



(12) 发明专利申请

(10) 申请公布号 CN 102787105 A

(43) 申请公布日 2012. 11. 21

(21) 申请号 201210245765. 2

*C12N 1/21* (2006. 01)

(22) 申请日 2012. 07. 16

*A01H 5/00* (2006. 01)

(71) 申请人 河北农业大学

地址 071001 河北省保定市灵雨寺街 289 号

(72) 发明人 马峙英 韩洁 王省芬 潘玉欣

张艳 张桂寅

(74) 专利代理机构 北京路浩知识产权代理有限

公司 11002

代理人 王朋飞 张庆敏

(51) Int. Cl.

*C12N 9/04* (2006. 01)

*C12N 15/53* (2006. 01)

*C12N 15/63* (2006. 01)

*C12N 5/10* (2006. 01)

*C12N 1/15* (2006. 01)

*C12N 1/19* (2006. 01)

权利要求书 1 页 说明书 8 页

序列表 15 页 附图 4 页

(54) 发明名称

棉花尿苷二磷酸葡萄糖脱氢酶基因 UGD6、其编码蛋白及应用

(57) 摘要

本发明涉及棉花尿苷二磷酸葡萄糖脱氢酶基因 UGD6、其编码蛋白及应用。棉花尿苷二磷酸葡萄糖脱氢酶具有如 SEQ ID No. 2 所示的氨基酸序列或该序列经替换、缺失或添加一个或几个氨基酸且具有同等功能的由 SEQ ID NO. 2 衍生的氨基酸序列。棉花尿苷二磷酸葡萄糖脱氢酶基因具有如 SEQ ID No. 1 所示的核苷酸序列。通过研究该基因在纤维发育不同时期的表达模式及转化拟南芥,表明该基因在形成细胞壁纤维素和半纤维素前体物中起着重要作用,而这些物质又直接参与棉纤维的形成,因此该基因为棉纤维品质改良提供了新的候选基因,具有重大的应用价值。

1. 一种棉花尿苷二磷酸葡萄糖脱氢酶,其特征在于,具有如 SEQID NO. 2 所示的氨基酸序列或该序列经替换、缺失或添加一个或几个氨基酸且具有同等功能的由 SEQ ID NO. 2 衍生的氨基酸序列。

2. 编码权利要求 1 所述棉花尿苷二磷酸葡萄糖脱氢酶的基因。

3. 根据权利要求 2 所述基因,其特征在于,具有如 SEQ ID No. 1 所示的核苷酸序列。

4. 含有权利要求 2 所述基因的载体。

5. 含有权利要求 2 或 3 所述基因或权利要求 4 所述载体的宿主细胞。

6. 权利要求 1 所述棉花尿苷二磷酸葡萄糖脱氢酶或权利要求 2 或 3 所述基因在提高棉纤维品质中的应用。

## 棉花尿苷二磷酸葡萄糖脱氢酶基因 UGD6、其编码蛋白及应用

### 技术领域

[0001] 本发明涉及基因工程领域,特别是涉及棉花尿苷二磷酸葡萄糖脱氢酶基因、其编码蛋白及应用。

### 背景技术

[0002] 棉纤维是已知纤维素纯度最高的天然资源之一。棉纤维品质主要包括纤维长度、强度、马克隆值等性状,这些性状主要受本身基因型控制。我国棉纤维存在内在品质相对较差,纤维强度比美棉低等问题。因此,开展棉纤维发育的分子生物学及相关基因表达调控的研究,对于利用基因工程改良棉花纤维品质具有重要意义。

[0003] 棉花纤维实质上是种皮毛,由位于子房内的胚珠外被上的表皮细胞发育分化而成。从形态上讲,棉花纤维细胞的发育过程是细胞超常伸长和细胞壁超常加厚的过程。棉花纤维细胞的形成一般可以分为四个相互有重叠的时期:纤维原始细胞分化突起、伸长期、次生壁加厚期、脱水成熟期。伸长期和次生壁加厚期是棉纤维品质形成的关键时期。初生细胞壁的合成是细胞伸长期的重要活动之一,与纤维长度直接相关。在伸长期结束之前,纤维细胞开始次生细胞壁的合成,主要是纤维素的淀积,纤维细胞次生壁的加厚决定纤维强力。

[0004] 棉纤维细胞壁主要包括纤维素和木葡聚糖、木聚糖、果胶多糖等非纤维素成分。纤维素含量约占次生壁干重的 95%;非纤维素物质则构成初生壁的主要成分。尿苷二磷酸葡萄糖(UDP-Glc)是纤维素合成的底物,UDP-Glc 在纤维素合酶作用下合成纤维素。非纤维素多糖由不同的单糖聚合而成,只有当它们被激活形成核苷糖后,才能作为非纤维素多糖合成的底物。核苷糖的合成由一整套酶来完成,研究发现几乎所有的核苷糖都是以 UDP-Glc 和 GDP-Man(鸟苷二磷酸甘露糖)为前体经过相关代谢酶催化合成的,因此,核苷糖转换相关的酶对棉纤维细胞壁纤维素和非纤维素的合成发挥着重要作用。

[0005] 尿苷二磷酸葡萄糖醛酸(UDP-GlcA)是许多核苷糖合成中的关键前体物,据推测,大约 50%的细胞壁物质来源于前体 UDP-GlcA,它是尿苷二磷酸葡萄糖脱氢酶(UDP-glucose dehydrogenase,UGD)催化 UDP-Glc 形成的产物。因此,尿苷二磷酸葡萄糖脱氢酶是形成非纤维素多糖和果胶的关键酶。在植物上已有多篇有关 UGD 基因的克隆报道。UGD 基因首先在大豆中克隆出来,随后,在拟南芥、甘蔗、玉米中分别克隆出此基因。研究表明,UGD 在不同的物种和组织中,其表达水平有较大差异。在杨树中 UGD 主要在幼叶和未成熟的木质部中表达,在成熟韧皮部中几乎无表达。然而在拟南芥幼苗中主要在根部表达,而成熟的植株中主要在维管系统中表达。苧麻 UGD 主要在茎中表达,其次在韧皮部和叶中,根中表达最少。大豆 UGD 在主根尖和侧根中有高量表达,在主根、成熟叶中则只有微量表达。说明 UGD 在幼嫩的组织有非常高的表达,而在植物成熟组织中表达量很少,这说明此酶在形成细胞壁纤维素和半纤维素前体物的过程中起着非常重要的作用,UGD 的表达将对细胞壁纤维素和非纤维素多糖含量产生影响,并进而影响到纤维素微纤丝和非纤维素交联。综上所述,UGD 基因对于揭示棉纤维品质形成的内在机理有重要意义,在纤维品质遗传改良中具有良好的应

用前景。UGD 在棉纤维合成过程中起着重要作用,但至今在棉花中仍无深入研究。

## 发明内容

[0006] 本发明目的是提供一种棉花尿苷二磷酸葡萄糖脱氢酶基因(UGD6)及其编码蛋白,以用于提高棉纤维品质。

[0007] 本发明另一目的是提供含有上述基因的载体。

[0008] 本发明又一目的是提供含有上述基因或载体的宿主细胞。

[0009] 本发明再一目的是提供棉花尿苷二磷酸葡萄糖脱氢酶或棉花 UGD6 基因在提高棉纤维品质中的应用。

[0010] 本发明所述棉花 UGD6 基因编码的蛋白,具有 SEQ ID NO. 2 所示的氨基酸序列或该序列经替换一个或几个氨基酸形成的具有同等功能的氨基酸序列。

[0011] 本发明所述棉花 UGD6 基因的核苷酸序列如 SEQ ID No. 1 所示,全长 1912bp, ORF 长 1443bp, 5' UTR 长 168bp, 3' UTR 长 301bp。

[0012] 具体地,本发明基于拟南芥四条尿苷二磷酸葡萄糖脱氢酶(At1g26570, AT3g29360, At5g15490 和 At5g39320)的编码序列,利用 BLASTN 在美国国立生物技术信息中心网站(<http://www.ncbi.nlm.nih.gov/>)棉花的表达序列标签数据库中搜索相似性较高的 EST 序列,结合电子拼接、RT-PCR 以及 RACE 技术克隆得到棉花尿苷二磷酸葡萄糖脱氢酶基因,命名为 UGD6。

[0013] 应当理解,考虑到密码子的简并性,例如可在其编码区,在不改变氨基酸序列的条件下,或在其非编码区在不影响蛋白表达的条件下,对编码上述蛋白的基因序列进行修改。因此,本发明还包含对编码上述蛋白的基因序列进行的替换、添加和 / 或缺失一个或多个核苷酸,具有与上述编码基因具有相同功能的核苷酸序列。本发明还包括基于所述基因的正义序列或反义序列,包括含有所述核苷酸序列或其片段的克隆载体或表达载体、含有所述载体的宿主细胞,利用所述宿主细胞制备棉花尿苷二磷酸葡萄糖脱氢酶的方法等。

[0014] 所述该序列经替换一个或几个氨基酸形成的具有同等功能的氨基酸序列,是指该序列的活性功能域之外的位点进行有限氨基酸的保护替换所得的序列仍能保持原来的活性。本发明氨基酸序列的活性功能域为 209~306 位,在此位点外可替换一个或几个氨基酸得到具有同等功能的氨基酸序列。比如,SEQ ID NO. 3 所示为海岛棉 UGD7 基因的核苷酸序列,SEQ ID NO. 4 所示为 UGD7 基因编码的氨基酸序列,与 UGD6 氨基酸序列相比,第 26、40、56、60、64、149、193、202、368 位发生了 9 个氨基酸的替换,且具有同样的功能效果。

[0015] 利用 Real-time PCR 检测 UGD6 基因在陆地棉“中棉所 8 号”和海岛棉“Pima90-53”棉纤维发育不同时期及幼苗根茎叶中的表达量,发现此基因在两个品种中表达量的变化趋势相似,在开花后 20~30d 的纤维中持续高效表达,30d 之后呈下降趋势,说明它对棉纤维次生壁的起始和合成具有一定作用。15d 幼苗中 UGD6 基因在“中棉所 8 号”和“Pima90-53”的幼苗根、下胚轴和叶的组织中,均有类似的表达量,表达量差异为:茎>根>叶。

[0016] 将 UGD6 基因与原核表达载体 pET-32a 连接,转化大肠杆菌 DH5 $\alpha$  感受态细胞,获得含有重组质粒的大肠杆菌工程菌株。在 IPTG 诱导条件下,表达融合蛋白,并进行 SDS-PAGE 电泳分析。pET32a-UGD6 重组质粒在大约 68.4kDa 的位置上出现特异条带,说明此基因可成功在原核细胞中表达蛋白。

[0017] 将 UGD6 基因与融合表达载体 pCamE::GFP 连接,转入大肠杆菌 DH5 $\alpha$  感受态细胞,获得含有重组质粒的大肠杆菌工程菌株。通过基因枪轰击洋葱表皮细胞。暗培养后在显微镜(蓝光激发)下观察,确定该基因编码的蛋白定位于细胞质膜系统。

[0018] 采用农杆菌浸花法将 S-UGD6 (正义)及 A-UGD6 (反义)基因表达载体转入 Columbia 型拟南芥中,转 S-UGD6 和 A-UGD6 拟南芥中的 UGD 酶活性相应增高和降低。转 A-UGD6 拟南芥植株纤维素含量与野生型相比显著增高,而转 S-UGD6 拟南芥则显著下降。转 A-UGD6 基因拟南芥茎长和茎粗均显著小于野生型,茎秆强度显著变小,而转 S-UGD 拟南芥表型无明显变化。

[0019] 提取野生型和转基因拟南芥 RNA,进行 Real-time PCR 分析,发现 UGD 基因的导入和敲除,显著影响了拟南芥内源 UGD 基因的表达水平,而 UGD 作为 UDP-葡萄糖代谢的关键酶,会对核苷糖合成乃至纤维素和果胶的合成产生影响。

[0020] 本发明的有益效果:

[0021] (1) 本发明提供的棉花 UGD6 基因为棉纤维品质改良基因工程提供了新的候选基因。

[0022] (2) 本发明提供的棉花 UGD6 基因在纤维强力不同的棉种中,表达量变化趋势有差异,UGD6 基因在棉纤维发育伸长期上调表达,其表达高峰在开花后 20~30 天,可用于初生壁的非纤维素多糖及次生壁纤维素合成研究。

[0023] (3) 本发明提供的棉花 UGD6 基因与棉花纤维发育有重要关系,可以利用基因工程手段进一步转化棉花从而提高棉纤维品质。

## 附图说明

[0024] 图 1 为本发明棉花 UGD6 基因的开放读码框(open reading frame, ORF)RT-PCR 扩增结果,其中, M 为 DL5000Marker, 1 为 UGD6 ORF 扩增产物。

[0025] 图 2 为本发明在 NCBI 上寻找棉花 UGD6 基因保守结构域的结果。

[0026] 图 3 为本发明棉花 UGD6 基因 5' RACE 及 3' RACE 扩增产物,其中, M 为 DL2000Marker, 1 为 5' RACE 扩增产物, 2 为 3' RACE 扩增产物。

[0027] 图 4 为本发明棉花 UGD6 基因在纤维发育不同时期的表达分析结果;横坐标 DPA 为开花后的不同天数,纵坐标为 UGD6 基因的表达量。

[0028] 图 5 为本发明棉花 UGD6 基因在根茎叶中的表达分析结果。

[0029] 图 6 为转棉花 UGD6 基因的 E. coli BL21(DE3)诱导前和诱导后目的融合蛋白表达的 SDS-PAGE 检测结果,箭头所示为目的融合蛋白,其中, M:蛋白标准分子量, 1:BL21(DE3)plys 空菌株未经 IPTG 诱导的总蛋白, 2:pET-32a(+) 空载体未经 IPTG 诱导的总蛋白; 3:UGD6-pET-32a(+) 未经 IPTG 诱导的总蛋白; 4:BL21(DE3)plys 空菌株经 IPTG 诱导 6 小时后的总蛋白; 5:pET-32a(+) 空载体经 IPTG 诱导 6 小时后的总蛋白; 6:UGD6-pET-32a(+) 经 IPTG 诱导 6 小时后的总蛋白。

[0030] 图 7 为本发明 UGD6 基因亚细胞定位的结果,其中, A:pCamE-GFP 蓝光激发下的细胞; B:A 的可见光下的细胞; C、D、F、G:pCamEUGD6::GFP 蓝光激发下的细胞; E:C 和 D 可见光下的细胞。 H:F 和 G 的可见光下的细胞。

[0031] 图 8 为本发明棉花 UGD6 基因真核表达载体构建示意图,其中, A:质粒 pBI121 结

构 ;B :pBI121-SUGD6 结构 ;C :pBI121-AUGD6 结构。

[0032] 图 9 为本发明棉花 UGD6 基因转化拟南芥的幼苗,其中,A :转 S-UGD6 拟南芥 ;B :转 A-UGD6 拟南芥。

[0033] 图 10 为生长 10 天的转 A-UGD6 T<sub>3</sub> 代与野生型拟南芥植株个体比较,其中,WT :野生型拟南芥对照,AU :转 A-UGD6 拟南芥植株。

[0034] 图 11 为生长 10 天的转 A-UGD6 基因 T<sub>3</sub> 代与野生型拟南芥的根长比较,其中,WT :野生型拟南芥对照,AU :转 A-UGD6 拟南芥植株。

[0035] 图 12 为生长 35 天的转 A-UGD6 基因 T<sub>3</sub> 代与野生型拟南芥的表型比较,其中,WT :野生型拟南芥对照,AU :转 A-UGD6 拟南芥植株。

[0036] 图 13 为转 S-UGD6 和 S-UGD6 拟南芥株系与野生型纤维素含量的比较图,其中,WT :野生型拟南芥对照,SU- :转 S-UGD6 拟南芥 T<sub>3</sub> 代不同株系,AU- :转 A-UGD6 拟南芥 T<sub>3</sub> 代不同株系。

[0037] 图 14 为转 S-UGD6 和 A-UGD6 拟南芥株系与野生型 UGD 酶活性的比较图,其中,WT :野生型拟南芥对照,SU- :转 S-UGD6 拟南芥 T<sub>3</sub> 代不同株系,AU- :转 A-UGD6 拟南芥 T<sub>3</sub> 代不同株系。

## 具体实施方式

[0038] 以下实施例用于说明本发明,但不用来限制本发明的范围。

[0039] 本发明的载体,菌株,质粒均可市售获得。

[0040] 实施例 1 棉花 UGD6 基因的克隆

[0041] (1) 首先利用拟南芥四条尿苷二磷酸葡萄糖醛酸脱羧酶 (At1g26570, AT3g29360, At5g15490 和 At5g39320) 的编码序列,利用 tBLASTN 在 NCBI 棉花表达序列标签数据库中搜索相似性较高的 EST 序列,利用 DNASTAR 软件对检索获得的 EST 序列进行电子拼接,获得具有完整开放阅读框的拼接序列 UGD6。

[0042] (2) 依据拼接序列设计特异性引物,同时在上、下游引物分别引入 KpnI 和 SalI 酶切位点(下划线部分)。UGD6 基因的上、下游引物分别为 :

[0043] 上游引物 :5' -GGTACCATGGTGAAGATCTGTTGC-3'

[0044] 下游引物 :5' -GTCGACTTATGCCACTGCAGGC-3'

[0045] (3) 利用设计的特异性引物分别以陆地棉“中棉所 8 号”和海岛棉“Pima90-53”(购自中国农业科学院棉花研究所国家棉花种质资源中期库) 开花后 20 天的棉纤维 cDNA 为模板进行 PCR 扩增,获得 UGD6 基因目的条带(图 1) ;

[0046] (4) 将回收得到的 UGD6 基因的 PCR 产物与 pGM-T 连接,采用天根生化科技(北京)有限公司的 pGM-T 克隆试剂盒克隆目的片段,连接产物采用热激转化法转入大肠杆菌 TOP10 感受态细胞(购自天根生化科技有限公司),并筛选阳性克隆,经上海生物工程有限公司测序,所测序列与预期结果一致,且两棉种间基因序列无差异。UGD6 基因 ORF 序列见 SEQ ID NO. 1 (长 1443bp),对应氨基酸序列为 SEQ ID NO. 2。

[0047] 实施例 2 棉花 UGD6 基因编码蛋白的相似性分析和保守结构域分析

[0048] 利用 DNAMAN 对棉花 UGD6 基因、棉花 GhUGD1~UGD5 以及 4 个拟南芥 UDP-葡萄糖脱氢酶基因编码蛋白的氨基酸序列进行相似性分析,结果显示棉花 UGD6 基因与棉花 UDP-葡

葡萄糖脱氢酶基因编码蛋白的氨基酸序列相似性分布在 89.0%~99.6% 之间,与拟南芥 UDP-葡萄糖脱氢酶基因编码蛋白的氨基酸序列相似性分布在 84.4%~92.7% 之间,说明棉花 UGD6 基因属于 UDP-葡萄糖酸脱氢酶基因家族。

[0049] 利用 NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) CDD 数据库进行在线保守域 BLAST 分析,发现了 UDP-葡萄糖 /GDP-甘露糖脱氢酶基因家族保守区(图 2),说明本发明的 UGD6 基因属于 UDP-葡萄糖脱氢酶基因家族。

[0050] 实施例 3 棉花 UGD6 基因 5' 和 3' 非翻译区(UTR)序列的克隆

[0051] 以棉花 UGD6 基因 ORF 序列为基础,按 Clontech RACE 试剂盒要求,分别设计 5' RACE 及 3' RACE 引物,具体引物序列如下:

[0052] 5' RACE 引物:

[0053] GSP :5' -CAGCTCCGAGCCCCGAGTCTTGGT-3'

[0054] NGSP :5' -TCACTGTTCCAGGCAGCAGATCCTAG-3'

[0055] 3' RACE 引物:

[0056] GSP :5' -ATATATGACCCGAGGTGACCGAAG-3'

[0057] NGSP :5' -GCAACCCATGAGTCCCACGACTGTC-3'

[0058] 以陆地棉“中棉所 8 号”和海岛棉“Pima90-53”开花后 20 天的棉纤维 cDNA 为材料,按 Clontech RACE 试剂盒操作步骤完成 5' 端及 3' 端非翻译区序列扩增,获得目的条带,UGD6 5' RACE 长 420bp,3' RACE 长 530bp(图 3)。对 PCR 产物进行测序,最后分析得 UGD6 5' UTR 长 168bp,3' UTR 长 301bp。将测序结果与 UGD6 基因开放阅读框拼接,获得全长 cDNA 序列,且两棉种间基因序列无差异(SEQ ID NO.1)。

[0059] 实施例 4 棉花 UGD6 基因的时空表达模式分析

[0060] 使用天根生化科技(北京)有限公司的 RNAlant 提取试剂盒提取“中棉所 8 号”和海岛棉“Pima90-53”开花当天胚珠以及开花后 5、10、15、20、25、30、35、40 天(days post anthesis,DPA)的纤维 RNA。棉花根、茎和叶 RNA 的提取使用北京奥莱博生物技术公司的植物总 RNA 提取试剂盒。

[0061] 采用大连宝生物公司的 PrimeScript® RT reagent Kit With gDNA Eraser(Perfect Real Time)试剂盒合成以上 9 个时期的单链 cDNA,浓度用 Beckman DU800 分光光度计测定。对棉花 UGD6 基因在棉纤维发育不同时期的表达进行 Real-time PCR 分析。内参选用 EF1 $\alpha$ ,其引物序列如下:

[0062] EF1 $\alpha$  F :5' -GCTGAGATGAACAAGAGGTCATTC-3'

[0063] EF1 $\alpha$  R :5' -GGAATCAATAATCAAAACAGCACAG-3'

[0064] Real-time PCR 结果(图 4)显示,棉花 UGD6 基因在 20~30DPA 的纤维中持续高效表达,30d 之后呈下降趋势,说明 UGD6 对棉纤维次生壁的起始和合成具有一定作用;15d 幼苗中 UGD6 基因在棉花根茎叶中均有表达,在茎中表达量最高,根中次之,叶中最低(图 5)。

[0065] 实施例 5 UGD6 基因原核表达载体的构建及诱导表达

[0066] (1) 原核表达载体的构建和转化子的获得

[0067] 利用 KpnI 和 SalI 分别对实施例 1 中克隆得到的 UGD6 基因的 ORF 和原核表达载体 pET-32a(+) 质粒(购自 Novagen 公司)进行双酶切并回收目的片段。

[0068] 将表达载体与目的基因片段进行连接,转化大肠杆菌 DH5 $\alpha$ ,经过 PCR 和酶切检

测,筛选含重组质粒的阳性克隆,利用热激法将含有阳性克隆的质粒转入 BL21 (DE3) 菌株(购自天根生化科技有限公司)中,并进一步挑取阳性克隆,经质粒提取、酶切检测获得含有重组质粒 pET-32a (+)-UGD6 的大肠杆菌工程菌株。

[0069] (2) 重组蛋白的诱导表达与鉴定

[0070] 将含有以上各类型的重组质粒和含有空质粒(对照)的 BL21 (DE3) 菌株在 28℃, IPTG 终浓度为 1.0mmol/L 的条件下诱导 6h 后,进行 SDS-PAGE 垂直式凝胶电泳,其中分离胶为 12%,浓缩胶为 5%。

[0071] 结果如图 6 所示,经 IPTG 诱导的含有去掉跨膜区重组质粒 pET-32a (+)-UGD6 的大肠杆菌 BL21(DE3)与含空质粒对照相比,在 68.4kDa 出现一条表达量明显增高的差异带。生物信息学预测可表达的 UGD6 目的融合蛋白包括 UGD6 蛋白(约 48kDa)和 pET-32a 标签蛋白(约 20.4kDa),分子量大约为 68.4kDa。因此,该差异带是 pET-32a (+)-UGD6 基因在大肠杆菌 BL21 (DE3) 中表达的目的融合蛋白。

[0072] 实施例 6UGD6 融合表达载体构建及融合蛋白的亚细胞定位

[0073] (1) 融合表达载体的构建

[0074] 设计用于构建融合表达载体的引物。在 UGD6 基因的 ORF 两端添加酶切位点 SalI 和 Kpn I (下划线部分),引物序列如下:

[0075] PUF :5' -GTCGACATGGTGAAGATCTGTTGC-3'

[0076] PUR :5' -GGTACCTGCCACTGCAGGCAT-3'

[0077] 以该引物进行 PCR 反应,并回收目的片段,与 pGM-T 载体连接,热击法转化大肠杆菌感受态细胞 DH5  $\alpha$ ,挑取阳性克隆,提取质粒进行酶切检测、测序,得到中间载体 pGM-UGD6。

[0078] 将中间载体 pGM-UGD6 和表达载体 pCamE :: GFP 分别用 SalI 和 Kpn I 双酶切,电泳检测酶切产物,回收目的片段进行连接,热击法转化大肠杆菌感受态细胞 DH5  $\alpha$ ,挑取阳性克隆,进行质粒提取,酶切检测和测序,获得融合表达载体 pCamUGD6 :: GFP。

[0079] (2) 基因枪介导的亚细胞定位

[0080] 在倒好的 MS 培养基平板上铺一层灭菌滤纸,将洋葱内表皮平铺在平板的中央,通过基因枪轰击转化融合表达载体。培养室暗培养 1-3d 后在显微镜(蓝光激发)下观察并照相。

[0081] 显微镜下观察基因枪轰击后的洋葱表皮细胞,发现不含外源基因的载体 pCamE-GFP 定位在细胞核和细胞膜上(图 7A-7B),与前人定位结果一致。PCamE-UGD6 :: GFP 绿色荧光出现在细胞膜、核膜、液泡膜以及内质网等结构上,另外观察到了明显的膜泡运动,说明 UGD6 蛋白定位在细胞膜系统上(图 7C-H)。

[0082] 实施例 7 棉花 UGD6 基因真核表达载体构建和拟南芥遗传转化

[0083] (1) 真核表达载体的构建

[0084] 真核表达载体构建示意图见图 8。设计用于构建植物表达载体的引物,在 UGD6 基因的 ORF 两端添加酶切位点 SacI 和 Xba I (下划线部分),序列如下:

[0085] SUF :5' -TCTAGAATGGTGAAGATCTGTTGC-3'

[0086] SUR :5' -GAGCTCTTATGCCACTGCAGG-3'

[0087] AUF :5' -GAGCTCATGGTGAAGATCTGTTGC-3'



[0088] AUR :5' -TCTAGATTATGCCACTGCAGGCAT-3'

[0089] 以该引物进行 PCR 反应,并回收目的片段,与 pGM-T 载体连接,热击法转化大肠杆菌感受态细胞 DH5  $\alpha$  (天根生化科技有限公司产品),挑取阳性克隆提取质粒,经酶切检测并测序,得到中间载体 pGM-SUGD6 和 pGM-AUGD6。

[0090] 将中间载体和真核表达载体 pBI121p35S 分别用 SacI 和 Xba I 双酶切,电泳酶切产物,回收目的片段,然后进行连接,热击法转化大肠杆菌感受态细胞 DH5  $\alpha$ ,挑取阳性克隆,进行质粒提取,经酶切检测和测序,获得真核表达载体 pBI121p35S::SUGD6 和 pBI121p35S::AUGD6。

[0091] (2) 拟南芥遗传转化及 UGD6 的功能分析

[0092] 1) 拟南芥遗传转化:将质粒 pBI121p35S::SUGD6 和 pBI121p35S::AUGD6 转化农杆菌菌株 GV3101 感受态细胞,采用农杆菌浸花法将 S-UGD6 和 A-UGD6 载体转化 Columbia 型拟南芥。将转化拟南芥植株花序浸泡于含有目的基因的转化介质,几分钟后取出,保湿暗培养 24h 后浇足营养液,恢复正常光照培养,待拟南芥种子成熟,收取种子进行干燥并春化后,种植于 MS 筛选培养基上(卡那霉素浓度为 100mg/L),筛选得到转基因阳性苗(图 9),移栽于蛭石中培养,并进行 PCR 检测进一步确定转基因阳性植株。之后,再经 2 轮的种植、筛选直至获得 T<sub>3</sub> 代阳性拟南芥转基因株系,进行表型和细胞壁成分分析。

[0093] 2) 表型分析:转基因拟南芥与野生型拟南芥相比,表型发生显著变化,转反义 UGD6 拟南芥发育迟缓,根长显著变短(图 10, 图 11),抽薹时间较野生型晚(图 12),茎长和茎粗均显著短于野生型,茎秆强度亦显著变小(表 1)。WT 为野生型拟南芥对照, AU- 为转 A-UGD6 拟南芥 T<sub>3</sub> 代不同株系。而转正义 UGD6 拟南芥表型无明显变化。

[0094] 表 1 转 A-UGD6 拟南芥株系与野生型表型指标的比较

[0095]

株系	WT	AU-2	AU-3	AU-5
茎长 /cm	26.78±1.07	14.98±0.80**	17.43±0.67**	18.25±1.29**
茎粗 /mm	1.05±0.05	0.73±0.08**	0.77±0.07**	0.77±0.08**
茎秆强度 /N	12.1±0.92	11.3±0.81**	9.9±0.68**	9.5±0.63**

[0096]

[0097] 3) 细胞壁成分分析:为进一步确定 UGD6 基因在细胞壁多糖转化中的功能,提取转基因和野生型拟南芥的细胞壁,采用蒽酮比色法测定拟南芥细胞壁中纤维素的含量。

[0098] 具体步骤为:采用乙酸、硝酸和水(8:1:2)的混合溶液沸水浴提取细胞壁中纤维素;ddH<sub>2</sub>O 和丙酮洗涤后真空抽干;用 72% 的硫酸重新溶解纤维素;加入蒽酮试剂并充分反应后测定 A<sub>620</sub>,对比标准曲线既得细胞壁纤维素含量。

[0099] 结果显示转反义 UGD6 拟南芥株系(AU)纤维素含量与野生型相比显著增高,而转正义 UGD6 拟南芥(SU)则显著下降(图 13)。

[0100] 4) UGD 酶活性分析:提取转基因及野生型拟南芥的 UGD 酶,按株系取生长六周的拟南芥,提取蛋白酶;材料称重,加入两倍体积的蛋白提取液(50mM Tris-HCl(pH7.5);2mM

EDTA,用前加 5mMDTT),充分研磨;4℃,11,200g 离心 2min,弃掉细胞壁碎片,上清再次离心 2min,即为粗酶液。UGD 酶的纯化使用上海生物工程有限公司的重力型 Sephadex 脱盐柱(BSP090)进行蛋白的脱盐纯化。纯化步骤严格按照说明书进行。

[0101] 转反义 UGD6 拟南芥株系酶活性与野生型相比显著降低,而转正义 UGD6 拟南芥植株则显著增高(图 14)。

[0102] UGD6 的功能分析结果表明:UGD6 基因的过量表达显著提高了拟南芥 UGD 酶活性,导致大量的尿苷二磷酸葡萄糖(UDP-Glc)被消耗,而细胞壁主要组分的纤维素是由葡聚糖结合所形成,UGD 的过量表达造成了纤维素含量的降低。

[0103] UGD 基因的抑制表达,一方面造成 UDP-Glc 的大量累积,引起纤维素含量的增高,另一方面也直接导致了由 UGD 催化葡萄糖的产物尿苷二磷酸葡萄糖醛酸(UDP-GlcA)含量的减少,在半纤维素和果胶中大部分糖(树胶醛糖、芹菜糖、半乳糖醛酸和木糖等)直接或间接来源于 UDP-GlcA,打破了原来植物细胞壁纤维素和非纤维素多糖含量的平衡,影响了细胞壁结构,从而导致一系列的表型变化。

[0104] 5) 转基因拟南芥内源 UGD 基因的表达分析:

[0105] 提取野生型和转基因拟南芥的 RNA,进行 Real-time PCR 反应,结果发现外源基因 S-UGD6 和 A-UGD6 均在拟南芥中高效转录,无论是转 S-UGD6 或 A-UGD6 基因,拟南芥内源 UGD 基因的转录水平都大幅降低。

[0106] 结果表明 UGD6 基因的过量表达对拟南芥内源 UGD 基因的表达产生了共抑制作用,但是由于 35S 启动子的强启动能力,外源 UGD6 基因得到了很好的表达,内源和外源基因的共同作用,并未引起植株表型的明显变化。UGD6 基因的抑制表达,则打破了拟南芥中 UGD 表达的平衡,造成一系列的表型差异。

[0107] 因此,UGD6 基因的过量和低量表达,导致棉纤维细胞壁中纤维素含量大幅变动,同时可能引起非纤维素多糖物质即木葡聚糖、木聚糖或果胶多糖含量的变化,形成类似于拟南芥的表型特征变化,进而影响到棉纤维长度和强度,影响棉纤维品质。

[0108] 虽然,上文中已经用一般性说明及具体实施方案对本发明作了详尽的描述,但在本发明基础上,可以对之作一些修改或改进,这对本领域技术人员而言是显而易见的。因此,在不偏离本发明精神的基础上所做的这些修改或改进,均属于本发明要求保护的范围。

[0001]

序列表

<110> 河北农业大学

<120> 棉花尿苷二磷酸葡萄糖脱氢酶基因 UGD6、其编码蛋白及应用

<130> KHP12H13417.0

<160> 18

<170> PatentIn version 3.3

<210> 1

<211> 1912

<212> DNA

<213> 人工序列

<400> 1

<220>

<221> CDS

<222> (169)..(1611)

<400> 1

tctaatacga ctcactatag ggcaagcagt ggtatcaacg cagagtacat gggatategt 60

tgaagaagca aagctectet cttettecte ttcattttcc tetgaattct ttgtctctat 120

ttaaaaaaaa agagcaattt tttgagggtg gtagtataca aaaacaaa atg gtg aag 177

Met Val Lys

I

atc tgt tgc att gga gct ggt tat gtt ggt ggg cct aca atg get gtg 225

Ile Cys Cys Ile Gly Ala Gly Tyr Val Gly Gly Pro Thr Met Ala Val

5

10

15

atc gcg ctt aaa tgt ccg gat att gaa gtg gct gtc gtt gat atc tet 273

Ile Ala Leu Lys Cys Pro Asp Ile Glu Val Ala Val Val Asp Ile Ser

[0002]

20	25	30	35	
gta cct agg atc gct gcc tgg aac agt gac cag ctc cct atc tat gag				321
Val Pro Arg Ile Ala Ala Trp Asn Ser Asp Gln Leu Pro Ile Tyr Glu				
	40	45	50	
cct ggt ctt gat ggt gtt gtg aag gag tgt cgt ggc aga aac ctt ttc				369
Pro Gly Leu Asp Gly Val Val Lys Glu Cys Arg Gly Arg Asn Leu Phe				
	55	60	65	
ttc agc act gat gta gag aaa cat gtg cgt gag gct gat ata gtc ttt				417
Phe Ser Thr Asp Val Glu Lys His Val Arg Glu Ala Asp Ile Val Phe				
	70	75	80	
gtt tet gtc aac act ccc acc aag act cgg ggg ctc gga gct ggc aaa				465
Val Ser Val Asn Thr Pro Thr Lys Thr Arg Gly Leu Gly Ala Gly Lys				
	85	90	95	
gct gca gac ctt aca tat tgg gaa agt gct gcc cgc atg atc gcg gac				513
Ala Ala Asp Leu Thr Tyr Trp Glu Ser Ala Ala Arg Met Ile Ala Asp				
100	105	110	115	
gta tcg aaa tct gac aag att gtt gtt gag aaa tca aca gtc cct gtc				561
Val Ser Lys Ser Asp Lys Ile Val Val Glu Lys Ser Thr Val Pro Val				
	120	125	130	
aaa act gct gag gea att gaa aag att ctc acc cac aac agc aag gga				609
Lys Thr Ala Glu Ala Ile Glu Lys Ile Leu Thr His Asn Ser Lys Gly				
	135	140	145	

[0003]

atc aaa ttc caa att ctt teg aac cca gaa ttt ctt gct gag gga act	657
Ile Lys Phe Gln Ile Leu Ser Asn Pro Glu Phe Leu Ala Glu Gly Thr	
150 155 160	
gca att caa gat ctt ttt aat cct gat cgt gtt ctt atc gga ggc cgg	705
Ala Ile Gln Asp Leu Phe Asn Pro Asp Arg Val Leu Ile Gly Gly Arg	
165 170 175	
gaa acc cca gaa ggc aat aag gct gtc caa gca ttg aag gag gtt tat	753
Glu Thr Pro Glu Gly Asn Lys Ala Val Gln Ala Leu Lys Glu Val Tyr	
180 185 190 195	
gct cat tgg gtt cct gag gag aga att ctt acc acc aat ctt tgg tet	801
Ala His Trp Val Pro Glu Glu Arg Ile Leu Thr Thr Asn Leu Trp Ser	
200 205 210	
gca gag ctt tcc aag ctt gct gcc aat gct ttc ttg gct cag agg atc	849
Ala Glu Leu Ser Lys Leu Ala Ala Asn Ala Phe Leu Ala Gln Arg Ile	
215 220 225	
tca tcc gtg aat gcc atg tct gct ctt tgt gag gca acc ggg gct gat	897
Ser Ser Val Asn Ala Met Ser Ala Leu Cys Glu Ala Thr Gly Ala Asp	
230 235 240	
gtt aag cag gtc tct tat gct gtt ggt aag gac aca agg att gga ccc	945
Val Thr Gln Val Ser Tyr Ala Val Gly Lys Asp Thr Arg Ile Gly Pro	
245 250 255	

[0004]



Val Gln Arg Asp Leu Ser Met Asn Lys Phe Asp Trp Asp His Pro Leu	
375	380
cac ttg caa ccc atg agt ccc acg act gtc aag caa gtc agc igt gtg	1377
His Leu Gln Pro Met Ser Pro Thr Thr Val Lys Gln Val Ser Cys Val	
390	395
tgg gat gct tat gag gca acc aaa gat gcc cat gga att tgc att ctg	1425
Trp Asp Ala Tyr Glu Ala Thr Lys Asp Ala His Gly Ile Cys Ile Leu	
405	410
acc gag tgg gac gaa ttc aag aaa ctt gac ttc aag agg ata tac gac	1473
Thr Glu Trp Asp Glu Phe Lys Lys Leu Asp Phe Lys Arg Ile Tyr Asp	
420	425
aac atg caa aag ccg gct ttt gtt ttc gat gga agg aac atc gtg aac	1521
Asn Met Gln Lys Pro Ala Phe Val Phe Asp Gly Arg Asn Ile Val Asn	
440	445
gcg gat gag ctg agg gag att ggt ttc ata gtg tac tca att ggt aag	1569
Ala Asp Glu Leu Arg Glu Ile Gly Phe Ile Val Tyr Ser Ile Gly Lys	
455	460
cca ctg gat cca tgg ctc aag gac atg cct gca gtg gca taa	1611
Pro Leu Asp Pro Trp Leu Lys Asp Met Pro Ala Val Ala	
470	475
atgtctatcg attacagatg gagtccgggt cagtttgatt catacctatt gatttttgtt	1671

[0006]

ttcgatgaag agttgtgttg taccttggtt tccccccagc cacaacataa ttggaatttt 1731  
 accaccataa tgaagattta aaaggctttg cgttgtgcct ttaattcttg tcgtctgta 1791  
 ggaaacttaa attgagttcg atcagaactt gtgtcatttg tagaatgttg aaatttcata 1851  
 aacgatttgt ctctttaatg caaaaaaaaa aaaaaagtac tctgcgttga taccactgct 1911  
 t 1912

<210> 2

<211> 480

<212> PRT

<213> 人工序列

<400> 2

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

[0007]



Ile Val Phe Val Ser Val Asn Thr Pro Thr Lys Thr Arg Gly Leu Gly			
	85	90	95
Ala Gly Lys Ala Ala Asp Leu Thr Tyr Trp Glu Ser Ala Ala Arg Met			
	100	105	110
Ile Ala Asp Val Ser Lys Ser Asp Lys Ile Val Val Glu Lys Ser Thr			
	115	120	125
Val Pro Val Lys Thr Ala Glu Ala Ile Glu Lys Ile Leu Thr His Asn			
	130	135	140
Ser Lys Gly Ile Lys Phe Gln Ile Leu Ser Asn Pro Glu Phe Leu Ala			
	145	150	155
Glu Gly Thr Ala Ile Gln Asp Leu Phe Asn Pro Asp Arg Val Leu Ile			
	165	170	175
Gly Gly Arg Glu Thr Pro Glu Gly Asn Lys Ala Val Gln Ala Leu Lys			
	180	185	190
Glu Val Tyr Ala His Trp Val Pro Glu Glu Arg Ile Leu Thr Thr Asn			
	195	200	205
Leu Trp Ser Ala Glu Leu Ser Lys Leu Ala Ala Asn Ala Phe Leu Ala			
	210	215	220
Gln Arg Ile Ser Ser Val Asn Ala Met Ser Ala Leu Cys Glu Ala Thr			
	225	230	235
Gly Ala Asp Val Thr Gln Val Ser Tyr Ala Val Gly Lys Asp Thr Arg			
	245	250	255
Ile Gly Pro Lys Phe Leu Asn Ala Ser Val Gly Phe Gly Gly Ser Cys			
	260	265	270
Phe Gln Lys Asp Ile Leu Asn Leu Val Tyr Ile Cys Glu Cys Asn Gly			
	275	280	285
Leu Pro Glu Val Ala Glu Tyr Trp Lys Gln Val Ile Lys Ile Asn Asp			
	290	295	300

[0008]

Tyr Gln Lys Asn Arg Phe Val Asn Arg Val Val Ser Ser Met Phe Asn  
 305 310 315 320  
 Thr Val Ser Asn Lys Lys Ile Ala Ile Leu Gly Phe Ala Phe Lys Lys  
 325 330 335  
 Asp Thr Gly Asp Thr Arg Glu Thr Pro Ala Ile Asp Val Cys Lys Gly  
 340 345 350  
 Leu Leu Gly Asp Lys Ala Arg Leu Ser Ile Tyr Asp Pro Gln Val Thr  
 355 360 365  
 Glu Asp Gln Val Gln Arg Asp Leu Ser Met Asn Lys Phe Asp Trp Asp  
 370 375 380  
 His Pro Leu His Leu Gln Pro Met Ser Pro Thr Thr Val Lys Gln Val  
 385 390 395 400  
 Ser Cys Val Trp Asp Ala Tyr Glu Ala Thr Lys Asp Ala His Gly Ile  
 405 410 415  
 Cys Ile Leu Thr Glu Trp Asp Glu Phe Lys Lys Leu Asp Phe Lys Arg  
 420 425 430  
 Ile Tyr Asp Asn Met Gln Lys Pro Ala Phe Val Phe Asp Gly Arg Asn  
 435 440 445  
 Ile Val Asn Ala Asp Glu Leu Arg Glu Ile Gly Phe Ile Val Tyr Ser  
 450 455 460  
 Ile Gly Lys Pro Leu Asp Pro Trp Leu Lys Asp Met Pro Ala Val Ala  
 465 470 475 480

<210> 3

<211> 1443

<212> DNA

<213> 人工序列

<400> 3

[0009]

atggtgaaga tctgttgcac tggagctggt tatgttggtg ggcctacaat ggcgtgtgatt	60
gcacttaaat gcccatctat tgaagtggct gttgtggata tetctgtgce tegtattaat	120
gcatggaata gtgaccagct tccaatctat gagectggte ttgatgctgt tgtgaageag	180
tgctgtggaa agaacccttt ctteagcact gatgtagaaa aacatgtgcg tgaggctgat	240
atagctctttg ttcccgtaaa cactccgacc aagactcgag ggcctggage tgcaagget	300
gcagatotta catattggga aagtgtgct ogaatgattg ctgatgtatc aaaatctgat	360
aagatigtgg tcgagaaatc aacagtcctt gtcaaaactg ctgaggcaat cgaaaagatt	420
ctcaccocaa atagcaaggg aatcaatttc caaattcttt cgaaccocaa gttctctgce	480
gagggaaactg ctattcaaga tctttttaac cctgaccgtg ttctcattgg agtctgggag	540
accccagaag gcaacaaage tgtccaagca ttgaaggatg ttatgtctca ttgggtaccg	600
gaagaccgaa ttctcaccac taatctttgg tctgcagaac ttccaagct tctgtccaat	660
gctttcttgg ctccagcgat ctctctctga aacgccatgt ctgcacttg tgaggcgact	720
gggctgatg ttacacaagt gtcgtatgct gttggttaagg atacaaggat tggacccaaa	780
ttcttgaatg ccagtgctg ttctggcgga tctgtttcc aaaaggacat tctgaatct	840
gtttacatct gtgagtgcaa tggccttctt gaggtagctg agtattggaa acaagtcatt	900
aagatcaatg actatcagaa aaaccgattt gttaccgtg ttgtttctct gatgtttaac	960
accgtttcaa acaaaaagat tctatctta gggttcctt ttaagaagga taccggtgac	1020
actagggaga cccctctat cgtatgatge aaaggactct tgggagacaa agctcgattg	1080
agcatatatg acccgcaggt gaacgaagac caggtccaac gggatcttct catgaacaag	1140
tttgactggg accatctct tcacttgcaa cccatgagtc ccacgactgt caagcaagtc	1200
agcigtgtgt gggatgctta tgaggcaacc aaagatgcc atggaatttg cattctgacc	1260
gagtgggacg aattcaagaa acttgacttc aagaggatat acgacaacat gcaaaagccg	1320
gcttttgttt tcgatggaag gaacatcgtg aacgcggatg agctgagga gattggtttc	1380
atagtgact caattgglaa gccactggat ccatgctca aggacatgcc tgcagtggca	1440
taa	1443

<210> 4

<211> 480

<212> PRT

[0010]

<213> 人工序列

<400> 4

```

Met Val Lys Ile Cys Cys Ile Gly Ala Gly Tyr Val Gly Gly Pro Thr
1           5           10           15
Met Ala Val Ile Ala Leu Lys Cys Pro Ser Ile Glu Val Ala Val Val
           20           25           30
Asp Ile Ser Val Pro Arg Ile Asn Ala Trp Asn Ser Asp Gln Leu Pro
           35           40           45
Ile Tyr Glu Pro Gly Leu Asp Ala Val Val Lys Gln Cys Arg Gly Lys
           50           55           60
Asn Leu Phe Phe Ser Thr Asp Val Glu Lys His Val Arg Glu Ala Asp
65           70           75           80
Ile Val Phe Val Ser Val Asn Thr Pro Thr Lys Thr Arg Gly Leu Gly
           85           90           95
Ala Gly Lys Ala Ala Asp Leu Thr Tyr Trp Glu Ser Ala Ala Arg Met
           100          105          110
Ile Ala Asp Val Ser Lys Ser Asp Lys Ile Val Val Glu Lys Ser Thr
           115          120          125
Val Pro Val Lys Thr Ala Glu Ala Ile Glu Lys Ile Leu Thr His Asn
           130          135          140
Ser Lys Gly Ile Asn Phe Gln Ile Leu Ser Asn Pro Glu Phe Leu Ala
145          150          155          160
Glu Gly Thr Ala Ile Gln Asp Leu Phe Asn Pro Asp Arg Val Leu Ile
           165          170          175
Gly Gly Arg Glu Thr Pro Glu Gly Asn Lys Ala Val Gln Ala Leu Lys
           180          185          190
Asp Val Tyr Ala His Trp Val Pro Glu Asp Arg Ile Leu Thr Thr Asn
           195          200          205

```

[0011]

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

[0012]

435 440 445  
 Ile Val Asn Ala Asp Glu Leu Arg Glu Ile Gly Phe Ile Val Tyr Ser  
 450 455 460  
 Ile Gly Lys Pro Leu Asp Pro Trp Leu Lys Asp Met Pro Ala Val Ala  
 465 470 475 480

<210> 5

<211> 24

<212> DNA

<213> 人工序列

<400> 5

ggtaccatgg tgaagatctg ttgc 24

<210> 6

<211> 22

<212> DNA

<213> 人工序列

<400> 6

gtcgaattat gccactgcag gc 22

<210> 7

<211> 25

<212> DNA

<213> 人工序列

<400> 7

cagctccgag cccccgagtc ttggt 25

[0013]

<210> 8

<211> 25

<212> DNA

<213> 人工序列

<400> 8

tcaactgttcc aggcagcgat cctag 25

<210> 9

<211> 25

<212> DNA

<213> 人工序列

<400> 9

atatatgacc cgcaggtgac cgaag 25

<210> 10

<211> 25

<212> DNA

<213> 人工序列

<400> 10

gcaacccatg agtcccacga ctgtc 25

<210> 11

<211> 24

<212> DNA

<213> 人工序列

[0014]

<400> 11	
gctgagatga acaagaggtc attc	24
<210> 12	
<211> 25	
<212> DNA	
<213> 人工序列	
<400> 12	
ggaatcaata atcaaaacag cacag	25
<210> 13	
<211> 24	
<212> DNA	
<213> 人工序列	
<400> 13	
gtcgacatgg tgaagatctg ttgc	24
<210> 14	
<211> 21	
<212> DNA	
<213> 人工序列	
<400> 14	
ggtacctgcc actgcaggca t	21
<210> 15	

[0015]



<211>	24	
<212>	DNA	
<213>	人工序列	
<400>	15	
tctagaatgg tgaagatctg ttgc		24
<210>	16	
<211>	21	
<212>	DNA	
<213>	人工序列	
<400>	16	
gagctcttat gccactgcag g		21
<210>	17	
<211>	24	
<212>	DNA	
<213>	人工序列	
<400>	17	
gagctcatgg tgaagatctg ttgc		24
<210>	18	
<211>	24	
<212>	DNA	
<213>	人工序列	
<400>	18	
tctagattat gccactgcag gcat		24

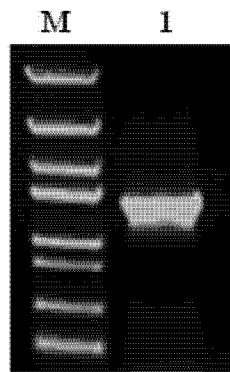


图 1

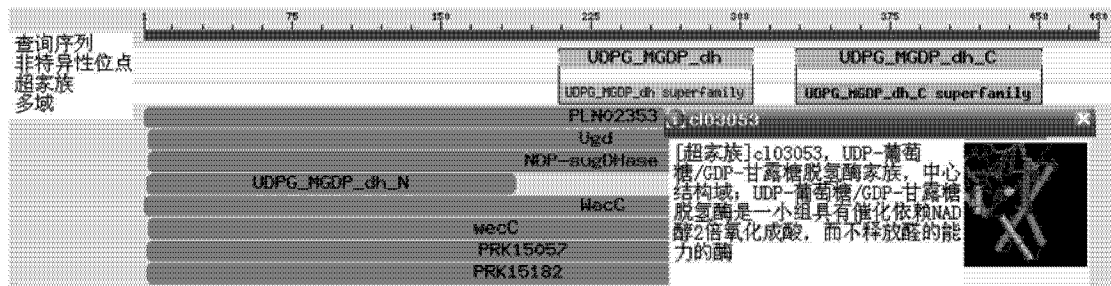


图 2

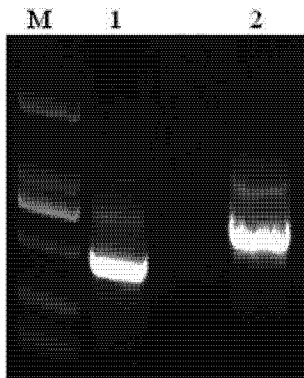


图 3

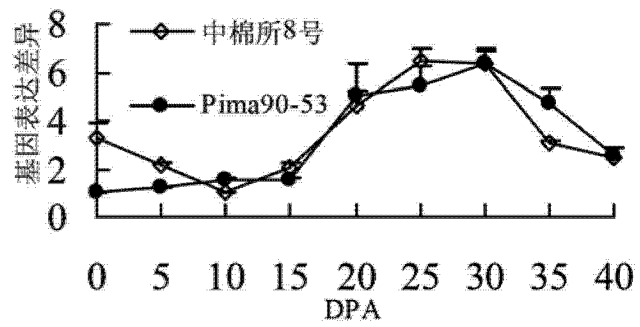


图 4

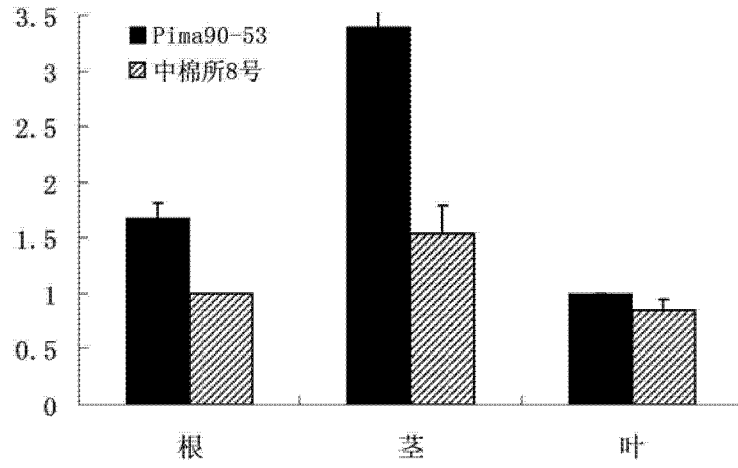


图 5

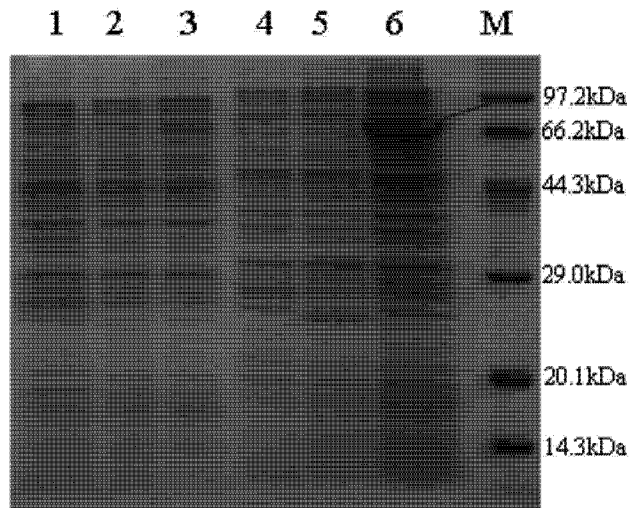


图 6

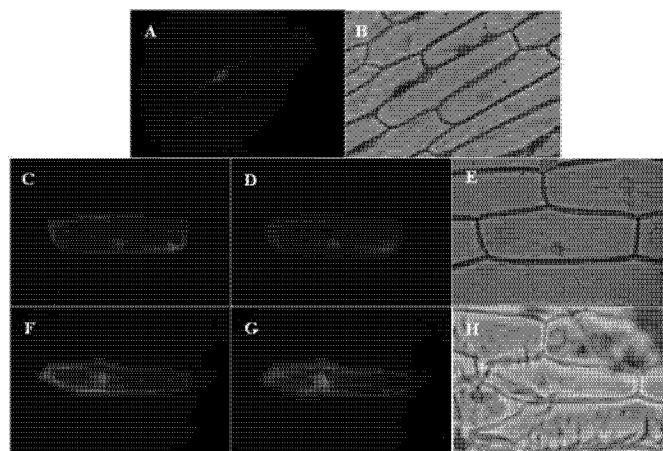


图 7

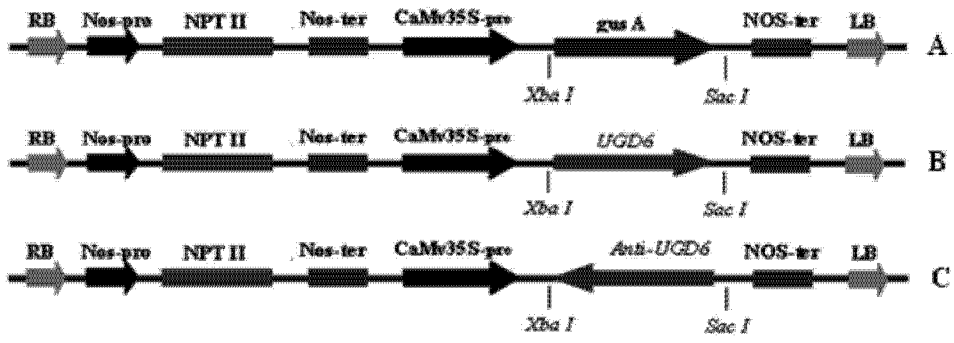


图 8

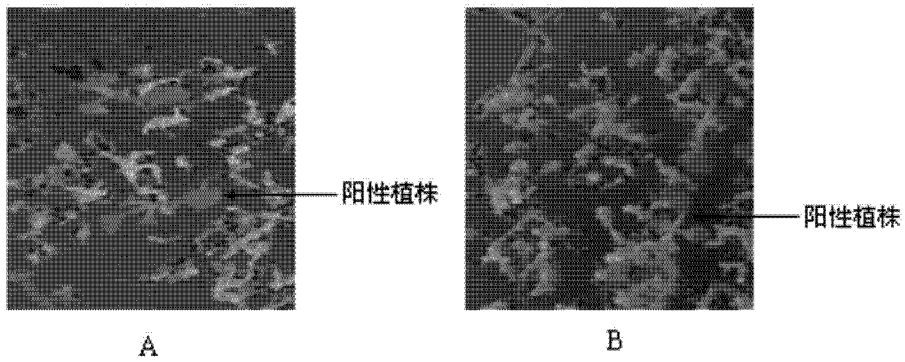


图 9

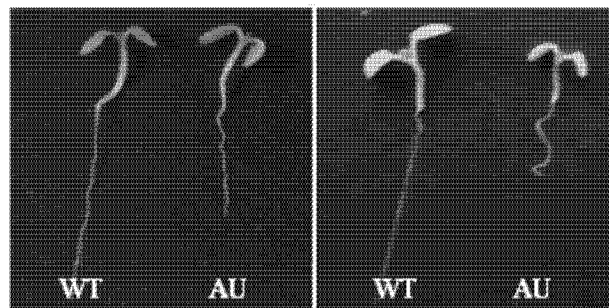


图 10



图 11

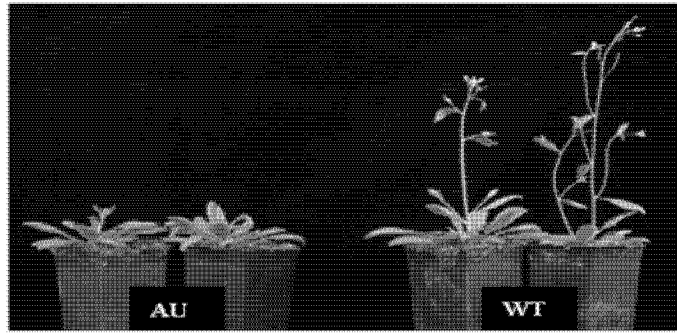


图 12

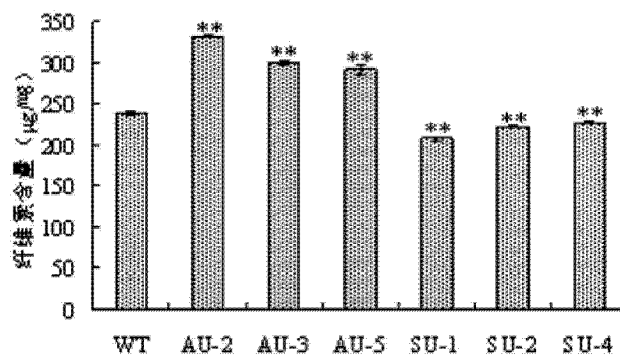


图 13

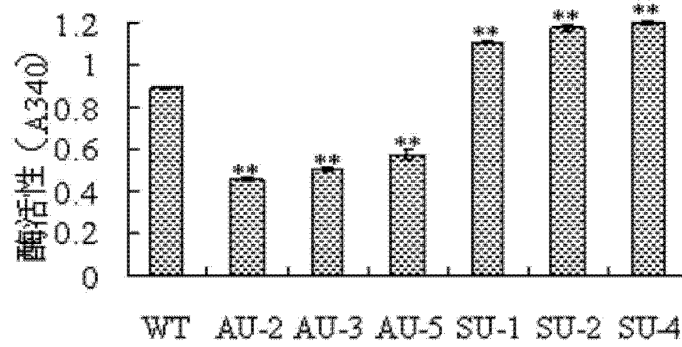


图 14