

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200610117721.6

[51] Int. Cl.

C07K 14/415 (2006.01)

C12N 15/29 (2006.01)

C12N 15/63 (2006.01)

[43] 公开日 2008年5月7日

[11] 公开号 CN 101173000A

[22] 申请日 2006.10.30

[21] 申请号 200610117721.6

[71] 申请人 中国科学院上海生命科学研究院

地址 200031 上海市岳阳路320号

[72] 发明人 何祖华 王二涛 李群 朱旭东

王建军 王林友

[74] 专利代理机构 上海专利商标事务所有限公司

代理人 徐迅

权利要求书1页 说明书21页 附图4页

[54] 发明名称

作物籽粒灌浆基因 G1F1 及其应用

[57] 摘要

本发明公开了一种作物籽粒灌浆基因 G1F1 及其应用，G1F1 基因可用于控制作物籽粒灌浆、提高作物的产量或品质、或提高作物籽粒的抗病性或耐储藏性。本发明还公开了一种改良作物的方法。G1F1 是一个在作物产量，品质，粮食储存和抗病上有重要应用价值的基因。

1. 一种分离的作物籽粒灌浆蛋白，其特征在于，该蛋白选自下组：
 - (a) 具有SEQ ID NO: 2氨基酸序列的多肽；或
 - (b) 将SEQ ID NO: 2氨基酸序列经过一个或多个氨基酸残基的取代、缺失或添加而形成的，且具有促进作物籽粒灌浆功能的由(a)衍生的多肽。
2. 一种分离的多核苷酸，其特征在于，该多核苷酸选自下组：
 - (i) 编码权利要求1所述的作物籽粒灌浆蛋白的多核苷酸；或
 - (ii) 与(i)中的多核苷酸互补的多核苷酸。
3. 如权利要求2所述的多核苷酸，其特征在于，该多核苷酸编码具有SEQ ID NO: 2所示氨基酸序列的多肽。
4. 如权利要求2所述的多核苷酸，其特征在于，该多核苷酸具有选自下组的序列：
 - (1) SEQ ID NO: 1所示的核苷酸序列；
 - (2) SEQ ID NO: 3所示的核苷酸序列；或
 - (3) 与(1)或(2)所示的核苷酸序列互补的核苷酸序列。
5. 一种载体，其特征在于，它含有权利要求2所述的多核苷酸。
6. 一种遗传工程化的宿主细胞，其特征在于，它含有权利要求5所述的载体；或它的基因组中整合有权利要求2所述的多核苷酸。
7. 权利要求1所述的作物籽粒灌浆蛋白或其编码基因的用途，其特征在于，用于：
 - 调节作物籽粒灌浆；
 - 调节作物籽粒的糖代谢或积累；或
 - 提高作物籽粒的抗病性或耐储藏性。
8. 一种改良作物的方法，其特征在于，该方法包括：
 - 增加所述作物中作物籽粒灌浆基因的表达。
9. 一种制备转基因植物的方法，其特征在于，它包括步骤：
 - 将权利要求2所述的多核苷酸导入植物细胞或组织中，培养所述的植物细胞或组织，再生成植物。
10. 一种权利要求1所述的作物籽粒灌浆蛋白或其编码基因的激动剂或拮抗剂。

作物籽粒灌浆基因 *GIF1* 及其应用

技术领域

本发明属于基因技术和植物学领域。更特别的，本发明涉及新的作物籽粒灌浆基因 *GIF1* 及其应用。

背景技术

当前，对于农作物产量改良的研究主要集中在以下几个方面：1. 增加作物的源，即加强作物的光合作用；2. 增加作物库的大小；3. 提高作物光合产物由源向库运输的能力。其中，增大库的容量和提高光合产物向库运输的能力是有效的育种途径。

尽管人们采取了许多方式来提高农作物的产量和对农作物进行改良，然而，目前还缺乏有效的手段。例如，我国的主要粮食作物水稻，当前的许多水稻高产栽培品种尤其超级杂交稻和大穗大粒品种存在籽粒灌浆不饱满的情况，这在很大程度上影响了水稻产量的进一步提高。

因此，本领域迫切需要找到有效的手段，解决农作物籽粒灌浆不饱满的问题，从而进一步改良农作物，实现农作物产量和品质的提高。

发明内容

本发明的目的在于提供一种新的作物籽粒灌浆基因及其应用。

在本发明的第一方面，提供一种分离的作物籽粒灌浆蛋白，所述蛋白选自下组：

- (a) 具有SEQ ID NO: 2氨基酸序列的多肽；或
- (b) 将SEQ ID NO:2氨基酸序列经过一个或多个氨基酸残基的取代、缺失或添加而形成的，且具有促进作物籽粒灌浆功能的由(a)衍生的多肽。

在本发明的第二方面，提供一种分离的多核苷酸，该多核苷酸选自下组：

- (i) 编码所述的作物籽粒灌浆蛋白的多核苷酸；或
- (ii) 与(i)中的多核苷酸互补的多核苷酸。

在本发明的另一优选例中，该多核苷酸编码具有SEQ ID NO: 2所示氨基酸

序列的多肽。

在本发明的另一优选例中，该多核苷酸具有选自下组的序列：

- (1) SEQ ID NO: 1所示的核苷酸序列；
- (2) SEQ ID NO: 3所示的核苷酸序列；或
- (3) 与(1)或(2)所示的核苷酸序列互补的核苷酸序列。

在本发明的第三方面，提供一种载体，它含有所述的多核苷酸。

在本发明的第四方面，提供一种遗传工程化的宿主细胞，它含有所述的载体；或它的基因组中整合有所述的多核苷酸。

在本发明的第五方面，提供所述的作物籽粒灌浆蛋白或其编码基因的用途，用于：

调节作物籽粒灌浆(更优选的，为促进作物籽粒灌浆)；

调节作物籽粒的糖代谢或积累；或

提高作物籽粒的抗病性或耐储藏性。

在本发明的第六方面，提供一种改良作物的方法，该方法包括：

增加所述作物中作物籽粒灌浆基因的表达。

在本发明的第七方面，提供一种制备转基因植物的方法，它包括步骤：

将所述的多核苷酸导入植物细胞或组织中，培养所述的植物细胞或组织，再生成植物。

在本发明的另一优选例中，所述的方法包括步骤：

(a) 提供携带表达载体的农杆菌，所述的表达载体含有所述的籽粒灌浆蛋白的编码基因；

(b) 将作物细胞、组织或器官与步骤(a)中的农杆菌接触，从而使该籽粒灌浆蛋白DNA编码序列转入作物细胞、组织或器官，并且整合到作物细胞的染色体上；

(c) 将转入了所述籽粒灌浆蛋白DNA编码序列的作物细胞、组织或器官再生成作物植株。

在本发明的第八方面，提供一种所述的作物籽粒灌浆蛋白或其编码基因的激动剂或拮抗剂。

本发明的其它方面由于本文的公开内容，对本领域的技术人员而言是显而易见的。

附图说明

图 1 显示了 *GIF1* 基因对于米质的影响，其中，图 1A 表示野生型植株的糙米颗粒，图 1B 表示突变型植株的糙米颗粒，图 1C 表示表示野生型植株的精米颗粒，图 1D 表示突变型植株的精米颗粒。

图 2 显示了 *GIF1* 基因对于种子活力的影响，其中，WT(左)表示 ZH11 野生型植株，*gif1*(右)表示 ZH11 突变型植株。

图 3 显示了 *GIF* 基因对植株穗部的影响，其中，WT 表示 ZH11 野生型植株的穗，F1 表示 ZH11 野生型与 ZH11 突变型杂交的第 1 代的穗，*gif1* 表示 WT 表示 ZH11 突变型植株的穗。

图 4 显示了含有突变型 *gif1* 基因的植株的种子上分离获得的代表性的仓储病害，图 4A 和图 4C(对图 4A 的放大)为分离得到的根霉菌(*Rhizopus sp.*)；图 4B 和图 4D (对图 4B 的放大)为分离得到的链格孢菌(*Alternaria sp.*)。

图 5 显示了 *GIF1* 基因的组织特异性。其中，A, 根；B, 节间；C, 伸长的节间；D, 2 天；E, 4 天；F, 6 天；G, 10 天；H, 15 天；I, 25 天；J, 颖壳；K, 10 天种子横切面；L, 突变体和野生型 *GIF1* 基因的表达情况；M, *GIF1* 基因在幼苗，叶，根，节间，穗的表达情况(其中，LS 代表水稻幼苗、L 代表水稻叶子、R 代表水稻根、I 代表水稻节间、P 代表水稻穗子)；N, *GIF1* 基因在开花后不同时间的表达情况；DAF(Day After Flowering)指开花后天数。

图 6 显示了 *GIF1* 基因对籽粒的糖代谢与积累的影响，其中总淀粉包括直链淀粉和支链淀粉。

具体实施方式

本发明人经过深入的研究，首次发现了一个控制作物籽粒灌浆的基因，该基因的不表达或降低表达可严重影响作物籽粒的灌浆，降低种子粒重。通过精细定位，本发明人克隆了该基因，并将之命名为作物籽粒灌浆基因 (*GIF1*, *Grain Incomplete Filling 1*)。试验证实，*GIF1* 基因正常表达的野生型植株的籽粒正常生长，而 *GIF1* 基因发生突变、不表达 *GIF1* 蛋白后，植株的米质差，种子活力低、对仓储病害的抗性差。在此基础上完成了本发明。

如本文所用，所述的“作物”包括但不限于：禾本科植物。更优选的，所述的禾本科植物包括但不限于：小麦、大麦、玉米、高粱等。

如本文所用，“分离的”是指物质从其原始环境中分离出来(如果是天然物质，原始环境即是天然环境)。如活体细胞内的天然状态下的多聚核苷酸和多肽是没有分离纯化的，但同样的多聚核苷酸或多肽如从天然状态中同存在的其他物质中分开，则为分离纯化的。

如本文所用，“分离的 GIF1 蛋白”或“分离的 GIF1 多肽”是指 GIF1 蛋白基本上不含天然与其相关的其它蛋白、脂类、糖类或其它物质。本领域的技术人员能用标准的蛋白质纯化技术纯化 GIF1 蛋白。基本上纯的多肽在非还原聚丙烯酰胺凝胶上能产生单一的主带。

本发明的多肽可以是重组多肽、天然多肽、合成多肽，优选的是重组多肽。本发明的多肽可以是天然纯化的产物，或是化学合成的产物，或使用重组技术从原核或真核宿主(例如，细菌、酵母、高等植物、昆虫和哺乳动物细胞)中产生。根据重组生产方案所用的宿主，本发明的多肽可以是糖基化的，或可以是非糖基化的。本发明的多肽还可包括或不包括起始的甲硫氨酸残基。

本发明还包括 GIF1 蛋白的片段、衍生物和类似物。如本文所用，术语“片段”、“衍生物”和“类似物”是指基本上保持本发明的 GIF1 蛋白相同的生物学功能或活性的多肽。本发明的多肽片段、衍生物或类似物可以是(i)有一个或多个保守或非保守性氨基酸残基(优选保守性氨基酸残基)被取代的多肽，而这样的取代的氨基酸残基可以是也可以不是由遗传密码编码的，或(ii)在一个或多个氨基酸残基中具有取代基团的多肽，或(iii)成熟多肽与另一个化合物(比如延长多肽半衰期的化合物，例如聚乙二醇)融合所形成的多肽，或(iv)附加的氨基酸序列融合到此多肽序列而形成的多肽(如前导序列或分泌序列或用来纯化此多肽的序列或蛋白原序列，或融合蛋白)。根据本文的定义这些片段、衍生物和类似物属于本领域熟练技术人员公知的范围。

在本发明中，术语“GIF1 蛋白”指具有 GIF1 蛋白活性的 SEQ ID NO: 2 序列的多肽。该术语还包括具有与 GIF1 蛋白相同功能的、SEQ ID NO:2 序列的变异形式。这些变异形式包括(但不限于): 若干个(通常为 1-50 个，较佳地 1-30 个，更佳地 1-20 个，最佳地 1-10 个，还更佳如 1-8 个、1-5 个)氨基酸的缺失、插入和/或取代，以及在 C 末端和/或 N 末端添加一个或数个(通常为 20 个以内，较佳地为 10 个以内，更佳地为 5 个以内)氨基酸。例如，在本领域中，用性能相近或相似的氨基酸进行取代时，通常不会改变蛋白质的功能。又比如，在 C 末端和/或 N 末端添加一个或数个氨基酸通常也不会改变蛋白质的

功能。该术语还包括 GIF1 蛋白的活性片段和活性衍生物。

多肽的变异形式包括：同源序列、保守性变异体、等位变异体、天然突变体、诱导突变体、在高或低的严紧度条件下能与 GIF1 蛋白 DNA 杂交的 DNA 所编码的蛋白、以及利用抗 GIF1 蛋白的抗血清获得的多肽或蛋白。本发明还提供了其他多肽，如包含 GIF1 蛋白或其片段的融合蛋白。除了几乎全长的多肽外，本发明还包括了 GIF1 蛋白的可溶性片段。通常，该片段具有 GIF1 蛋白序列的至少约 20 个连续氨基酸，通常至少约 30 个连续氨基酸，较佳地至少约 50 个连续氨基酸，更佳地至少约 80 个连续氨基酸，最佳地至少约 100 个连续氨基酸。

发明还提供 GIF1 蛋白或多肽的类似物。这些类似物与天然 GIF1 蛋白的差别可以是氨基酸序列上的差异，也可以是不影响序列的修饰形式上的差异，或者兼而有之。这些多肽包括天然或诱导的遗传变异体。诱导变异体可以通过各种技术得到，如通过辐射或暴露于诱变剂而产生随机诱变，还可通过定点诱变法或其他已知分子生物学的技术。类似物还包括具有不同于天然 L-氨基酸的残基(如 D-氨基酸)的类似物，以及具有非天然存在的或合成的氨基酸(如 β 、 γ -氨基酸)的类似物。应理解，本发明的多肽并不限于上述列举的代表性的多肽。

修饰(通常不改变一级结构)形式包括：体内或体外的多肽的化学衍生形式如乙酰化或羧基化。修饰还包括糖基化。修饰形式还包括具有磷酸化氨基酸残基(如磷酸酪氨酸，磷酸丝氨酸，磷酸苏氨酸)的序列。还包括被修饰从而提高了其抗蛋白水解性能或优化了溶解性能的多肽。

在本发明中，“GIF1 蛋白保守性变异多肽”指与 SEQ ID NO:2 的氨基酸序列相比，有至多 20 个，较佳地至多 10 个，更佳地至多 5 个，最佳地至多 3 个氨基酸被性质相似或相近的氨基酸所替换而形成多肽。这些保守性变异多肽最好根据表 1 进行氨基酸替换而产生。

表 1

氨基酸残基	代表性的取代	优选的取代
Ala (A)	Val; Leu; Ile	Val
Arg (R)	Lys; Gln; Asn	Lys
Asn (N)	Gln; His; Lys; Arg	Gln
Asp (D)	Glu	Glu
Cys (C)	Ser	Ser
Gln (Q)	Asn	Asn
Glu (E)	Asp	Asp
Gly (G)	Pro; Ala	Ala

His (H)	Asn; Gln; Lys; Arg	Arg
Ile (I)	Leu; Val; Met; Ala; Phe	Leu
Leu (L)	Ile; Val; Met; Ala; Phe	Ile
Lys (K)	Arg; Gln; Asn	Arg
Met (M)	Leu; Phe; Ile	Leu
Phe (F)	Leu; Val; Ile; Ala; Tyr	Leu
Pro (P)	Ala	Ala
Ser (S)	Thr	Thr
Thr (T)	Ser	Ser
Trp (W)	Tyr; Phe	Tyr
Tyr (Y)	Trp; Phe; Thr; Ser	Phe
Val (V)	Ile; Leu; Met; Phe; Ala	Leu

本发明还提供了编码本发明 GIF1 蛋白或其保守性变异多肽的多核苷酸序列。

本发明的多核苷酸可以是 DNA 形式或 RNA 形式。DNA 形式包括 cDNA、基因组 DNA 或人工合成的 DNA。DNA 可以是单链的或是双链的。DNA 可以是编码链或非编码链。编码成熟多肽的编码区序列可以与 SEQ ID NO: 1 或 SEQ ID NO: 3 所示的编码区序列相同或者是简并的变异体。如本文所用, “简并的变异体”在本发明中是指编码具有 SEQ ID NO:2 的蛋白质, 但与 SEQ ID NO: 1 或 SEQ ID NO: 3 所示的编码区序列有差别的核酸序列。

编码 SEQ ID NO:2 的成熟多肽的多核苷酸包括: 只编码成熟多肽的编码序列; 成熟多肽的编码序列和各种附加编码序列; 成熟多肽的编码序列(和任选的附加编码序列)以及非编码序列。

术语“编码多肽的多核苷酸”可以是包括编码此多肽的多核苷酸, 也可以是还包括附加编码和/或非编码序列的多核苷酸。

本发明还涉及上述多核苷酸的变异体, 其编码与本发明有相同的氨基酸序列的多肽或多肽的片段、类似物和衍生物。此多核苷酸的变异体可以是天然发生的等位变异体或非天然发生的变异体。这些核苷酸变异体包括取代变异体、缺失变异体和插入变异体。如本领域所知的, 等位变异体是一个多核苷酸的替换形式, 它可能是一个或多个核苷酸的取代、缺失或插入, 但不会从实质上改变其编码的多肽的功能。

本发明还涉及与上述的序列杂交且两个序列之间具有至少 50%, 较佳地至少 70%, 更佳地至少 80%相同性的多核苷酸。本发明特别涉及在严格条件下与本发明所述多核苷酸可杂交的多核苷酸。在本发明中, “严格条件”是指:(1)

在较低离子强度和较高温度下的杂交和洗脱，如 $0.2 \times \text{SSC}$ ， $0.1\% \text{SDS}$ ， 60°C ；或(2)杂交时加有变性剂，如 $50\%(\text{v/v})$ 甲酰胺， 0.1% 小牛血清/ $0.1\% \text{Ficoll}$ ， 42°C 等；或(3)仅在两条序列之间的相同性至少在 90% 以上，更好是 95% 以上时才发生杂交。并且，可杂交的多核苷酸编码的多肽与 SEQ ID NO:2 所示的成熟多肽有相同的生物学功能和活性。

本发明还涉及与上述的序列杂交的核酸片段。如本文所用，“核酸片段”的长度至少含 15 个核苷酸，较好是至少 30 个核苷酸，更好是至少 50 个核苷酸，最好是至少 100 个核苷酸以上。核酸片段可用于核酸的扩增技术(如 PCR)以确定和/或分离编码 GIF1 蛋白的多聚核苷酸。

应理解，虽然本发明的 GIF1 基因优选获自水稻，但是获自其它植物的与水稻 GIF1 基因高度同源(如具有 80% 以上，如 85% 、 90% 、 95% 、甚至 98% 序列相同性)的其它基因也在本发明考虑的范围之内。比对序列相同性的方法和工具也是本领域周知的，例如 BLAST。

本发明的 GIF1 蛋白核苷酸全长序列或其片段通常可以用 PCR 扩增法、重组法或人工合成的方法获得。对于 PCR 扩增法，可根据本发明所公开的有关核苷酸序列，尤其是开放阅读框序列来设计引物，并用市售的 cDNA 库或按本领域技术人员已知的常规方法所制备的 cDNA 库作为模板，扩增而得有关序列。当序列较长时，常常需要进行两次或多次 PCR 扩增，然后再将各次扩增出的片段按正确次序拼接在一起。

一旦获得了有关的序列，就可以用重组法来大批量地获得有关序列。这通常是将其克隆入载体，再转入细胞，然后通过常规方法从增殖后的宿主细胞中分离得到有关序列。

此外，还可用人工合成的方法来合成有关序列，尤其是片段长度较短时。通常，通过先合成多个小片段，然后再进行连接可获得序列很长的片段。

目前，已经可以完全通过化学合成来得到编码本发明蛋白(或其片段，或其衍生物)的 DNA 序列。然后可将该 DNA 序列引入本领域中已知的各种现有的 DNA 分子(或如载体)和细胞中。此外，还可通过化学合成将突变引入本发明蛋白序列中。

本发明也涉及包含本发明的多核苷酸的载体，以及用本发明的载体或 GIF1 蛋白编码序列经基因工程产生的宿主细胞，以及经重组技术产生本发明所述多肽的方法。

通过常规的重组 DNA 技术 (Science, 1984; 224: 1431), 可利用本发明的多聚核苷酸序列可用来表达或生产重组的 GIF1 蛋白。一般来说有以下步骤:

- (1). 用本发明的编码 GIF1 蛋白的多核苷酸(或变异体), 或用含有该多核苷酸的重组表达载体转化或转导合适的宿主细胞;
- (2). 在合适的培养基中培养的宿主细胞;
- (3). 从培养基或细胞中分离、纯化蛋白质。

本发明中, GIF1 蛋白多核苷酸序列可插入到重组表达载体中。术语“重组表达载体”指本领域熟知的细菌质粒、噬菌体、酵母质粒、植物细胞病毒、哺乳动物细胞病毒或其他载体。总之, 只要能在宿主体内复制和稳定, 任何质粒和载体都可以用。表达载体的一个重要特征是通常含有复制起点、启动子、标记基因和翻译控制元件。

本领域的技术人员熟知的方法能用于构建含 GIF1 蛋白编码 DNA 序列和合适的转录/翻译控制信号的表达载体。这些方法包括体外重组 DNA 技术、DNA 合成技术、体内重组技术等。所述的 DNA 序列可有效连接到表达载体中的适当启动子上, 以指导 mRNA 合成。表达载体还包括翻译起始用的核糖体结合位点和转录终止子。

此外, 表达载体优选地包含一个或多个选择性标记基因, 以提供用于选择转化的宿主细胞的表型性状, 如真核细胞培养用的二氢叶酸还原酶、新霉素抗性以及绿色荧光蛋白(GFP), 或用于大肠杆菌的卡那霉素或氨苄青霉素抗性。

包含上述的适当 DNA 序列以及适当启动子或者控制序列的载体, 可以用于转化适当的宿主细胞, 以使其能够表达蛋白质。

宿主细胞可以是原核细胞, 如细菌细胞; 或是低等真核细胞, 如酵母细胞; 或是高等真核细胞, 如植物细胞。代表性例子有: 大肠杆菌, 链霉菌属、农杆菌; 真菌细胞如酵母; 植物细胞等。

本发明的多核苷酸在高等真核细胞中表达时, 如果在载体中插入增强子序列时将会使转录得到增强。增强子是 DNA 的顺式作用因子, 通常大约有 10 到 300 个碱基对, 作用于启动子以增强基因的转录。

本领域一般技术人员都清楚如何选择适当的载体、启动子、增强子和宿主细胞。

用重组 DNA 转化宿主细胞可用本领域技术人员熟知的常规技术进行。当宿主为原核生物如大肠杆菌时, 能吸收 DNA 的感受态细胞可在指数生长期后收获,

用 CaCl_2 法处理, 所用的步骤在本领域众所周知。另一种方法是使用 MgCl_2 。如果需要, 转化也可用电穿孔的方法进行。当宿主是真核生物, 可选用如下的 DNA 转染方法: 磷酸钙共沉淀法, 常规机械方法如显微注射、电穿孔、脂质体包装等。转化植物也可使用农杆菌转化或基因枪转化等方法, 例如叶盘法、水稻幼胚转化法等。对于转化的植物细胞、组织或器官可以用常规方法再生成植株, 从而获得籽粒灌浆能力或籽粒品质发生改变的植物。

获得的转化子可以用常规方法培养, 表达本发明的基因所编码的多肽。根据所用的宿主细胞, 培养中所用的培养基可选自各种常规培养基。在适于宿主细胞生长的条件下进行培养。当宿主细胞生长到适当的细胞密度后, 用合适的方法(如温度转换或化学诱导)诱导选择的启动子, 将细胞再培养一段时间。

在上面的方法中的重组多肽可在细胞内、或在细胞膜上表达、或分泌到细胞外。如果需要, 可利用其物理的、化学的和其它特性通过各种分离方法分离和纯化重组的蛋白。这些方法是本领域技术人员所熟知的。这些方法的例子包括但不限于: 常规的复性处理、用蛋白沉淀剂处理(盐析方法)、离心、渗透破菌、超处理、超离心、分子筛层析(凝胶过滤)、吸附层析、离子交换层析、高效液相层析(HPLC)和其它各种液相层析技术及这些方法的结合。

重组的 GIF1 蛋白或多肽有多方面的用途。例如用于筛选促进或对抗 GIF1 蛋白功能的抗体、多肽或其它配体。用表达的重组 GIF1 蛋白筛选多肽库可用于寻找有价值的能抑制或刺激 GIF1 蛋白功能的多肽分子。

本发明的多核苷酸的一部分或全部可作为探针固定在微阵列(microarray)或 DNA 芯片(又称为“基因芯片”)上, 用于分析组织中基因的差异表达分析。用 GIF1 蛋白特异的引物进行 RNA-聚合酶链反应(RT-PCR)体外扩增也可检测 GIF1 蛋白的转录产物。

本发明还涉及一种改良作物的方法, 该方法包括增加所述植物中 GIF1 基因或其同源基因的表达。

增加 GIF1 基因或其同源基因表达的方法是本领域周知的。例如, 可通过用强启动子驱动从而增强 GIF1 基因或其同源基因的表达。或者通过增强子(如水稻 waxy 基因第一内含子、Actin 基因第一内含子等)来增强该 GIF1 基因的表达。适用于本发明方法的强启动子包括但不限于: 35s 启动子、水稻、玉米的 Ubi 启动子等。

作为本发明的一种优选方式, 所述的改良作物的方法包括步骤:

(1) 提供携带表达载体的农杆菌，所述的表达载体含有 *GIF1* 蛋白 DNA 编码序列；

(2) 将植物细胞或组织或器官与步骤(1)中的农杆菌接触，从而使该 *GIF1* 蛋白 DNA 编码序列转入植物细胞，并且整合到植物细胞的染色体上；

(3) 选择出转入所述 *GIF1* 蛋白 DNA 编码序列的植物细胞或组织；和

(4) 将步骤(3)中的植物细胞或组织再生成植株。

其中，可采用任何适当的常规手段，包括试剂、温度、压力条件等来实施此方法。

此外，本发明还涉及利用作物籽粒灌浆性状作为一种基因转化植株后代的追踪标记。此外，还可利用该基因的籽粒灌浆特性作为杂交制种过程中真杂种的指示标记。

在本发明的一个实例中，提供了一种 *GIF1* 基因，其基因组序列为 6840bp (SEQ ID NO: 3)，其中开放阅读框(ORF)位于 2380-2594, 3723-4605, 4994-5152, 5903-6168, 6276-6364, 6651-6840 位，全长 cDNA(SEQ ID NO: 1)为 1797bp，编码一个含有 598 个氨基酸的蛋白质(SEQ ID NO: 2)。所述的 *GIF1* 基因可为作物的品种改良提供新的途径，因而具有巨大的应用前景。

本发明的主要优点在于：

(1) 首次分离得到一种新的作物籽粒灌浆基因 *GIF1*，该基因具有控制种子灌浆的功能，同时也有调控品质的功能。

(2) 作物籽粒灌浆基因 *GIF1* 可以作为控制作物籽粒灌浆，提高产量和品质的一个基因，应用于作物品种的改良。

下面结合具体实施例，进一步阐述本发明。应理解，这些实施例仅用于说明本发明而不用来限制本发明的范围。下列实施例中未注明具体条件的实验方法，通常按照常规条件如 Sambrook 等人，分子克隆：实验室指南(New York: Cold Spring Harbor Laboratory Press, 1989)中所述的条件，或按照以下文献中公布的方法：Carl W. Dieffenbach 和 Gabriela S. Devksler eds. PCR Primer: A Laboratory Manual. Cold Spring Harbor Laboratory Press, 1995。或按照制造厂商所建议的条件。

材料和方法:

1. pCAMBIA1301-*GIF1* 的克隆和转化

GIF1 用 MunI、BamHI 酶切后, 连接到用 BamHI 和 EcoRI 酶切的 pBluescript sk+ 中(购自 Stratagene 公司), 得到的克隆再用 Hind III 和 Bam HI 酶切, 并克隆到用相同酶切的 pCAMBIA1301(参见 http://www.bios.net/daisy/cambia/585.html#dsy585_gus_intron) 中, 得到 pCAMBIA1301- *GIF1*。

将 pCAMBIA1301-*GIF1* 转入农杆菌:

1. 把 1 微克的 pCAMBIA1301- *GIF1* 的质粒加入到农杆菌 EHA105 (参见 Hood, E. E., Gelvin, S. B., Melchers, L. S. 和 Hoekema, A., *Transgenic Res.*, 1993, 2, 208 - 218) 的感受态中, 冰上放置 30 分钟;
2. 液氮冷冻 1 分钟;
3. 放到 37°C 的温水中直到液体融化;
4. 加入 1 毫升 YEP 在 28 度培养 2-4 小时;
5. 取 200 微升涂在含有抗生素的 YEP 平板上;
6. 28°C 放置两天直到阳性克隆出现, 选出阳性克隆。

2. 水稻成熟胚愈伤组织的诱导及转化

ZH11 和突变体 (*gif1*) 的种子去壳, 70% 乙醇浸泡 1min, 20%(v/v) NaClO 浸泡 20min, 并不停摇晃, 用无菌水冲洗 5-6 次, 使无异味且米粒呈现乳白色。无菌滤纸吸干后, 在 NBD/N6 培养基上诱导愈伤, 26°C 暗培养一周后, 剥愈伤组织, 剔去胚乳, 胚芽, 胚根等。

愈伤组织继代于 NBD/N6 培养基 (sigma 公司) 上暗培养, 每 2-3 周继代一次作为受体。

3. 农杆菌介导水稻愈伤组织的转化

1. 愈伤组织接种在 NBD/N6 培养基上 25-26°C 暗培养 4 天;
2. 配制 YEB CM 培养基待用;
3. 将含有重组质粒 pCAMBIA1301-*GIF1* 的农杆菌 EHA105 划 YEB 培养基 (含 Kan 50ul/ml, Rif 20ug/ml) 上, 28°C 200rpm 36hr;

4. 养至 OD660 为 1.0-1.5;
5. 将愈伤组织转移到一无菌三角瓶;
6. 将适量前述培养的农杆菌 EHA105 菌液倒入三角瓶 (保证浸没所有愈伤材料);
7. 室温放置 20min, 适量晃动;
8. 倒去菌液,用无菌滤纸吸去多余菌液,然后转移至 NBD 培养基 (+AS100) 上;
9. 20-25°C, 共培养 2-3 天;
10. 将共培养后的愈伤组织转移到无菌三角瓶,用含有 500mg/L 羧苄青霉素(Carbenicillin)的无菌水洗去农杆菌,2-3 次;
11. 然后转移到筛选培养基 (含有 NBD+Timent(特美丁) 200mg/l + 潮霉素(Hyg)50mg/l)上进行转化细胞的筛选,3 周为一周期,筛选 2-3 轮;
12. 2-3 周后,可将经过预分化的愈伤组织移到分化培养基(含有 NB + BAP 2mg/l + NAA 0.5mg/l) 上培养,16hr 光照,8hr 暗,26°C;
13. 2-3 周后,可将抗性再生植株转移至生根培养基(含有 1/2MS + NAA 0.5mg/l) 上壮苗生根;
14. 3 周后,再生抗性植株洗净琼脂,移栽于温室收种子并作为分子鉴定材料。

实施例 1 群体的构建、基因的克隆及功能研究

本发明人从水稻中花 11(ZH11)诱变突变体库中,发现了一个水稻突变体,该突变体的种子的灌浆受到较为严重的影响,营养生长和野生型相比没有变化,但籽粒灌浆受到严重影响,种子千粒重减少 15-30%,并且米质变差,本发明人将之命名为 *gif1* (*grain incomplete filling 1*),说明 *GIF1* 基因是一个重要的通过灌浆控制产量和米质的基因。

本发明人通过 *gif1* 突变体和珍汕 97 杂交得到基因定位群体。利用 BSA(Bulked Segregant Analysis)分析法,以均匀分布在水稻 12 条染色体上的 130 对 SSR 引物对 LIF1 位点进行扫描,将 *Lif1* 位点初步定位在 4 号染色体长臂 SRD5 附近,最终将 LIF1 基因定位在 caps-4 和 caps-8 之间的 32Kb 中,经预测这段区域共包含三个基因,本发明人通过对突变体和野生型的 DNA 的测序,发现在突变体 *gif1* 中有一个碱基的缺失(*GIF1* 基因组 DNA 序列第 4588 位)。

通过对 *GIF1* 的精细定位, 通过测序和功能验证, 得到野生型 *GIF1* 的基因组 (DNA) 序列见 SEQ ID NO: 3 (包括驱动子); *GIF1* 编码区的序列 (cDNA 序列) 见 SEQ ID NO: 1; *GIF1* 蛋白的序列见 SEQ ID NO: 2。

当发生前述的突变 (*GIF1* 基因组 DNA 序列第 4588 位缺失) 后, 突变型不表达 *GIF1* 蛋白。

实施例 2 *GIF1* 基因对种子灌浆和产量的影响

本发明人对比了从 *gif1* 突变型和 *GIF1* 野生型的中花 11 (ZH11) 植株获得的米粒的各项表型, 结果见表 2。

表 2 *gif1* 突变体对于米粒表型的影响

	ZH11	<i>gif1</i>	<i>gif1</i> / ZH11	显著性
穗子数目(个)/植物	11.00±2.30	9.96±2.73	0.90	非
种子数目(个)/穗子	122.12±33.71	124.82±30.70	1.02	非
种子数目(个)/植物	1343.29±372.12	1279.38±239.05	0.95	非
灌浆不饱满种子数目(个)/穗子	35.56±10.83	34.33±15.68	0.97	非
种子重(g)/穗子	2.90±0.73	2.45±0.63	0.84	是
种子重(g)/植物	32.14±8.60	25.09±4.44	0.78	是
千粒种子重(g)	24.00±0.01	19.71±0.01	0.82	是
千粒糙米重(g)	21.33±0.1	16.15±0.15	0.76	是

因此, *gif1* 突变体影响种子灌浆, 降低产量, 而转基因互补可以消除该影响, 并控制水稻的灌浆, 从而可实现提高产量。

实施例 3 *GIF1* 基因对于米质的影响

具有野生型 *GIF1* 基因的糙米和精米的颗粒大且饱满, 而 *GIF1* 基因不表达的突变型的颗粒小且不饱满。本发明人构建了含有 *GIF1* 基因的 pCAMBIA1301-*GIF1* 重组质粒, 并转化突变体 (*gif1*) 的愈伤组织, 得到突变型的

转基因植株，发现转基因植物能恢复到野生型的状态。

结果见图 1。由结果可见，具有野生型 *GIF1* 基因的糙米和精米的颗粒大且饱满，而 *GIF1* 基因不表达的突变型的颗粒小且不饱满。

因此，突变型 *gif1* 突变体降低米质，而转基因互补可以消除该影响，提高品质。

此外，发明人还进行了过量表达 *GIF1* 的植株与野生型植株的比较，发现过量表达 *GIF1* 可使籽粒比野生型更饱满。

实施例 4 *GIF1* 基因对于种子活力的影响

具有野生型 *GIF1* 基因的植株根系发达、叶子生长迅速，而具有突变型 *gif* 基因的植株根系明显不发达、叶子生长明显缓慢。本发明人构建了含有 *GIF1* 基因的 pCAMBIA1301-*GIF1* 重组质粒，并转化突变体(*gif1*)的愈伤组织，得到突变型的转基因植株，转基因植物的根系，叶子能恢复到野生型的状态。

结果见图 2，具有野生型 *GIF1* 基因的植株根系发达、叶子生长迅速，而具有突变型 *gif* 基因的植株根系明显不发达、叶子生长明显缓慢。

由结果可见，*gif1* 突变体降低种子活力，而转基因互补可以消除该影响。

实施例 5 *GIF1* 基因对于种子穗部和仓储病害的抗性

本发明人构建了含有 *GIF1* 基因的 pCAMBIA1301-*GIF1* 重组质粒，并转化突变体(*gif1*)的愈伤组织，得到突变型的转基因植株，转基因植物对仓储病害恢复到野生型的状态。

结果见图 3 和图 4。由图 3 可知，从含有野生型 *GIF* 基因的植株获得的麦穗颗粒饱满；而从含有突变型 *gif* 基因的植株获得的麦穗颗粒明显干瘪，且其穗子对病害更敏感。图 4 为从含有突变型 *gif* 基因的植株的种子上分离获得的代表性的仓储病害，图 4A 为分离得到的根霉菌 (*Rhizopus sp.*)；图 4B 为分离得到的链格孢菌 (*Alternaria sp.*)；此外，还分离到其它一些仓储病害。

由结果可见，*gif1* 突变体降低对穗部和仓储病害的抗性，缩短粮食储藏期；而转基因互补可以消除该影响，并提高抗性。

实施例 6 *GIF1* 基因的组织特异性

本发明人构建 *GIF1* 基因的启动子连 β -葡糖醛酸酶基因 (GUS) 的克隆，并转

化水稻 ZH11, 观察 *GIF1* 启动子启动 GUS 报告基因和 *GIF1* 基因的组织特异性表达。具体的构建方法如下:

用正向引物 `tataagcttgatcggccatactcc` (SEQ ID NO: 4) 和反向引物 `taggatcccttgctctcacacttg` (SEQ ID NO: 5), 以 *GIF1* 的基因组 DNA 为模板, PCR 得到 *GIF1* 基因的启动子, 并克隆载体 pBI101 (购自 Clonetechnology 公司, 带有 GUS), 再用 EcoR I、Hind III 酶切, 回收片断, 将该片断连入用相同酶切的 pCAMBIA1300 便得到了 pCAMBIA1300+ 启动子 +GUS 的相应克隆 (组织显色方法见 Jefferenson, RA (1987) Plant Mol Biol Rep)。

结果见图 5A-K, 可知 *GIF1* 基因在植株的根部、节间、种子的背腹部的微管束有特异性表达。

此外, 采用 RT-PCR 的方法, 从所述的野生型或突变型植株中抽提 mRNA, 通过 RT-PCR 分别扩增 *GIF1*, 将获得的扩增产物进行琼脂糖电泳检测。结果见图 5L-N, 其中, 图 5L 是突变体 *gif1* 和野生型的 mRNA 的 RT-PCR 检测结果; 图 5M 是水稻不同组织 mRNA 的 RT-PCR 检测结果; 图 5N 是不同时期 (DAF) 水稻穗部组织 mRNA 的 RT-PCR 检测结果。

实施例 7 *GIF1* 对籽粒的糖代谢与积累的影响

本发明人构建了含有 *GIF1* 基因的 pCAMBIA1301-*GIF1* 重组质粒, 并转化突变体 (*gif1*) 的愈伤组织, 得到突变型的转基因植株, 观察转基因植株的糖代谢与积累, 并与野生型的 ZH11 进行比较。

结果见图 6, 可见 *GIF1* 能够调节籽粒的糖代谢与积累, 从而调节米质。

在本发明提及的所有文献都在本申请中引用作为参考, 就如同每一篇文献被单独引用作为参考那样。此外应理解, 在阅读了本发明的上述讲授内容之后, 本领域技术人员可以对本发明作各种改动或修改, 这些等价形式同样落于本申请所附权利要求书所限定的范围。

序列表

<110> 中国科学院上海生命科学研究院

<120> 作物籽粒灌浆基因 GIF1 及其应用

<130> 065311

<160> 5

<170> PatentIn version 3.3

<210> 1

<211> 1797

<212> DNA

<213> 稻属 (*Oryza sativa* L.)

<400> 1

```

atgggagttc ttgtagtag ggtcgttgg gcatggctgg tccagctgct gctgctccag      60
cagctcgccg gagcgtcgca cgctgtctac gacgacctcg agctgcaggc ggctgctacc      120
acageggacg gcgtgccgcc gtccatcgct gactctgagc tccggactgg gtatcacttc      180
cagccacca agaactggat caatgatccg aacgcgccga tgtactaaa ggggtggtac      240
catctgttct accagtaca cccaagggc gccgtgtggg ggaacatcgt gtgggcgcac      300
tcagtgtcac gtgacctcat caactgggtg gcgctcaagc cggccatcga gccagcatc      360
agggccgaca agtacggctg ctggtcgggg tcggcgacga tgatggccga cgggacgccg      420
gtgatcatgt acaccggcgt caaccgcccc gacgtcaact accaggtgca gaacgtggcg      480
ctgccgagga acgggtcgga cccgctgctg cgcgagtggg tgaagcccgg ccacaacccg      540
gtgatcgtgc ccgagggcgg catcaacgcg acgcagtcc gcgaccgcac caccgcgtgg      600
cgcggggccg acggccactg gcggctgctc gtcggcagcc tcgcggggca gtcccgcggc      660
gtggcgtacg tgtaccggag cagggacttc cggcgggtga cgcgcgcggc gcagccgctg      720
cactcggcgc ccacggggat gtgggagtgc ccggacttct acccggtcac cgcggacggc      780
cgccgcgagg gcgtcgacac ctctgcccgc gtcgtcgacg ccgccgcctc ggcgcgcgtc      840
aagtacgtgc tcaagaacag cctcgacctg cgcgggtacg actactacac cgtcggaacg      900
tacgaccgga aggcgagcgc gtacgtgccg gacgaccccg ccggcgacga gcaccacatc      960
cgctacgact acggcaactt ctacgcctcc aagacgttct acgaccggc gaagcgcgcg      1020
cgcatectct ggggatgggc caacgagtcc gacaccgcg ccgacgacgt ggccaagggc      1080
tgggceggaa tccaggcgat tccaggaaa gtgtggctgg acccaagtgg gaagcaactg      1140
ttgcagtggc caatcgagga ggtcgagagg ctgagagggg agtggccggt cattctcaag      1200
gacagggtgg tcaagccagg ggaacacgtc gaggtgaccg ggctacaaac tgcacaggct      1260
gacgttgagg tgagcttcga ggtggggagc ctggaggcgg cggagcggct ggaccggcg      1320
atggcgtacg acgcgcagcg gctgtgcagc gcgcggggcg ccgacgcgag gggcggcgtg      1380
gggccgttcg gcctgtgggt gctcgcgtcc gcggggctgg aggagaagac cgcctgtttc      1440
ttcagggtgt tcaggccggc ggcgcgcggc ggcggcgcgg gcaagcccgt cgtgctcatg      1500
tgcaccgacc ccaccaagtc atcgcgcaac ccgaacatgt accagccgac gtttcgaggg      1560
ttcgttgaca cggacatcac caacgggaag atatctctga ggagcctgat cgacaggtcg      1620
gttgttgaga gcttcggggc tggaggaaag gcgtgcatcc tgtcgagggt gtaccgctcg      1680
ctggccatcg gcaagaacgc gcgcctttac gttttcaata acgggaaggc ggagatcaag      1740
gtgtcgcagc tcaccgcgtg ggagatgaag aagccggctc tgatgaatgg agcctaa      1797

```

<210> 2

<211> 598
 <212> PRT
 <213> 稻属 (*Oryza sativa* L.)

<400> 2

```

Met Gly Val Leu Gly Ser Arg Val Ala Trp Ala Trp Leu Val Gln Leu
1          5          10          15
Leu Leu Leu Gln Gln Leu Ala Gly Ala Ser His Val Val Tyr Asp Asp
          20          25          30
Leu Glu Leu Gln Ala Ala Ala Thr Thr Ala Asp Gly Val Pro Pro Ser
          35          40          45
Ile Val Asp Ser Glu Leu Arg Thr Gly Tyr His Phe Gln Pro Pro Lys
          50          55          60
Asn Trp Ile Asn Asp Pro Asn Ala Pro Met Tyr Tyr Lys Gly Trp Tyr
65          70          75          80
His Leu Phe Tyr Gln Tyr Asn Pro Lys Gly Ala Val Trp Gly Asn Ile
          85          90          95
Val Trp Ala His Ser Val Ser Arg Asp Leu Ile Asn Trp Val Ala Leu
          100          105          110
Lys Pro Ala Ile Glu Pro Ser Ile Arg Ala Asp Lys Tyr Gly Cys Trp
          115          120          125
Ser Gly Ser Ala Thr Met Met Ala Asp Gly Thr Pro Val Ile Met Tyr
          130          135          140
Thr Gly Val Asn Arg Pro Asp Val Asn Tyr Gln Val Gln Asn Val Ala
145          150          155          160
Leu Pro Arg Asn Gly Ser Asp Pro Leu Leu Arg Glu Trp Val Lys Pro
          165          170          175
Gly His Asn Pro Val Ile Val Pro Glu Gly Gly Ile Asn Ala Thr Gln
          180          185          190
Phe Arg Asp Pro Thr Thr Ala Trp Arg Gly Ala Asp Gly His Trp Arg
          195          200          205
Leu Leu Val Gly Ser Leu Ala Gly Gln Ser Arg Gly Val Ala Tyr Val
          210          215          220
Tyr Arg Ser Arg Asp Phe Arg Arg Trp Thr Arg Ala Ala Gln Pro Leu
225          230          235          240
His Ser Ala Pro Thr Gly Met Trp Glu Cys Pro Asp Phe Tyr Pro Val
          245          250          255
Thr Ala Asp Gly Arg Arg Glu Gly Val Asp Thr Ser Ser Ala Val Val
          260          265          270
Asp Ala Ala Ala Ser Ala Arg Val Lys Tyr Val Leu Lys Asn Ser Leu
          275          280          285
Asp Leu Arg Arg Tyr Asp Tyr Tyr Thr Val Gly Thr Tyr Asp Arg Lys
          290          295          300
Ala Glu Arg Tyr Val Pro Asp Asp Pro Ala Gly Asp Glu His His Ile
305          310          315          320
Arg Tyr Asp Tyr Gly Asn Phe Tyr Ala Ser Lys Thr Phe Tyr Asp Pro
          325          330          335
Ala Lys Arg Arg Arg Ile Leu Trp Gly Trp Ala Asn Glu Ser Asp Thr
          340          345          350
Ala Ala Asp Asp Val Ala Lys Gly Trp Ala Gly Ile Gln Ala Ile Pro

```


agagtctata	tagaaataca	attaagaaaa	aaaagaaatt	cggaattaaa	aaataaggaa	780
tattagaaat	agagtataga	gtctatatag	aaatacaatt	aagaaaaaaa	aatagaaatt	840
cggaattaaa	aaaatggaat	attagaatta	gagtatagag	tccatatagg	aatttaaaac	900
taactaaaat	ttggaataaa	cataataaaa	ttaaagtaga	gtttagagtc	cgtataaaaa	960
tacaatttac	aaataactaa	aattcgagat	taagaaaaat	atgggaagaa	gagtttaag	1020
tcaatataga	aatgcaattt	agaagtaact	gaaattcgaa	attaataaatt	aaagaatatt	1080
gaaagataag	tttagagtcc	acatagaaat	acaattaaaa	ataataaaaa	ttcagaaata	1140
aaaataaata	atattggaag	aagagcatag	agtctatata	gaaatacaat	ttacagaaaa	1200
ttcggaatta	aaaaatata	attaaaagac	gagtctagag	tgcatatagg	aatatatata	1260
atltacaaat	aattaaaaat	tgatattaaa	ataattaata	actaacacgt	atataaaata	1320
caatatgaat	attaccatt	agtagtttcg	taaagttatt	gcaaaaatta	aaattatatt	1380
gtcaatttaa	tatatitgaa	caatatattg	agaaaacata	tatgctatta	tatgagagaa	1440
aatataatag	ttcatagtga	attgtgaaca	ctgatttaaa	aacaaacaga	ttaacaacca	1500
catcgtttg	cttattcgtg	gaataagcta	aacggcata	ttgcaaacga	aaagtaattt	1560
gtgaataaaa	ttttatata	cgtgttctta	gcaatctaaa	atcaaagagt	gaaaaataaa	1620
cttcgatgaa	aaaaaccaa	aatcagcttc	aaatttaaa	ttaaaaattt	aagtttagct	1680
gataagtata	agtataaacg	aaaagatgat	gccgtaattt	tctagacatg	aaagatcaat	1740
agaacggatt	gacattttcg	taatggcctg	tagatagaga	tataagccac	gagaaggagc	1800
agtgctgctc	tcttctttac	aagctaactt	cagtgttttt	atacatgaaa	aatcgattg	1860
atatctttgg	gttggttcgg	attgataacg	atctccattc	aacttttgc	ttttgtttcc	1920
caatacgtta	tagatggctt	tgtgctccat	gcggtgacca	tgttaggaaa	aaaaagtata	1980
tggttttagg	cttgittgag	ctccagctta	taaatacagag	ccaatatgaa	gacaaatccc	2040
aatatatata	tatatatata	tatatatata	tatatatata	tatatatata	tatatatata	2100
aaaaacatga	gagaacatg	caggtcacgg	acgtcccagg	gcacaaataa	tgttaggctg	2160
tccacacagt	aatgcggccc	gcccggcct	gccattcttc	tgcccttttg	tccgactcag	2220
caatctttga	aatctcaacg	cttaggggaa	aaaaaacagc	gtcttttcca	taatgcatct	2280
cttcaccttc	gcagtataaa	taggaccccc	tctcctctgc	tcctgctcat	cctcctctcc	2340
tcttctcgtc	ctcacttctt	gtgcccag	gtgagagcaa	tgggagttct	tggtagtagg	2400
gtcgtttggg	catggtggt	ccagctgctg	ctgctccagc	agctcggcgg	agcgtcgcac	2460
gtcgtctacg	acgacctcga	gctgcaggcg	gctgctacca	cagcggacgg	cgtgcccgg	2520
tccatcgtcg	actctgagct	ccggactggg	tatcacttcc	agccaccaa	gaactggatc	2580
aatggtaatg	tgaactaact	gaaatgttgc	caacttgcca	ttgttcatgc	cagaacgccg	2640
gtcaggccgt	atgatttga	ggcataggg	caccacttgt	ggttgtggat	actggataga	2700
tgagcaaagg	gaacagagtg	ctctgttctt	gagaattgag	acgcagaatc	gtgcagagta	2760
actagtagag	ttttgacgac	gttgttgtgt	agaacatcac	ctgaactaaa	tggtcaact	2820
tgagtaattt	atagtcagag	ttgaaaatat	tgacatcata	gtcatatcaa	atgtttggca	2880
cacaacataa	attacggaca	gtacactaag	gcatcagttt	ttatgtccat	tttgtcgggt	2940
cagctagtag	agtcaacgtt	agcaccacag	cggtcacgct	gaaagaagta	gcttcagaag	3000
catctcacag	taaactactg	agagtttgcc	atctcttttt	catgaagctc	acacttagtc	3060
ccttcgaact	gttaacagat	gtactacctt	gttctacttt	tcttgctaat	gattcttgtg	3120
acaaggctta	gtcctaaccg	gcaattttct	tgtgcaatta	tttgggtggg	gtgtgctctg	3180
ctctacactg	tgattgctgc	tgctcatca	acattggaaa	cccgcagatc	cgaacggtac	3240
gtcgttttcc	cacctttat	aatatatcct	gtcacgaatc	tctgtctact	agtagtagta	3300
gtagtagtac	tagaactttt	atgccttgca	acttgcaatt	tcgttgtacg	ggagaggact	3360
gtagtttagt	acgccttca	tggtaggatt	aaaggttcaa	agcacatttt	agcacaaaa	3420
tggtaggcgc	actgggactc	cacatgcagg	cttgcttgtc	gaccgtgggg	tacctagccc	3480
ctaccacggc	tgatgaccac	aaagttcaga	aaatcttaac	ttctctcag	aaaggaatt	3540
agccaaaagt	tcacctttt	ctcgtacgaa	atgaagcatc	tatagttcta	taattaatcg	3600
tgagcagtgt	agagaaaaat	gcaatgtaca	cgcgcgatta	aactgaaatg	gtaattgatt	3660
tcaatgtact	actaagactg	aagatcattt	cttgatttgg	tgaaactgaa	cgggtgcatg	3720
cagcggcgt	gtactacaag	gggtgtacc	atctgttcta	ccagtacaac	cccaaggcgc	3780

ccgtgtgggg gaacatcgtg tgggcgcact cagtgtcacg tgacctcatc aactgggtgg 3840
 cgctcaagcc ggccatcgag cccagcatca gggccgacaa gtacggctgc tggcgggggt 3900
 cggcgacgat gatggccgac gggacgccgg tgatcatgta caccggcgtc aaccgccccg 3960
 acgtcaacta ccaggtgcag aacgtggcgc tgccgaggaa cgggtcggac ccgctgctgc 4020
 gcgagtgggt gaagcccggc cacaaccggg tgatcgtgcc cgagggcggc atcaacgcga 4080
 cgcagttccg cgacccgacc accgcgtggc gcggggcccga cggccactgg cggtgctcg 4140
 tcggcagcct cgcggggcag tcccgcggcg tggcgtacgt gtaccggagc agggacttcc 4200
 ggggtggac gcgcggcg cagccgctgc actcggcgc caggggatg tgggagtgcc 4260
 cggacttcta cccggtcacc gcgacggcc gccgcgaggg cgtcgacacc tcgtccgccg 4320
 tcgtcgacgc cgccgctcg gcgcgctca agtacgtgt caagaacagc ctcgacctgc 4380
 gccggtacga ctactacacc gtcggaacgt acgaccgaa ggccgagcgg tacgtgccgg 4440
 acgacccgc cggcgacgag caccacatcc gctacgacta cggcaacttc tacgcctcca 4500
 agacgttcta cgaccggcg aagcggccgc gcatcctctg gggatggcc aacgagtccg 4560
 acaccgccgc cgacgacgtg gccagggtt gggccggaat ccaggtaat aaccgcacgt 4620
 cctgactgca tacgtgcatg ccatttacgt gtccacatg catgctgcca tcttcagata 4680
 gtcaataatca ccataactc cctccgttct aaaatgttta acaccatga ctttttagca 4740
 catgtttgac cgttcgtctt attaaaaaaaa tatgaaatat ataaaactat atgtatacat 4800
 aaaagtatat ttaacaatga atcaaatgat atgaaaagaa caaataatta cttaaatttt 4860
 ttgaataaga cgaatgggtg caagtatttt gaaaaagag agtatatctt aaaagtcaaa 4920
 tggaacaaca ctagcagctc aattttgctg gtaatctttg attgaaatcgt gtgtttgtga 4980
 tgtgatgttt taggcgattc cgaggaaagt gtggctggac ccaagtggga agcaactggt 5040
 gcagtggcca atcgaggagg tcgagaggct gagagggaag tggccggtca ttctcaagga 5100
 cagggtggtc aagccagggg aacacgtcga ggtgaccggg ctacaaaactg cacaggtatt 5160
 cctttttgca tctgtaattc tgtaaaacta ttttttttac cccaaaaggg cattcgaata 5220
 aaactgctca cacatccatg gttctgtgca tgacagtagt aattattaat aagtatcct 5280
 gtttgttttg ctgtgtcctg gaccgatctt tatcttatct ggccagcctg aagtgtgtc 5340
 cagtgtgcag tgcccactga acaccaccta ctacgtgtgc cgtgtcgtt tcttctcgtc 5400
 cctttttacc atctctgca cactttgctc gtacttaact gatctcactg attctctcgt 5460
 catccgcgca tgtcacgtac aacttccagg ttgcagcgtg attagtgcac atatcactaa 5520
 gacactaac aacataatta gagagatagt taaggagctc aattaatgtg ctttgttgg 5580
 gacgtacgtg agtaggagct gtgatctctg atagcaagtt taatagtata gctaactact 5640
 ggctctaaat tatctatagt caatctaata ataaattcat ataatagtta cctataaaca 5700
 tataactaat aattaataca tggttccaca tgcatacac atatgcatct taaagtccgt 5760
 actataattt gctgtaaate tatagcttgt tgtttttctc tctctcttt tatctctcgt 5820
 atcgaaatgt gtttatagct ggcttatagt gtgctattgt ccctggctcgt atgaagtgat 5880
 catgcattct gtttgggtgg gtgcaggctg acgtggagggt gagcttcgag gtggggagcc 5940
 tggaggcggc ggagcggctg gaccggcga tggcgtacga cgcgcagcgg ctgtgcagcg 6000
 cgcgggggcg cgacgcgagg ggcggcgtgg ggccgttcgg cctgtgggtg ctcgcgtccg 6060
 cggggctgga ggagaagacc gccgtgttct tcagggtgtt caggccggcg gcgcgcggcg 6120
 gcggcgccgg caagcccgtc gtgctcatgt gcaccgacc caccaagtac gtgcggcttt 6180
 tgcactttat cggtgattga tcgactaca caataaacia aatattgctt tgactccgtt 6240
 tactgatttt ttggtatggt gcgtatgctg gcaggtcatc gcgcaaccg aacatgtacc 6300
 agccgacgtt tgcagggttc gttgacacgg acatcaccaa cggaagata tctctgagga 6360
 gcctggtagc taataggacc aaattatcgg gaaaaagga aatgtttgca tgacggtatc 6420
 ccgttcggat aaaattatac ctcttaataa ttgtccgata cctaataaat attaattggc 6480
 taataaacta ttgtaatggg atgatatctt tgaggatcgt tctgatacct atctgatagg 6540
 tacctcatag gtatcacctc gtccgaacgg gttactgttt tataacattc atctggaaaa 6600
 ggttcataaa ttgtagaata tgttttgata tcttgtgtct ctcttgtgca gatcgacagg 6660
 tcggttgttg agagcttcgg ggctggagga aaggcgtgca tctgtcagag ggtgtacccg 6720
 tcgctggcca tcggcaagaa cgcgcgctt tacgttttca ataacgggaa ggcggagatc 6780
 aagggtgcgc agctcaccgc gtgggagatg aagaagccgg tcatgatgaa tggagcctaa 6840

<210> 4
<211> 24
<212> DNA
<213> 人工序列

<220>
<221> misc_feature
<223> 引物

<400> 4
tataagcttg atcgccata ctcc

24

<210> 5
<211> 25
<212> DNA
<213> 人工序列

<220>
<221> misc_feature
<223> 引物

<400> 5
taggatccct ttgctctcac acttg

25

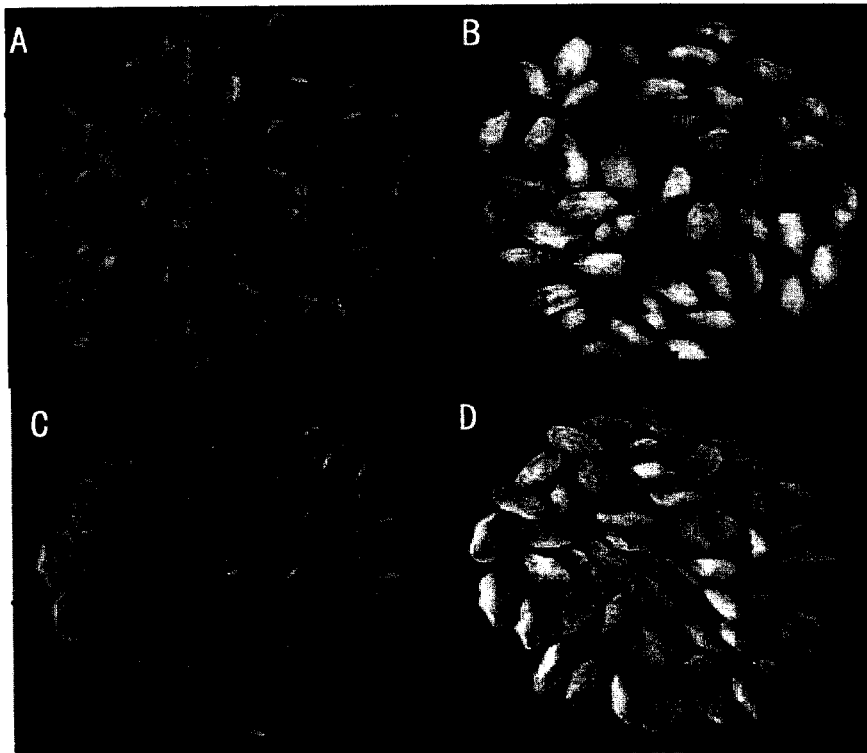


图 1

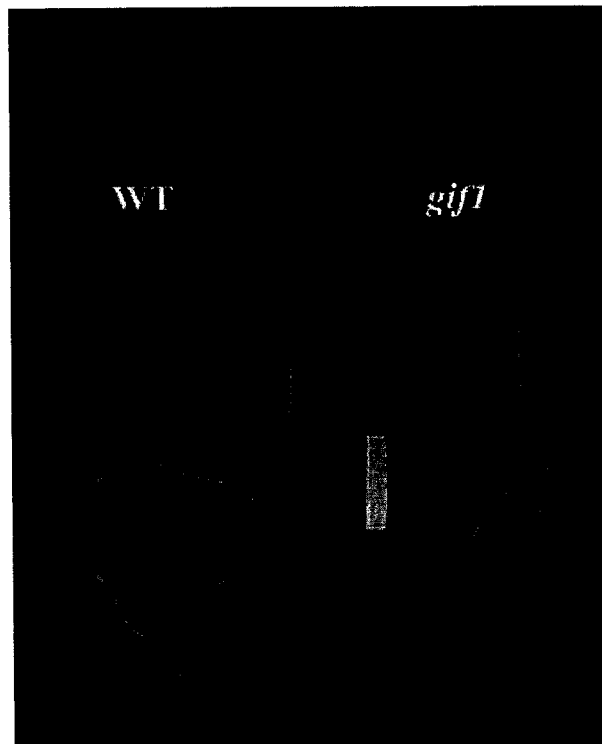


图 2

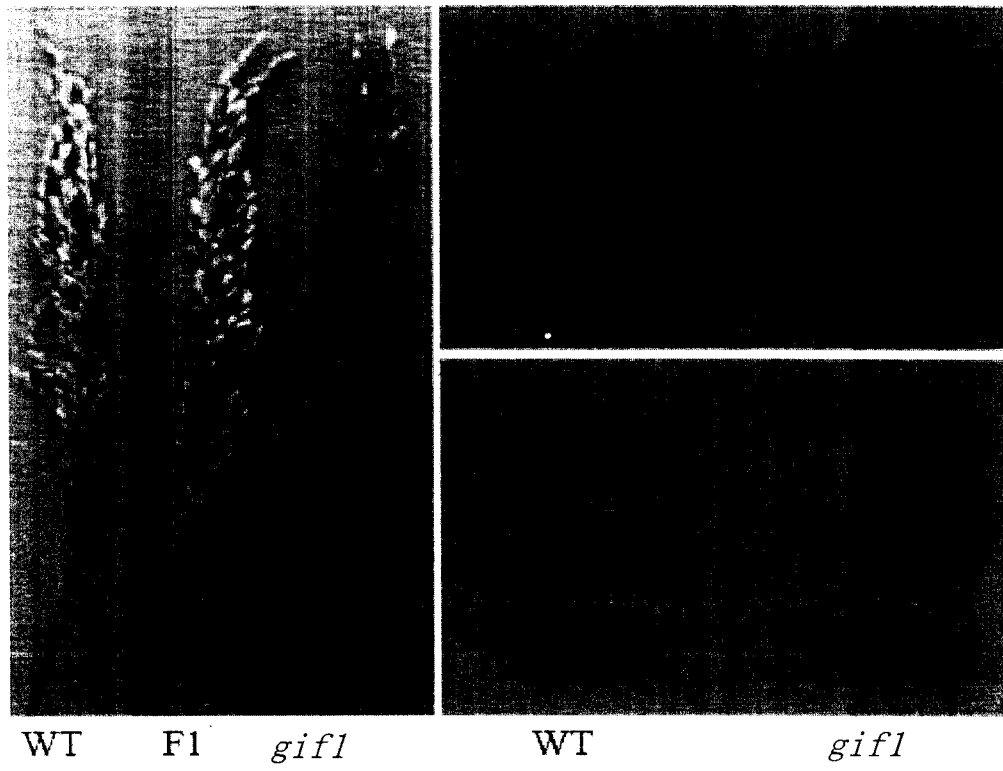


图 3

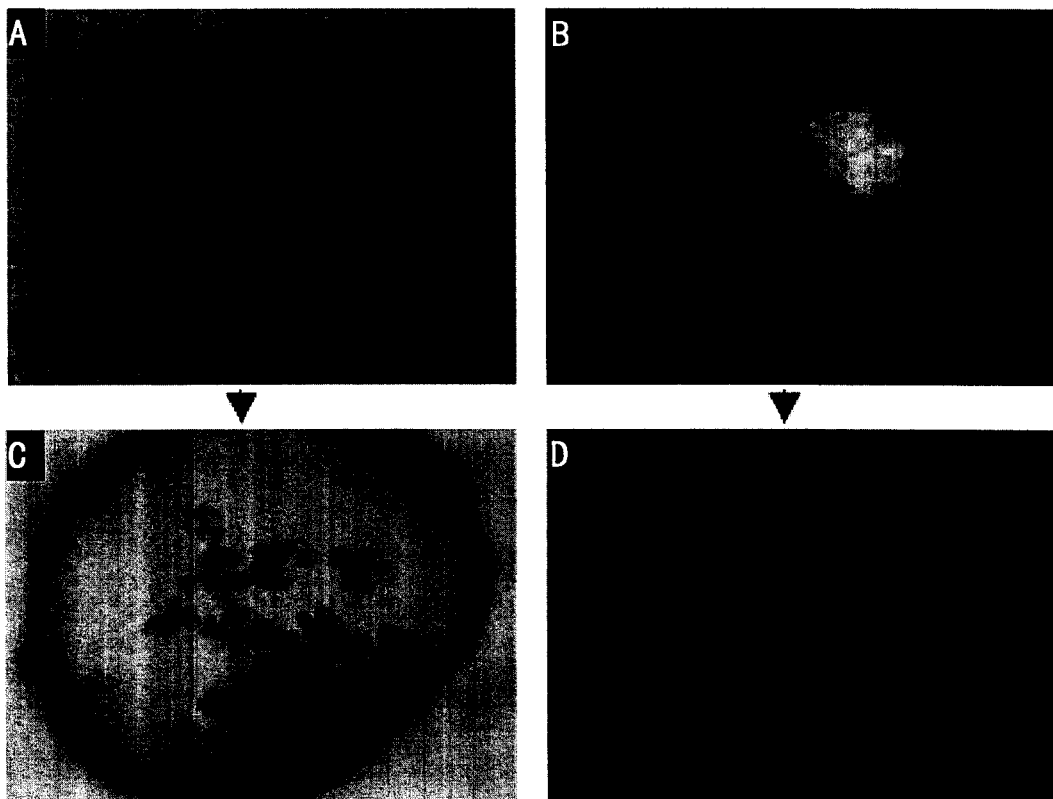


图 4

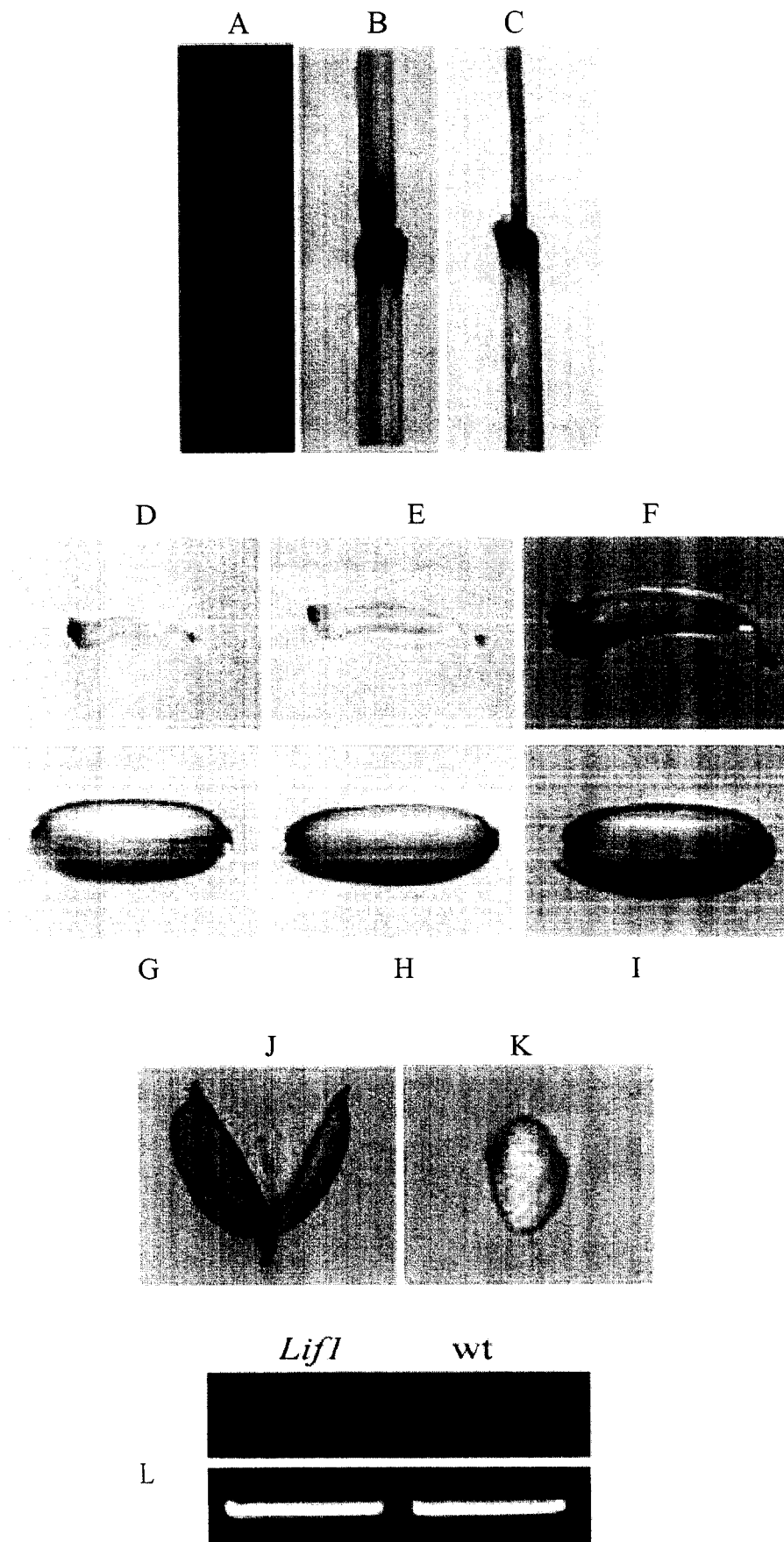


图 5

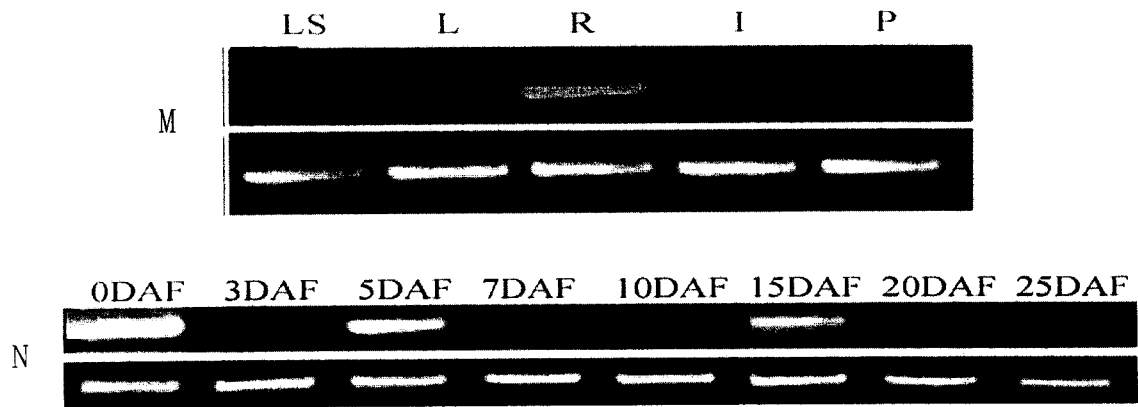


图 5(续)

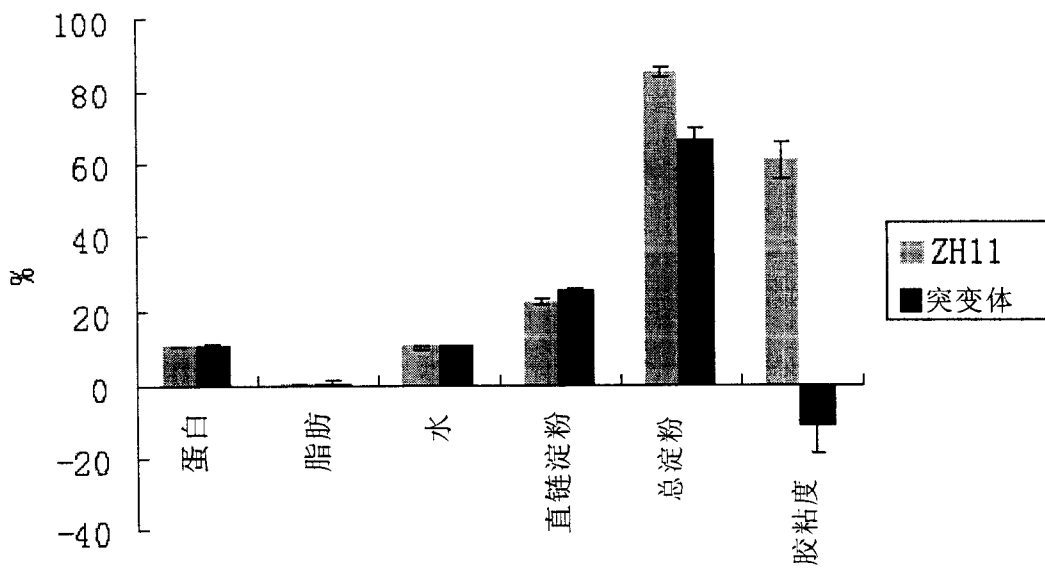


图 6