



(12) 发明专利申请

(10) 申请公布号 CN 101809149 A

(43) 申请公布日 2010.08.18

(21) 申请号 200880109102.X

C12N 15/82(2006.01)

(22) 申请日 2008.06.05

(30) 优先权数据

07113263.3 2007.07.26 EP

60/952,046 2007.07.26 US

(85) PCT申请进入国家阶段日

2010.03.26

(86) PCT申请的申请数据

PCT/NL2008/050352 2008.06.05

(87) PCT申请的公布数据

W02009/014430 EN 2009.01.29

(71) 申请人 西方种子国际有限公司

地址 荷兰纳尔德韦克

(72) 发明人 A·W·范赫于斯登 B·戈尔盖

(74) 专利代理机构 永新专利商标代理有限公司

72002

代理人 左路 林晓红

(51) Int. Cl.

C12N 9/24(2006.01)

权利要求书 2 页 说明书 22 页 序列表 5 页

附图 6 页

(54) 发明名称

新的多聚半乳糖醛酸酶和它们的用途

(57) 摘要

本发明涉及新的茄属多聚半乳糖醛酸酶的核苷酸序列。所述核苷酸序列可以用于标记辅助育种、TILLING 或转基因植物中以产生具有因不开裂花药所致的部位型不育表型的植物。

1. 核酸分子,其包含编码具有多聚半乳糖醛酸酶活性的多肽的核苷酸序列,其中所述核苷酸序列选自以下一组:

(a) 编码与 SEQ ID NO:2 的氨基酸序列具有至少 60% 序列相同性的氨基酸序列的核苷酸序列;

(b) 与 SEQ ID NO:1 的核苷酸序列具有至少 55% 序列相同性的核苷酸序列;

(c) 其互补链与 (a) 或 (b) 的核苷酸序列杂交的核苷酸序列;和

(d) 其序列因遗传密码简并性而不同于核苷酸序列 (c) 的序列的核苷酸序列。

2. 核酸分子,其包含来自 SEQ ID NO:1 的至少 26 个连续核苷酸的核苷酸序列。

3. 用于检测、分离、扩增和 / 或分析植物中 DPG 等位基因的方法,所述方法包括步骤:提供包含该植物核酸的样品并将该植物核酸与包含来自如权利要求 1 中所定义核苷酸序列的至少 10 个连续核苷酸的核苷酸序列的核酸分子杂交。

4. 权利要求 3 的方法,其中 DPG 等位基因是 ps-2- 等位基因。

5. 权利要求 4 的方法,其中 ps-2- 等位基因是以 C、A 或 T 作为 DPG 基因第五外显子 3' 端的最末核苷酸的等位基因,优选地该等位基因以 C 作为 DPG 基因第五外显子 3' 端的最末核苷酸。

6. 包含来自如权利要求 1 中所定义核苷酸序列的至少 10 个连续核苷酸的核苷酸序列的核酸分子在标记辅助育种中的用途。

7. 权利要求 6 的用途,其中标记辅助育种包括检测 ps-2- 等位基因。

8. 权利要求 7 的用途,其中 ps-2- 等位基因是以 C、A 或 T 作为 DPG 基因第五外显子 3' 端的最末核苷酸的等位基因,优选地该等位基因以 C 作为 DPG 基因第五外显子 3' 端的最末核苷酸。

9. 用于产生具有不开裂花药的植物的方法,其中该方法包括步骤:

a) 将第一植物与 ps-2- 等位基因纯合型的第二植物杂交;

b) 将 F1 代和后续世代与作为回归亲本的第一植物回交至少 2 代;并且,

c) 将 b) 中获得的最远回交世代自交至少一代;

其中在步骤 b) 和 c) 的至少一个步骤中使用分子标记以选择 ps-2- 等位基因纯合型的植物。

10. 权利要求 9 的方法,其中分子标记是对 DPG 等位基因特异的标记,所述标记存在于植物的基因组内部,距离如权利要求 1 中所定义的编码具有多聚半乳糖醛酸酶活性的多肽的核苷酸序列或其部分不超过 100kb。

11. 权利要求 10 的方法,其中所述分子标记是作为 DPG 基因第五外显子 3' 端的最末核苷酸的 C、A 或 T 或所述分子标记检测作为 DPG 基因第五外显子 3' 端的最末核苷酸的 C、A 或 T,其中最优选的是 C 作为 DPG 基因第五外显子 3' 端的最末核苷酸。

12. 用于产生在 DPG- 等位基因中具有突变的植物的方法,其中该方法包括步骤:

a) 诱变植物复组的种子;

b) 生长在 a) 中所获得诱变种子的植物;

c) 任选地,将 b) 中获得的植物回交至少一代;和

d) 对 b) 或 c) 中所获得植物筛选 DPG- 等位基因中突变的存在。

13. 权利要求 12 的方法,其中 DPG- 等位基因中的突变导致该等位基因成为 ps-2- 等位

基因。

14. 权利要求 13 的方法, 其中 ps-2- 等位基因是以 C、A 或 T 作为 DPG 基因第五外显子 3' 端的最末核苷酸的等位基因, 优选地该等位基因以 C 作为 DPG 基因第五外显子 3' 端的最末核苷酸。

15. 用于产生具有不开裂花药的转基因植物的方法, 其中该方法包括步骤: 用核酸构建体转化植物细胞, 所述核酸构建体至少包含如权利要求 1 中所定义的编码 DPG 的核苷酸序列的片段或其互补序列, 其中该核酸构建体在植物细胞中的存在降低 DPG 活性的表达至产生部位型不育性和不开裂花药的水平。

16. 权利要求 15 的方法, 其中所述核苷酸序列有效连接至用于在植物细胞中表达的启动子并且其中该核苷酸序列的表达通过 RNA 干扰降低 DPG 活性的表达。

17. 权利要求 15 的方法, 其中核酸构建体是用于同源重组的构建体并且其中核苷酸序列包含这样的突变, 所述突变降低 DPG 活性的表达至产生部位型不育性和不开裂花药的水平。

18. 核酸构建体, 其至少包含如权利要求 1 中所定义的编码 DPG 的核苷酸序列的片段或其互补序列, 所述片段或其互补序列有效连接至用于在植物细胞中表达的启动子。

19. 权利要求 18 的核酸构建体, 其中所述片段包含来自一核苷酸序列的 30 个连续核苷酸的序列, 所述核苷酸序列与如权利要求 1 中所定义的编码 DPG 的核苷酸序列或其互补序列具有至少 60% 序列相同性。

新的多聚半乳糖醛酸酶和它们的用途

发明领域

[0001] 本发明涉及植物育种领域,尤其番茄植物育种。本发明延伸至经典及分子植物遗传学领域并涉及新的多聚半乳糖醛酸酶的序列和它们在标记辅助育种中和转基因植物中的用途,例如产生具有因不开裂花药所致的部位型不育表型的植物。

[0002] 发明背景

[0003] 在高等植物中,成熟花粉通过开裂从花药释放,所述开裂包括绒毡层、隔膜和最终在裂口处依次发生的一系列细胞破坏。在绒毡层退化后,药室内壁细胞膨大并且木质化纤维带沉积,导致细胞壁增厚。随后,包围子囊腔的药室内壁和结缔组织细胞的脱水与皱缩在裂口处产生最终导致花粉释放的破坏力 (Keijzer 1987)。

[0004] 在理解花药开裂分子控制方面取得了最新进展,这主要涉及发现茉莉酸 (JA) 和乙烯的意义。在拟南芥属 (*Arabidopsis*) 中鉴定了影响花药内 JA 合成中各种步骤的几种突变体。观察到的表型导致花药开裂延迟 (Scott 等人,2004 综述)。Rieu 等人 (2003) 重点指出了乙烯信号传导在这种现象中的角色,他们观察到乙烯不敏感烟草植物的花药开裂延迟。

[0005] 多聚半乳糖醛酸酶类 (PG) 属于最大水解酶家族之一 (Torki 等人,2000 ;Markovic 和 Janecek,2001)。已经证实 PG 活性与广泛的植物发育程序相关 (Hadfield 和 Bennett 1998 综述),在所述活性中,已经在玉米、烟草、油菜籽油菜和拟南芥的花药开裂区内观察到 PG 的花药开裂活性 (Dubald 等人 1993 ;Sander 等人 2001)。然而,从未详细地研究它们在开裂过程中的作用。

[0006] 我们最近已经将引起番茄中不开裂花药的部位型不育性 -2 基因 (ps-2 基因) 精确标定 (Gorguet 等人 2006)。本发明的目的意在提供 ps-2 基因的核苷酸和氨基酸序列,以及其中这些序列用于标记辅助育种或用于转基因植物中,例如旨在产生具有因不开裂花药所致的部位型不育表型的植物和 / 或旨在改变果实成熟过程的方法。

[0007] 发明描述

[0008] 定义

[0009] 在本说明书中,除非另外指出,本文中所用的术语和定义是 (孟德尔) 遗传学中使用的术语和定义,对于所述术语和定义,参考 M.WStrickberger, *Genetics*, 第 2 版 (1976)、尤其第 113-122 和 164-177 页。如本文中所提及,“基因”一般地意指决定一种生物体 (即番茄植物) 的生物学特征的遗传性因子,“等位基因”是 (二倍体) 番茄植物内存在的成对基因中的单个基因。就此而言,可理解如本文中所用的术语 ps-2- 基因或 ps-2- 等位基因指能够产生或造成本发明部位型不育性表型的开裂多聚半乳糖醛酸酶 (DPG) 等位基因。如本文中所用的 ps-2- 基因或 ps-2- 等位基因因而可以涉及能够产生或造成本发明部位型不育性表型和 / 或不开裂花药的 (DPG-) 功能等位基因的任何丢失。DPG 基因的优选例子是本文所述的番茄开裂多聚半乳糖醛酸酶 (TDPG) 基因。ps-2- 基因或 ps-2- 等位基因的优选例子是本文所述的以 C 作为 TDPG 基因第五外显子 3' 端的最末核苷酸的特定番茄 ps-2 等位基因。

[0010] 多聚半乳糖醛酸酶 (EC 3.2.1.15) 在本文中理解为催化果胶酸和其他聚半乳糖醛酸中 1,4- α -D- 半乳糖醛酸键随机水解的酶。多聚半乳糖醛酸酶也称作果胶解聚酶或果胶酶。如先前所述 (Parenicova 等人, 1998, Eur. J. Biochem. 251 :72-80), 通过比色检测还原糖释放定量地测定多聚半乳糖醛酸酶活性, 因而 1 单位多聚半乳糖醛酸酶活性定义为能够从作为模式底物的多聚半乳糖醛酸 (Sigma) 在 30°C 于含有 0.25% (w/v) 底物浓度的 50mM 乙酸钠缓冲液 (pH 4.2) 中每分钟释放 1 μ mole 还原糖末端的酶量。

[0011] 当植物含有基因的相等等位基因时, 称此植物对该基因是“纯合型”的, 而当植物含有所述基因的两个不同等位基因时, 称此植物对该基因是“杂合型”的。大写字母的用途表示显性 (形式的) 基因, 而小写字母的用途表示隐性基因: “X, X” 因而代表基因或属性 X 的纯合子显性基因型; “X, x” 和 “x, X” 代表杂合子基因型; 并且 “x, x” 代表纯合子隐性基因型。如通常所知, 仅纯合子隐性基因型一般会提供相应的隐性表型 (即产生显示属性或性状 “x” 的植物), 而杂合子的和纯合子显性基因型一般会提供相应的显性表型 (即产生显示属性或性状 “X” 的植物), 除非其他基因和 / 或因子如多个等位基因、抑制基因、共显性等 (也) 在决定表型中发挥作用。

[0012] 作为一般原则, 通过杂交往往属于不同品系的两种不同亲本番茄植物获得杂交种子。使用本身已知的栽培技术和植物育种技术, 可以提供具有高度特性所需属性的此类杂种, 这使“设计”杂种即赋予杂交植物预定的可遗传特征成为可能。这通常借助适当选择两个亲本系 (的属性), 使其杂交以提供杂交种子而实现。这些亲本系通常是经过几个世代自我受精 (自我授粉) 所获得的近交系, 并且此类近交系通常会由育种者再次专门“设计”, 从而与另一个 - 通常预定的 - 近交亲本系杂交时提供具有所需属性的杂交后代。通常, 此类亲本系在遗传学上是纯合型的和相同的 (即, 作为育种的结果), 从而它们可以以稳定和可靠的方式提供在遗传学上均一的 - 即便是杂合子 - 的杂交系组合, 后者可以联合亲本系属性。这样做的目的在一方面是将源自亲本系的某些属性尽可能纯粹地杂交到种子中, 而在另一方面是利用已知的杂种优势或近交生长效应, 所述效应可以在植物和果实生长方面并因而在产量方面提供改良的属性等等。当或因为所用亲本系就某些遗传属性方面不相关 (即当亲本系在遗传学上“隔离存在”) 时, 获得这种杂种优势效应。对于使用经典栽培技术 (包括形成杂种) 的一般植物 (且尤其是番茄) 育种技术的其他描述, 参考已知的手册, 通过引用的方式将所述手册的内容并入本文。

[0013] 如本文中所用, 术语“植物”包括完整植物或其任意部分或衍生物, 如植物细胞、植物原生质体、从中可以再生番茄植物的植物细胞组织培养物、植物愈伤组织、植物团块和在植物或植物部分中完好的植物细胞如胚、花粉、胚珠、果实 (收获的番茄)、花、叶、种子、根、根尖等。

[0014] 植物学术语: 林奈被视为植物学分类之父。虽然他首次将现代番茄划分为茄属 (*Solanum*) 植物, 不过现代番茄的科学名多年来是番茄 (*Lycopersicon esculentum*)。类似地, 现代番茄的野生近亲在番茄属 (*Lycopersicon*) 内部分类, 如潘那利番茄 (*L. pennellii*)、多毛番茄 (*L. hirsutum*)、秘鲁番茄 (*L. peruvianum*)、智利番茄 (*L. chilense*)、小花番茄 (*L. parviflorum*)、克梅留斯基番茄 (*L. chmeilewskii*)、契斯曼尼番茄 (*L. cheesmanii*)、樱桃番茄 (*L. cerasiforme*) 和醋栗番茄 (*L. pimpinellifolium*)。新提出的现代番茄科学名是番茄 (*Solanum lycopersicum*)。类似地, 可以改变野生物种

的名称,潘那利番茄可以变成 *Solanum pennellii*,多毛番茄可以变成 *S. habrochaites*,秘鲁番茄可以细分成 *S. N. peruvianum* 和 *S. Callejon de Huayles*、*S. peruvianum* 和 *S. corneliomuelleri*,小花番茄可以变成 *S. neorickii*,克梅留斯基番茄可以变成 *S. chmielewskii*,智利番茄可以变成 *S. chilense*,契斯曼尼番茄可以变成 *S. cheesmaniae* 或 *S. galapagense* 并且醋栗番茄可以变成 *S. pimpinellifolium* (茄科 (*Solanaceae*) 基因组网络 (2005) Spooner 和 Knapp; http://www.sgn.cornell.edu/help/about/solanum_nomenclature.html)。

[0015] 包含 ps-2 或 DPG 基因和等位基因的核酸序列或片段也可以通过它们与 SEQ ID NO:1 “杂交”、优选在中等严格杂交条件下或更优选在严格杂交条件下与之杂交的能力而定义。严格杂交条件在本文中定义为这样的条件,其允许至少约 25 个、优选约 50 个核苷酸、75 或 100 个并最优选约 200 个或更多个核苷酸的核酸序列在约 65°C 温度在包含约 1M 盐的溶液、优选 6×SSC 或具有相当的离子强度的任何其他溶液中杂交并在 65°C 于包含约 0.1M 或更少盐的溶液、优选 0.2×SSC 或具有相当的离子强度的任何其他溶液中洗涤。优选地,杂交过夜进行,即至少持续 10 小时,并洗涤优选地进行至少 1 小时,同时至少更换 2 次洗涤液。这些条件通常将引起具有约 90% 或更高序列相同性的序列的特异性杂交。中等条件在本文中定义为这样的条件,其允许至少 50 个核苷酸、优选约 200 个或更多个核苷酸的核酸序列在约 45°C 温度在包含约 1M 盐的溶液、优选 6×SSC 或具有相当的离子强度的任何其他溶液中杂交并在室温于包含约 1M 盐的溶液、优选 6×SSC 或具有相当的离子强度的任何其他溶液中洗涤。优选地,杂交过夜进行,即至少持续 10 小时,并洗涤优选地进行至少 1 小时,同时至少更换 2 次洗涤液。这些条件通常将引起具有至多到 50% 序列相同性的序列的特异性杂交。本领域技术人员将能够调节这些杂交条件以特异性鉴定相同性在 50% 与 90% 之间变动的序列。

[0016] 发明详述

[0017] 我们最近已经将引起番茄中不开裂花药的部位型不育性 -2 基因精确标定 (Gorguet 等人 2006)。我们现在已经通过定位克隆分离出 ps-2 基因。ps-2 基因的后续表征表明该野生型基因编码一种新的多聚半乳糖醛酸酶 (PG) 并且基因编码序列中的单一突变是造成因不开裂花药所致的部位型不育性这种表型的原因。我们发现这个突变影响该基因的内含子剪接识别位点之一,从而产生缺少一个外显子的异常 mRNA。我们还发现这个新 PG 基因 (此后命名为开裂 PG (DPG)) 也在成熟果实中表达。茉莉酸和乙烯在控制 DPG 表达中发挥作用并且 DPG 参与果实成熟过程。

[0018] 因而在第一方面,本发明涉及编码具有多聚半乳糖醛酸酶活性的多肽的核苷酸序列的核酸分子、优选分离的核酸分子。多聚半乳糖醛酸酶活性和这种活性可以如本文以上所定义那样测定。编码具有多聚半乳糖醛酸酶活性的多肽的核苷酸序列可以选自以下核苷酸序列组成的组:(a) 编码与 SEQ ID NO:2 的氨基酸序列具有至少 60、70、80、90、95、98、99 或 100% 序列相同性的氨基酸序列的核苷酸序列;(b) 与 SEQ ID NO:1 的核苷酸序列具有至少 55、60、70、80、90、95、98、99 或 100% 序列相同性的核苷酸序列;(c) 其互补链与 (a) 或 (b) 的核苷酸序列杂交的核苷酸序列;和 (d) 其序列因遗传密码简并性而不同于核苷酸序列 (c) 的序列的核苷酸序列。

[0019] 本发明的优选核苷酸序列来自茄属内部的物种。更优选地,该核苷酸序列

来自番茄复组 (complex) 内部的物种, 包括例如番茄、克梅留斯基番茄、多毛番茄 (*S. habrochaites*)、醋栗番茄、小花番茄 (*S. neorickii*) 和潘那利番茄 (*S. pennellii*)。或者, 该核苷酸序列可以来自茄子 (*Solanum melongena*)。虽然本发明的优选核苷酸序列编码具有多聚半乳糖醛酸酶活性的多肽, 然而本发明中明确地包括了包含一个或多个核苷酸替换、插入和缺失中的至少一种情况并且从其中不会表达出活性多聚半乳糖醛酸酶的等位基因, 包括该核苷酸序列的工程化和诱变形式。

[0020] 本发明的又一优选核酸分子是这样的分子, 其包含编码具有多聚半乳糖醛酸酶活性的多肽的核苷酸序列或其无活性等位基因的一部分。该核酸分子优选地包含来自编码具有多聚半乳糖醛酸酶活性的多肽的核苷酸序列或其无活性等位基因或更优选来自 SEQ ID NO :1 的至少 10、12、14、16、18、20、22、24、26、28、30、35、40 或 50 个连续核苷酸。可理解该核酸分子可以包含多于一个节段的此类连续核苷酸, 所述节段可以直接地连接在一起或被其他序列隔开。

[0021] 在第二方面, 本发明用于检测、分离、扩增和 / 或分析植物中、优选茄属植物中、优选番茄复组的植物中或茄子植物中 DPG 等位基因的方法。该方法优选地至少包括步骤: 提供包含该植物核酸的样品并将植物核酸与包含来自编码具有多聚半乳糖醛酸酶活性的多肽的核苷酸序列中至少 10 个连续核苷酸的核苷酸序列的核酸分子杂交, 所述核苷酸序列可以选自以下核苷酸序列组成的组: (a) 编码与 SEQ ID NO :2 的氨基酸序列具有至少 60、70、80、90、95、98、99 或 100% 序列相同性的氨基酸序列的核苷酸序列; (b) 与 SEQ ID NO :1 的核苷酸序列具有至少 55、60、70、80、90、95、98、99 或 100% 序列相同性的核苷酸序列; (c) 其互补链与 (a) 或 (b) 的核苷酸序列杂交的核苷酸序列; 和 (d) 其序列因遗传密码简并性而不同于核苷酸序列 (c) 的序列的核苷酸序列。可理解至少 10 个连续核苷酸的核苷酸序列也可以从如 (a)–(d) 中所定义核苷酸序列的包含一个或多个核苷酸替换、插入和缺失中的至少一种情况并且从其中不会表达出活性多聚半乳糖醛酸酶的等位基因取得, 所述的等位基因包括该核苷酸序列的工程化和诱变形式。更优选地, 用于与取样的植物核酸杂交的核酸分子包含来自如上文所定义核苷酸序列的至少 12、14、16、18、20、22、24、26、28、30、35、40 或 50 个连续核苷酸的核苷酸序列。可理解在用于的核酸分子是单链分子的情况下, 该核酸分子也可以包含互补序列, 如上文所定义核苷酸序列的相对链。

[0022] 用于检测、分离、扩增和 / 或分析植物中、优选茄属植物中特定核酸序列如 DPG 等位基因的方法依赖于包含预定序列的核酸分子 (即用于杂交的核酸分子) 与样品中待检测、分离、扩增和 / 或分析的核酸的序列特异性杂交。在该方法中, 用于杂交的核酸分子因而可以是能够与 DPG 等位基因、优选茄属 DPG 等位基因杂交的任意核酸分子。该分子可以是探针, 例如杂交探针, 或可以是延伸或扩增反应中由聚合酶待延伸的引物。

[0023] 通常, 用于检测、分离、扩增和 / 或分析特定核苷酸序列的方法是本领域熟知的, 其中所述方法依赖于包含预定序列的核酸分子的序列特异性杂交 (见例如 Sambrook 和 Russell, 2001, “Molecular Cloning :A Laboratory Manual”, 3 版, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, New York)。更具体地, 用于检测、分离、扩增和 / 或分析本发明特定 DPG 等位基因序列的方法可以包括 PCR 扩增 (美国专利号 4,683,195; 和 4,683,202; 《PCR 技术:DNA 扩增原理和应用》(PCR Technology : Principles and Applications for DNA Amplification), H.A.Erlich 编 著, Freeman

Press, NY, N. Y., 1992)、等位基因-特异性 PCR(Gibbs, 1989, Nucleic Acid Res. 17: 12427-2448)、等位基因特异性寡核苷酸筛选方法(ASO; Salki 等人, 1986, Nature 324: 163-166)、连接酶介导的等位基因检测方法如寡核苷酸连接测定法(OLA; 1988, Landegren 等人, Science 241:107-1080)或连接扩增反应(Wu 等人, 1989, Genomics 4:560-569; Barany, 1990, Proc. Nat. Acad. Sci. 88:189-193)、变性梯度凝胶电泳(在 1992 年 Erlich 编著《PCR 技术:DNA 扩增原理和应用》(PCR Technology:Principles and Applications for DNA Amplification)的第 7 章, W. H. Freeman and Co., New York)、温度梯度凝胶电泳(TGGE)、单链构象多态性分析(Orita 等人, 1989, Proc. Nat. Acad. Sci. 85:2766-2770)、错配碱基对差异性化学切除法(Grompe 等人, 1991, Am. J. Hum. Genet. 48:212-222)、错配碱基对酶切除法(Nelson 等人, 1993, Nature Genetics 4:11-18)、系统如 TaqMan™(Perkin Elmer)、包括依赖于荧光催化性释放的等温扩增的入侵者测定法(Invader Assay)(在 www.twt.com 的 ThirdWave Technology)、(例如在 GS 20 基因组测序仪上进行的 Roche/454LifeSciences 的)巨量平行测序以产生数千个体中给定基因座的序列信息, 其任选地与 3D 汇合策略联合(见例如在 www.keygene.com 的 Keypoint™ 技术)等。

[0024] 本发明的优选方法是用于检测、分离、扩增和/或分析 ps-2- 等位基因的方法, 其中 ps-2- 等位基因优选地是以 C、A 或 T 作为 DPG 基因第五外显子 3' 端的最末核苷酸的等位基因, 优选地该等位基因以 C 作为 DPG 基因第五外显子 3' 端的最末核苷酸。该位置对应于 SEQ ID NO:1 第 3772 位置处的 G。或者, 可以开发针对 DPG 等位基特异的分子标记。“分子标记”在本文中理解为意指核酸序列或一组核酸序列, 其中所述核酸序列(直接或间接地)指示特定等位基因(例如, 如本文所定义 DPG 等位基因或 ps-2 等位基因)的存在或不存在。该分子标记的存在或不存在可以在类型广泛的分子测定法或检验法中检测到, 包括例如限制性片段长度多态性(RFLP)、随机扩增多态性 DNA(RAPD)、随机引发聚合酶链反应(AP-PCR)、DNA 扩增指纹法(DAF)、序列表征扩增区域(SCAR), 和扩增片段长度多态性(AFLP)。DPG 等位基因特异的分子标记在本文中理解为这样的标记, 所述标记在植物、优选茄属植物的基因组内部距离如本文以上所定义的编码具有多聚半乳糖醛酸酶活性的多肽的核苷酸序列或其部分不超过 100、50、20、10、5 或 2kb。此类分子标记可以用于检测特定 DPG 等位基因, 包括 ps-2- 等位基因。

[0025] 在第三方面, 本发明涉及如本文以上定义的用于杂交的核酸分子在标记辅助育种中的用途。优选地, 标记辅助育种包括检测 ps-2- 等位基因。ps-2- 等位基因在本文中理解为植物、优选茄属(Solanum)植物、更优选番茄复组植物中或茄子植物中的 DPG 等位基因, 其中所述 DPG 等位基因纯合型时产生因不开裂花药所致的部位型不育性表型。ps-2- 等位基因通常将是这样的等位基因, 其中在对于所述等位基因纯合型的植物内从所述等位基因表达不充分的 DPG 活性以防止部位型不育性和不开裂的花药。更优选地, 标记辅助育种包括检测以 C、A 或 T 作为 DPG 基因第五外显子 3' 端的最末核苷酸, 最优选 C 作为 DPG 基因第五外显子 3' 端的最末核苷酸的 ps-2- 等位基因。

[0026] 在第四方面, 本发明涉及用于产生具有不开裂花药的植物的方法。优选地, 该植物是茄属植物、更优选地是番茄复组植物或茄子植物。该植物优选地具有部位型不育性表型。该方法优选地包括步骤:(a) 将第一植物与 ps-2- 等位基因纯合型的第二植物杂交;(b) 将 F1 代和后续世代与作为回归亲本的第一植物回交至少一代(优选至少 2 代);并且, (c) 将

b) 中获得的最远回交世代自交至少一代(优选至少 2 代);其中在步骤 b) 和 c) 的至少一个步骤中使用分子标记以选择 ps-2- 等位基因纯合型的植物。优选地在该方法中,分子标记是对如本文定义的 DPG 等位基因特异的标记,该标记更优选是对如本文定义的 ps-2- 等位基因特异的。最优选地,分子标记是或检测作为 DPG 基因第五外显子 3' 端的最末核苷酸的 C、A 或 T,最优选 C 作为 DPG 基因第五外显子 3' 端的所述最末核苷酸。

[0027] 在第五方面,本发明涉及用于产生在 DPG- 等位基因中具有突变或遗传修饰的植物的方法。优选地,该植物是茄属植物、更优选地是番茄重组植物或茄子植物。突变或遗传修饰可以使用 TILLING 技术(定向诱导基因组中的局部损伤)导入编码 DPG 的核苷酸序列/基因的基因座中。用于 TILLING 的方法是本领域熟知的(McCallum 等人,2000 Nat Biotechnol. 18(4):455-7;Stemple,2004, Nat Rev Genet. 5(2):145-50 综述)。TILLING 诱变技术用来产生和/或鉴定以及用来逐渐分离 DPG 基因(包括 ps-2- 等位基因)的诱变变体。TILLING 也允许选择携带此类突变变体的植物。TILLING 联合了高密度诱变法与高通量筛选法。在 TILLING 中一般遵循的步骤是:(a) 诱变,例如甲磺酸乙酯(EMS)或 N-甲基-N-亚硝脒(ENU);(b) 从诱变的种子培育植物和任选地将该植物回交至少一代;(c) 从该植物的组织样品制备 DNA 和任选地,从个体汇集 DNA;(d) 例如使用本文以上定义的引物,PCR 扩增目的区域,即 DPG 基因;(e) 检测突变 PCR 产物;(f) 鉴定突变个体;和(h) 任选地对突变 PCR 产物测序。最初在该方法中,通过变性和复性以导致异双链体形成和随后通过 DHPLC(变性高效液相色谱)检测异双链体来开展步骤(e)中和任选(f)中突变 PCR 产物的检测(上文 McCallum 等人,2000)。通过利用与鉴定突变的基于凝胶的系统组合的限制酶 Ccl-I 使该方法具有更大的高通量(Colbert 等人,2001, Plant Physiol. 126(2):480-4)。用于检测或鉴定可以用于 TILLING 方法内的靶 DPG 基因中单碱基突变的其他方法例如包括如已经描述的 DNA 再测序(见例如 Slade 等人,2005, Nat Biotechnol. 23(1):75-81)或(例如在 GS 20 基因组测序仪上进行的 Roche/454 Life Sciences 的)巨量平行测序以产生数千个体中给定基因座的序列信息,其任选地与 3D 汇合策略联合(见例如在 www.keygene.com 的 Keypoint™ 技术)。诱变可以通过辐射或用化学诱变剂如 EMS(Lightner 和 Caspar, 1998, "Seed mutagenesis of Arabidopsis" 在: 方 Molecular Biology: Arabidopsis Protocols(J. M. Martinez-Zapater 和 J. Salinas 编著), 第 91-103 页. Totowa, New Jersey: Humana Press) 或 ENU(Draper 等人,2004, Methods Cell Biol. 2004;77:91-112) 进行。用于产生在 DPG- 等位基因中具有突变的植物的优选方法优选地至少包括步骤:(a) 诱变植物的种子;(b) 培育 a) 中所获得的诱变种子的植物;(c) 任选地将 b) 中获得的植物回交至少一代;并且对 b) 或 c) 中获得的植物筛选 DPG- 等位基因中突变的存在。优选地,该植物是茄属植物、更优选地是番茄重组植物或茄子植物。优选地, DPG- 等位基因中的突变导致该等位基因成为 ps-2- 等位基因。更优选地,该 ps-2- 等位基因是以 C、A 或 T 作为 DPG 基因第五外显子 3' 端的最末核苷酸的等位基因,最优选地该等位基因以 C 作为 DPG 基因第五外显子 3' 端的最末核苷酸。

[0028] 在第六方面,本发明涉及用于产生具有不开裂花药的转基因植物、优选茄属植物的方法。具有不开裂花药的转基因植物可以以各种方法获得,所述方法包括例如通过将反义、有义阻抑或 RNAi(RNA 干扰)核酸构建体导入植物来沉默 DPG 基因或通过同源重组敲除 DPG 基因。该方法因而至少包括步骤:用包含编码如本文以上所定义 DPG 的核苷酸序列的至

少一个片段的核酸构建体转化植物细胞,其中该核酸构建体在植物细胞中的存在降低 DPG 活性的表达至产生部位型不育性和不开裂花药的水平。具有不开裂花药的植物可以通过技术人员本身已知的方法从转化的植物细胞或从包含此类转化细胞的植物衍生。

[0029] 因而在优选施方案中,该核酸构建体中的核苷酸序列有效连接至用于植物细胞(例如茄属植物细胞)中表达的启动子,并且该核苷酸序列的表达通过 RNA 干扰降低 DPG 活性的表达。用于基因沉默的方法和核酸构建体在 US20070130653、US20070074311 和其中引用的参考文献中描述。本发明因而涉及作为构建体的核酸分子,其中所述构建体包含编码如本文以上所定义 DPG 的核苷酸序列的至少片段,所述片段有效连接至用于在植物细胞中表达的启动子。在核酸构建体中,所述片段优选地包含自与编码如本文以上所定义 DPG 的核苷酸序列具有至少 60%、70%、80%、90%、95% 或 100% 序列相同性的核苷酸序列或其互补序列的 30、60、100、200 或 500 个连续核苷酸的序列,更优选地,编码 TDPG 的核苷酸序列是 SEQ ID NO :1。

[0030] 在另一个优选的施方案中,具有不开裂花药的转基因植物可以通过同源重组敲除 DPG 基因。在该方法中,核酸构建体构建是用于同源重组的构建体并且核苷酸序列包含这样的突变,其降低 DPG 活性的表达至产生部位型不育性和不开裂花药的水平。这种突变可以是最广泛意义上的突变,包括例如编码序列和 / 或启动子序列完全缺失或插入选择标记或其他序列以失活 DPG 基因。同源重组允许在基因组中限定选择的位置处导入选择的核酸。已经对模式植物 (*Offringa* 等, 1990 *EMBO J* 9(10) :3077-84) 和作物植物例如稻 (*Iida* 和 *Terada*, *Curr Opin Biotechnol.* 2004 年 4 月 ;15(2) :132-8) 描述了用于植物中开展同源重组的方法。待靶向的核酸因而优选地是用来替代内源性 DPG 基因的缺陷型 DPG 等位基因,例如 ps-2- 等位基因。

[0031] 具有不开裂花药和 / 或部位型不育表型的本发明植物具备优势,即通过避免或减少自我受精,它们导致更有经济地产生杂交植物。杂种的产生可能涉及利用雄性不育如 CMS、遗传不育性或部位型不育性。优选部位型不育性,原因是一方面它减少或防止不希望的自我受精,而在另一方面如果需要,它通过(手工)开启或破坏花药允许自我受精。本发明的 DPG 等位基因因而可以用作防止杂种产生期间近交的工具,所述的杂种可以或可以不具有因不开裂花药所致的部位型不育表型(这取决于父本的表型)。此外,不开裂的花药和 / 或部位型不育表型可以有利地在无种子果实的生产中用于避免或减少(自我)受精并因而改善无种子表型。

[0032] 虽然通过番茄植物例举阐述了本发明,可理解本发明包括 DPG 核酸、ps-2 等位基因和用于产生全部商业重要作物的(转基因)植物的方法,所述包括商业重要作物例如菊科 (*Asteraceae*) (包括粮食作物莴苣、菊芋、刺菜蓟 (*globe artichoke*)、葵花、菊薯 (*yacón*)、红花)、葫芦科 (*Cucurbitaceae*) (通常称作葫芦 (*gourd* 或 *cucurbit*) 并且包括作物如黄瓜、南瓜(包括笋瓜 (*pumpkin*))、丝瓜、甜瓜和西瓜)、芸苔属 (*Brassica*) (包括蕉青甘蓝、芜菁、球茎甘蓝、卷心菜、球芽甘蓝、花椰菜、青花椰菜、芥菜和油菜籽油菜)、豆科作物 (*Leguminous Crop*) (包括干菜豆、干蚕豆、干豌豆、鹰嘴豆 (*chickpea*)、鹰嘴豆 (*garbanzo*)、鹰嘴豆 (*bengal gram*)、干豇豆、中国豆 (*black-eyed pea*、*blackeye bean*)、木豆 (*pigeon pea*)、toor、树豆 (*cajan pea*)、congo bean、扁豆、班巴拉地豆、*earth pea vetch*、*common vetch* *Lupins*、大豆、落花生)、藜科 (*Chenopodiaceae*) (菠菜)、百合科

(Liliaceae) (洋葱、韭菜)、伞形科 (Apiaceae) (胡萝卜)、谷物作物 (稻、大麦、小麦、燕麦、玉米)、茄科 (Solanaceae) (包括番茄、辣椒、茄子)。

[0033] 在本文件及其权利要求中,以非限制性意义使用动词“包含”和其词形变化来意指包括该词后的成分,不过不排除没有具体提到的成分。此外,借助不定冠词“a”或“an”对某要素的称谓不排除多于一个所述要素存在的可能,除非上下文清楚地要求存在及仅存在一个所述要素。非限定冠词“a”或“an”因而通常意指“至少一个”。

[0034] 附图简述

[0035] 图 1 :ps-2ABL 和野生型 (Moneymaker) 的花药的显微镜观察结果。

[0036] a :用甲苯胺蓝染色的处于早花期前 (PA)、晚期花芽 (FB) 处的花药的横截面。箭头指可以在 WT 和 ps-2ABL 的隔膜细胞中见到的晶体。s :隔膜 ;st :裂口 ;p :花粉。

[0037] b :用甲苯胺蓝染色的处于花期的花药的横截面。箭头指示 WT 中花药的开口和 ps-2ABL 中仍关闭的裂口。

[0038] c :通过 SEM 观察到的处于花期的花药囊的纵剖面。箭头指示 WT 中花药的纵向开口和 ps-2ABL 中未打开的花药。

[0039] 图 2 :ps-2 基因的克隆,从遗传图至候选基因

[0040] a :重组 F2 种群中产生的遗传图 (ps-2ABL× 醋栗番茄 ;Gorguet 等人,2006)。每种标记之间的重组植物数为白色字。

[0041] b :物理图 :以虚线表示的箭头代表 BAC 的计算性定位。以实线表示的箭头代表通过分子标记定位 BAC。(S :SP6 ;T :T7)。

[0042] c :ORF 在 BAC 克隆 143M15 上的遗传学位置。

[0043] d :候选基因 ORF4 的结构。带罗马数字的有色圆柱形代表外显子。推定性 TATA 框和 PolyA 的位置由序列开始处的倒蓝色三角形和序列结尾处的绿色菱形显示。

[0044] 图 3 :在来自 ps-2ABL 和栽培品种 Moneymaker 的花期后花药囊的 RNA 上进行的 RT-PCR。

[0045] 图 4 :Moneymaker (WT) 和 ps-2ABL 中花药内 ORF4 外显子 IV 与 VI 之间的内含子剪接过程。

[0046] 图 5 :20 种被子植物 PG 氨基酸序列和 1 种裸子植物 PG 氨基酸序列的进化系统树。该进化图是 3296 个树形图 (使用树加权) 的 50% 简单多数规则共有区的结果。PG 序列划分成被鉴定作充分表征的进化枝 A 和 B 的两种主要进化枝。候选蛋白质 (TDPG) 是进化枝 B 的部分。

[0047] 图 6 :TFPG、ORF4 和 ps-2 (ORF4 突变体) 的 Pfam GH28 结构域的比对结果。彩色代码仅对 TFPG 和 ORF4 之间的比较有效。相同的氨基酸以反色显示。保守替换以灰色显示。如 Rao 等人 (1996) 的多聚半乳糖醛酸酶的 4 个保守结构域以粗体加下划线方式显示。

[0048] 图 7 :ORF4 的组织特异性表达。使用专用于定量 PCR 分析引物,将源自多种组织的 Moneymakers 的 cDNA 进行 PCR 扩增。使用 Moneymaker 的基因组 DNA 作为污染对照。

[0049] 1 :叶离区 ;2 :花离区 ;3 :花期花药 ;4 :成熟期果实。

[0050] 图 8 :激素控制其中 DPG 分离的番茄内花药开裂和果实成熟的简化模型。

实施例**[0051] 1. 实施例 1****[0052] 1.1 材料与方法****[0053] 1.1.1 植物材料**

[0054] 从 ps-2ABL (ps-2 突变纯合型的一种纯种高级育种系番茄) 和醋栗番茄之间的杂交发展而来的 F₂ 重组次级种群用于遗传作图。对 ps-2 分离的该种群由 ps-2 基因座区中的 146 个 F₂ 植物重组体组成 (Gorguet 等人, 2006)。

[0055] 另外 176 个 ABL 用来检验所鉴定的 SNP 与 ps-2 基因座之间的关联, 其中 8 个 ABL 是 ps-2/ps-2。

[0056] ps-2ABL 和栽培品种 Moneymaker 的花药用显微镜观察。

[0057] 1.1.2 显微镜观察

[0058] 植物材料在 4℃ 于含有 4% 多聚甲醛的 0.1M 磷酸盐缓冲液 pH 7.0 中固定 24 小时。如 Dornelas 等人 (2000) 中所述处理扫描电子显微镜样品和使用 Orion Framegrabber (Matrox Electronic System, Unterhaching, 德国) 获得数字图像。将光学显微镜的样品包埋在 Technovit 7100 (Heraeus Kulzer, Wehrheim, 德国) 中, 用甲苯胺蓝染色并封装在 Euparal (Chroma-Gesellschaft, Kongen, 德国) 中。

[0059] 1.1.3 BAC 文库筛选和重叠群构建

[0060] 我们使用从栽培品种号 Heinz 1706 的基因组 DNA 构建的番茄 HindIII BAC 文库。Heinz 文库是一个具有平均大小 114.5kb 的 15 基因组等同物 (Budiman 等人, 2000)。通过 PCR 扩增筛选该 BAC 文库, 首先在平板汇集物上并随后在个体进行。分离阳性 BAC 克隆的质粒 DNA 并进一步分析。

[0061] 阳性 BAC 克隆的 BAC 末端序列从 SGN 数据库 (Mueller 等人 2005) 获得。将 BAC 末端序列如 Gorguet 等人 (2006) 所述转变成 CAPS 或 dCAPS 标记。BAC 末端序列衍生的 PCR 标记的详细内容呈现于表 1 中。

[0062] 表 1: 用于遗传连锁图中的 PCR 标记

[0063]

标记名称	用途	引物序列	大小(bp)	限制酶
从 BAC 末端序列衍生的标记				
67F23-T	CAPS	Fw: CTACTCTTCCGCCATAACTG Rv: GATCCAAACGAACAAAAGTCA	599	HincII
67F23-S	CAPS	Fw: TCATTCCGTTTGCTGAATGAGA Rv: ATAACTTATATCACTCCCAATCA	413	DraI
69C22-T	dCAPS	Fw: TCTTTTCGATATTTTTCAGAACTAA Rv: TGAGATGTTTGCAATAACATTCT	200	DdeI
143M15-S	CAPS	Fw: CATCGAAGTAACAGAGATATTA Rv: CCATAGGGATTATGATGTGTA	369	MwoI
114C15-T	CAPS	Fw: GCACTGAAGAATGGATAGACTC Rv: GGAATTGACCAAAAAGGATAGC	457	MnII
118A17-T	CAPS	Fw: GGCATGGTGAAGTCCACATT	739	HaeIII

[0064]

		Rv: GTGTCACAGGTTTGGTTCAT		
15N23-T	CAPS	Fw: GGCAGATATCTGCAATACGT	576	TaqI
		Rv: ATCATGAACAGCAAAACAACCA		
从 BAC 143M15 序列衍生的标记				
ORF1	CAPS	Fw: CTGTATCTATGACGAGGAGA	625	HaeIII
		Rv: GATCCTGAAGCTGAAGCTT		
ORF2	CAPS	Fw: AATATTTTCAACTTTCAAATCTCTT	206	MnII
		Rv: ACGAAGGCATGATTGTCGTTA		
ORF3	CAPS	Fw: GTTGAACCTTATACCACTAGGA	937	NdeI
		Rv: GTGCGGTCTCATCAACTCAA		
ORF4	dCAPS	Fw: GAACACTTAGGTTAAAATATAGC	217	AluI
		Rv: CCTACTATCCTTCTTGTAATCT		
ORF5	CAPS	Fw: CTTAAAGGCACACTTAGATTCA	962	HpyCH4IV
		Rv: CTGAGAATTCTCTTGACTGCA		
ORF4(1)	SCAR	Fw: GCTTTATTTCATAGTAAATTCTGT	805/885	
		Rv: TCAGACAAATCATCGTATATTGA		
ORF4(2)	SCAR	Fw: TCCATTTGTAGTTTCATAAAGC	465 / 515	
		Rv: CCAAGCGGATAATTAATGTCA		
开发用于番茄内部 ps-2ABL 等位基因鉴定的标记				
ps-2 标记	CAPS	Fw: CAAATTGGATGAGAGTTTTGAA	695	HpyCH4IV
		Rv: CATTTTACAAGTGTAACAACCTTG		

[0065] 基本上如 Brugmans 等人 (2006) 所述, 使用 Hind III/TaqI 酶组合产生各个 BAC 克隆的指纹模式。

[0066] 1.1.4BAC DNA 测序与分析

[0067] 通过脉冲场凝胶电泳估计 BAC 克隆 143M15 的大小。由 Greenomics (荷兰) 通过鸟枪测序方法将 BAC 克隆 143M15 测序。如 Gorguet 等人 (2006) 所述, 从在 BAC 序列上所鉴定的候选基因衍生 PCR 标记。为鉴定推定性基因, 使用 SGN 的 TBLASTX 界面 (Mueller 等人, 2005), 显著性阈值 E-10, 将最终 BAC DNA 序列针对源自 SGN 的番茄 Unigene 数据库 (Mueller 等人, 2005) 和源自 TAIR 的拟南芥基因模式数据库 (<http://www.Arabidopsis.org> Huala 等人, 2001) 扫描。开发基于推定性 ORF 序列的 PCR 标记并如 Gorguet 等人 (2006) 所述对该重组种群筛选。这些 PCR 标记的细节在表 1 中给出。

[0068] 1.1.5 候选基因分析和系统发育分析

[0069] 使用产生约 900bp 产物的几个连续性重叠引物对, 扩增并测序 Moneymaker 和 the ps-2ABL 中的 ORF4 的完整基因组 DNA 序列以及分别含有启动子和基因终止子的上游及下游序列。Moneymaker 和 ps-2ABL 的所得 DNA 序列用 DNASTar 装配。使用 Softberry 基因搜索软件来鉴定候选基因的推定性外显子和内含子。

[0070] 选择具有已知组织表达的番茄多聚半乳糖醛酸酶蛋白质序列以及用该候选蛋白

质所得的蛋白质 BLAST 搜索最佳命中来实施系统发育分析。所选的蛋白质序列中,仅 Pfam 糖基化水解酶 28 结构域用于该分析。用 NCBI 的蛋白质 Blast 界面鉴定每个蛋白质序列的 Pfam GH28 结构域。使用 ClustalW 多重序列比对软件 (Higgins 等人,1994) 比对 Pfam GH28 结构域的所选氨基酸序列。

[0071] 使用 PAUP 软件包版本 4 (Swofford 2002) 来构建简单多数规则共有区进化系统,使用最大简约性 (1000 个自助重复和 250 个添加序列重复)。使用雪松 (Cedar) PG 蛋白质序列并将其定义为外群 (outgroup)。

[0072] 氨基酸序列和它们的蛋白质鉴定编号是:猕猴桃果 (AAC14453;Atkinson 和 Gardner,1993)、葡萄浆果 (AAK81876;Nunan 等人,2001)、大豆荚 (AAL30418;Christiansen 等人,2002)、梨果 (CAA54448;Lester 等人,1994)、苹果果实 (AAA74452;Atkinson, 1994)、桃果实 (BAC22688;Hiwasa 等人,2003)、拟南芥开裂区 ADPG1 (CAA05525;Sander 等人,2001)、油菜籽油菜开裂区 RDPG1 (CAA65072;Petersen 等人,1996)、油菜籽油菜荚 (CAA90272;Jenkins 等人,未发表)、芜菁长角果瓣开裂 (CAD21651;Rodriguez-Gacio 等人,2004)、灯笼椒果实 (BAE47457;Ogasawara S 和 Nakajima T,未发表)、番茄果实 TFPG (CAA32235;Bird 等人,1988)、番茄 pistil (AAC70951;Hong 和 Tucker,2000)、番茄离区 TAPG1,2,4,5 (AAC28903, AAB09575, AAB09576, AAC28906;Hong 和 Tucker,1998)、番茄受伤叶 (AAD17250;Bergey 等人,1999) 和番茄种子 (AAF61444;Sitrit 等人,1999)。

[0073] 1.1.6 总 RNA 分离、cDNA 合成和定量 PCR 分析

[0074] 使用 RNeasy 植物微量试剂盒 (Qiagen, Hilden, 德国) 分离总 RNA。每个 RNA 分离反应使用 50mg 和 100mg 之间的每种组织 (花药、果实、花离区和叶离区)。用 DNA 酶 I 处理 (Boehringer Mannheim), 每个样品仅 1 μ g 总 RNA 用于 cDNA 合成。使用随机六聚体作为引物并使用 Multiscribe™ 逆转录酶 (Applied Biosystems) 合成第一链 cDNA 模板。使用位于第一外显子上的始于起始密码子下游 47 个核苷酸的正向引物: TAGCTCCAAAGCTATCCACAT 和位于最末外显子上的止于终止密码子上游 100 个核苷酸的反向引物: TGGAGAATGTGAAATTGTTAGG, 扩增 ORF4 的准完整编码序列以研究内含子剪接过程。用标准 PCR 反应 (55°C 复性温度和 35 个循环) 扩增 ORF4 的准完整 CDS。

[0075] 使用 SYBR 绿 PCR 主混合物试剂盒 (Applied Biosystems), 在 iCyclerMyiQ 检测系统 (Bio-Rad) 中开展实时实验。引物序列是: 正向引物 5' -TTTGGCATTGCCATTGATA-3', 反向引物 5' -TGTGGTGTCCCAGAACAGA-3' (ORF4); 和正向引物 5' -ACCACTTTCCGATCTCCTCTC-3', 反向引物 5' -ACCAGCAAATCCAGCCTTCAC-3' (β -肌动蛋白)。通过采用公式 $2^{-\Delta Ct}$, 用 β -肌动蛋白内对照计算 ORF4 转录物水平的相对定量。用解链曲线验证 PCR 产物的纯度。所述反应在杜普洛进行。在缺少所添加的逆转录酶的情况下进行 PCR 对照以确保 RNA 样品避免 DNA 污染。

[0076] 2. 结果

[0077] 2.1ps-2 表型的显微镜观察

[0078] 制备野生型 (Moneymaker) 和突变体 (ps-2ABL) 的花药的横截切片并用甲苯胺蓝染色以鉴定突变体中花药发育 / 开裂被阻断的阶段。制备用于电子显微镜的处于花期的花药囊的纵向切片。未见发育阶段的差异直至花期。在晚期花芽处, 在突变体和野生型的隔膜细胞中观察到晶体; 并且花粉发育似乎正常 (图 1a)。突变体中隔膜的断裂在类似于野

生型的阶段上出现。然而,在花期,突变体裂口不退化并且花粉仍留在花药中(图1-b,c)。此外,药室内壁增厚没有在突变体中发生并且因而外皮细胞在花期缺少在裂口上产生破坏力的刚性。

[0079] 2.2 物理作图和候选基因鉴定

[0080] 我们先前已经将ps-2基因座标定至染色体4短臂上由COS衍生的CAPS标记T0958和T0635限定的1.65cM间隔(Gorguet等人,2006)。使用Heinz BAC文库建立ps-2基因座区域的物理图谱。用相对于ps-2基因座亲缘关系最近的标记,通过PCR扩增并使用那些标记的序列,借助计算工具筛选该文库。随后通过将BAC末端转变成PCR标记(图2)并通过在重组种群中筛选这些标记而把阳性BAC标定至该遗传图。也比较BAC指纹以估计源自相同重叠群的BAC的重叠情况并验证BAC是否作为相同重叠群的部分。相对于ps-2基因座的最接近BAC末端随后用于第二轮筛选。逐渐地,完整重叠群覆盖从COS衍生CAPS标记T1070至BAC末端15N23-T的1.70cM(32个重组体)(图2)。BAC 143M15覆盖ps-2基因座并因而进行测序。

[0081] 为了鉴定BAC克隆143M15中存在的基因,使用BLASTN界面,将BAC DNA序列再次针对番茄SGN Unigene数据库扫描,并使用BLASTN界面,针对拟南芥基因模式数据库扫描。2个番茄编码序列和5个拟南芥基因匹配该BAC克隆序列(其中的两个匹配所述两个番茄编码序列;表2)。这5个候选基因命名为ORF1至ORF5。另外,也鉴定到逆转座子家族的6个基因,不过没有在进一步的研究中考虑这6个基因。5个相应拟南芥基因的位置在拟南芥基因组中不是连续的,这凸显这两个物种之间针对该区域缺乏同线性(synteny)。另外,它们均不与(Gorguet等人,2006中列出的)拟南芥中所鉴定的功能雄性不育性基因之一同源。

[0082] 表2:鉴定基于拟南芥基因模式的候选基因

[0083]

基因	番茄	拟南芥		染色体号	AGI 坐标(碱基)	TBLASTX	
	Unigenes 登录号	登录号	基因功能			E 值	评分
ORF1	U323899	AF014399	镁螯合酶	1	2696415-2700961	0	214
ORF2		NM112785	转录因子 B3 家族	3	6548875-6551847	3E-21	105
ORF3	U317249	AF326883	Remorin 家族蛋白	2	17477944-17480014	9E-37	110
ORF4		NM111676	多聚半乳糖醛酸酶	3	2541012-2543438	1E-79	108
ORF5		AB017502	糖基解酶家族 3	5	7107378-7111311	0	276

[0084] 为了定位高分辨率图中的候选基因,我们将推定性基因序列(或附近序列)转换成PCR标记并将它们在重组种群中作图。在高分辨率连锁图中的不同位置处标出每个候选基因,且因而我们可以基于遗传图中它们的位置轻易鉴定最可能的ps-2候选物。推定性多聚半乳糖醛酸酶基因ORF4的位置与ps-2基因座最接近(图2)。随后,使用Softberry的FGENSH软件鉴定推定性内含子和外显子。候选基因ORF4由9个外显子和8个内含子组成,覆盖从推定性起始密码子至终止密码子的6716个核苷酸的基因组距离,对应于1179个核苷酸的编码序列。用来开发PCR标记ORF4的SNP位于第二推定内含子内。我们开发了两个额外PCR标记并将它们作图,其中所述两个标记之一基于醋栗番茄等位基因[ORF4(1)]

第一内含子中的 76bp 缺失,并且另一个标记基于醋栗番茄等位基因 [ORF4(2)] 第六内含子中的 38bp 插入。3 个 PCR 标记 ORF4(1)、ORF4 和 ORF4(2) 标定在两个标记之间的一个重组体的间隔区处,这表明至少两个重组已经在重组种群中的候选基因 ORF4 内部出现(图 2)。ORF4(2) 随高分辨率图上的 ps-2 基因座共分离。

[0085] 2.3ORF4 中的突变和分子标记开发

[0086] 为证实 ORF4 对应于 ps-2 基因的假设,我们检索了 ps-2ABL 等位基因中的序列变异。对栽培品种 Moneymaker 和 ps-2ABL 测定从 ORF4 的推定性启动子至推定性终止子的完整 9kb 序列。在 ORF4 的第五推定性外显子的最末核苷酸处鉴定到一个单一突变,其中鸟嘌呤核苷酸被胞嘧啶替换。为检验 ORF4 中鉴定的 SNP 与 ps-2 性状之间的关系,我们以如此方式开发了基于该 SNP 的分子标记,从而可能在凝胶上轻易区分番茄中的 ps-2ABL 等位基因和野生型等位基因。对其中 176 个 ABL 群体检验该标记,所述群体当中 8 个 ABL 是 ps-2/ps-2。这 7 个功能性不育植物显示与其他 ABL 明显不同的相同标记模式,这证实 SNP 存在于所检验的其他 ps-2 品系中。这个标记现在可以轻易用于分子辅助导入 ps-2 性状至现代番茄品系中。

[0087] 2.4 替代性内含子剪接

[0088] 因为 ORF4 的 ps-2ABL 等位基因中的序列突变位于诸内含子识别剪接位点之一中,我们假设这个突变可能影响该基因的前体 mRNA(pre-mRNA) 剪接。为验证该假设,我们设计引物来扩增 ORF4 的准全长 cDNA 克隆(1179nt 中的 1032nt):正向引物基于 ORF4 的第一外显子设计并且反向引物基于最末的外显子(表 3)。在从 ps-2ABL 和栽培品种 Moneymaker 的花期后花药囊的 RNA 制得的 cDNA 上进行 RT-PCR。对栽培品种 Moneymaker 观察到的扩增产物具有预期大小(1032bp),这通过测序予以证实。ps-2ABL 的扩增产物明显比 Moneymaker 产物更小,这提示内含子剪接过程中的变异(图 3)。在 ps-2ABL 的 cDNA 上所扩增片段的测序显示 cDNA 序列中遗漏了其存在 SNP 的第五外显子。这个外显子连同两个侧翼内含子在前体 mRNA 成熟过程期间被除去(图 4)。这个外显子-内含子剪接接合处的野生型 5' 序列是 CAG/GTATCG,其与马铃薯(*Solanum tuberosum*)中鉴定到的剪接接合序列的 5' 序列(Brown,1986)相同。ps-2ABL 等位基因中找到的突变诱导以下序列:CAC/GTACG,该序列不存在于内含子剪接识别位点的名单中。第五外显子在成熟 mRNA 中的缺乏代表 208 核苷酸缺失,这诱导下游其余编码序列的移码。这种移码引起翻译在第 14 氨基酸后过早终止,原因在于新近符合可读框的终止密码子。与野生型蛋白质的 392 氨基酸相比,完整的推定性突变蛋白质因而长 154 个氨基酸并且很可能是无功能的。

[0089] 2.5ORF4 序列分析

[0090] 用 ORF4 的推定性候选蛋白质序列在 NCBI 蛋白质数据库中的 BLAST 搜索产生一系列来自几种植物的 PG 蛋白质。鉴定的白质具有果实成熟和长角果/荚开裂中的功能。已经发现它们当中的 ADPG1 蛋白在拟南芥长角果的开裂区以及花药开裂区中表达(Sander 等人,2001)。最佳 BLAST 命中的 Pfam GH28 结构域的氨基酸一起与候选蛋白质的序列和具有已鉴定功能的已知番茄 PG 蛋白及一种裸子植物 PG(雪松)比较。在最终比对结果上进行系统发育分析以将候选蛋白质置于一个参考 PG 进化枝中。将 ORF4 鉴定为进化枝 B 的一种 PG(图 5)。进化枝 B 由编码果实和开裂区中所表达的 PG 的全部克隆基因组成,如先前 Hadfields 和 Bennett(1998)表征。在本发明的进化系统树中也观察到该情况。作为已知在

果实中表达的唯一番茄 PG 的 TFPG 也是同一个进化枝的部分。在图 6 中显示 ORF4 和 TFPG 的 Pfam GH28 结构域的比对结果。ORF4 和 TFPG 在完整蛋白质序列上具有 59% 的相似性。

[0091] ORF4 的推定性衍生蛋白含有作为 PG 蛋白质特征的 4 个保守结构域, 如 Rao 等人提出 (1996 ; 图 6)。第一保守结构域位于第五外显子上, 其中在突变蛋白中遗漏所述第五外显, 并且其余 3 个结构域进一步位于该突变蛋白序列上并且因而不符合该突变序列的可读框。因而该突变蛋白不含有在该蛋白质的功能中发挥重要作用的 4 个保守结构域之任意者。分析 ORF4 的起始密码子上游 2000 个核苷酸表明在相对于起始密码子的第 -667、-700 和 -1955 位置处存在 3 个乙烯应答元件 (ERE ; AWTTCAAA), 在第 -1632 位置处存在一个 bZIP 蛋白结合基序 (TGACG) 并在第 -1329 位置处存在一个 G- 框 (CACGTG)。ORF4 启动子序列中 ERE 基序、bZIP 蛋白结合基序和 G- 框的存在表明 ORF4 的转录受乙烯和茉莉酸酯刺激。

[0092] 2. 6ORF4 在花药、果实和其他组织中的表达

[0093] 我们检验几种组织中 ORF4 转录物的存在性, 所述组织包括叶和花的离区 \ 成熟果实和处于花期的花药。在所述离区中未检测到 ORF4 转录物 (图 7)。在花药以及在果实中证实存在 ORF4 转录物。为研究 ORF4 转录水平在花药和果实发育的不同阶段期间的演化, 我们在花药的 4 个发育阶段 : 花芽 ; 前花期 ; 花期 ; 后花期和果实从授粉后 5 日 (授粉后日) 至授粉后 57 日 (成熟果实) 的 8 个发育阶段上开展 ORF4 的定量表达分析 ; 授粉后 47 日对应于开裂期 (数据未显示)。在花药中, 对 MoneyMaker 和 ps-2ABL 检测到 ORF4 的转录水平, 而在果实中, 仅对 MoneyMaker 检测到 ORF4 的转录水平。在授粉后 37 日前的果实中未检测到 ORF4 转录物。从授粉后 37 日, 检测到 ORF4 转录物并且该转录物随时间推移显著增加至成熟期 (授粉后 57 日) 达到最大值。在 MoneyMaker 的花药中, 已经在花芽期检测到 ORF4 转录物。在花期前, ORF4 转录物的水平类似于花芽。ORF4 转录物聚集在花期增加并在花期后达到最大。在 ps-2ABL 的花药中, 除花芽期之外, 也检测到 ORF4 转录物。然而, ps-2ABL 花药中该转录物的水平显著低于花期和花期后 MoneyMaker 花药中该转录物的水平。

[0094] 3. 讨论

[0095] ps-2 纯种高级育种系产生不发生开裂的花药。在本研究中, 我们证实 ps-2ABL 的花药发育 / 开裂在最终阶段中被阻断。裂口不退化并且药室内壁不增厚。在不存在这两种生理变化的情况下, 花药仍保持关闭并且表皮细胞缺少可能逐渐破坏裂口并释放花粉的刚性。

[0096] 这种表型突变是隐性的并处于一个单基因座的控制下。在先前研究中, 我们将 ps-2 基因精确标定在染色体 4 的短臂上 (Gorguet 等人, 2006)。这里, 我们报道了 ps-2 基因的分离和功能表征。我们发现 ps-2 表型是在一个由 9 个外显子组成的迄今未知的多聚半乳糖醛酸酶基因中单核苷酸突变的结果。这个单核苷酸突变位于第五外显子 3' 端的最末核苷酸处, 影响内含子剪接识别位点, 该内含子剪接识别位点从 CAG/GTATCG 变成 CAC/GTATCG (外显子 3' / 内含子 5')。尽管植物中在这个特定位置处的 11% 内含子剪接位点胞嘧啶碱基中存在胞嘧啶碱基, 然而在植物的任意剪接位点处从未检测到外显子 3' 末端处的组合“CAC” (Brown 1986)。第五外显子连同两个侧翼内含子被剪接去掉。分析剪接位点周围具有突变的拟南芥突变体已经揭示植物剪接过程中外显子遗漏的几个例子 (Brown 和 Simpson, 1998 综述)。这些突变的大部分位于该识别剪接位点的内含子部分。据我们所知, 在植物中, 迄今显示因识别剪接位点的内含子部分中突变所致外显子遗漏的唯一突

变体是拟南芥中的 *spy-1* 突变体 (Jacobsen 等人, 1996), 其中第八外显子末端处的 CAG/GTTTGA (外显子' / 内含子 5') 识别剪接位点突变成 CAA/GTTTGA。在 ORF4 的突变等位基因中观察到的外显子遗漏导致该序列的其余部分移码, 这进一步导致产生提前 14 个氨基酸终止密码子的后果。与野生型长度 392 氨基酸相比, 完整突变蛋白因而长度为 154 个氨基酸, 并且不含有作为 PG 特征的 4 个结构域之任何一个。

[0097] 3. *lps-2* 是进化枝 B 的 PG

[0098] 导致 *ps-2* 表型的分离的基因是迄今未知的一种 PG。我们提出缩写 DPG, 其代表开裂多聚半乳糖醛酸酶或在番茄基因 TDPG 的情况下代表番茄开裂多聚半乳糖醛酸酶。TDPG 的系统发育分析揭示与如 Hadfield 和 Bennett (1998) 所定义进化枝 B 的 PG 的密切相似性。进化枝 B 在本文由果实 PG 组成, 果实 PG 当中包括番茄果实 PG 以及长角果或荚开裂 PG。其他番茄 PG 聚类于不同进化枝中。根据论断: PG 基因家族趋异发生在被子植物物种的隔离之前 (Hadfield 和 Bennett, 1998), TDPG 在本文中比源自其他进化枝的番茄 PG 更密切地相关于源自其他物种的相同进化枝的基因。其余番茄 PG 的大部分相关于脱落 (TAPG)。在我们的研究中, 没有检测到花及叶离区中 TDPG 的表达。然而, 除花药组织之外, 我们在果实中检测到 TDPG 的 mRNA 转录物。

[0099] 3. 2TDPG 转录物聚集随花药发育增加

[0100] 我们测量了花药中不同阶段上 TDPG 转录物的相对水平。已经在 Moneymaker 的花药中花芽阶段上检测到 TDPG 转录物。该转录物水平随阶段推移增加至花期后花药裂开时达到最大。这种 TDPG 转录物聚集增加与隔膜和裂口退化以及花药中观察到的药室内壁细胞壁增厚平行。从花期前伊始检测 *ps-2ABL* 花药中 TDPG 转录物水平, 不过该转录物水平与 Moneymaker 相比在花期和花期后保持极低。突变 mRNA 极有可能被识别为无意义并由无义密码子介导的 mRNA 降解作用 (NMD) 降解。NMD 作为质量控制机制发挥作用以消除异常转录物 (Lejeune 和 Maquat 2005)。

[0101] 3. 2TDPG 很可能处于乙烯和茉莉酸酯控制下

[0102] TDPG 启动子区域内乙烯和茉莉酸酯应答元件的存在表明 DPG 基因的转录可能由两种激素诱导。乙烯已经参与烟草中花药开裂的时间调控 (Rieu 等人, 2003)。最近在矮牵牛中已经证实乙烯调节花药开裂与花开启同步 (Wang 和 Kumar, 2006)。此外, 众多研究已经将茉莉酸酯鉴定为花药开裂过程和时间调控中的重要化合物。已经为研究该现象鉴定了 JA 生物合成酶的几个突变体。Scott 等人 (2004) 提出乙烯和 JA 可能在控制花药开裂中冗余地发挥作用, 这会解释在雄蕊内不能合成 JA 的拟南芥突变体如 *dde-1* 或乙烯不敏感 *Tetr* 突变体为何逐渐发生花药开裂。

[0103] 在本研究中, 我们证实 *ps-2ABL* 的花药仍不开裂, 这不同于其他突变体中的延迟开裂, 这支持以下假设: DPG 在控制花药开裂于乙烯和 JA 的下游发挥作用并且 DPG 是这个过程的主要角色。

[0104] 3. 3TDPG 也可能在番茄果实成熟中发挥作用

[0105] 迄今鉴定的唯一番茄果实多聚半乳糖醛酸酶 TFPG 已经被表征为果实软化过程中的主要角色之一。反义抑制 TFPG 已经导致产生具有更长货架期的番茄果实, 不过这种果实的确经历成熟过程, 这表明其他原物质也在果实软化过程中发挥重要作用 (Smith 等人, 1988)。DPG 也可以因造成果实细胞壁退化而是这些原物质之一。在 Moneymaker 中的果实

发育晚期检测到 TDPG 转录物 (图 8)。在成熟期 (授粉后 57 日) 找到最多的转录物。类似地, 仅在始于成熟绿色期或果实花端着色期的成熟期检测到 TFPG (Thompson 等人, 1999; Eriksson 等人, 2004)。还发现 TFPG 的水平在栽培品种 AC 和 Liberto 中随时间从果实花端着色期至成熟果实提高 (Thompson 等人, 1999)。

[0106] 在番茄中, 虽然乙烯对 PG mRNA 聚集的调节作用持续不确定的长时间, 然而已经证实果实成熟过程中 TFPG 聚集是乙烯依赖性的 (Sitrit 和 Bennett, 1998)。因而, 已经在 TFPG 的启动子中找到 ERE 基序 (Montgomery 等人, 1993)。乙烯作为果实成熟控制中的主要植物激素存在 (Giovannoni, 2004 综述)。表达针对乙烯的生物合成途径中限速酶的反义 RNA 抑制番茄中的果实成熟并因而下调 TFPG 的产生 (Oeller 等人, 1991; Sitrit 和 Bennett, 1998)。

[0107] 在本研究的较早时期, 我们提出 DPG 因启动子序列中存在 ERE 基序而受乙烯控制。阻遏番茄果实中的乙烯有可能抑制 TFPG 和 TDPG 并因而完全阻止果实成熟过程。图 8 中呈现花药开裂和果实成熟中激素控制 DPG 的简化模型。

[0108] 不清楚 DPG 在果实成熟和货架期中的作用是否具有与 TFPG 相似的重要性。手工开启花药后的 ps-2ABL 与 Moneymaker 或任何其他正常番茄品系之间的比较是不可靠的, 原因在于遗传背景差异。通过反义 RNA 或 RNAi 在正常番茄品系中敲除 DPG 可能回答上述问题。可能测量手工授粉至成熟果实期和果实货架期之间的时间并将其与未转化的对照比较以估计 DPG 在果实中的作用。敲除其他植物物种中 TDPG 的同源物以验证在最末期控制花药开裂是否在物种之间保守也是有价值。

[0109] 表 3. 野生型序列 (WT) 和突变序列的比对结果

[0110] WT : ATGGAGAAATTCAATGAAGAAGAAGATCAAGCTAAGGTTACAACAATTA 60

[0111] 突变体 : ATGGAGAAATTCAATGAAGAAGAAGATCAAGCTAAGGTTACAACAATTA 60

[0112] Pt WT : M E K F N E E E D Q A K V T T I N V D S

[0113] Pt Mt : M E K F N E E E D Q A K V T T I N V D S

[0114]  ▼

[0115] WT :  AAGGTGATGGAAGTATAGATGATACAAATGCATTTCAAAAAGCATGGAAA 120

[0116] 突变体 :  AAGGTGATGGAAGTATAGAT GATACAAATGCATTTCAAAAAGCATGGAAA 120

[0117] Pt WT : F G A K G D G S I D D T N A F Q K A W K

[0118] Pt Mt : F G A K G D G S I D D T N A F Q K A W K

[0119] WT : GAAGCTTGTTTCATCTTCACATGTTGTGAATTTTGTGGTGTCCCAGAACAAGAAATATCTT 180

[0120] 突变体 : GAAGCTTGTTTCATCTTCACATGTTGTGAATTTTGTGGTGTCCCAGAACAAGAAATATCTT 180

[0121] Pt WT : E A C S S S H V V N F V V S Q N K K Y L

[0122] Pt Mt : E A C S S S H V V N F V V S Q N K K Y L

[0123] ▼

[0124] WT : CTCAAACCAATCAAATTTTATGGGCCATGCAAATCTTCCATTACAATGCAGATTTATGGA 240

[0125] 突变体 : CTCAAACCAATCAAATTTTATGGGCCATGCAAATCTTCCATTACAATGCAGATTTATGGA 240

[0126] Pt WT : L K P I K F Y G P C K S S I T M Q I Y G

[0127] Pt Mt : L K P I K F Y G P C K S S I T M Q I Y G

[0128] WT: ACCCTATTAGCATCTGATGATACTTCAGATTACAAGAAGGATAGTAGGCACTGGCTTATT 300
[0129] 突变体:ACCCTATTAGCATCTGATGATACTTCAGATTACAAGAAGGATAGTAGGCACTGGCTTATT 300
[0130] Pt WT: T L L A S D D T S D Y K K D S R H W L I
[0131] Pt Mt: T L L A S D D T S D Y K K D S R H W L I
[0132] WT: TTTGATAGTGTTCAAAAATTGGTTGTTGGAGGAGCTGGAGTTATCAATGGCAATGGCAAA 360
[0133] 突变体:TTTGATAGTGTTCAAAAATTGGTTGTTGGAGGAGCTGGAGTTATCAATGGCAATGGCAAA 360
[0134] Pt WT: F D S V Q K L V V G G A G V I N G N G K
[0135] Pt Mt: F D S V Q K L V V G G A G V I N G N G K
[0136]

▼

[0137] WT: ATTTGGTGGCAACATTCTTGCAAAATTAATAAAAAATTGCCATGCAAGGTAGCACCCACG 420
[0138] 突变体:ATTTGGTGGCAACATTCTTGCAAAATTAATAAAAAATTGCCATGCAAGGTAGCACCCACG 420
[0139] Pt WT: I W W Q H S C K I N K K L P C K V A P T
[0140] Pt Mt: I W W Q H S C K I N K K L P C K V A P T
[0141] WT: GCCCTGACATTTTACAAGTGTAAACAACCTGAAAGTGAAGGACCTTAAATAGAAAATGCA 480
[0142] 突变体:-----
[0143] PtWT: A L T F Y K C N N L K V K D L K I E N A
[0144] Pt Mt: - - - - -
[0145] WT: CAACAAATACATTTGCTAATTGAGAAGTGTGTTGGTGTGAAGTTACAAAATTGGTAGTG 540
[0146] 突变体:-----
[0147] Pt WT: Q Q I H L L I E K C V G V E V T K L V V
[0148] Pt Mt: - - - - -
[0149] WT: ACTTCTCCAGAAAATAGCCCTAATACTGATGGAATCCATATAACTAGCACTCAAAATATT 600
[0150] 突变体:-----
[0151] Pt WT:T S P E N **S P N T D G I H** I T S T Q N I
[0152] Pt Mt:- - - - -
[0153] 结构域 I

[0154]

▼

[0155] WT: CAAATTTCTGATTCCACCATTGCCACAGGTGATGATTGCATCTCAATTGTGGATGGATCT 660
[0156] 突变体:-----GTGATGATTGCATCTCAATTGTGGATGGATCT 452
[0157] Pt WT:Q I S D S T I A T **G D D C** I S I V D G S
[0158] Pt Mt:- - - - -V M I A S Q L W M D L
[0159] 结构域 II

[0160]

▼

[0161] WT: CAGAAGGTCTTAGCCACTGGCATTACTTGTGGACCAGGTCATGGAATTAGTATTGGAAGT 720
[0162] 突变体:CAGAAGGTCTTAGCCACTGGCATTACTTGTGGACCAGGTCATGGAATTAGTATTGGAAGT 512
[0163] Pt WT:Q K V L A T G I T **C G P G H G I S I G S**
[0164] Pt Mt:R R S 终止 结构域 III

- [0165] WT : TTGGGAGGTGGAATTCAGAAGCTCATGTGTCTGATATTCATGTAAATGGAGCTAAGCTT 780
- [0166] 突变体 :TTGGGAGGTGGAATTCAGAAGCTCATGTGTCTGATATTCATGTAAATGGAGCTAAGCTT 572
- [0167] Pt WT : **L** **G** G N S E A H V S D I H V N G A K L
- [0168] ▼
- [0169] WT : TATGAACTACAAATGGACTTAGGATTAAGACTTGGCCGGGAGGATTTGGAAGTGCAAGC 840
- [0170] 突变体 :TATGAACTACAAATGGACTTAGGATTAAGACTTGGCCGGGAGGATTTGGAAGTGCAAGC 632
- [0171] Pt WT : Y E T T N G L **R I K** T W P G G F G S A S
- [0172] 结构域 IV
- [0173] WT : AATATTAAGTATCAAAATGTGGTTATGAATAATGTCAAAAATCCAATAATTATAGACCAA 900
- [0174] 突变体 :AATATTAAGTATCAAAATGTGGTTATGAATAATGTCAAAAATCCAATAATTATAGACCAA 692
- [0175] Pt WT : N I K Y Q N V V M N N V K N P I I I D Q
- [0176] ▼
- [0177] WT : AATTATTGTGATCAAGCTGATGGTCCATGCAAAGCTGAGACTGACTCGGCAGTTGAAGTG 960
- [0178] 突变体 :AATTATTGTGATCAAGCTGATGGTCCATGCAAAGCTGAGACTGACTCGGCAGTTGAAGTG 752
- [0179] Pt WT : N Y C D Q A D G P C K A E T D S A V E V
- [0180] WT : AAAAATGTGATTTATCAAAATATCAAAGGCACAAGTGCAACAAATGATGCAATAAGTATC 1020
- [0181] 突变体 :AAAAATGTGATTTATCAAAATATCAAAGGCACAAGTGCAACAAATGATGCAATAAGTATC 812
- [0182] Pt WT : K N V I Y Q N I K G T S A T N D A I S I
- [0183] 
- [0184] WT : AAGTGCAGCAAAAAAATTCATGTGAAGGAATTTGA  A 1080
- [0185] 突变体 :AAGTGCAGCAAAAAAATTCATGTGAAGGAATTTGA  A 872
- [0186] Pt WT : K C S K K I P C E G I L M E N V K L L G
- [0187] WT : GGAAATGGTGAACTCCAAATGGTATTTGGGGAAATATCAATAATCTTACGTGCAAAAAT 1140
- [0188] 突变体 :GGAAATGGTGAACTCCAAATGGTATTTGGGGAAATATCAATAATCTTACGTGCAAAAAT 932
- [0189] Pt WT : G N G E T P N G I W G N I N N L T C K N
- [0190] WT : GTTTTACCAGAATGTCAAAAAAATCAAAAATTTGTATAA 1179
- [0191] 突变体 :GTTTACCAGAATGTCAAAAAAATCAAAAATTTGTATAA 971
- [0192] Pt WT : V L P E C Q K N S K I V. 终止
- [0193] Pt WT :野生型 (Moneymaker) 的蛋白质序列
- [0194] Pt Mt :突变体 (ps-2ABL) 的蛋白质序列
- [0195] - 粗体的核苷酸“G”指示突变的位置
- [0196] - 保守的蛋白质结构域以粗体下划线方式显示。
- [0197] - 内含子的位置用序列上方的黑色箭头▼指示。
- [0198] - 用来扩增准完整编码序列的引物序列的位置以显示。
- [0199] 参考文献 :
- [0200] Atkinson R G 和 Gardner R C(1993)A polygalacturonase gene from

kiwifruit(*Actinidiadeliciosa*)*Plant Physiol.* 103 :669-670

[0201] Atkinson R G(1994)A cDNA clone for endopolygalacturonase from apple.
Plant Physiol. 105 :1437-1438

[0202] Bergey D R, Orozco-Cardenas M, de Moura D S, Ryan C A(1999)A wound-and systemin-inducible polygalacturonase in tomato leaves *Proc. Natl. Acad. Sci. U. S. A.* 96(4), 1756-1760(1999)

[0203] Bird C R, Smith C J, Ray J A, Moureau P, Bevan M W, Bird A S, Hughes S, Morris P C, Grierson D, Schuch W(1988)The tomato polygalacturonase gene and ripening-specific expression in transgenic plants *Plant Mol. Biol.* 11 :651-662

[0204] Brown J W S(1986)A catalogue of splice junction and putative branch point sequences from plant introns. *Nucleic acids research* 14 :9549-9559

[0205] Brown J W S 和 Simpson C G(1998)Splice site selection in plant pre-mRNA splicing. *Annual review of plant physiology and plant molecular biology* 49 : 77-95

[0206] Brugmans B, Hutten R G B, Rookmaker A N O, Visser R G F, Van Eck H J(2006) Exploitation of a marker dense linkage map of potato for positional cloning of a wart disease resistance gene. *Theoretical and Applied Genetics* 112 :269-277

[0207] Budiman M A, Mao L, Wood T C, Wing R A(2000)A deep-coverage tomato BAC library and prospects toward development of an STC framework for genome sequencing. *Genome Res* 10 :129-136

[0208] Christiansen L C, Dal Degan F, Ulvskov P, Borkhardt B(2002)Examination of the dehiscence zone in soybean pods and isolation of a dehiscence-related endopolygalacturonase gene *Plant Cell Environ.* 25 :479-490

[0209] Colbert T, Till B J, Tompa R, Reynolds S, Steine M N, Yeung A T, McCallum C M, Comai L, Henikoff S. High-throughput screening for induced point mutations. *Plant Physiol.* 2001 Jun ;126(2) :480-4

[0210] Dornelas M, van Lammeren A, 和 Kreis M(2000). *Arabidopsis thaliana* SHAGGY-related protein kinases (AtSK11 and 12) function in perianth and gynoecium development. *Plant Journal* 21 :419-429

[0211] Dubald M, Barakate A, Mandaron P, Mache R(1993)The ubiquitous presence of exopolygalacturonase in maize suggests a fundamental cellular function for this enzyme. *The plant journal* 4 :781-791

[0212] Gorguet B, Schipper D, van Heusden A W, Lindhout P(2006)High resolution fine mapping of ps-2, a mutated gene conferring functional male sterility in tomato due to non-dehiscent anthers. *Theoretical and Applied Genetics.* 113 : 1437-1448

[0213] Hadfield K A, Bennett A B(1998)Polygalacturonases : Many genes in search of a function. *Plant physiology* 117 :337-343

[0214] Higgins D, Thompson J, Gibson T, Thompson J D, Higgins D G, Gibson T J(1994)

CLUSTAL W:improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22 :4673-4680

[0215] Hong S B 和 Tucker M L (1998) Genomic organization of six tomato polygalacturonases and 5' upstream sequence identity with tap1 and win2 genes *Mol. Gen. Genet.* 258 :479-487 Hong S-B 和 Tucker M L, (2000) Molecular characterization of a tomato polygalacturonase gene abundantly expressed in the upper third of pistils from opened and unopened flowers. *Plant Cell Reports* 19 : 680-683

[0216] Huala E, Dickerman A W, Garcia-Hernandez M, Weems D, Reiser L, LaFond F, Hanley D, Kiphart D, Zhuang M, Huang W, Mueller L A, Bhattacharyya D, Bhaya D, Sobral B W, Beavis W, Meinke D W, Town C D, Somerville C, Rhee S Y (2001) The Arabidopsis Information Resource (TAIR) : a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. *Nucleic Acids Res.* 29 :102-105

[0217] Jacobsen S E, Binkowski K A, Olszewski N E (1996) SPINDLY, a tetratricopeptide repeat protein involved in gibberellin signal transduction in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* 93 :9292-9296

[0218] Jenkins E S, Paul W, Coupe S A, Bell S, Davies L, Roberts J A. Characterization and expression of an mRNA encoding polygalacturonase during pod development in oilseed rape (*Brassica napus* L.) Unpublished

[0219] Keijzer C J (1987) The processes of anther dehiscence and pollen dispersal. I. The opening mechanism of longitudinally dehiscing anthers. *New Phytologist* 105 :487-498

[0220] Lejeune F, Maquat L E (2005) Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Current Opinion in Cell Biology* 17 :309-315

[0221] Lester D R, Speirs J, Orr G, Brady C J (1994) Peach (*Prunus persica*) endopolygalacturonase cDNA isolation and mRNA analysis in melting and nonmelting peach cultivars *Plant Physiol.* 105 :225-231

[0222] Markovic O, Janecek S (2001) Pectin degrading glycoside hydrolases of family 28: sequence-structural features, specificities and evolution. *Protein engineering* 14 :615-631

[0223] McCallum CM, Comai L, Greene EA, Henikoff S. Targeted screening for induced mutations. *Nat Biotechnol.* 2000 Apr; 18(4) :455-7

[0224] McCallum CM, Comai L, Greene EA, Henikoff S. Targeting induced local lesions in genomes (TILLING) for plant functional genomics. *Plant Physiol.* 2000a Jun; 123(2) :439-42

- [0225] Montgomery J, Pollard V, Deikman J, Fischer R L (1993) Positive and negative regulatory regions control the spatial-distribution of polygalacturonase transcription in tomato fruit pericarp. *The plant cell* 5:1049-1062
- [0226] Mueller L A, Solow T H, Taylor N, Skwarecki B, Buels R, Binns J, Lin C, Wright M H, Ahrens R, Wang, Y, Herbst E V, Keyder E R, Menda N, Zamir D, Tanksley S D (2005) The SOL Genomics Network. A comparative resource for Solanaceae biology and beyond. *Plant Physiol* 138:1310-1317
- [0227] Nunan K J, Davies C, Robinson S P 和 Fincher G B (2001) Expression patterns of cell wall-modifying enzymes during grape berry development. *Planta* 214:257-264
- [0228] Oeller P W, Min-Wong L, Taylor L P, Pike D A, Theologis A (1991) Reversible inhibition of tomato fruit senescence by antisense RNA. *Science* 254:437-439
- [0229] Ogasawara S 和 Nakajima T. A gene coding for bell pepper endo-polygalacturonase is expressed during fruit ripening. 未发表
- [0230] Petersen M, Sander L, Child R, van Onckelen H, Ulvskov P, Borkhardt B (1996) Isolation and characterisation of a pod dehiscence zone-specific polygalacturonase from *Brassica napus*. *Plant Mol. Biol.* 31:517-527
- [0231] Rao M N, Kembhavi A A, Pant A (1996) Implication of tryptophan and histidine in the active site of endo-polygalacturonase from *Aspergillus ustus*: elucidation of the reaction mechanism. *Biochim. Biophys. Acta* 1296:167-173
- [0232] Rieu I, Wolters-Arts M, Derksen J, Mariani C, Weterings K (2003) Ethylene regulates the timing of anther dehiscence in tobacco. *Planta* 217:131-137
- [0233] Rodriguez-Gacio M, Nicolas C, Matilla A (2004) Cloning and analysis of a cDNA encoding an endo-polygalacturonase expressed during the desiccation period of the silique-valves of turnip-tops (*Brassica rapa* L. cv. Rapa). *Plant Physiol.* 161:229-236
- [0234] Rouster J, Leah R, Mundy J, Cameron-Mills V (1997) Identification of a methyl jasmonate-responsive region in the promoter of a lipoxygenase 1 gene expressed in barley grain. *The plant journal* 11:513-523
- [0235] Sander L, Child R, Ulvskov P, Albrechtsen M, Borkhardt B (2001) Analysis of a dehiscence zone endo-polygalacturonase in oilseed rape (*Brassica napus*) and *Arabidopsis thaliana*: evidence for roles in cell separation in dehiscence and abscission zones, and in stylar tissues during pollen tube growth. *Plant Mol. Biol.* 46:469-479
- [0236] Scott R J, Spielman M, Dickinson H G (2004) Stamen structure and function. *The plant cell* 16:S40-S60
- [0237] Sitrit Y, Hadfield K A, Bennett A B, Bradford K J, Downie A B (1999) Expression of a polygalacturonase associated with tomato seed germination. *Plant Physiol.* 121:419-428

- [0238] Smith C J S, Watson C F, Ray J, Bird C R, Morris P C, Schuch W, Grierson D(1988)Anti-sense RNA inhibition of polygalacturonase gene expression in transgenic tomatoes. *Nature* 334 :724-726
- [0239] Swofford D L(2002)PAUP*.Phylogenetic Analysis Using Parsimony(*and Other Methods).Version 4. Sinauer Associates, Sunderland, Massachusetts
- [0240] Torki M, Mandaron P, Mache R, Falconet D(2000)Characterization of a ubiquitous expressed gene family encoding polygalacturonase in *Arabidopsis thaliana*. *Gene* 242 :427-436
- [0241] Wang Y, Kumar P P(2006)Characterization of two ethylene receptors PhERS1 and PhETR2 from petunia :PhETR2 regulates timing of anther dehiscence. *Journal of Experimental Botany* doi :10.1093/jxb/erl229

序列表

<110> 西方种子国际有限公司

<120> 新的多聚半乳糖醛酸酶和它们的用途

<130>P6015073PCT

<150>EP 07113263.3

<151>2007-07-26

<150>US 60/952,046

<151>2007-07-26

<160>2

<170>PatentIn version 3.3

<210>1

<211>6877

<212>DNA

<213>Solanum lycopersicum

<220>

<221>misc_feature

<222>(4386)..(4386)

<223>n is a, c, g, or t

<400>1

```
aataattagt agattccaaa gatataataa tatggagaaa ttcaatgaag aagaagatca    60
agctaagggtt acaacaatta atgtggatag ctttggagct aaaggtgatg gaagtataga   120
tgatacaaat gtaagtatta gttttctttt aaaaaaaaaat tctctttcga actattatta   180
aaaacacgtg aatttcttat gaaaatttct cagacaaatc atcgtatatt gacttgtaa    240
aagcttttca gacagacacg ctccattca aaaaaaataa ttaaaattca catgttttta   300
atagtgaac aatatatgta tactaccttg atttgattta atttatttgg cgatatttga   360
tttaagttaa aattaaat aaatgaagat tttaaactt atgatgttaa ttagatgtcc    420
tataattttt caatggttgt aaaggettaa atgtatatat aaaaatagga tgtatataaa   480
gttacattat tttaaatat aaataaaata tttttattag tcatacatc aaaaatgaaat   540
aaaagaagta tataaagtta cattatttta aaatataaat aaaatatatt tattagtcac   600
acatatataa tgaaataaaa gaagtatatt attagatata aaagtttgtg aagaaatgct   660
```

ttttttttta	aaaagaat	ttttgtatg	caacgataag	ttgataactg	ataacatatg	720
tcacatcttt	gcggttgat	ttgcattagt	tattggtggt	aaaatcaaat	aatcaat	780
atccggtctt	ataattatta	aattaaagta	aatctaaaat	aatatatatt	tatcgatttg	840
attgttattg	gtttcagttt	gattagatta	tttgattact	aatcatatat	aaaattaaaa	900
gaaataat	attaaaaacg	tgaaaagagc	ggtaacaatt	tattcaaaat	gaaaaatagt	960
tagaatgact	aaaattaaat	aattttcaat	tacagaat	actatgaata	aagcttcaaa	1020
atttactatt	ttgacgtaga	agaaagagac	aataacat	tttgtcttaa	aaaaatgatc	1080
ataacggtta	tatacatgaa	accaatataa	gtcaaatgtt	ctcaccttga	ggattattgt	1140
ttttgcgtaa	gtacaaatca	taaat	tgaaaacaaa	tataatatct	ttaaaaatcg	1200
tctacaaaaa	taatata	tatatgtatg	ataaaatata	atatatgtat	atgatttate	1260
gatttgattc	agttatttct	ttgatatttt	tctagcaaaa	tcacaatcaa	actaaatact	1320
attgat	aaaaatccaa	aatcaaatca	aaccataatc	caaattcaat	ttgtatcaat	1380
tttatccaaa	cgtgaacact	cctagtcaca	taacttaaaa	attaattata	agtaaccatc	1440
aataagaaat	gaccgaacta	gttaattaat	ataagtatac	acacacatta	attaattaaa	1500
catctttttt	cattcgatat	tgatattgaa	gctcaattgt	ttcgactttg	tcctgcgtaa	1560
ggcctcat	ggcaaaaaaa	tactctttag	caagaat	ttttatatgc	aaaactcaaa	1620
tctgagactc	cgaattaaga	taggagaatt	aaacatctca	tttcttggtg	gaaaataaag	1680
attatgtcat	acttagaagc	caattattat	catgggtgaat	atatgtagca	ggttgggatt	1740
acttctacga	cagaatttta	tcagcatcat	aagcaattta	attagttttt	tttaataaat	1800
taattaatat	attaaaataa	atttatctta	atttcaggca	tttcaaaaag	catggaaaga	1860
agcttggttca	tcttcacatg	ttgtgaattt	tgtgggtgtcc	cagaacaaga	aatatcttct	1920
caaaccaatc	aaattttatg	ggccatgcaa	atcttccatt	acaatgcagg	ttaggttgac	1980
atatcattat	gtaaaaaaa	aactcgaagt	aacgataaaa	ttattctcgt	cctaactcact	2040
ataaagtatg	gatgttctga	gtcatgggtca	agtaattaga	tcttgaggct	tgcacaggg	2100
taggctgttt	gtgtcatggt	gtggctcttt	ccaaaactta	tatgaaagtg	agatgctttg	2160
tgcacccgat	tatcgattca	aactaacaga	ataaaagt	tttttatacg	attaggttat	2220
tttaaaatat	gataaataat	cttattcgac	gagttgaatt	gcatcgtgta	tgtggatcta	2280
tagtgtaatc	taaacattta	cgcgacttta	ttgaatagta	tccaaaaaat	attaaatcag	2340
aaattgatat	tgttatataa	acggataagt	ttcaaatctt	gaatacacct	ttgataatat	2400
gaaaaactta	ttacattgtc	tgcgcacgga	tataagttaa	ttaaactctg	ctattattat	2460
atatatgtag	agaaataagc	tctatgtata	gtgacaaagc	tatatccgtt	gaacgatgat	2520
tagttgaaca	cttaggttaa	aatataacct	taattttata	tattgcgaca	cctttcgtga	2580
aaattttaaa	tctcgtcacc	acttttagtt	acatttggaa	ataaataata	ggttggagt	2640
tagaatatta	ttttgacact	tttgttcaat	ttttttttat	tgtagattta	tggaaacctta	2700
ttagcatctg	atgatacttc	agattacaag	aaggatagta	ggcactggct	tatttttgat	2760
agtgttcaaa	aattgggtgt	tggaggagct	ggagttatca	atggcaatgg	caaaatttgg	2820
tggcaacatt	cttgcaaaat	taataaaaaa	ttggtataaa	tacatactta	attttgaaag	2880
ggtcgagaat	cttgaaacaa	tcctaaagtt	tatagggtcac	cgtcagttcg	aatcgtgtgc	2940
aaataattac	tattatttgc	actaggatag	actgtctaca	ttatacgcca	atcaagtga	3000

accttcccaa	accttgcata	agtgcattgat	gttttatgag	ccggggccgg	aaagcccttt	3060
tgctcctttac	atacttaatt	ttaaaaggag	agtcttggag	taatgataaa	attgtctatg	3120
tgcatgtgat	tataggttac	tgtecgttca	atccgtgtag	gaacaattac	taatacatgt	3180
tttagaatag	actatctata	ttacactcca	atcaagtgc	gttctttccg	aaccttgc	3240
aaacacatga	tgttttatgc	aacggggcgt	gggcaaccct	tttgcctttt	gcaaacttta	3300
tttattttatt	tatatatata	ttcctatttg	tttcgtttgt	ttttatgata	tataactaat	3360
tttgtcta	ttttatctgt	attgtctact	ttttttttta	ttttgcagcc	atgcaaggta	3420
gcaccacagg	tacgtttatt	tgatttgatt	tcttattttt	caaaagaaca	aaacaattaa	3480
ttattacgtc	atttttcttt	aataattttt	ttgttgatag	agatttttta	ttttatttta	3540
ttggtgacat	ttaaattttt	tcaggccctg	acattttaca	agtgtataca	cttgaaagtg	3600
aaggacctta	aaatagaaaa	tgacacaaca	atacatttgc	taattgagaa	gtgtgttggt	3660
gttgaagtta	caaaattggg	agtgcattct	ccagaaaaata	gccctaatac	tgatggaatc	3720
catataacta	gcactcaaaa	tattcaaatt	tctgattcca	ccattgccac	aggtatcgat	3780
aaattaacga	ttttaatttt	tataacttta	tagtaaaaaa	taattatatt	atcaagttgt	3840
atgtataaat	taaaggtaaa	catgctttta	ataagtga	cttatgataa	gttattagcc	3900
tgaaaatatg	aaagattaaa	attttcgaca	attgatcatc	atcgtctcat	cgaccacttg	3960
gtctttgacg	tagggctttt	cctaattccc	gaatgaactt	gataagtcaa	ctcaaagata	4020
aatatatgtt	tagatcatga	gtaacgtgtg	atacccttagg	taaatttcac	ttcgatatac	4080
atataagaaa	tgttgggtga	taggaaaaca	caatttagac	aatagtgtca	atgtgggttg	4140
atttaaaagg	tgcgattggg	ataaaggaaa	acatttttca	gaatcaactc	attttcctca	4200
aaattaagga	taatgaattc	ccttcaaaaa	ttgaggaaaa	catttttcaa	aactctcatc	4260
caatttgaaa	ttatatattt	tttaaagaaa	acatcgattt	cgaaaaatat	attcaatttc	4320
aaaattttat	atattcacca	ccaccccacc	ccacccgac	atgccttcac	ccccccccc	4380
ccccanaaaa	aaaatagatt	ttttttttta	aaaaatattt	ttaatttcaa	aatttcattt	4440
ctccatccct	accctttacc	cccccccct	accaaccctc	tacccccctc	ccaaaaaaga	4500
aaattaagtt	tgctctaaaa	aataattttc	acttcaaatt	tttatttttt	cacccttacc	4560
tcgacccccg	ccccccaaca	cccaccccac	cccactccat	ctttaaagat	aaattttcat	4620
tttgtttttt	taaaaatatt	tttaacctca	aatttttatt	ttttctaccc	ctacgtcccc	4680
ctctccagtc	cccttaccat	ccctccagcc	ccaccaaaaa	aatatacttt	aatttttaaa	4740
aaatatatttc	aacttttaaat	ttttattttt	tcactccctaa	cctcccgcac	ccaccacccc	4800
ccgccagccc	cccaccccc	aaaaaaaaatt	attttcaatt	tcaaaaatta	gtttgtactc	4860
tagtaaaaaat	aaaagatatt	tctcaaaaaat	atttttcatt	cataaattaa	acactaaaaa	4920
atatttttcta	ctcatcaacc	aaacatgaaa	aaacaaatca	taaatctact	tatttttctaa	4980
aaaaaaaaaca	ttttccta	taccaaacac	accgaaaatc	ttgaaagtgt	agttccaagc	5040
ggataattaa	tgctactata	gtgagttgta	ctattataga	ttgttgaaat	aatattttaca	5100
aatatttaaat	tatactaatt	gtacccaaaa	aaaatgacat	taaatttttt	tattttactaa	5160
ctatgtatatt	tggaataaat	ttttgtggta	ggtgatgatt	gcatctcaat	tgtggatgga	5220
tctcagaagg	tcttagccac	tggcattact	tgtggaccag	gtcatggaat	taggtaaaaa	5280
caaaataatt	tcattatatt	agtattatta	tgctagtatt	atcttatttt	ttattatatt	5340

```

attagtggtta ttctctaat ttttattatc ttacttttca ttacttttct ttctttaata 5400
tctttatcat tttatTTTTg tatatctttt gaactgttct gaaaaaattg tttttttatc 5460
gtttatatga aacaatctcc taccacacca aggtaagaaa taagtttttg ggtcaaacct 5520
cacttgcatg attatTTTct ttatcatatt cagattatTT tatcaaaatt tataatttaa 5580
ttatgtaatt tttttttatt attattatTT gtgcttagta ttggaagttt gggaggtgga 5640
aattcagaag ctcatgtgtc tgatattcat gtaaattggag ctaagcttta tgaaactaca 5700
aatggactta ggattaagac ttggccggta aattaattaa ttaattatTT ttattaattt 5760
acaatatatt tttataatta atatacatat aattatatTT ttgaattatt ttgtataggg 5820
aggatttgga agtgcaagca atattaagta tcaaaatgtg gttatgaata atgtcaaaaa 5880
tccaataatt atagaccaa attatttgta tcaagctgat ggtccatgca aagctgaggt 5940
aattgaaaat gaatttaact tatatatatt aactatacga aaaaataatc tactttgtta 6000
aagcagatat cttatcgtat ttttagata attaattcgt tccattatgg gttgttagtt 6060
aagctatctt tagatagcca aatatggtaa attttttaat aatgtcgaca tatgtatgtt 6120
ataaaaatat ataacttaaa ctgaacaaa aatgaatttt atttatataa catacttaat 6180
gtatagaata tttttatacc attagtgcaa ttgctacat catttgga caatagtggt 6240
ttaacttata aacaaaaata atttctcccc ttcaactttc ttaaaattca ataaaatccc 6300
catacaaaat attgtataat caaatttacc ccactttgaa aatgtaacat ataacgacat 6360
attgtacatt ataaagtTgt gacggaatat ctatgacgat aatattaatg acagagttaa 6420
tatggtcgat tcataatatg attcaatcac aaacttgTga taatttcatt gatctcttaa 6480
tcacaaaaat atattttatac atttatagac tgactcggca gttgaagtga aaaatgtgat 6540
ttatcaaaat atcaaaggca caagtgcac aaatgatgca ataagtatca agtgcagcaa 6600
aaaaattcca tgtgaaggaa ttttgatgga gaatgtgaaa ttgttaggag gaaatggtga 6660
aactccaaat ggtatttggg gaaatatcaa taatcttacg tgcaaaaatg ttttaccaga 6720
atgtcaaaaa aactcaaaaa ttgtataatt aagtttgata gagttaatta atgtttaata 6780
ggatattgta tagttataca tatgacaaat taagtgttag aatacataat tagcaaatga 6840
tttgtgataa atgatattct ttgtacacat gaaataa 6877

```

<210>2

<211>319

<212>PRT

<213>Solanum lycopersicum

<400>2

```

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

```

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

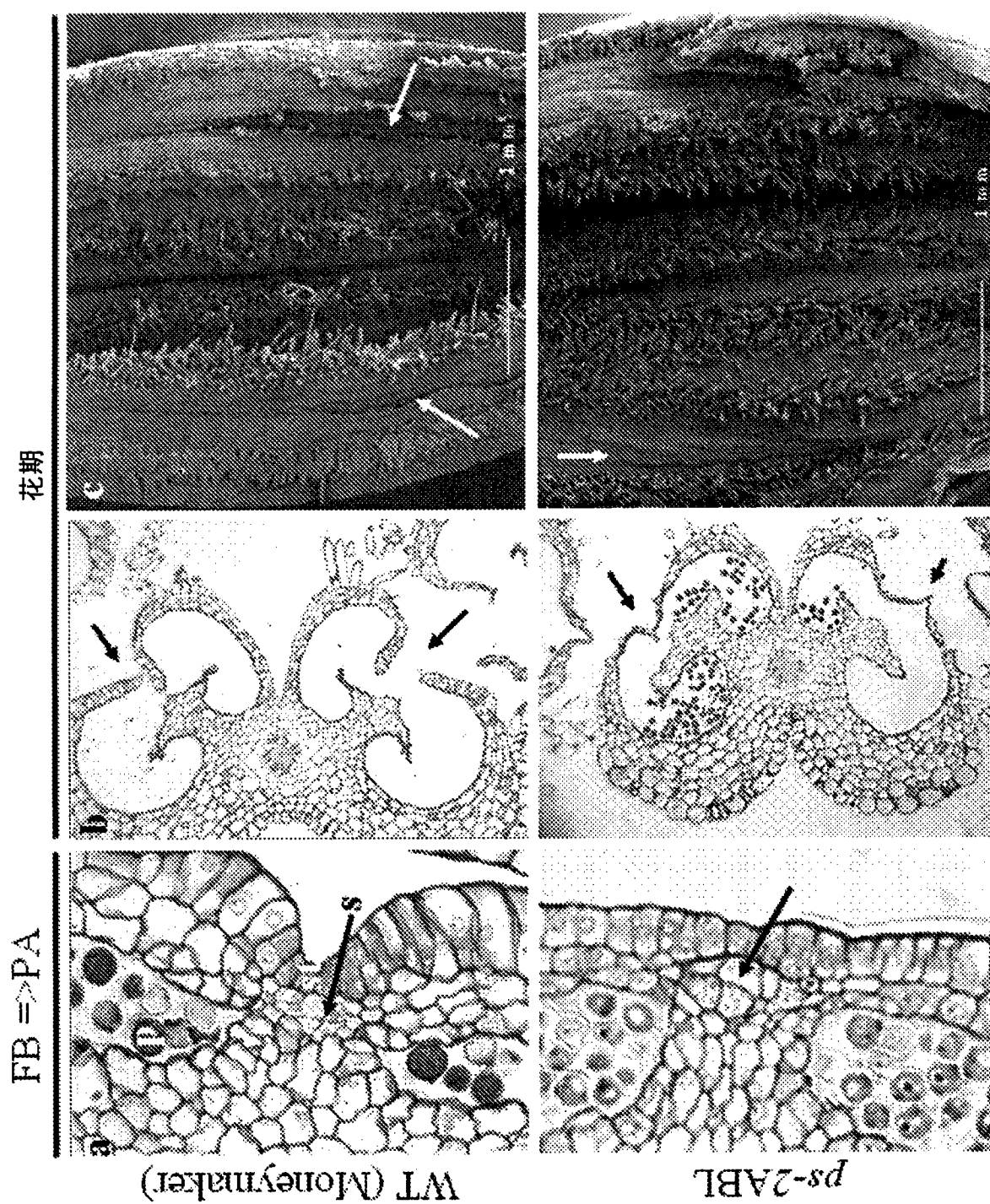


图 1

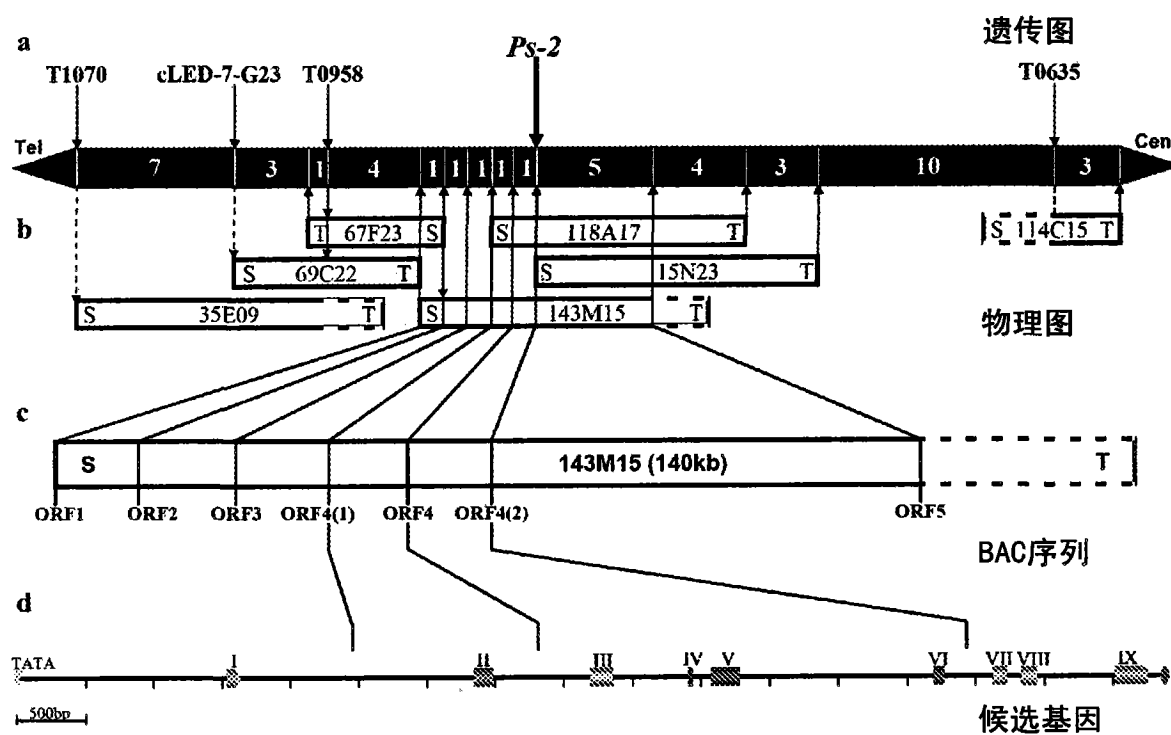


图 2

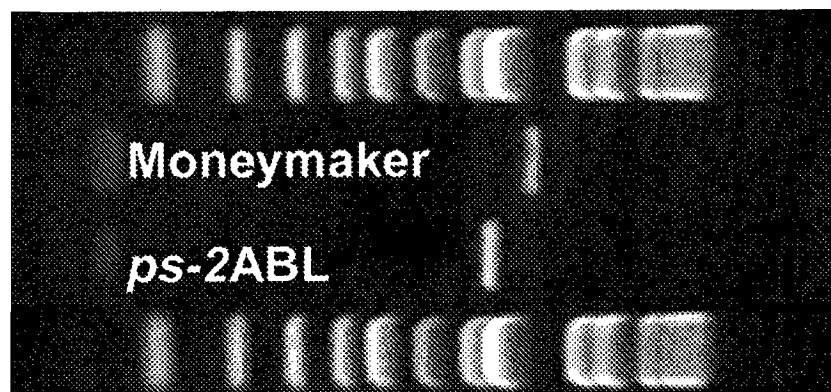


图 3

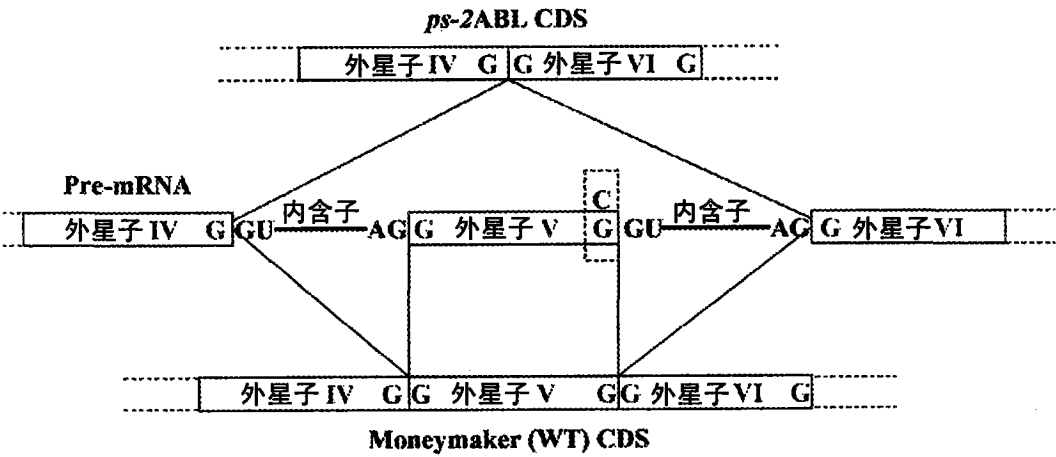


图 4

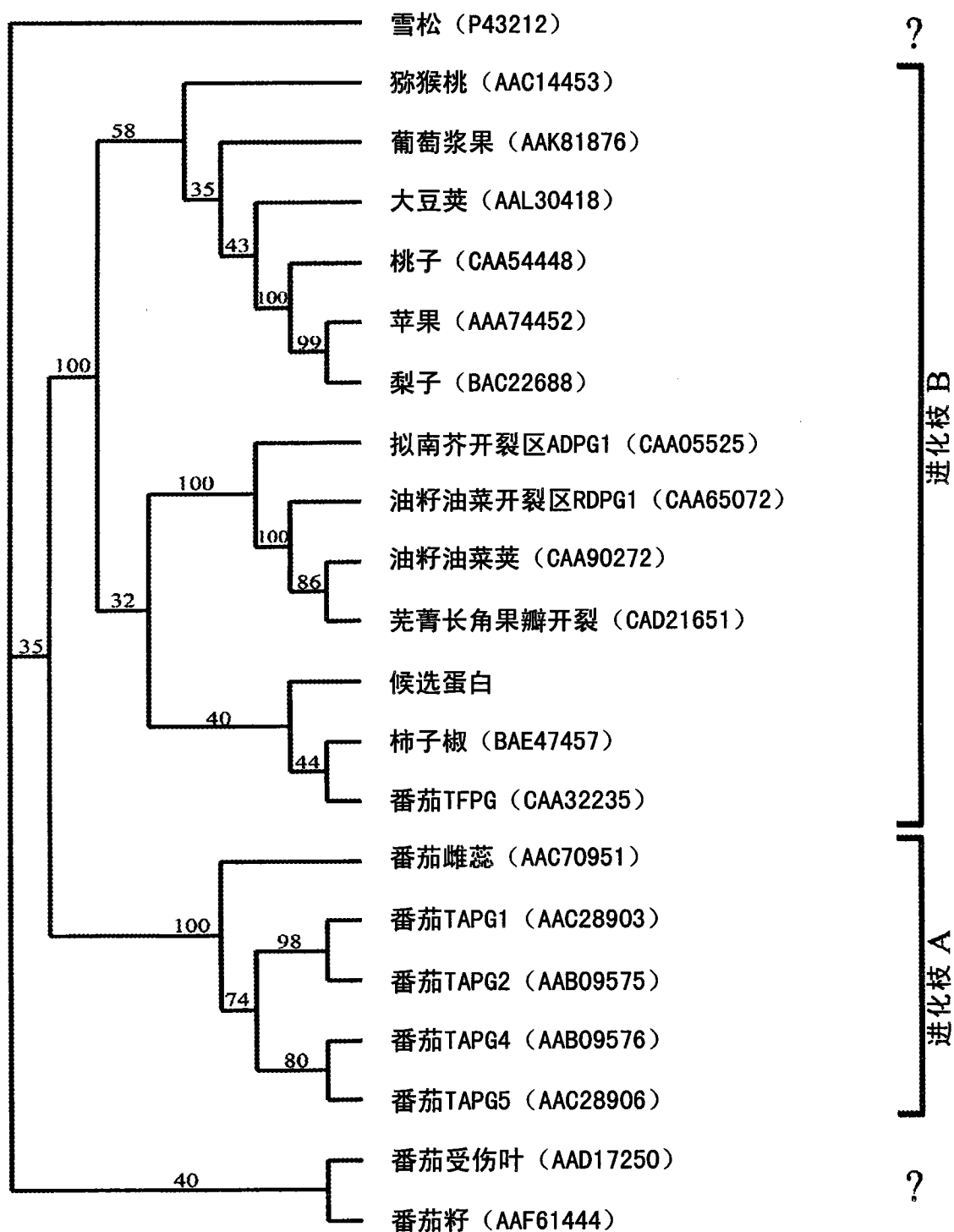


图 5

```

TEPG:  CSSRTPVQFVVPKNNKNNYLLKQIIFESGFCPSSISVKIIFGSLFASSKIISDYK-DRRLWIAFD 59
ORF4:  CSSSHVVNFVVSQNKKYLLKPIKIFYGFCSSITMQIYGLLASDDTSDYKSDSRHWIIFD 60
ps-2:  CSSSHVVNFVVSQNKKYLLKPIKIFYGFCSSITMQIYGLLASDDTSDYKSDSRHWIIFD 60

TEPG:  SVQNLVVGCGGTINNGQVWMPSSCKINKSLPCRDAPTALTETWCKNLKVNMLKSKNAQQ 119
ORF4:  SVQKLVVGCGAGVINNGKINWQHSCCKINKKLPCQKVAPTALIEYCKNLKVKELKIEAQQ 120
ps-2:  SVQKLVVGCGAGVINNGKINWQHSCCKINKKLPCQKVAPTVMIASQLWMDLRRS stop 112

TEPG:  IHHKTESCTNVVASNMENINASAKSPNTDGVHVSNTQYIQISDIHIGTGDDCISIVSGSQN 179
ORF4:  IHHLEKQVGVGVTKIIVTSPENSPNTDGHITSTQNIQISDSTIATGDDCISIVDGSQK 180

TEPG:  VQATNLTCCGPGHGISIGSLGSGNSEAYVSNVTVNEAKIIGAENGVRIKFTWQGGSGQASNI 239
ORF4:  VLAATGTTCCGPGHGISIGSLGSGNSEAHVSDIHVNGAKHYETTINGIRIKFTWPGGFGSASNI 240

TEPG:  KEILNVEMQDVKYPIIIDQNYCDRVE-PCIQQFS-AVQVKNVWYENIKGTSATKVAIKEDC 297
ORF4:  KYQNVVANNVKNPIIIDQNYCDOADGECKAETDSAVEVKNVVIYQNIKGTSATNDALSIKC 300

TEPG:  STNFPCEGLIMENININIGESGKPSATCKNVH 329
ORF4:  SKKTPCEGLIMENVRLIGG----- 319

```

图 6

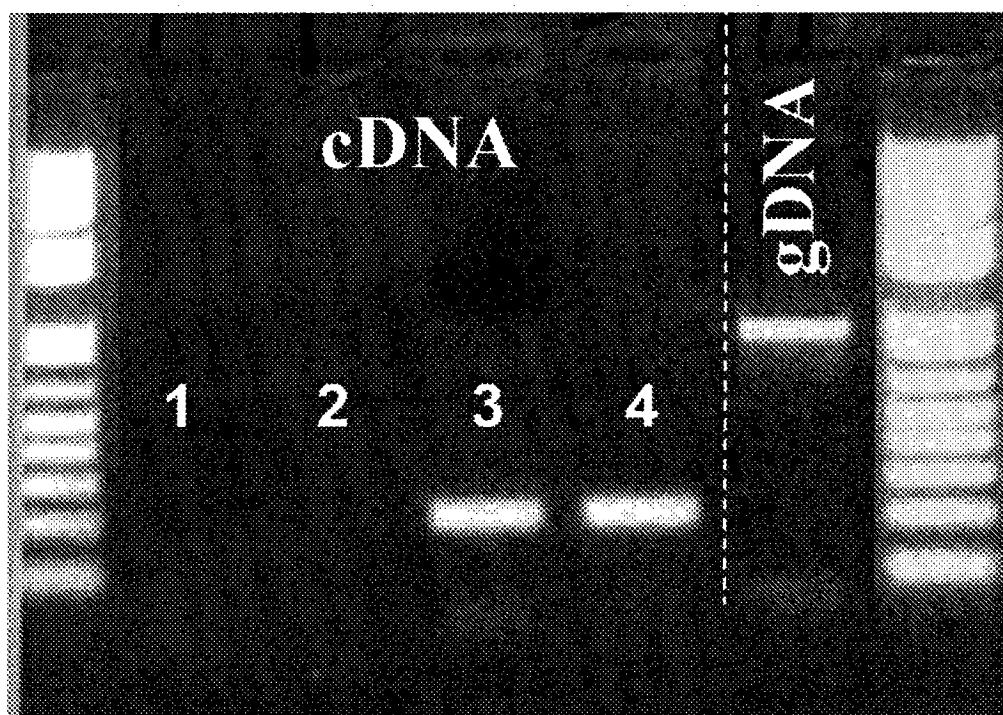


图 7

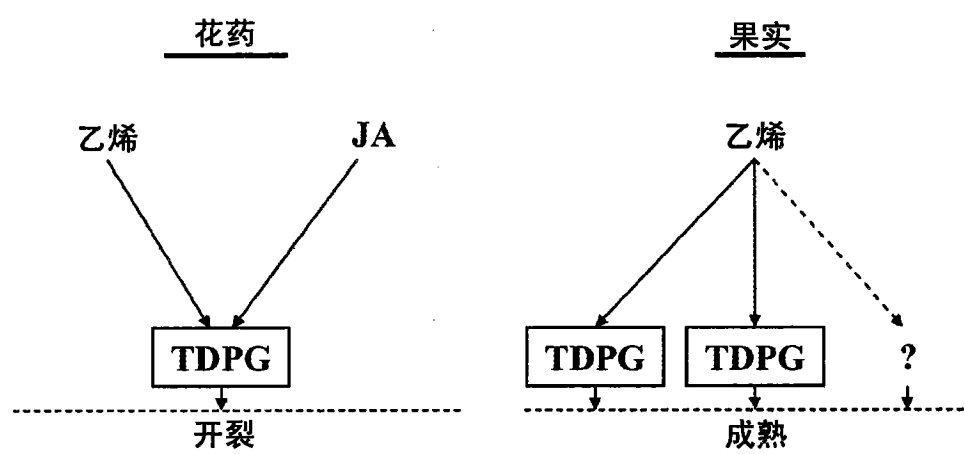


图 8