标准材料计算的校准曲线。

[0211] 结果,在WT番茄的根中未检测到7-脱氢胆固醇,而在T3代的#3株系(3-14-29)和#7株系(7-1-15)中分别检测到了7-脱氢胆固醇,浓度为1.3 μ g/g和1.0 μ g/g(图28)。

[0212] 实施例5.3.T2代中无抗生素对象的选择

[0213] 为了在实施例4中获得的T2代的#3-14株系(D100-3-14)和#7-1株系(D100-7-1)中选出不具有抗生素抗性的对象,通过RT-PCR确认了对象的mRNA中抗生素抗性基因的表达。

[0214] 结果确认#3-14株系(D100-3-14)中的5个对象、#7-1株系(D100-7-1)中的11个对象是无抗生素抗性对象(图23)。

[0215] 在选择为无抗生素抗性对象的对象中,对#3-14-29和#7-1-15对象进行了DWF5基因编辑检查。

[0216] 结果如图24和图25所示,确认了该对象的核苷酸序列与野生型(WT)番茄不同(图24),并且由于对应于外显子6的序列缺失,因此氨基酸序列也与野生型(WT)番茄不同(图25)。

[0217] 实施例5.4.T2代的表型及种子生成的确认

[0218] 检查实施例4中获得的T2代的#3-14株系(D100-3-14)和#7-1株系(D100-7-1)的表型和种子产生。具体地,当仅缺失主要在果实中表达的DWF5基因(DWF5-1,SEQ ID NO:63)时,与野生型番茄相比,果实、叶和茎的表型没有显著差异(图27)。因此,在本发明中,通过仅缺失DWF5-1(SEQ ID NO:63)基因,可以产生能够弥补种子生成问题同时诱导维生素D3原富集的番茄。

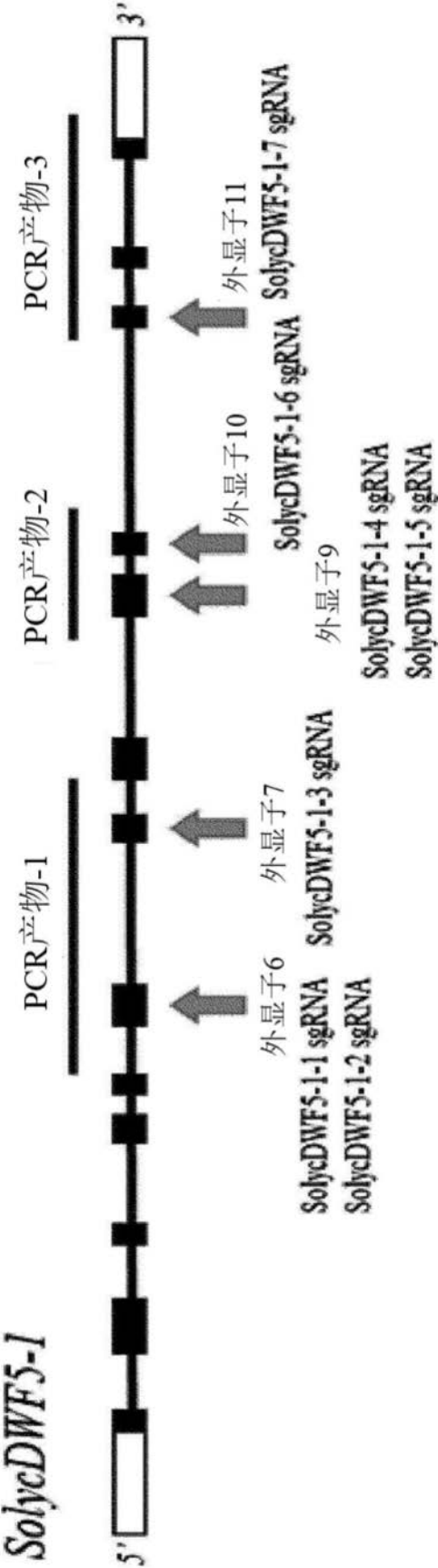


图1

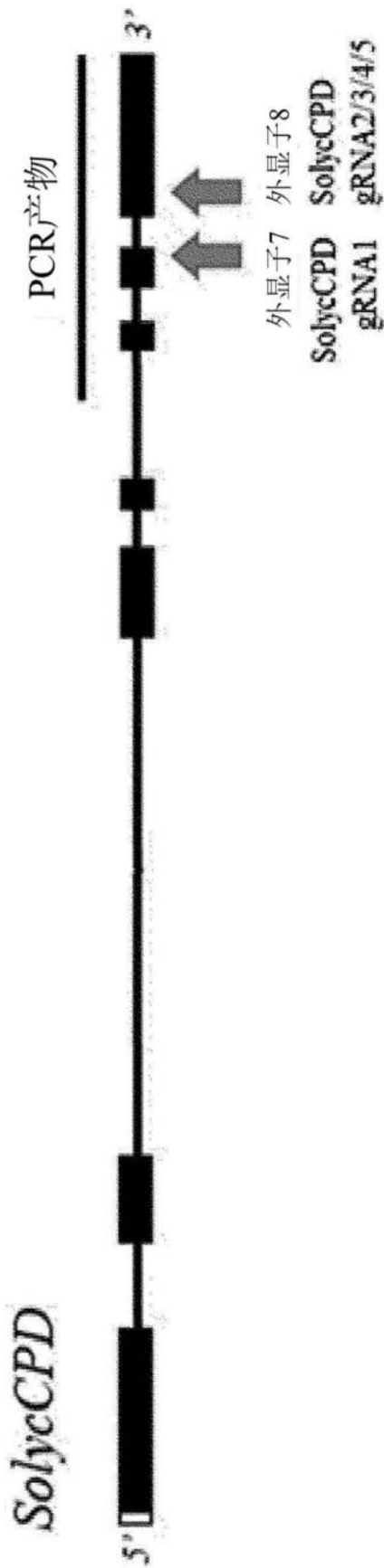


图2

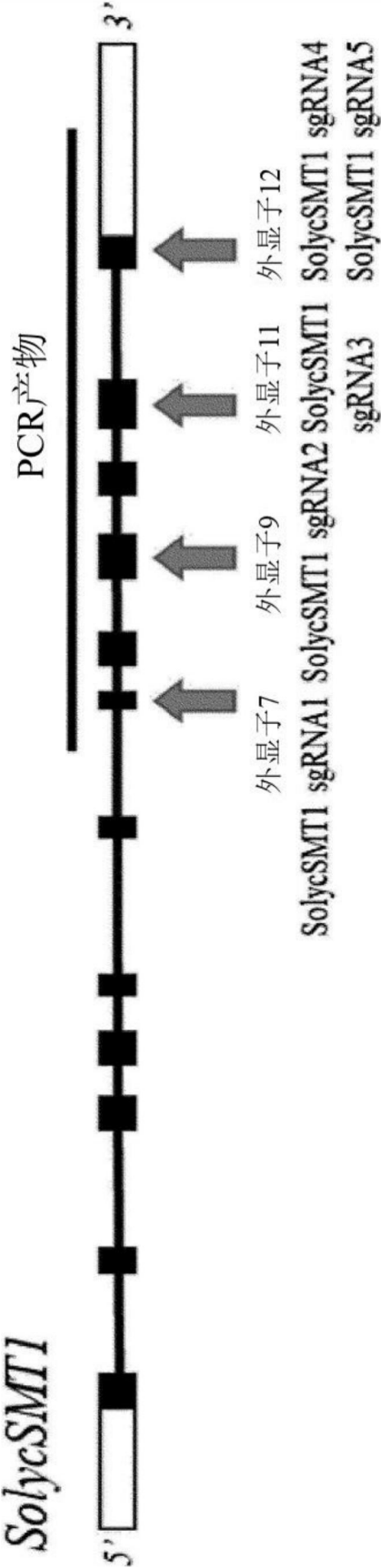


图3

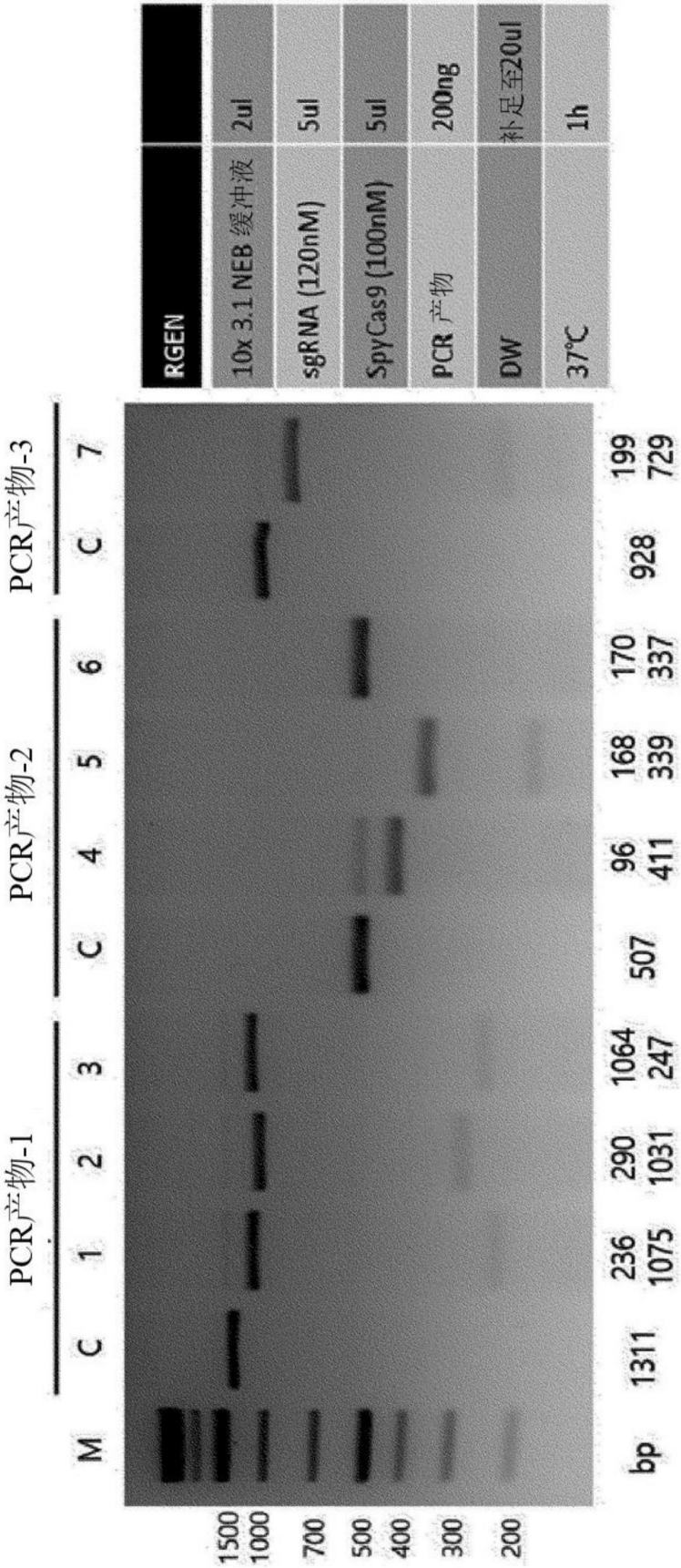


图4

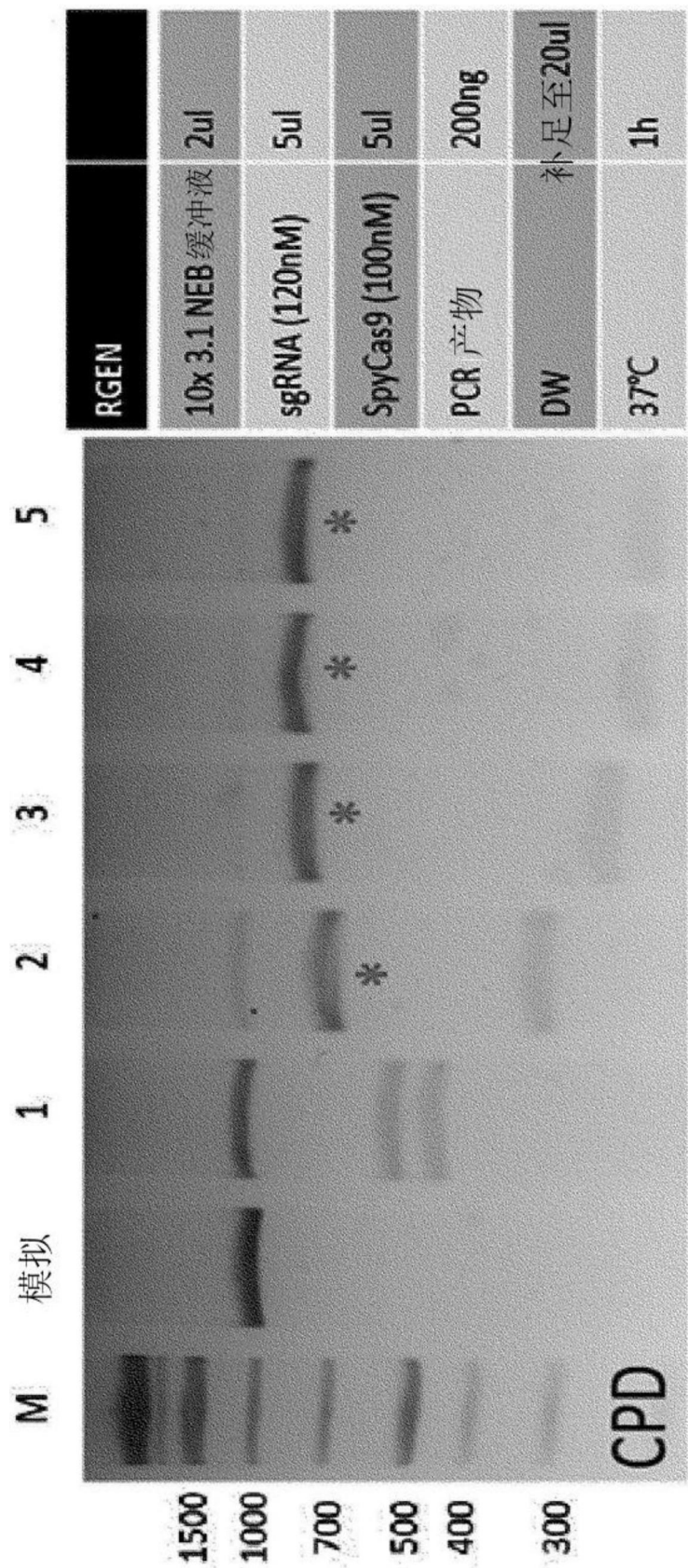


图5

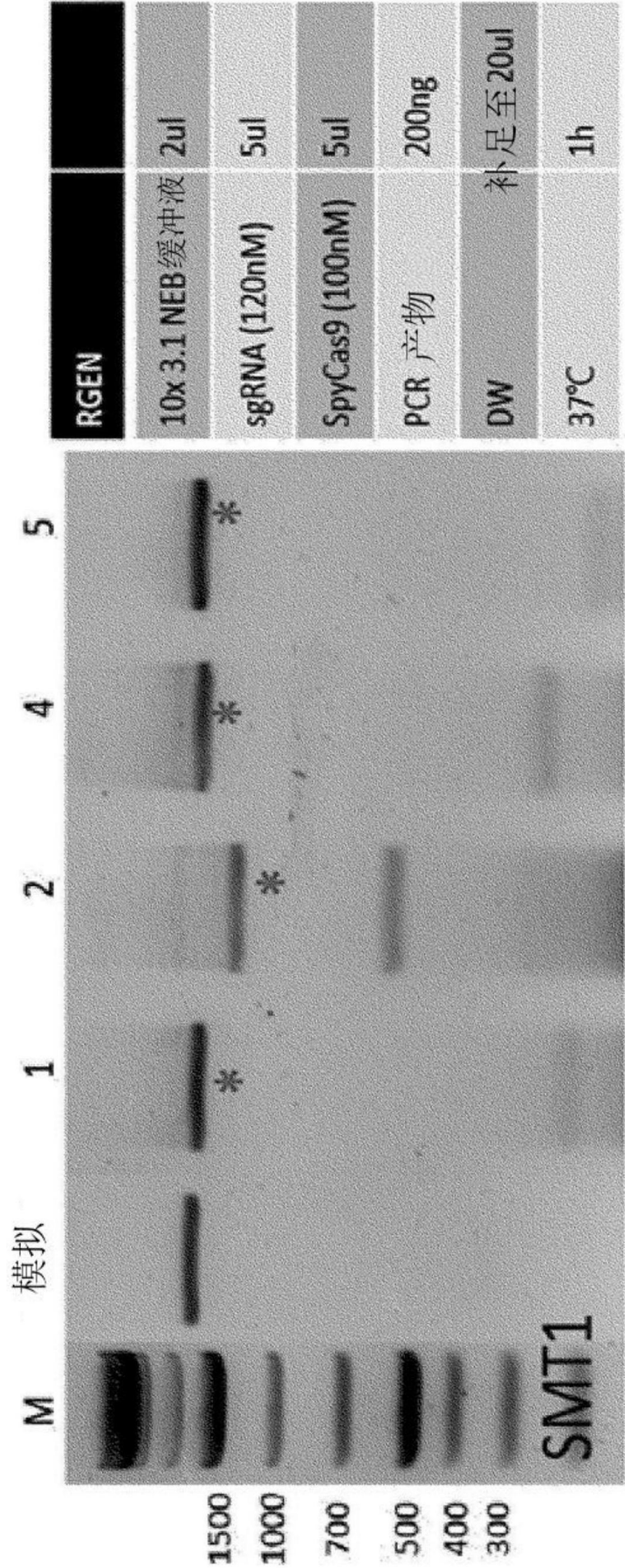


图6

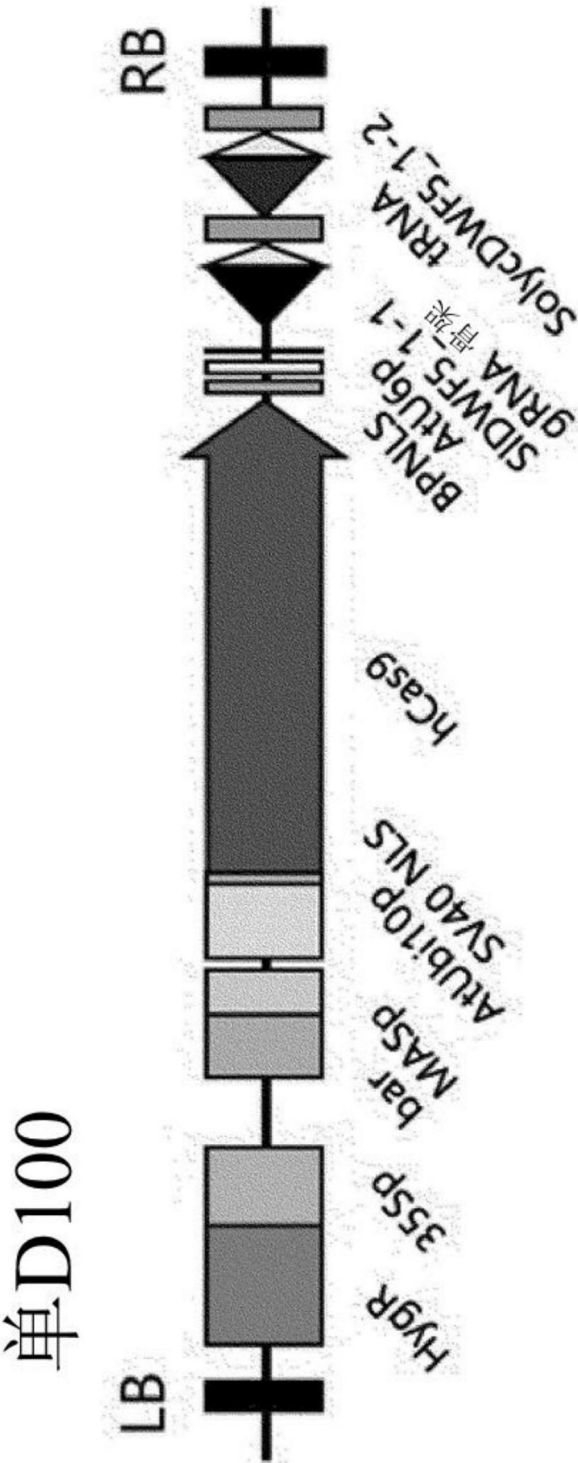


图7

ED121

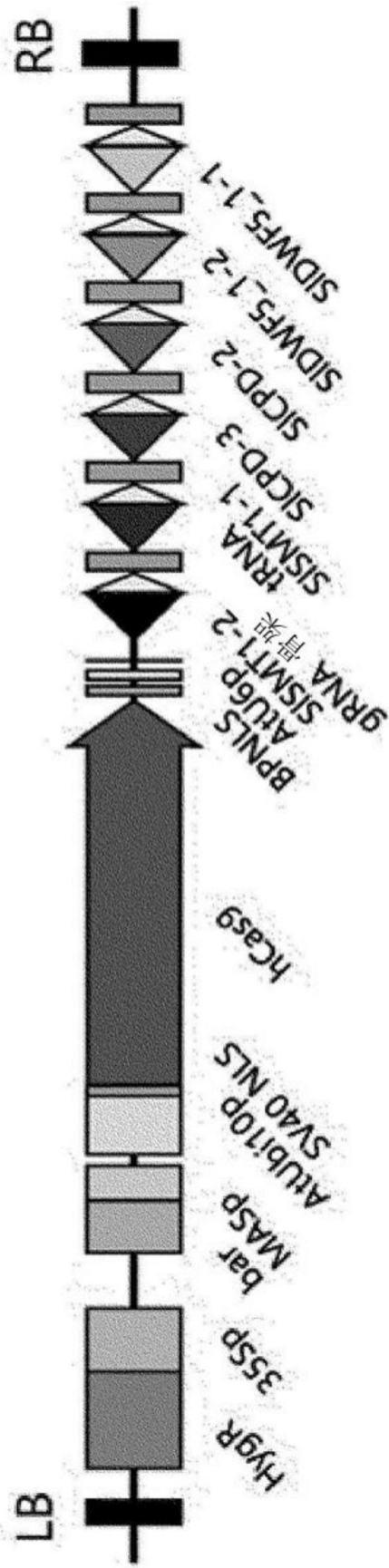


图9

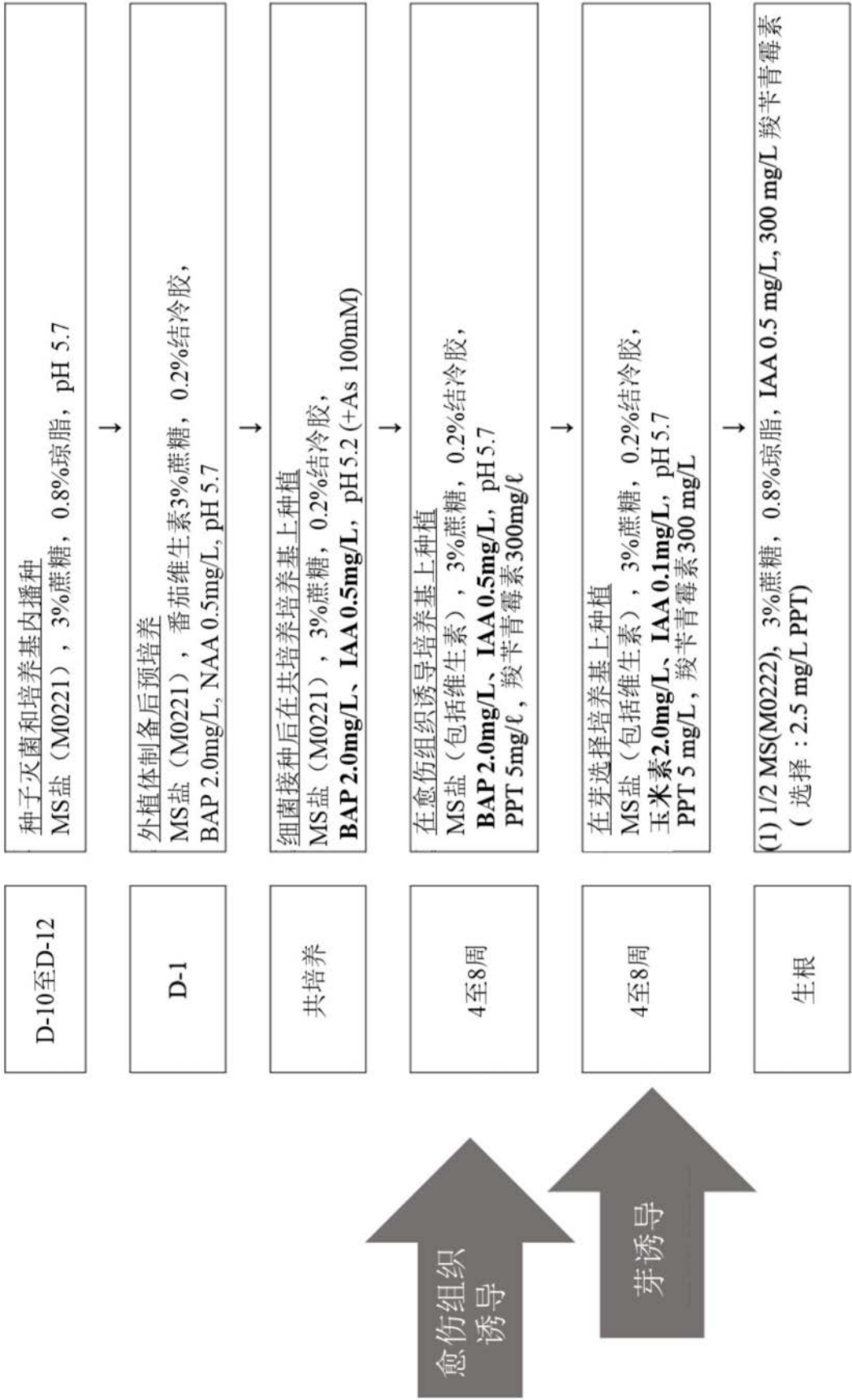


图10

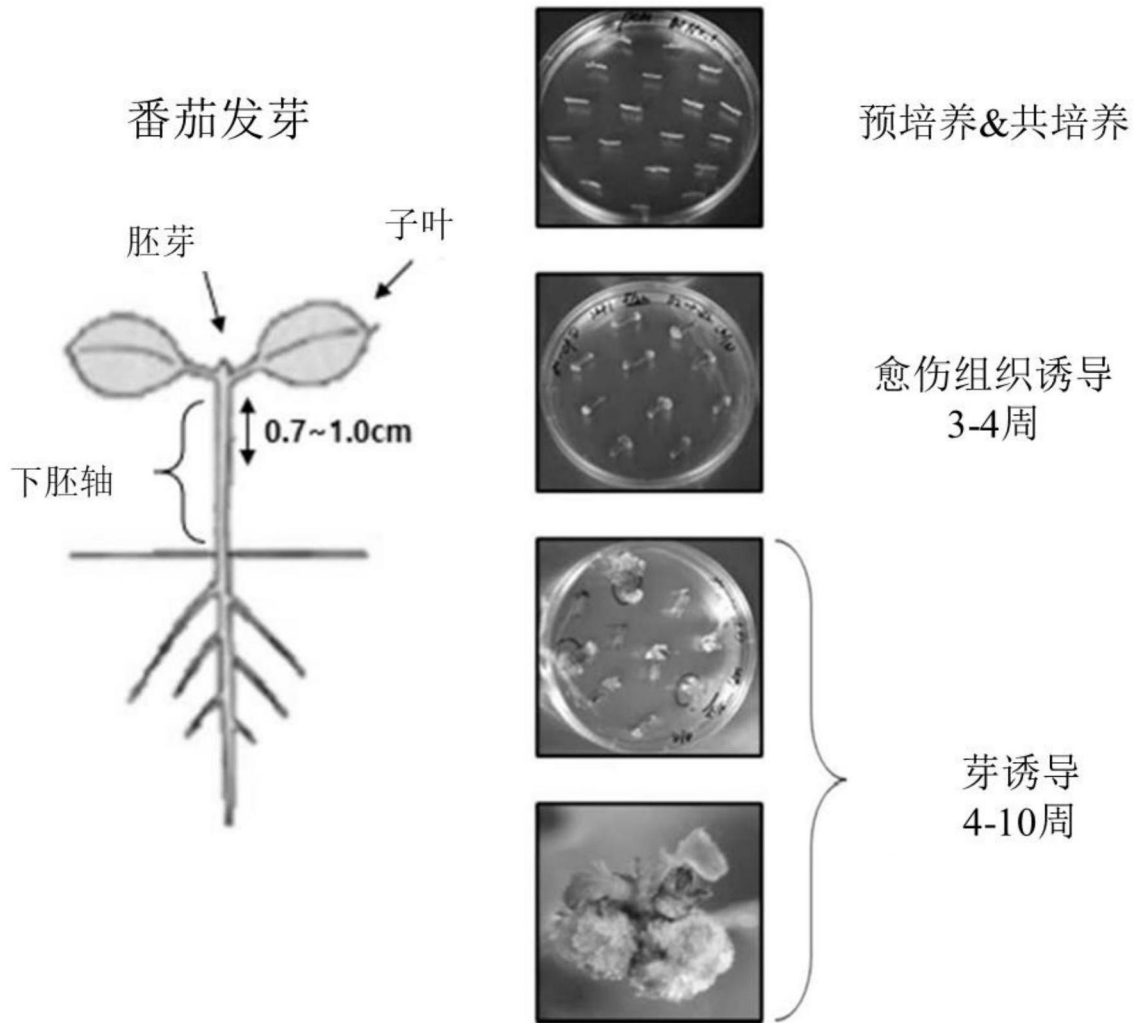


图11

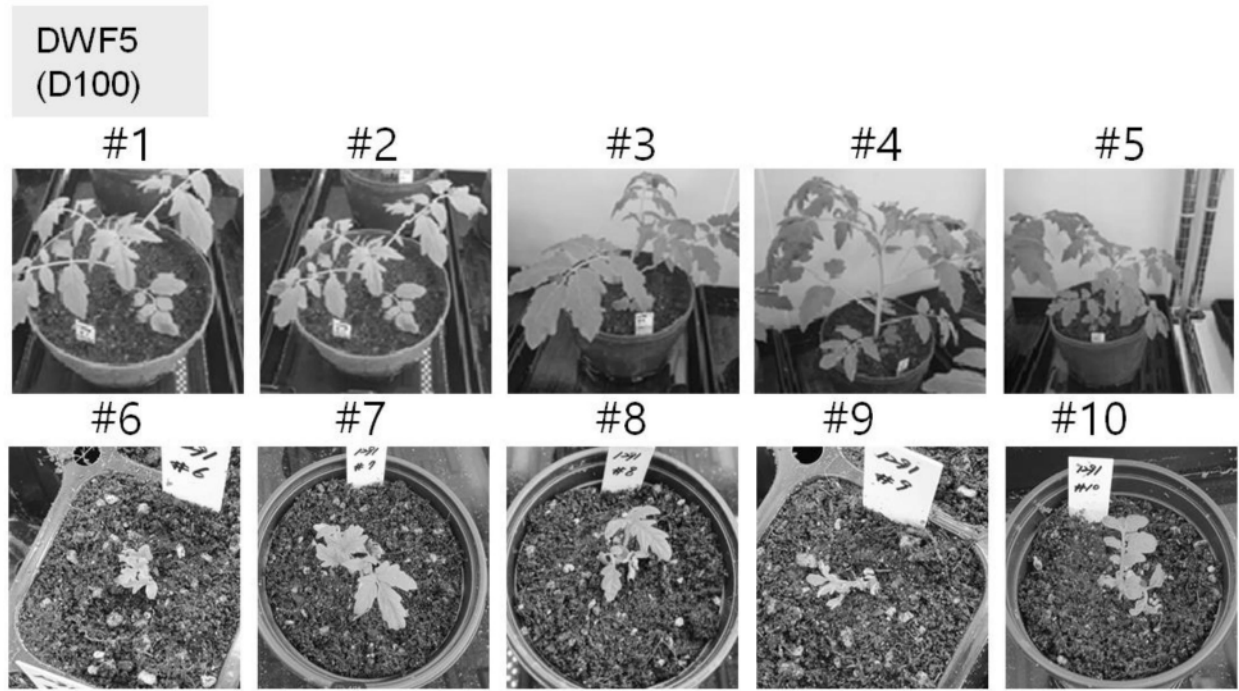
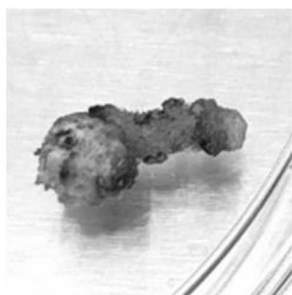


图12

DWF5
CPD
(D120)

#1



#2



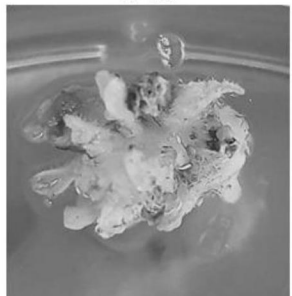
#3



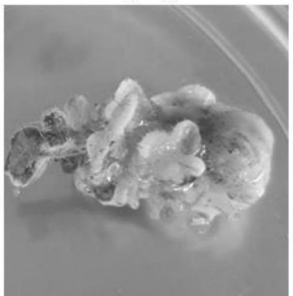
#4



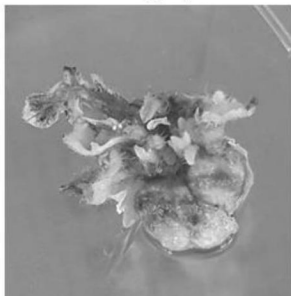
#5



#6



#7



#8

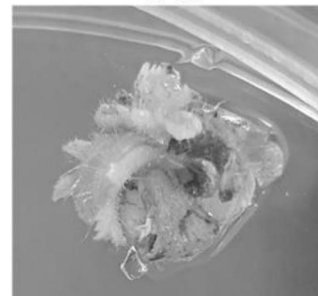


图13

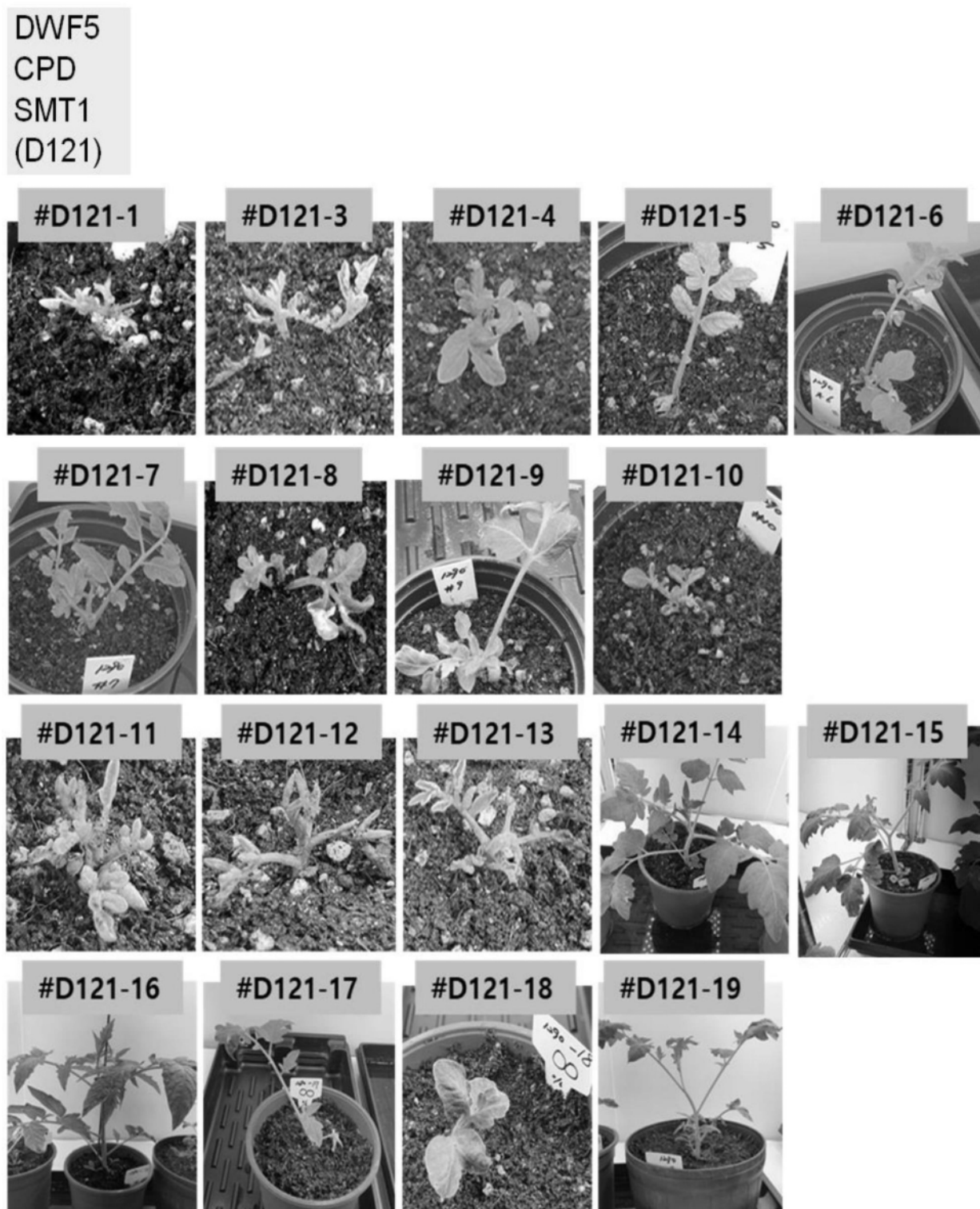


图14

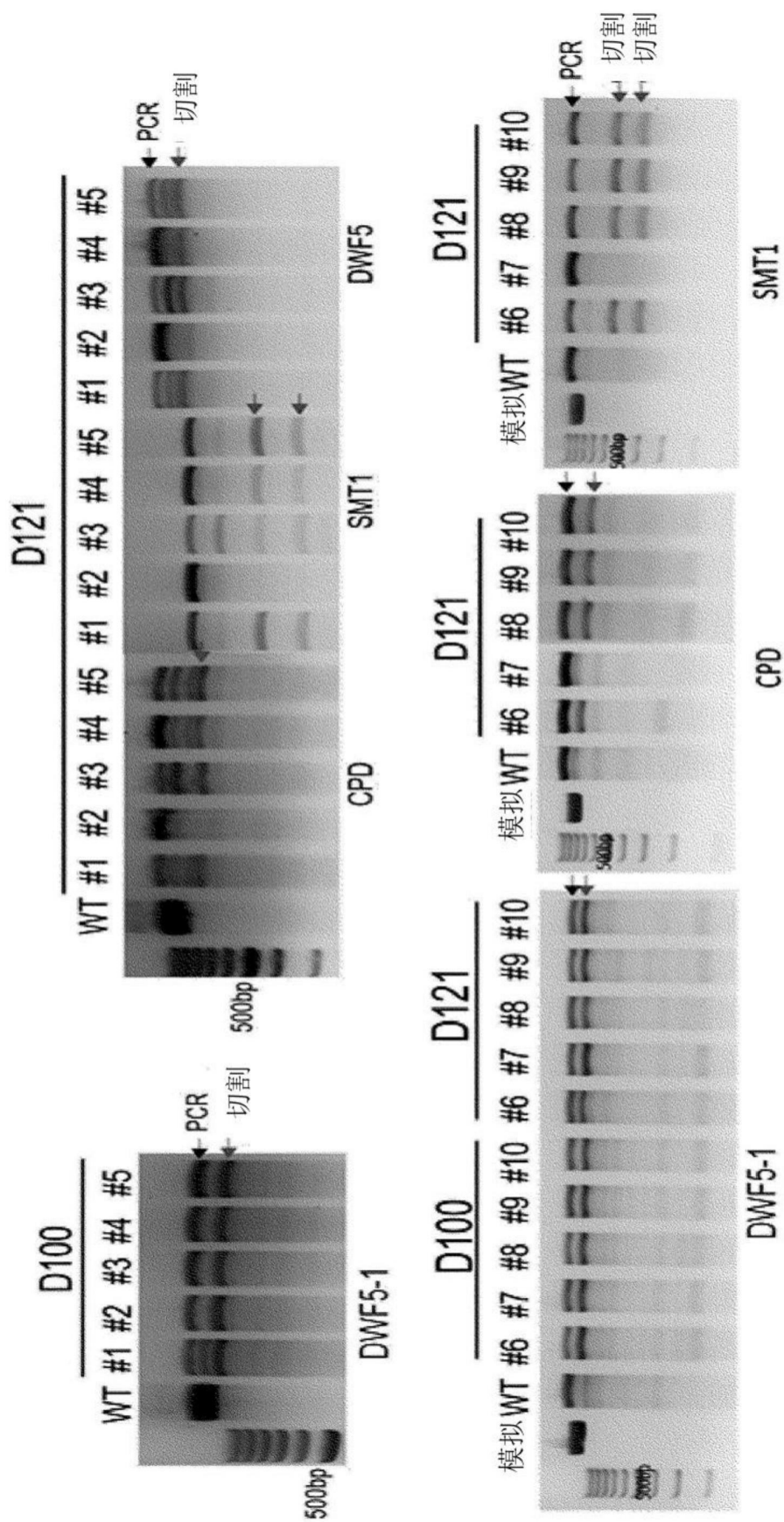


图15

D100-1		D100-2	D100-3	D100-4	D100-5	D100-6	D100-7
DWF5-1	突变	0/-159	+1/-4	+1/-4	+1/-4	+1/-4, 0/-1,	0/-1, +1/-4
		+1/-4	-8/-1, -8/-4	-2/-1	-6/-1	-4/-1, -5/-1,	-4/-1, -17/-1
		-0/-11				-33/-1, -11/-1	
接合性		同	同	同	同	同	同

D100-8		D100-10	D100-12	D100-13	D100-14	D100-15	D100-16
DWF5-1	突变	+1/-4	+1/-4	+1/-4	+1/-4	+1/4	+1/-4
		0/-1	+1/-4	-8/-1	-5/-1	-5/-1	-8/-1
			-5/-1				
接合性		同	同	同	同	同	同

图16

T0	T1	总读段 (raw fastq)	SgRNA 1			SgRNA 2		
			插入	缺失	Indel频率	插入	缺失	Indel频率
D100-1	D100_1-1	87344	86370	340	86710 (99.8%)	0	86913	86913 (99.9%)
	D100_1-2	54277	53495	29	53524 (99.9%)	0	53629	53629 (100.0%)
	D100_1-4	54192	53220	51	53271(99.9%)	0	53357	53357 (100.0%)
D100-3	D100_1-5	69409	68498	167	68665 (99.9%)	0	68821	68821 (100.0%)
	D100_3-5	50717	44	50256	50300 (99.9%)	0	50376	50376 (100.0%)
	D100_3-6	71229	70528	65	70593 (99.8%)	0	70775	70775 (99.9%)
	D100_3-11	54964	17	54531	54548 (100.0%)	0	54615	54615 (100.0%)
	D100_3-14	44027	16	42986	43002 (100.0%)	0	43042	43042 (100.0%)
D100-4	D100_3-15	53998	53473	70	53543 (100.0%)	0	53663	53663 (100.0%)
	D100_3-18	53317	49	52868	52917 (100.0%)	0	52978	52978 (100.0%)
	D100_4-4	83511	386	82710	83096 (100.0%)	0	82938	82938 (100.0%)
	D100_4-8	83297	65	82819	82884 (100.0%)	0	82757	82757 (100.0%)
D100-5	D100_4-9	79034	108	78516	78624 (100.0%)	0	78519	78519 (100.0%)
	D100_5-4	60994	60531	60	60591 (99.9%)	0	60737	60737 (100.0%)
	D100_5-9	71134	70473	134	70607 (99.9%)	0	70699	70699 (99.9%)
D100-7	D100_7-1	89757	89183	54	89237(99.9%)	0	89421	89421 (100.0%)
	D100_7-5	99082	58	98701	98759 (100.0%)	0	98803	98803 (100.0%)
D100-8	D100_8-1	71913	37	39	76 (0.1%)	0	71658	71658 (100.0%)
	D100_8-3	76121	75656	81	75737 (99.9%)	0	75867	75867 (100.0%)
	D100_8-4	113584	85	111676	111761 (100.0%)	0	111869	111869 (100.0%)
D100-12	D100_12-3	88658	88001	163	88164 (99.9%)	0	88331	88331 (100.0%)
	D100_12-7	110408	109131	623	109754 (99.9%)	0	109946	109946 (100.0%)
	D100_12-8	7532	7377	13	7390 (99.9%)	0	7412	7412 (100.0%)

图17

标签	总读段	SgRNA 1		SgRNA 2		T2种子量 (g)
		插入	缺失	插入	缺失	
D100-1-4	54192	53220	51	0	53357	0.37
D100-7-1	89757	89183	54	0	89421	1.06
D100-12-8	7532	7377	13	0	7412	0.34
D100-3-11	54964	17	54531	0	54615	0.13
D100-3-14	44027	16	42986	0	43042	0.45
D100-4-8	83297	65	82819	0	82757	0

SolycDWF5_1-1

WT GGTATGGAGCTTTATCCTCGTA-TCGGGAAACACTTCGATATTAAAGCTTCACAAACTGCAGATTGGTATGATGCTTGGCAGTTC (SEQ ID NO: 66)

#3-11 GGTATGGAGCTTTATCCTC--A-TCGGGAAACACTTCGATATTAAAGCTTCACAAACTGCAGATTGGTATGAT-TCTTGGGCAGTTC (-2/-1)(SEQ ID NO: 67)

#3-14 GGTATGGAGCTTTATCCTC--A-TCGGGAAACACTTCGATATTAAAGCTTCACAAACTGCAGATTGGTATGAT-TCTTGGGCAGTTC (-2/-1)(SEQ ID NO: 68)

#4-8 GGTATGGAGCTTTATCCTC--A-TCGGGAAACACTTCGATATTAAAGCTTCACAAACTGCAGATTGGTATGAT-TCTTGGGCAGTTC (-2/-1)(SEQ ID NO: 69)

#1-4 GGTATGGAGCTTTATCCTCGTAATCGGGAAACACTTCGATATTAAAGCTTCACAAACTGCAGATTGGTAT----TCTTGGGCAGTTC (+1/-4) (SEQ ID NO: 70)

#7-1 GGTATGGAGCTTTATCCTCGTAATCGGGAAACACTTCGATATTAAAGCTTCACAAACTGCAGATTGGTAT----TCTTGGGCAGTTC (+1/-4) (SEQ ID NO: 71)

#12-8 GGTATGGAGCTTTATCCTCGTAATCGGGAAACACTTCGATATTAAAGCTTCACAAACTGCAGATTGGTAT----TCTTGGGCAGTTC (+1/-4) (SEQ ID NO: 72)

SolycDWF5_1-2

图18a

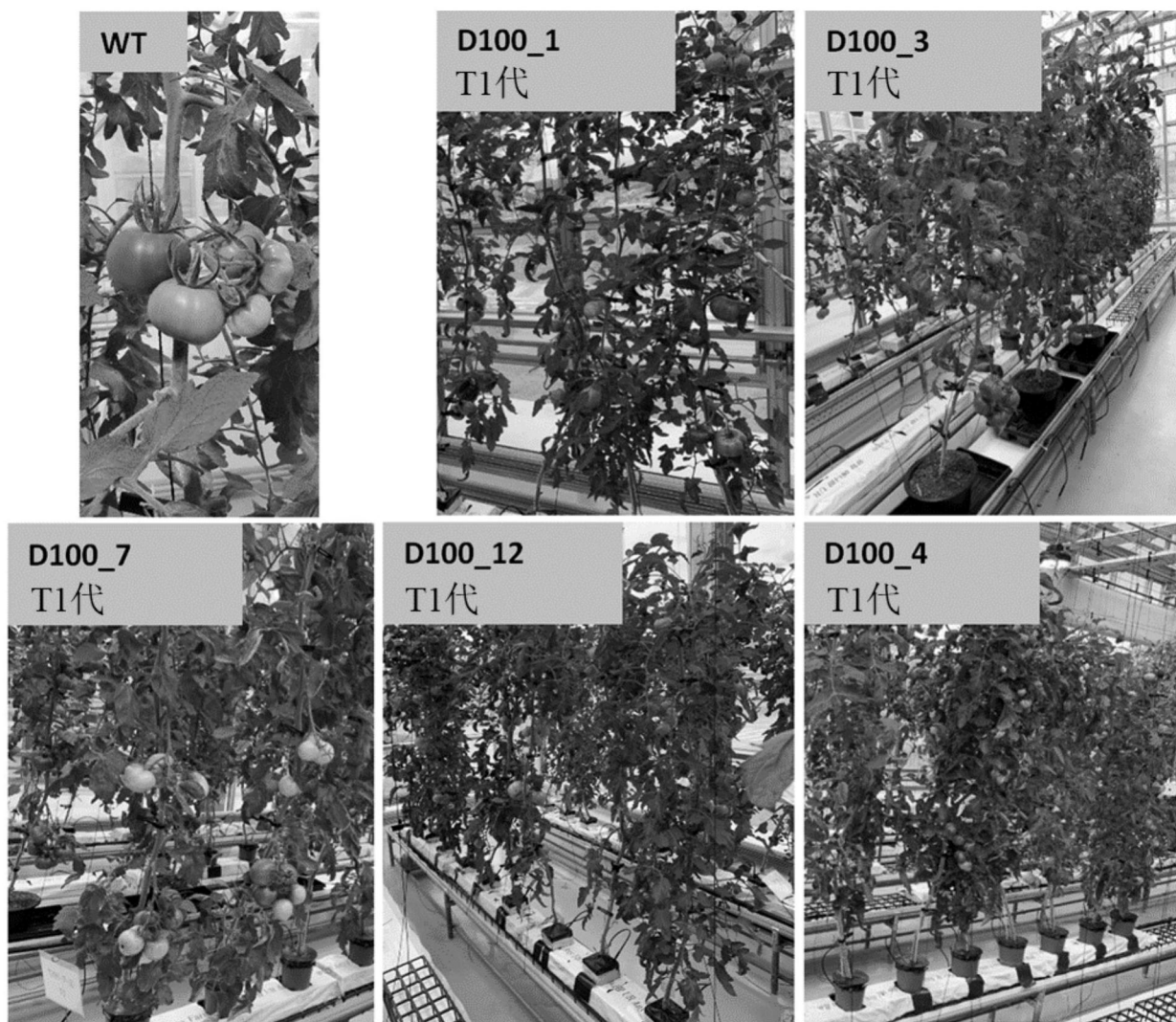


图18b

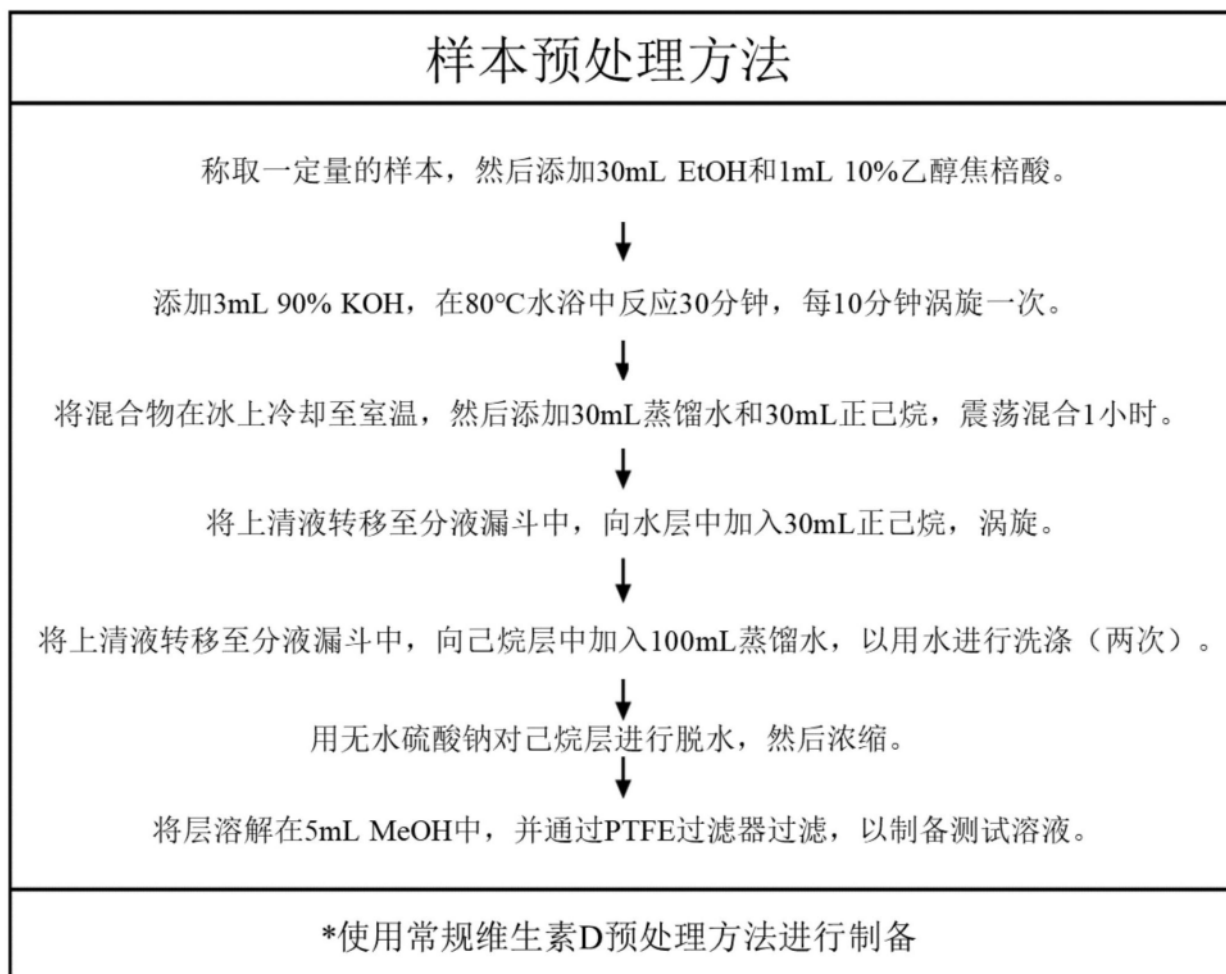


图19

仪器	HPLC-MS/MS		
离子源类型	APCI		
注射体积	20 μ L		
柱	Waters HSS T3 (1.8 μ m, 2.1 $\times$ 100mm)		
流动相	A : D.W中0.1%甲酸		
	B : MeOH中0.1%甲酸		
	Time(min)	A(%)	B(%)
	0	90	10
	1	0	100
	10	0	100
	10.1	90	10
	14	90	10
流速	0.2 mL/min		
运行时间	14 min		
炉温	30 $^{\circ}$ C		
阳离子放电电流	4 (μ A)		
外壳气压	50 (Arb)		
辅助气压	10 (Arb)		
扫气压力	0 (Arb)		
离子转移管温度	325 $^{\circ}$ C		
SRM	亲代 MS - 385.275		
	产物 MS-	定量离子	- 367.437
		定性离子	-201.325 / 241.325

图20

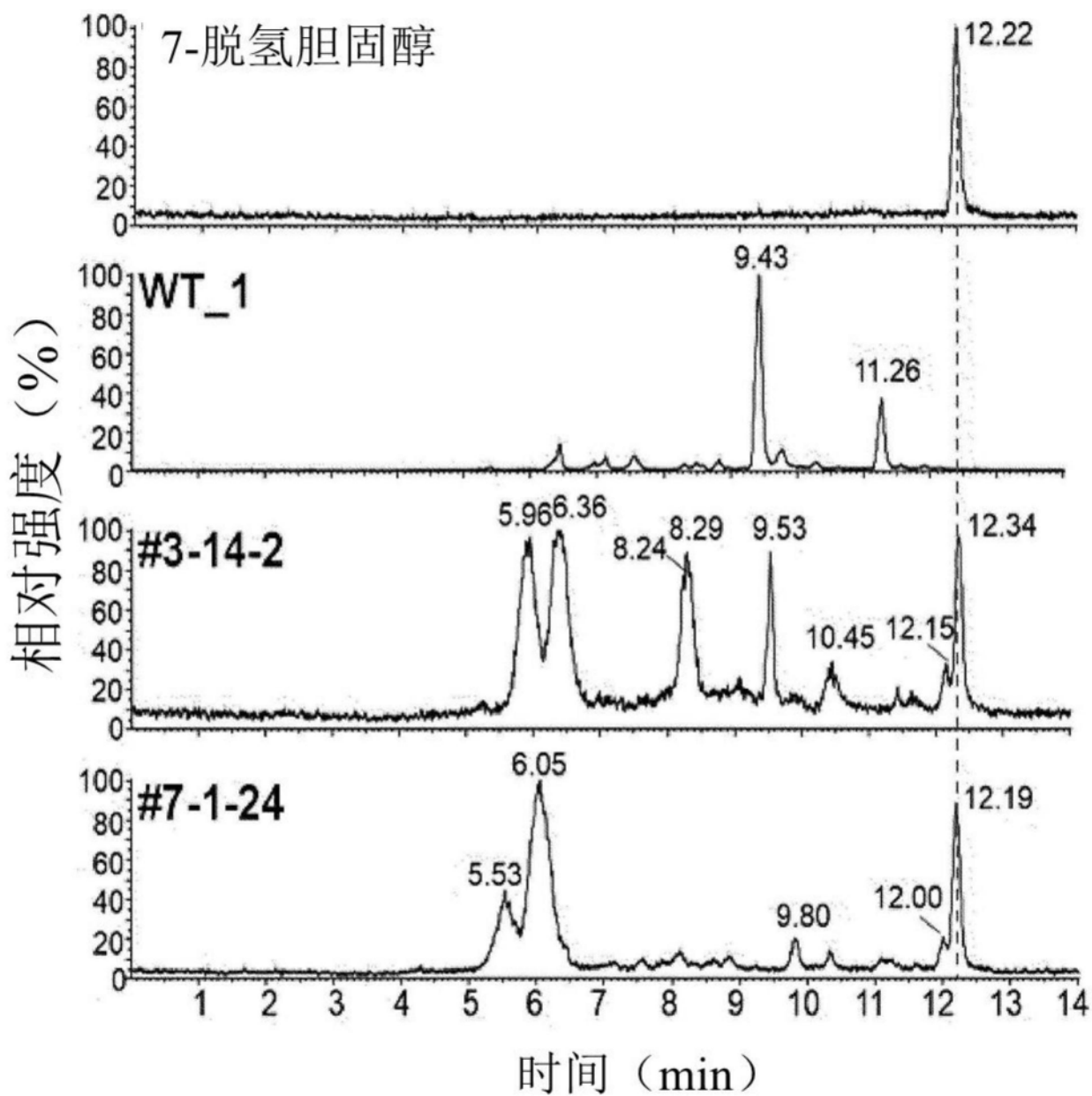


图21

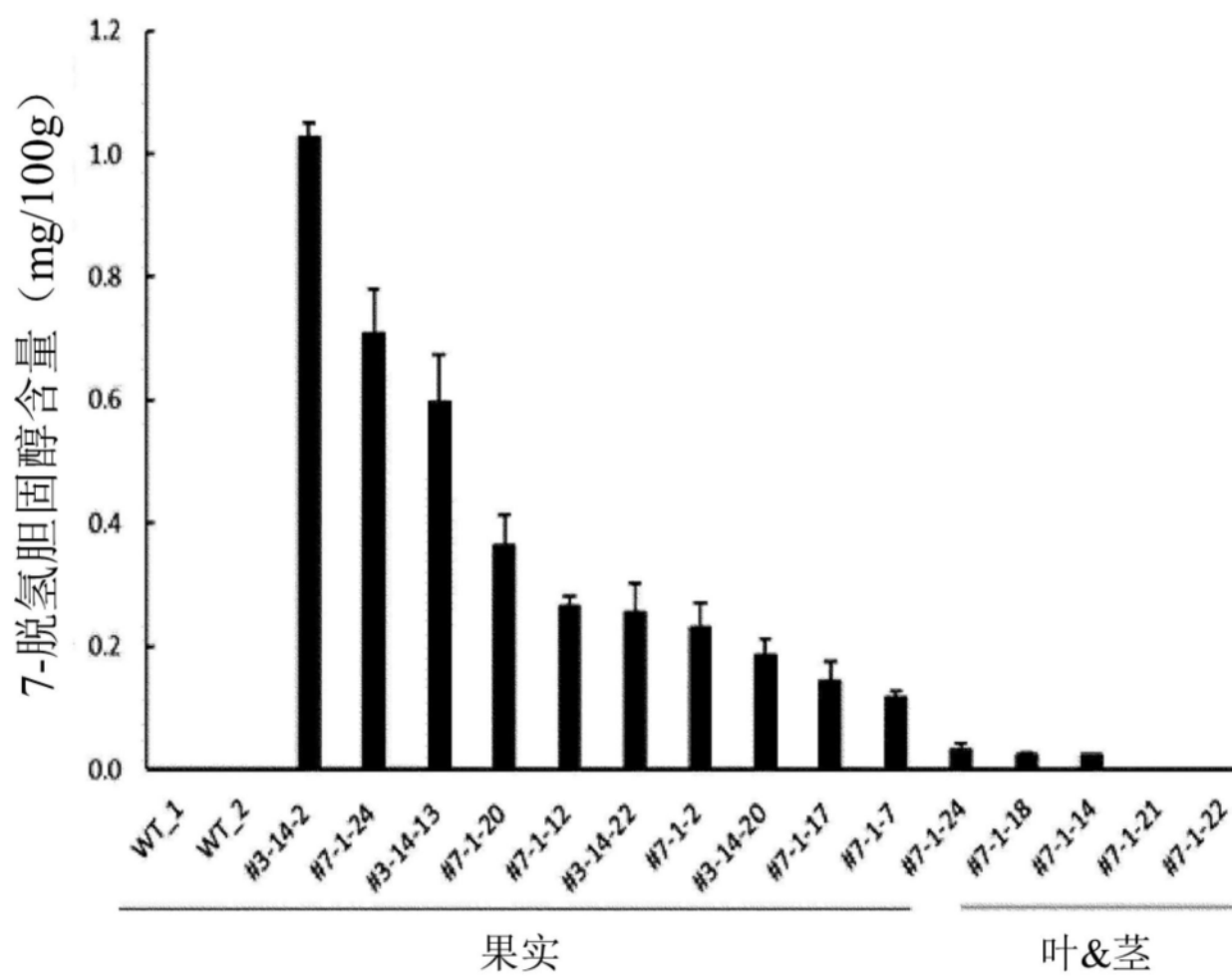


图22

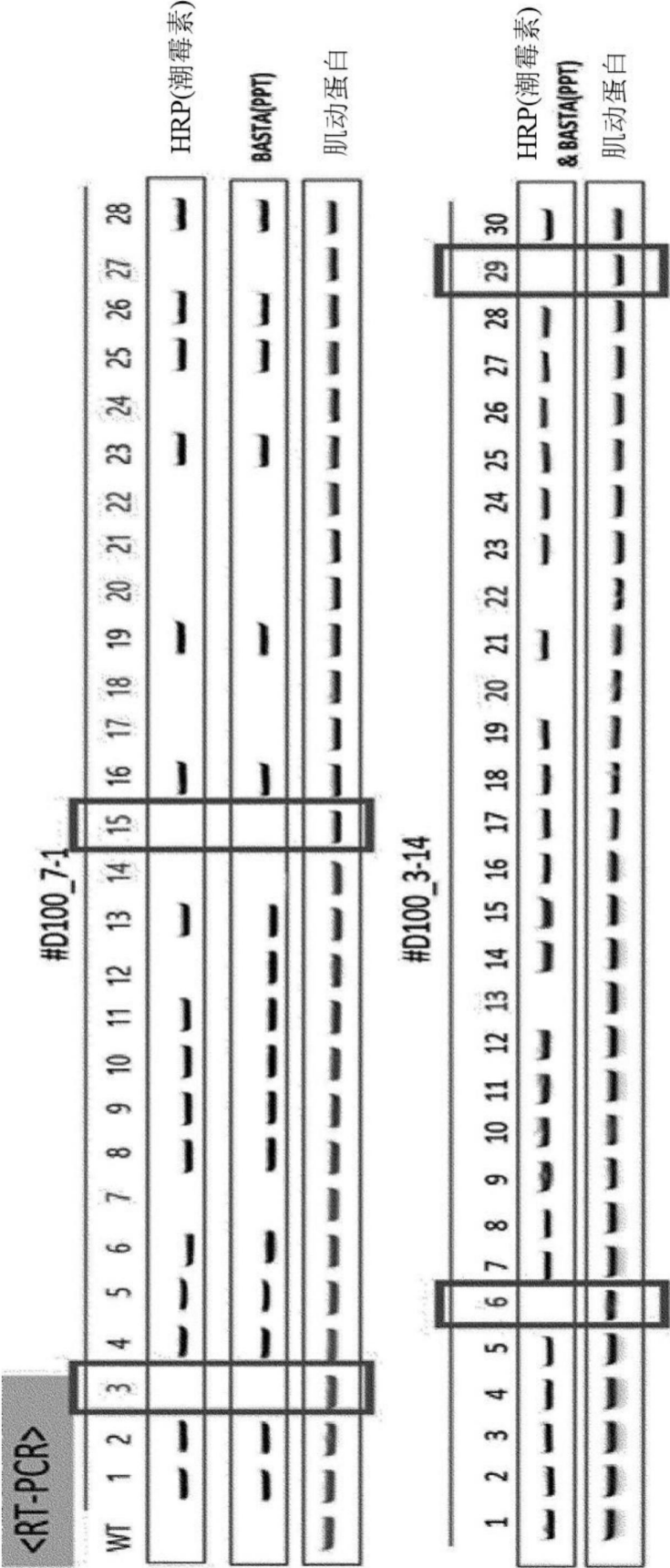


图23

一致 番茄WT D100_3-14-29 D100_7-1-15	ATGGAGCTTTATCCTCGT ATCGGGAACACTTCGATATTAAAGTCTTCACAACTGCAGATTGGTATGAT TCTTGGGCAGTCTCTGCTGCAGATTGCATAAAAC			
	ATGGAGCTTTATCCTCGT -ATCGGGAACACTTCGATATTAAAGTCTTCACAACTGCAGATTGGTATGAT -TCTTGGGCAGTCTCTGCTGCAGATTGCATAAAAC			
	ATGGAGCTTTATCCTCGT -ATCGGGAACACTTCGATATTAAAGTCTTCACAACTGCAGATTGGTATGAT -TCTTGGGCAGTCTCTGCTGCAGATTGCATAAAAC			
	ATGGAGCTTTATCCTCGT -ATCGGGAACACTTCGATATTAAAGTCTTCACAACTGCAGATTGGTATGAT -TCTTGGGCAGTCTCTGCTGCAGATTGCATAAAAC			
	ATGGAGCTTTATCCTCGT -ATCGGGAACACTTCGATATTAAAGTCTTCACAACTGCAGATTGGTATGAT -TCTTGGGCAGTCTCTGCTGCAGATTGCATAAAAC			
一致 番茄WT D100_3-14-29 D100_7-1-15	AGCATGAGAAATATGGGCGTGTATCTGATTCATGCTGTAAATAGCATATTGATGTTGGTGTATGTCACAAAATCTTTTGGTGGGAGCTGGGTACTGGAAACAATG			
	AGCATGAGAAATATGGGCGTGTATCTGATTCATGCTGTAAATAGCATATTGATGTTGGTGTATGTCACAAAATCTTTTGGTGGGAGCTGGGTACTGGAAACAATG			
	AGCATGAGAAATATGGGCGTGTATCTGATTCATGCTGTAAATAGCATATTGATGTTGGTGTATGTCACAAAATCTTTTGGTGGGAGCTGGGTACTGGAAACAATG			
	AGCATGAGAAATATGGGCGTGTATCTGATTCATGCTGTAAATAGCATATTGATGTTGGTGTATGTCACAAAATCTTTTGGTGGGAGCTGGGTACTGGAAACAATG			
	AGCATGAGAAATATGGGCGTGTATCTGATTCATGCTGTAAATAGCATATTGATGTTGGTGTATGTCACAAAATCTTTTGGTGGGAGCTGGGTACTGGAAACAATG			
一致 番茄WT D100_3-14-29 D100_7-1-15	GATATTGCACATGACAGAGCTGGCTTTTATATATGTTGGGGATGCTTGGT			
	GATATTGCACATGACAGAGCTGGCTTTTATATATGTTGGGGATGCTTGGT			
	GATATTGCACATGACAGAGCTGGCTTTTATATATGTTGGGGATGCTTGGT			
	GATATTGCACATGACAGAGCTGGCTTTTATATATGTTGGGGATGCTTGGT			
	GATATTGCACATGACAGAGCTGGCTTTTATATATGTTGGGGATGCTTGGT			

图24



图25

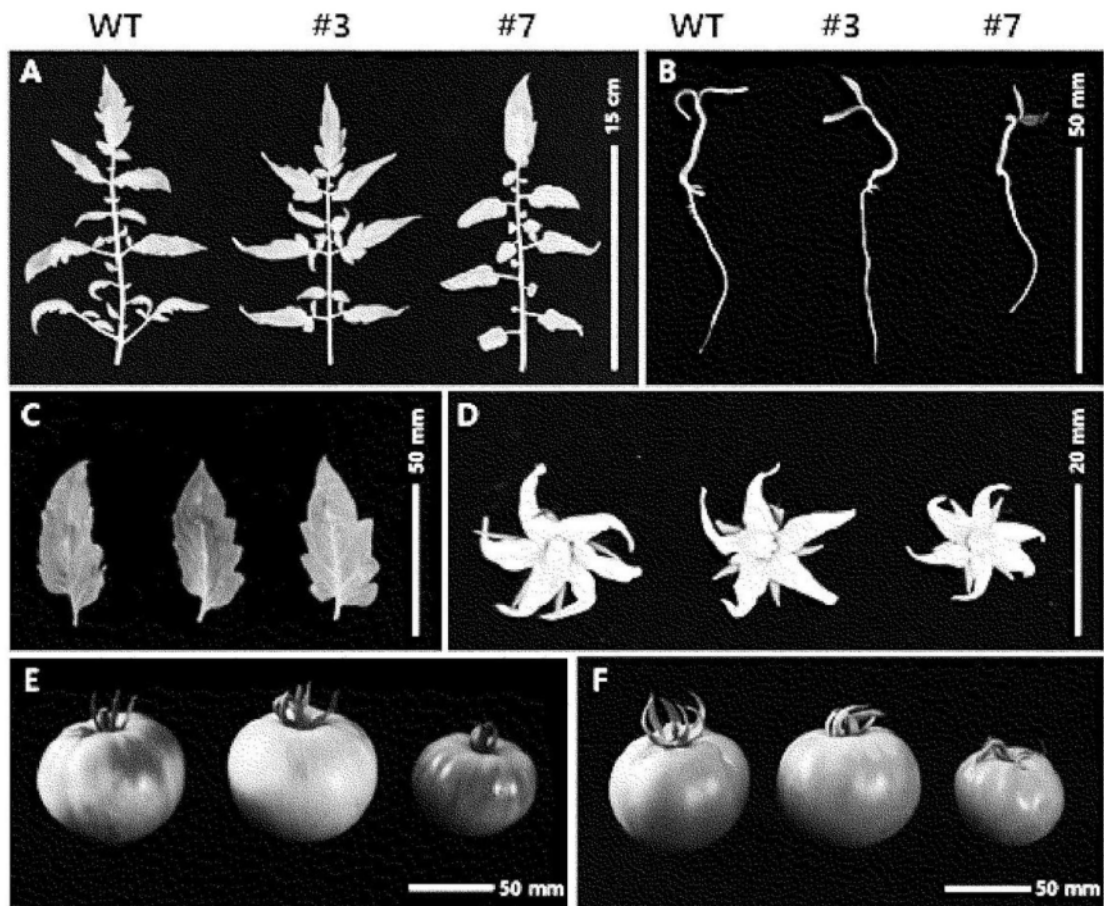


图26

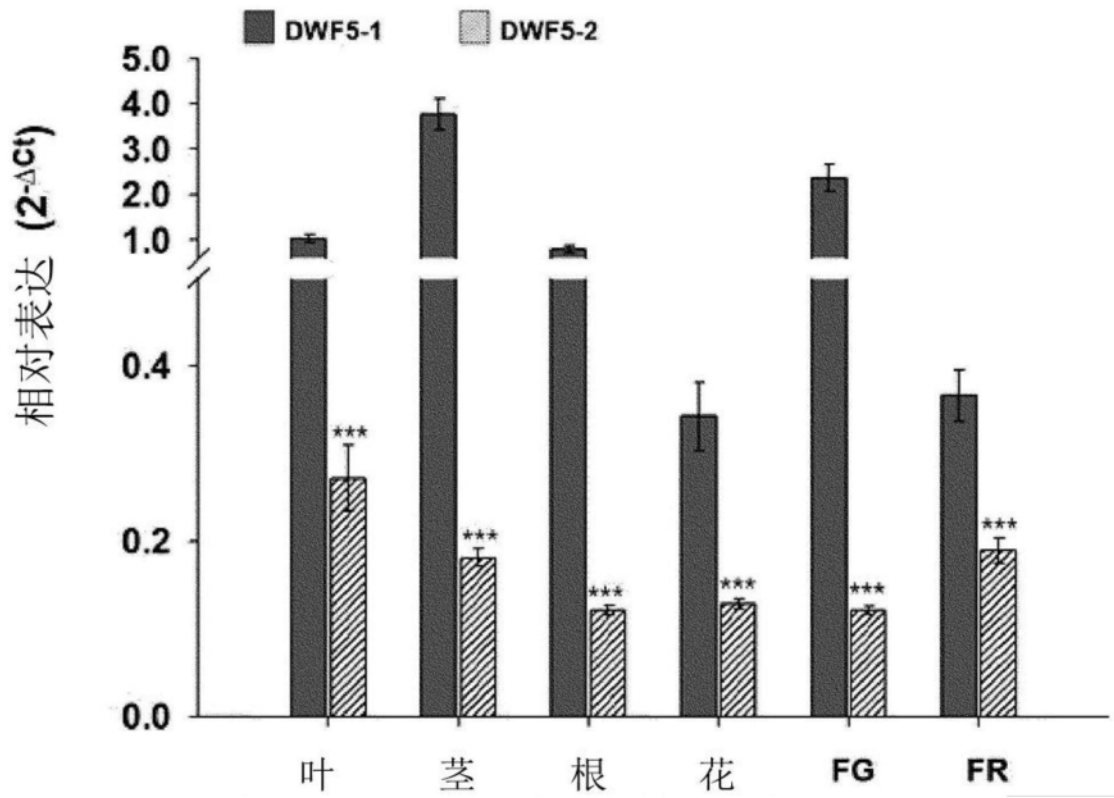


图27

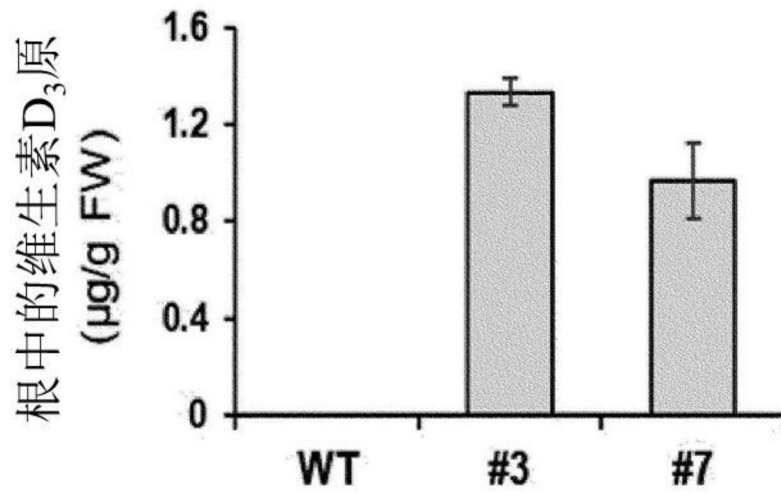


图28