



(12) 发明专利申请

(10) 申请公布号 CN 103975841 A

(43) 申请公布日 2014. 08. 13

(21) 申请号 201310711601. 9

(74) 专利代理机构 中国专利代理(香港)有限公司 72001

(22) 申请日 2006. 09. 07

代理人 权陆军

(30) 优先权数据

60/714,779 2005.09.07 US

(51) Int. Cl.

60/722,493 2005.09.30 US

A01H 1/02 (2006.01)

C12Q 1/68 (2006.01)

(62) 分案原申请数据

200680041478.2 2006.09.07

(83) 生物保藏信息

ATCC PTA-6892 2005.07.28

权利要求书2页 说明书41页
序列表53页 附图9页

ATCC PTA-6893 2005.07.28

(71) 申请人 孟山都技术有限公司

地址 美国密苏里州

(72) 发明人 K. 吴 T. 霍尔西 N. 布林格 J. 杨

D. 佩 R. 赖特 J. 拜伦

(54) 发明名称

农学上优良的具有高 β -伴大豆球蛋白含量的大豆

(57) 摘要

本发明涉及农学上优良的具有高 β -伴大豆球蛋白含量的大豆。本发明如下克服了本领域的缺陷：通过提供农学上优良的大豆植物，其具有选自 Gy1、Gy2、Gy3、Gy4 和 Gy5 的至少 2 种大豆球蛋白亚基的非转基因突变，例如在种子中赋予 Gy3 和 Gy4 无效表型和提高的 β -伴大豆球蛋白含量。本发明也提供了这些植物的衍生物和植物部分及其用途。作为本发明的部分，也提供了标记辅助的选择大豆品种的方法，所述大豆品种包含赋予降低的 Gy1、Gy2、Gy3、Gy4 和 Gy5 表型的非转基因突变。也提供了生产另外脂加氧酶和 / 或 Kunitz 胰蛋白酶抑制剂无效的植物的方法，和由此生产的植物。本发明的重要之处是，来自这样的植物的大豆是优选的饮食添加剂，并提供重要的健康益处。

1. 植物育种的方法,其包含下述步骤:

(a) 测定大豆植物中在离非转基因突变 Gy1、Gy2、Gy3 或 Gy4 等位基因 50cM 内的大豆植物基因组区域中至少第一种多态性的存在,所述等位基因赋予降低的大豆球蛋白和提高的 β -伴大豆球蛋白含量;

(b) 选择包含所述多态性的至少第一种大豆植物,以选择所述等位基因;和

(c) 使包含所述多态性的第一种大豆植物与第二种大豆植物杂交,以产生包含至少第一种非转基因突变 Gy1、Gy2、Gy3 或 Gy4 等位基因的子代植物,所述等位基因赋予降低的大豆球蛋白和提高的 β -伴大豆球蛋白含量。

2. 权利要求 1 的方法,还包含下述步骤:

(d) 使用步骤 (c) 的子代植物作为原材料,重复步骤 (a)-(c) 至少约 2-10 次,以生产其它子代植物。

3. 权利要求 1 的方法,其中所述第二种大豆植物属于农学上优良的品种。

4. 权利要求 1 的方法,其包含选择包含所述多态性和农学上优良特征的大豆植物。

5. 权利要求 1 的方法,其中所述突变 Gy1 等位基因包含缺失。

6. 权利要求 1 的方法,其中所述非转基因突变 Gy1、Gy2、Gy3 或 Gy4 等位基因对应着在系 B2G2 中发现的突变等位基因,所述系 B2G2 的种子的代表样品在 ATCC 登记号 PTA-6893 下保藏。

7. 权利要求 1 的方法,其中所述突变 Gy4 等位基因包含取消翻译起始密码子的点突变。

8. 权利要求 1 的方法,还包含选择关于所述多态性纯合的第一种大豆植物。

9. 权利要求 1 的方法,还包含选择包含至少 2 种多态性的第一种大豆植物。

10. 权利要求 1 的方法,其包含测定在 Gy1 和 Gy4 等位基因 50cM 内的多态性。

11. 权利要求 1 的方法,其包含测定在 Gy2 和 Gy4 等位基因 50cM 内的多态性。

12. 权利要求 1 的方法,其包含测定与 SEQ ID NO:165 的核苷酸 682 相对应的 Gy4 等位基因中的多态性。

13. 权利要求 1 的方法,还包含选择包含至少 3 种多态性的第一种大豆植物。

14. 权利要求 1 的方法,还包含选择包含在非转基因突变 Gy3 等位基因 50cM 内的多态性的第一种大豆植物。

15. 权利要求 14 的方法,其中选择第一种大豆植物包含,检测选自下述的 Gy3 等位基因中的至少第一种多态性:在位置 848-851 处的插入,在位置 1083 处的 SNP,在位置 1120 处的 SNP,和在位置 1866 处的 SNP;其中所述位置是与 SEQ ID NO:164 相对应的所述等位基因中的位置。

16. 权利要求 1 的方法,还包含选择第一种大豆植物,其包含在非转基因突变 Gy5 等位基因 50cM 内的多态性。

17. 权利要求 16 的方法,其中选择植物包含,检测选自下述的 Gy5 等位基因中的至少第一种多态性:在位置 363 处的单核苷酸多态性 (SNP),在位置 612 处的 SNP,在位置 447-453 处的缺失,和在位置 519-524 处的缺失,其中所述位置是与 SEQ ID NO:166 相对应的所述等位基因中的位置。

18. 权利要求 1 的方法,其包含检测选自 NS0199002、NS0199003 和 NS0199008 的标记。

19. 权利要求 1 的方法,还包含选择第一种大豆植物,其包含在非转基因突变 1x2 等位

基因 50cM 内的多态性。

20. 权利要求 17 的方法,其中所述多态性包含选自下述的 1x2 等位基因中的多态性 : 在位置 323 处的单核苷酸多态性 (SNP),在位置 439 处的 SNP,在位置 1390 处的 SNP,在位置 1431 处的 SNP,在位置 1458 处的 SNP,在位置 2486-2487 处的缺失,和在位置 2542 处的 SNP ;其中所述位置是与 SEQ ID NO :158 相对应的所述等位基因中的位置。

21. 权利要求 1 的方法,其中测定大豆植物的多态性包含 PCR。

22. 权利要求 1 的方法,其中通过用标记的核苷酸探针进行杂交,检测所述多态性。

23. 权利要求 1 的方法,其中通过 DNA 测序检测所述多态性。

24. 权利要求 1 的方法,其中步骤 (a) 还包含,测定大豆植物中在离 Kunitz 胰蛋白酶抑制剂 (KTI) 无效等位基因 50cM 内的大豆植物基因组区域中至少第一种多态性的存在。

25. 权利要求 1 的方法,其中步骤 (b) 的选择至少第一种大豆植物还包含,选择第一种大豆植物,其包含在 Kunitz 胰蛋白酶抑制剂 (KTI) 无效等位基因 50cM 内的多态性,以选择所述无效等位基因。

26. 权利要求 24 的方法,其中所述 Kunitz 胰蛋白酶抑制剂 (KTI) 无效等位基因 50cM 内的多态性选自 :在位置 622-623 处的 2-bp 缺失,和在位置 624 处的突变,其中所述位置是与 SEQ ID NO :167 相对应的所述等位基因中的位置。

27. 权利要求 24 的方法,其中测定大豆植物的多态性包含 PCR。

28. 权利要求 24 的方法,其中通过用标记的核苷酸探针进行杂交,检测所述多态性。

29. 权利要求 24 的方法,其中通过 DNA 测序检测所述多态性。

30. 植物育种的方法,其包含下述步骤 :

(a) 测定大豆植物在 1ox2 等位基因中选自 SNP1390、SNP1431、SNP1458 和 INDEL2486-2487 的至少第一种多态性的存在,其中所述等位基因与降低的脂加氧酶 -2 含量有关 ;

(b) 选择包含所述多态性的至少第一种大豆植物,以选择所述等位基因 ;和

(c) 使第一种大豆植物与第二种大豆植物杂交,以产生包含所述多态性和所述等位基因的子代植物。

31. 权利要求 30 的方法,还包含下述步骤 :

(d) 使用步骤 (c) 的子代植物作为步骤 (a) 的原材料,重复步骤 (a)-(c) 至少约 2-10 次,以生产其它子代植物。

32. 权利要求 30 的方法,还包含选择关于所述多态性纯合的第一种大豆植物。

33. 权利要求 30 的方法,还包含选择包含至少 2 种所述多态性的第一种大豆植物。

34. 权利要求 30 的方法,还包含选择包含至少 3 种所述多态性的第一种大豆植物。

农学上优良的具有高 β -伴大豆球蛋白含量的大豆

[0001] 本申请是国际申请日为 2006 年 9 月 7 日的国际申请 PCT / US2006 / 034495 进入中国、申请号为 200680041478.2 的题为“农学上优良的具有高 β -伴大豆球蛋白含量的大豆”的发明专利申请的分案申请。

[0002] 发明背景

[0003] 本申请要求 2005 年 9 月 7 日提交的美国临时申请系列号 60 / 714,779 和 2005 年 9 月 30 日提交的美国临时申请系列号 60 / 722,493 的优先权。

[0004] 1. 技术领域

[0005] 本发明总的来说涉及植物育种和分子生物学领域。具体地，本发明涉及农学上优良的具有高 β -伴大豆球蛋白 (conglycinin) 含量的大豆品种和用于制备这样的植物的材料。

[0006] 2. 背景技术

[0007] 大豆球蛋白和 β -伴大豆球蛋白是大豆中的两种主要贮存蛋白，占总蛋白的约 70%，或种子总重量的 40%。大豆球蛋白 (11s 球蛋白) 由 5 种不同的亚基组成，这些亚基分别命名为 A1aB2、A2B1a、A1bB1b、A5A4B3、A3B4。每个亚基由通过二硫键共价连接的 2 个多肽组成，所述多肽一个是酸性的，一个是碱性的。2 个多肽链源自大豆球蛋白原 (proglycinin) 前体的翻译后切割，该步骤发生在前体进入蛋白体之后 (Chrispeels 等，1982)。已经鉴别出编码这些多肽亚基的 5 个主要基因。它们分别命名为 Gy1、Gy2、Gy3、Gy4 和 Gy5 (Nielsen 等，1997)。另外，也报道了假基因 gy6 和次要基因 Gy7 (Beilinson 等，2002)。各个研究组已经报道了这些基因的遗传作图 (Diers 等，1993, Chen 和 Shoemaker 1998, Beilinson 等，2002)。Gy1 和 Gy2 的位置相隔 3kb，且作图在连锁群 N (Nielsen 等，1989)，Gy3 作图在连锁群 L (Beilinson 等，2002)。Gy4 和 Gy5 分别作图在连锁群 O 和 F。所有这些基因都使用 RFLP 探针在 DNA 印迹上作图。

[0008] 另一方面， β -伴大豆球蛋白由 α (约 67kDa)、 α' (约 71kDa) 和 β (约 50kDa) 亚基组成，且每个亚基通过共翻译或翻译后的修饰来加工 (Ladin 等，1987; Utsumi, 1992)。 β -伴大豆球蛋白亚基分别由基因 Cgy1、Cgy2 和 Cgy3 编码。遗传分析表明，Cgy2 与 Cgy3 紧密连锁，而 Cgy1 与另外 2 个独立地分离。 β -伴大豆球蛋白基因家族含有至少 15 个成员，它们分成 2 个主要组，分别编码 2.5kb 和 1.7kb 胚 mRNA (Harada 等，1989)。

[0009] 具有提高的 β -伴大豆球蛋白水平和降低的大豆球蛋白水平的大豆植物会提供相当大的益处。一个原因是， β -伴大豆球蛋白是一种可溶蛋白，而大豆球蛋白的溶解度低得多。还已经发现，与大豆球蛋白相比， β -伴大豆球蛋白、尤其 α' 亚基具有明显更高的营养价值和对人类健康的积极影响 (Baba 等，2004)。许多使用动物模型的实验已经表明，来自大豆 β -伴大豆球蛋白的 α' 亚基会降低血浆甘油三酯，且也会增加 LDL (“坏”胆固醇) 从血液中的去除 (Duranti 等，2004, Moriyama 等，2004, Adams 等，2004, Nishi 等，2003)。因此，具有提高的 β -伴大豆球蛋白含量的大豆品种具有比传统品种更高的价值，且将适用于营养饮料和其它食品中。

[0010] 令人感兴趣地，大豆球蛋白基因中的突变对大豆种子中的 β -伴大豆球蛋白含量

具有直接影响。具有降低的大豆球蛋白含量的突变大豆植物具有提高的 β -伴大豆球蛋白含量。但是,由于多个大豆球蛋白等位基因参与大豆球蛋白亚基生产,已经证实难以培育多个 Gy 亚基表达降低的植物,因为这样的植物具有其它特征,例如低产量、过度倒伏和绿种子,这些使它们商业上不合算。以前的测定导致降低的大豆球蛋白含量的突变遗传的方法,不能实现在导入农学上优异的特征的同时选择这些表型所需的高通量技术。例如,以前的 Gy 遗传的评估,依赖于蛋白表达的分析,这是高成本的、劳动密集的,且不能跟踪隐性突变的遗传。不顾劳动生产这样的植物的可能性也是未知的,这是由于其它的复杂因素,例如与渐渗突变 Gy 等位基因的尝试有关的连锁累赘 (linkage drag) 和上位性。在得到的表型方面,等位基因的组合也是不可预知的。因而,本领域长期需要(但是没有得到满足)农学上优良的多个 Gy 蛋白亚基表达降低的大豆植物和生产这样的植物的方法。

[0011] 脂加氧酶是催化多不饱和脂肪酸的双加氧的酶。大豆种子含有 3 种脂加氧酶同工酶 - 脂加氧酶 1、2 和 3。这些同工酶有助于生产大豆种子的令人厌恶的风味。在缺乏这些同工酶、尤其缺乏脂加氧酶 -2 的种子中,这种令人厌恶的风味不存在或不太显著。因此,需要将缺乏一种或多种脂加氧酶同工酶的大豆种子用于制备饮料和食品。脂加氧酶 1、2、和 3 缺陷系的遗传研究表明,各自的不存在分别是由于单个隐性等位基因 $-1x1$ 、 $1x2$ 和 $1x3$ 。由 $1x1$ 和 $1x2$ 定义的基因座紧密连锁,且在遗传上不与 $1x3$ 连锁 (Kitamura, 1984; Kitamura 等, 1985; Haiika 等, 1992; Hildebrand 等, 1982)。已经克隆了编码脂加氧酶 1、2、和 3 的结构基因,且分别命名为 Lox1、Lox2 和 Lox3 (Shibata 等, 1987; Shibata 等, 1988; Yenofsky 等, 1988)。

[0012] Kunitz 胰蛋白酶抑制剂 (KTI) 是大豆中的一种抗营养且变应原性因子,当存在于饮食中时,它干扰蛋白的消化和吸收。因而,具有 KTI- 无效突变性状的大豆品种具有比传统品种更高的商业价值。已经对大豆系中的 KTI 生产进行了遗传和生化研究(例如 de Moraes 等, 2006; Natarajan 等, 2006),且已经鉴别出 3 个相关的基因,其中 KTI3 在培养的大豆基因型中编码优势的 Kunitz 胰蛋白酶抑制剂蛋白 (Natarajan 等, 2006)。已经报道了与某些大豆系中与丧失 KTI 生产有关的一些特定 DNA 标记 (de Moraes 等, 2006)。

发明内容

[0013] 在一个实施方案中,本发明提供了农学上优良的大豆植物,其具有赋予 Gy3 和 Gy4 无效表型和提高的种子 β -伴大豆球蛋白含量的非转基因突变。因而,本发明的植物在一个方面包含具有低大豆球蛋白含量和高 β -伴大豆球蛋白含量的种子。在某些实施方案中,本发明植物的种子 β -伴大豆球蛋白含量是总蛋白含量的约或至少约 34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50% 或更高。在有些实施方案中,本发明的植物的种子大豆球蛋白含量是总蛋白的约或小于约 15、14、13、12、11、10、9、8、7、6、5、4、3、2、1 或 0%。在有些情况下,本发明植物可以包含突变 Gy4 等位基因。例如,突变 Gy4 等位基因可以包含在核苷酸 682 处的点突变,从而取消翻译起始密码子。在其它实施方案中,本发明提供的植物可以包含 Gy1 和 Gy2 无效等位基因。在本发明的一个实施方案中,Gy1、Gy2、Gy3 和 / 或 Gy4 等位基因中的任意一个或多个可以与系 B2G2 中的无效等位基因相同,所述系 B2G2 是已经在 ATCC 登记号 PTA-6893 下保藏的种子的代表样品。

[0014] 在本发明的某些实施方案中,提供了大豆植物,其另外包含赋予降低水平的

Gy1 / Gy2 蛋白的突变。本文使用的“降低水平的 Gy1 / Gy2 蛋白”是指,与缺乏突变的具有相同遗传背景的植物相比,包含非转基因突变的植物的种子具有降低的 Gy1 / Gy2 蛋白水平。例如,包含赋予降低的 Gy1 / Gy2 的非转基因突变的植物的 Gy1 / Gy2 蛋白含量可以小于总种子蛋白的约 3.1%。在某些情况下,赋予降低的 Gy1 / Gy2 蛋白含量的突变可以是非转基因突变。在本发明的有些方面,本发明的植物包含突变 Gy1 等位基因。例如,突变 Gy1 等位基因可以包含跨越上游启动子区、外显子 I 和内含子 I 的缺失。

[0015] 在某些实施方案中,本发明的植物可以另外包含赋予降低的 Gy5 蛋白水平的突变。在某些情况下,赋予降低的 Gy5 蛋白含量的突变是非转基因突变。因而在有些方面,本发明植物包含赋予降低的 Gy1、Gy2、Gy3、Gy4 和 Gy5 蛋白水平的突变。在某些方面,本发明植物可以包含赋予降低的 Gy1、Gy2Gy3、Gy4 和 Gy5 表型的非转基因突变。这些植物的种子的大豆球蛋白含量可以是总蛋白的约或小于约 15、14、13、12、11、10、9、8、7、6、5、4、3、2、1 或 0%。因而,包含降低的 Gy1、Gy2、Gy3、Gy4 和 Gy5 表型的本发明植物可以包含这样的种子,其 β -伴大豆球蛋白含量是总蛋白含量的约或至少约 34、35、36、37、38、39、40、41、42、43、44、45、46、47、48、49、50% 或更多。

[0016] 在某些实施方案中,本发明的植物可以另外包含赋予 1x1、1x2、和 / 或 1x3 表型的突变。在某些情况下,赋予 1x1、1x2、和 / 或 1x3 表型的突变是非转基因突变。因而在有些方面,本发明植物包含赋予降低的 Gy3 和 Gy4 表型的突变,和赋予 1x1、1x2、和 / 或 1x3 表型中的一个或多个的突变。在本发明的一个实施方案中,这样的植物可以另外包含赋予降低的 Gy1、Gy2Gy3 和 Gy5 表型的突变。

[0017] 本发明也提供了植物部分。本发明的植物部分包括,但不限于,花粉、胚珠、分生组织、细胞和种子。本发明的细胞还可以包含可再生的细胞,例如胚分生细胞,花粉,叶,根,根尖和花。因而,这些细胞可以用于再生本发明的植物。

[0018] 本文也提供了本发明植物的种子部分。因而,从本发明种子制备的粉碎的种子和粗粉或面粉,也作为本发明的部分来提供。本发明还包含制备大豆粗粉或面粉的方法,其包含,粉碎或研磨本发明的种子。根据本发明的这种大豆面粉或粗粉可以包含本发明植物的基因组材料。在一个实施方案中,可以将食物定义为包含这样的植物的基因组。在其它实施方案中,本发明的大豆粗粉或面粉可以定义为,与从具有相同遗传背景的植物种子制备的粗粉或面粉相比,包含提高的 β -伴大豆球蛋白和降低的大豆球蛋白含量,但是不包含非转基因的突变 Gy3 和 Gy4 无效等位基因。

[0019] 在本发明的另一个方面,提供了生产大豆种子的方法,其包含,使本发明植物与它本身或与第二种大豆植物杂交。因而,该方法可以包含,通过使本发明植物与第二种不同的大豆植物杂交,制备杂种大豆种子。

[0020] 本发明的另一个方面是生产用于人或动物消费的食品的方法,其包括:(a) 得到本发明的植物;(b) 栽培该植物至成熟;和(c) 从该植物制备食品。在本发明的某些实施方案中,食品可以是蛋白浓缩物、蛋白分离物、粗粉、油、面粉或者大豆壳。在有些实施方案中,食品可以包含饮料例如豆浆和其它营养饮料,泡制食物、沙司、调味品、色拉调料、果汁、糖浆、甜食、糖衣和饼馅、软的冷冻产品、糖食或者半成品食物。从本发明植物生产的食物可以包含提高的 β -伴大豆球蛋白含量,且因而具有比用典型大豆品种制备的食物更大的营养价值。另外,包含降低的大豆球蛋白含量的本发明植物可以用于需要低不溶蛋白量的食物。

组合物中。

[0021] 在其它实施方案中,本发明的植物可以另外包含转基因。例如,植物可以包含赋予除草剂耐受性、抗病性、昆虫和害虫抗性、改变的脂肪酸、蛋白或糖类代谢、增加的谷粒产量(grain yield)、改变的植物成熟度和 / 或改变的形态学特征的转基因。例如,除草剂耐受性转基因可以包含草甘膦抗性基因。

[0022] 在某些实施方案中,本发明的植物可以定义为通过下述方法制备,其中使包含赋予Gy3 和 Gy4 无效表型和提高的 β -伴大豆球蛋白含量的非转基因突变的植物与包含农学上优良特征的植物杂交。可以测定该杂交的子代的农学上优良的特征和 Gy3 和 Gy4 蛋白含量,并基于这些特征选择子代植物,从而产生本发明的植物。因而,在某些实施方案中,可以如下生产本发明的植物:使选择的起始品种与包含农学上优良的特征的第二种大豆植物杂交。在有些实施方案中,本发明的植物可以定义为通过下述方法制备,其中使包含赋予 1x1、1x2、和 / 或 1x3 表型的非转基因突变的植物与包含降低的 Gy3 和 Gy4 蛋白含量和提高的 β -伴大豆球蛋白含量的植物杂交。

[0023] 本发明也提供了植物育种的方法,其中测定植物的与 Gy1 / Gy2 和 Gy4 等位基因有关的大豆植物基因组区域中多态性的存在,其包含,选择植物,和使该植物与第二种大豆植物杂交,以生产子代。在有些实施方案中,本发明的方法可以包含如下选择子代植物:测定植物的与低 Gy1 / Gy2 或 Gy4 表型有关的多态性,和使该植物与第二种大豆植物杂交,以生产其它子代植物。在本发明的某些实施方案中,第二种大豆植物可以包含农学上优良的特征。本发明的方法也可以另外包含,选择包含所述多态性和农学上优良的特征的大豆植物。因而,本发明使得能将赋予 Gy1 / Gy2 和 / 或 Gy4 表型和提高的种子 β -伴大豆球蛋白含量的非转基因突变导入农学上优良的大豆植物中。本发明的方法可以根据需要重复 1、2、3、4、5、10、15、20 或更多次,以在每个步骤中选择具有指示在 Gy1 / Gy2 和 / 或 Gy4 等位基因处的非转基因突变的多态性的农学上优良的子代。在本发明的某些实施方案中,第一种大豆植物可以是系 B2G2 的植物,所述系 B2G2 是已经在 ATCC 登记号 PTA-6893 下保藏的种子的代表样品。在另一个实施方案中,本发明的方法可以另外包含,选择包含指示在 Gy1 / Gy2 和 Gy4 等位基因中的非转基因突变的多态性的植物。

[0024] 在有些实施方案中,本发明的方法可以另外包含,选择具有指示降低的 Gy3 和 / 或 Gy5 含量的标记的植物。因而,根据本发明的标记辅助的植物育种方法可以用于培育具有降低的 Gy1、Gy2、Gy3、Gy4 和 Gy5 含量的大豆。

[0025] 在本发明的有些实施方案中,赋予降低的 Gy1、Gy2 或 Gy4 表型的非转基因突变可以包含在 Gy1、Gy2、Gy3 或 Gy4 等位基因中的突变。在某些实施方案中,使用在 Gy 等位基因的 50cM 内包含多态性的遗传标记,检测突变 Gy 等位基因。在本发明的其它方面,具有降低的 Gy1 / Gy2 表型的植物包含突变 Gy1 等位基因。在有些情况下,突变 Gy1 等位基因包含缺失,例如缺失启动子区域、外显子 I 和内含子 I。在其它实施方案中,可以使用标记 NS0199002 或 NS0199008 来检测突变 Gy1 等位基因。在本发明的某些方面,可以用 Gy2 的标记检测突变 Gy1 等位基因,因为这 2 个基因紧密连锁。因而,在本发明的其它方面,可以用 Gy1 的标记检测突变 Gy2 等位基因。本文表明,具有突变 Gy1,2 等位基因的植物也可以具有降低的 Gy3 表型。因而,在本发明的有些方面,可以用 Gy1 或 Gy2 的标记选择包含降低的 Gy3 表型的植物,因为这些标记已经被表明与 B2G2 衍生植物中降低的 Gy3 表型有关。因此,

Gy1 或 Gy2 的标记,例如 NS0199002 或 NS0199008,可以用于测定突变 Gy1、Gy2 和 Gy3 等位基因的存在。在本发明的某些实施方案中,表型上 Gy4 无效的植物包含突变 Gy4 等位基因。在其它实施方案中,突变 Gy4 等位基因包含点突变,例如取消翻译起始密码子的 SNP。在本发明的其它方面,可以用 NS0199003 标记检测 Gy4 无效等位基因。可以检测 SNP 标记,例如使用荧光标记的寡核苷酸。

[0026] 在有些实施方案中,本发明的方法可以另外包含,选择具有指示降低的脂加氧酶 1、2、和 / 或 3 含量的标记的植物。因而,根据本发明的标记辅助的植物育种的方法可以用于生产具有降低的脂加氧酶 1、2、和 / 或 3 含量的大豆。

[0027] 在本发明的具体实施方案中,赋予 1x1、1x2、和 / 或 1x3 表型的非转基因突变可以包含在 Lox1、Lox2、和 / 或 Lox3 等位基因中的突变。在本发明的一个实施方案中,使用包含在 Lox 等位基因的 50cM 内的多态性的遗传标记,检测赋予 1x1、1x2、和 / 或 1x3 表型的突变等位基因。在某些实施方案中,使用如下面表 14 所示的 INDEL178–180、SNP326、SNP363、SNP380、SNP713、SNP1196、SNP1253、SNP1372、SNP1388、SNPR1527、SNP1554、SNP2267、SNP3088、SNP3125、SNP3139、INDEL3832–3905、SNP4043、SNP4057、SNP4193、SNP4225、SNP4247、SNP4267、或 SNP4439 中的一个或多个,检测 1x1 等位基因。在有些实施方案中,使用如下面表 15 所示的 SNP323、SNP439、SNP1390、SNP1431、SNP1458、INDEL2486–87、或 SNP2542 中的一个或多个,可以检测 1x2 等位基因。SNP2542 也称作 NS0203296 标记。

[0028] 在有些实施方案中,本发明的方法可以另外包含,选择具有指示 KTI- 无效或 KTI- 减少的性状的标记的植物。因而,根据本发明的标记辅助的植物育种的方法可以用于生产具有减少的或不可检测的 Kunitz 胰蛋白酶抑制剂含量的大豆。

[0029] 在本发明的某些实施方案中,赋予 KTI 无效表型的突变可以包含 KTI 编码基因中的突变。在具体的实施方案中,赋予 KTI 无效表型的突变包含也称作“KTIA”的 KTI3 基因中的突变。在本发明的一个实施方案中,使用包含在 KTI 等位基因的 50cM 内的多态性的遗传标记,检测赋予 KTI 无效表型的突变等位基因。在某些实施方案中,使用位于 KTI3 基因内的一个或多个 INDEL 或 SNP,例如如下面表 18 中所示的那些,检测 KTI 等位基因。这样的选择因而可以基于标记信息(植物基因型),而不是胰蛋白酶活性的酶分析或 KTI 含量分析。

[0030] 关于本文所述的任何其它方法或组合物,可以采用在本发明的方法和 / 或组合物的上下文中讨论的实施方案。因而,属于一种方法或组合物的实施方案,也可以应用于本发明的其它方法和组合物。

[0031] 在说明书或权利要求书中使用的“一个 / 种”可以指一个 / 种或多个 / 种。如本文在权利要求书中使用的,当与词语“包含”一起使用时,词语“一个 / 种”可以指一个 / 种或超过一个 / 种。本文使用的“另一个 / 种”可以指至少第二个 / 种或多个 / 种。

[0032] 从下面的详述中,可以明白本发明的其它目的、特征和优点。但是,应当理解,详述和具体实施例尽管指明了本发明的优选实施方案,但它们仅仅作为解释来给出,因为本领域技术人员从该详述中可以明白在本发明精神和范围内的多种变化和修饰。

附图说明

[0033] 下面的附图构成本说明书的一部分,且用于进一步证实本发明的某些方面。参考

这些附图中的一幅或多幅,以及本文所述的具体实施方案的详述,可以更好地理解本发明。

[0034] 图 1 :通过 SDS-PAGE,分离从指出的大豆品种的种子提取的蛋白,并通过考马斯染色显现。指出了每个基因编码的酸性大豆球蛋白亚基的迁移率。凝胶分离不足以分离 Gy1 编码的蛋白亚基和 Gy2 编码的蛋白亚基。

[0035] 图 2A-B :基于种子的 Gy1 和 Gy2 编码的蛋白含量,将包含突变 Gy1 和 Gy2 等位基因的子代植物分配成 2 个表型组。图 2A. 该图显示了与由 Gy1 和 Gy2 编码的酸性蛋白组成的总蛋白的百分比 (x 轴) 相比,F2 植物的数目 (y 轴)。图 2B. 来自 F2 子代植物的数据表明,来自 B2G2 大豆的突变 Gy1 和 Gy2 等位基因是隐性的。对 2 个类别中具有 Gy1 和 Gy2 编码的蛋白水平的子代植物的数目进行卡方分析,且在每种情况下测定概率值。

[0036] 图 3A-B :基于 Gy3 种子中总蛋白的百分比,将“11S 无效”植物的子代分配成 2 个群体。图 3A. 该图显示了与由 Gy3 编码的蛋白组成的总蛋白的百分比 (x 轴) 相比,F2 植物的数目 (y 轴)。图 3B. 来自子代植物的数据表明,来自 B2G2 大豆的突变 Gy3 等位基因是隐性的。对 2 个类别中具有 Gy3 编码的蛋白水平的子代植物的数目进行卡方分析,且在每种情况下测定概率值。

[0037] 图 4 :Gy3 编码的蛋白的量与 Gy1 和 Gy2 编码的蛋白的量正相关。图 4A. 该图绘制了观察到的 Gy1 和 Gy2 编码的酸性蛋白的量 (x 轴) 对 Gy3 编码的蛋白的量 (y 轴) 的图。图 4B. 该表显示了 Gy1、Gy2、Gy3、Gy4 和 Gy5 编码的蛋白的表达水平之间的相关系数。

[0038] 图 5 :基于 Gy4 种子中总蛋白的百分比,将 11S 无效植物的子代分配成 2 个群体。图 5A. 该图显示了与由 Gy4 编码的酸性蛋白组成的总蛋白的百分比 (x 轴) 相比,F2 植物的数目 (y 轴)。图 5B. 来自子代植物的数据表明,来自 B2G2 大豆的突变 Gy4 等位基因是隐性的。对 2 个类别中具有 Gy4 编码的蛋白水平的子代植物的数目进行卡方分析,且在每种情况下测定概率值。

[0039] 图 6A-B :大豆种子中大豆球蛋白的降低的表达与 β -伴大豆球蛋白的提高的表达相关。图 6A. 该图绘制了 Gy 等位基因编码的总蛋白的百分比 (x 轴) 对 Cgy 等位基因编码的总蛋白的百分比 (y 轴) 的图。图 6B. 该表显示了 Gy 编码的蛋白的表达和 Cgy1-4 编码的蛋白的表达之间的相关系数。

[0040] 图 7A-B :基因组 Gy 标记有效地选择生产具有高 β -伴大豆球蛋白表达的种子的常规植物。图 7A. 该图绘制了种子的总 β -伴大豆球蛋白蛋白含量 (y 轴) 对总种子大豆球蛋白含量 (x 轴) 的图。菱形指示通过大豆球蛋白亚基的蛋白分析选择的植物。图 7B. 该图与图 7A 相同,但是菱形指示通过标记 NS0199002、NS0199003 和 NS0199008 选择的植物。

[0041] 图 8 :为在位置 2542 处的 SNP 设计的 Taqman 测定的 Allelogram(标记 NS0203296)。如图 8 所示,该标记允许清楚地区分来自 1x2 突变体的“A”等位基因和来自如实施例 13 和表 15 所述野生型的“T”等位基因。

[0042] 图 9 :Kunitz 胰蛋白酶抑制剂的序列比对显示了 Kunitz 无效突变系中的缺失 / 插入。

具体实施方式

[0043] 本发明提供了植物和生产植物的方法,所述植物包含赋予 Gy3 和 Gy4 无效表型和农学上优良特征的非转基因突变。这些突变赋予突变植物种子低大豆球蛋白和高 β -伴大

豆球蛋白含量。因而，本发明植物将具有巨大价值，因为 β -伴大豆球蛋白提供比大豆球蛋白改善的营养特征和溶解度。另外，本文提供的植物包含农学上优良特征，使得能够实现商业上相当大产量的高 β -伴大豆球蛋白、低大豆球蛋白的大豆。在本发明的某些方面，具有提高的 β -伴大豆球蛋白含量的植物包含非转基因的 Gy3 和 / 或 Gy4 无效等位基因，且因此与在这些基因座处含有对应的转基因等位基因的大豆品种相比，具有减少的政府管理的额外优点。也提供了另外包含非转基因的 Gy1 和 Gy2 无效等位基因且也提供了这样的益处的植物。

[0044] 本发明也提供了植物和生产植物的方法，所述植物包含赋予脂加氧酶-2 无效表型的非转基因突变。脂加氧酶-2 无效和大豆球蛋白无效表型的组合，提供提高的含量的高功能性和健康的 β -伴大豆球蛋白蛋白。 β -伴大豆球蛋白尤其含有负责降低胆固醇和体重管理（通过饱满感效应和脂肪沉积的减少）益处的生物活性肽。

[0045] 根据本发明的另一个有价值的组合是脂加氧酶-2 无效的且中等油酸的含量（例如 40–65% 油酸）。这样的大豆将产生低水平的臭气，因为它缺乏脂类氧化的主要催化剂（脂加氧酶-2），且具有低得多水平的底物（亚油酸）。另外，这样的大豆将具有改良的（更低的） ω -6 / ω -3 脂肪酸比，这对心血管健康是有益的。

[0046] 本文还提供了培育大豆植物的方法，所述大豆植物包含赋予降低的 Gy1、Gy2、Gy3 和 Gy4 表型和农学上优良特征的非转基因突变。下面详述的研究鉴别出可以用于鉴别具有降低的 Gy1、Gy2、Gy3 和 Gy4 蛋白含量的植物的多态性。在本文中鉴别出的 3 个标记，NS0199002、NS0199003 和 NS0199008，可以用于准确预测大豆植物的降低的 Gy1、Gy2、Gy3 和 Gy4 表型。如下面所证实的，Gy1 和 Gy2 的遗传在遗传上连锁，因而 Gy1(NS0199008) 或 Gy2(NS0199002) 的标记可以用于跟踪降低的 Gy1 和 Gy2 表型的遗传。另外，如图 4 所示，降低的 Gy1,2 表型与降低的 Gy3 表型密切相关。因而，Gy1,2 的标记（例如 NS0199008 和 NS0199002）可以另外用于选择子代植物的降低的 Gy3 表型。还表明，标记可以用于鉴别 Gy4 编码的蛋白在表型上无效的植物。例如，NS0199003 标记用于研究中，来精确测定大豆植物的 Gy4 表型。因而，通过使用 Gy1,2,3 标记与 Gy4 遗传标记的组合，本发明使得能实现具有赋予降低的 Gy1、Gy2、Gy3 和 Gy4 表型和高种子 β -伴大豆球蛋白含量的非转基因突变的植物的高通量筛选和标记辅助育种。在本文中还采用测序研究，其已经鉴别出可以用于测定降低的 Gy5 表型的遗传和直接选择降低的 Gy3 表型的标记。

[0047] 还提供了培育大豆植物的方法，所述大豆植物包含赋予降低的脂加氧酶-2 表型的非转基因突变。例如，下面描述的研究鉴别出大豆中与脂加氧酶-2 无效 (1x2) 表型有关的序列变异。已经为 1x2 表型的这些序列变异开发出了分子标记。使用这些与脂加氧酶-2 无效性状有关的标记，育种人可以作出基于标记信息或基因型的选择，而不是基于通过 SDS-PAGE 进行的脂加氧酶分析。标记数据是更节省成本的、快速且可靠的，使得能够测试更大的数目和鉴别具有多种性状（例如脂加氧酶-2 无效和大豆球蛋白-无效）的优良系。

[0048] I. 本发明的植物

[0049] 本发明首次提供了大豆品种植物和其衍生物，其组合了赋予 Gy3 和 Gy4 无效表型和提高的 β -伴大豆球蛋白含量的非转基因突变与农学上优良的表型。在有些实施方案中，这样的植物可以另外包含非转基因的 Gy1 和 Gy2 无效等位基因。这样的植物可以定义

为具有商业上相当大产量，例如，定义为对照系 (check line) AG2703 和 DKB23-51 的至少 103% 的产量。在某些其它实施方案中，提供了包含非转基因的 Gy1-4 突变等位基因和提高的 β -伴大豆球蛋白含量和这些系的至少约 90%、94%、98%、100%、105% 或约 110% 的谷粒产量的植物。在本发明的某些实施方案中，可以将这样的植物定义为在至少 10 种环境中具有超过每英亩约 35、37、39、41、43 或 45 蒲式耳的产量。在某些实施方案中，本发明植物种子的 β -伴大豆球蛋白含量可以大于约 34%、35%、36%、37%、38%、39%、40%、41%、42%、43%、44%、45%、46%、47%、48%、49%、或甚至 50% 或其中可推导出的任意范围。在某些实施方案中，本发明植物可以另外包含赋予 1x1、1x2、和 / 或 1x3 表型的突变。

[0050] 本发明的一方面因此涉及前述植物和其部分和使用这些植物和植物部分的方法。植物部分包括，但不限于，花粉、胚珠和细胞。本发明还提供了这些植物的可再生细胞的组织培养物，所述培养物再生能够表达起始品种的所有生理学和形态学特征的大豆植物。此类可再生的细胞可以包括胚、分生细胞、花粉、叶、根、根尖或者花，或者从其得到的原生质体或愈伤组织。本发明还提供了从此类组织培养物再生的大豆植物，其中该植物能够表达用以得到所述可再生细胞的起始植物品种的所有生理学和形态学特征。

[0051] II. 用于产生具有非转基因的突变 Gy 等位基因和农学上优良表型的大豆品种的标记辅助的选择

[0052] 本发明提供了遗传标记和将非转基因的突变 Gy 等位基因导入农学上优良的大豆植物的方法。本发明因此首次允许产生组合这些赋予高种子 β -伴大豆球蛋白含量的突变 Gy 等位基因与商业上相当大产量和农学优良的遗传背景的植物。使用本发明的方法，赋予“11S 无效”表型的基因座可以导入所需大豆遗传背景中，例如，在产生具有商业上相当大产量和高种子 β -伴大豆球蛋白含量的新品种中。

[0053] 标记辅助的渐渗现象涉及将来自一种种质的一种或多种标记定义的染色体区域转移到第二种种质。该方法中的最初步骤是通过基因作图对性状进行定位，基因作图是通过连锁分析确定基因相对于其它基因和遗传标记的位置的方法。连锁作图的基本原理是染色体上的两个基因越近，它们一起遗传的可能性越高。简言之，通常在两种遗传上相容的但是相对于研究中的性状趋异的亲本之间进行杂交。然后可以用遗传标记跟踪杂交所得的子代，通常为回交 (BC1), F_2 , 或者重组近交群体中所研究的性状的分离。

[0054] 术语数量性状基因座或者 QTL 用于描述基因组的区域，该区域对表型显示出数量或者加性效应。Gy 基因座代表示例性的 QTL，因为多个突变 Gy 等位基因导致种子大豆球蛋白总含量的日益降低和 β -伴大豆球蛋白含量的重要的伴随增加。本文鉴别出了非转基因的突变 Gy 等位基因的遗传标记，其使得能够用农学上优异的植物培育包含非转基因的突变 Gy 等位基因的大豆植物，并选择遗传了突变 Gy 等位基因的子代。因而，本发明允许使用分子工具来组合这些 QTL 与所需农学特征。

[0055] 还鉴定了非转基因的突变 1x1 和 1x2 等位基因的遗传标记，其使得能培育包含非转基因的突变 1x1 和 1x2 等位基因之一或二者的大豆植物，并选择遗传了一个或多个突变 1x 等位基因的子代。具有减少的脂加氧酶的大豆植物本身是有用的，且可以与其它农学特征组合使用。例如，降低的脂加氧酶和降低的大豆球蛋白是有价值的性状组合。另一种有价值的组合是降低的脂加氧酶和中等油酸含量（例如 40–65% 油酸）。因而，本发明允许使用分子工具来在如上所述的农学上优异的植物中组合 1x 等位基因与例如突变 Gy 等位基因。

[0056] A. 连锁的遗传标记的开发和应用

[0057] 可以对样品第一植物群体确定遗传的遗传标记的基因型以形成基因型数据库。本文使用的“遗传的遗传标记”是单个基因座上的等位基因。基因座是染色体上的位置，并且等位基因指这些基因座上基因的状况，即不同的核苷酸序列。每个基因座的标记等位基因组成可以是纯合或者杂合的。为了在杂交中从遗传标记得到信息，标记必须是多态的；即它必须以不同的形式存在，从而携带突变基因的染色体可以通过它携带的标记的形式与具有正常基因的染色体区分开来。

[0058] 通过直接观察来自样品植物的人工或者自然自花传粉的子代的一种或多种性状或者通过定量评估样品植物的组合能力可完成基因型数据库的形成。例如，植物系可以与一种或多种测交品系 (tester) 杂交或者被一种或多种测交品系杂交。测交品系可以是近交系、单交杂种、双交杂种或者多交杂种，或者通过受控制的或者自由交配产生或维持的植物的任何其它集聚，或者其任意组合。对于一些自花传粉植物，优选直接评估，而不用子代测试。

[0059] 可以在测交世代和作图的标记基因座中确定标记基因型。为了通过连锁方法对特定性状作图，必须在特定染色体基因座的遗传和性状的遗传之间确立正相关。对于复合遗传，如数量性状，具体包括大豆球蛋白含量和产量，连锁通常将更难以辨别。在该情况下，可能需要统计学方法来确立表型和基因型之间的相关。这还必须检查来自特定杂交的许多子代，因为单个基因座可以具有对总体表型的小贡献。

[0060] 特定性状和标记的共同遗传 (coinheritance) 或者遗传连锁提示，它们在染色体上物理上相互接近。通过分析杂交中基因和标记的遗传模式来确定连锁。遗传图距离的单位是厘摩 (cM)，它随着重组的增加而增加。两个标记如果在减数分裂中在它们必须重组的每 100 次机会中重组约一次，那么它们相距 1 厘摩。厘摩是一种遗传度量，不是物理度量。位于与第二种基因座小于 50cM 的那些标记被认为是遗传连锁的，因为它们不相互独立地遗传。从而，每代在基因座之间观察到的重组百分比将小于 50%。在本发明的具体实施方案中，使用的标记可以定义为与基因座相距小于约 45、35、25、15、10、5、4、3、2、或 1 或更小的 cM。在本发明的某些实施方案中，可以使用检测有贡献的基因座自身内的多态性、并从而对于所述基因座位于 0cM 处的标记，例如，包含 Gy1、Gy2、Gy3、Gy4 或 Gy5 编码序列或者调节元件内的突变。

[0061] 在减数分裂期间，同源染色体对聚到一起并且在称作重组的过程中交换片段。标记越是远离基因，基因和该标记之间重组的机会越大。在连锁分析中，在特定杂交中跟踪标记和基因或性状的共同遗传。计算它们观察到的遗传模式可以仅仅偶然发生，即它们完全不连锁的概率。重复该计算，假定具体的连锁程度，并确定两种概率（无连锁对特定的连锁程度）的比率。该比率表达了是（及不是）所述连锁程度的可能性，并且因为使用了比率的对数，所以它被称为可能性对数，例如，lod 评分。例如，等于或者大于 3 的 lod 评分用于证实基因和标记连锁。这代表两个基因座连锁的 1000 : 1 可能性。通过使用利用程序的统计学分析极大地促进了连锁计算。

[0062] 通过基因作图模型可以确立标记分子的遗传连锁，所述基因作图模型诸如但不限于，Lander 和 Botstein (1989) 报导的侧翼标记模型，和基于 Lander 和 Botstein (1989) 描述的最大似然法并且用软件包 MAPMAKER / QTL 实现的区间作图。其它的软件包括 Qgene，版

本 2.23(1996) (Department of Plant Breeding and Biometry, 266 Emerson Hall, Cornell University, Ithaca, NY)。

[0063] B. 遗传标记

[0064] 遗传标记包括两个或更多个植物携带的遗传信息中检测得到的差异（多态性）。用遗传标记对基因座的遗传作图通常需要两个基本的组分：可检测的多态性等位基因和那些等位基因的重组或者分离。在植物中，测量的重组事实上总是减数分裂的，且因此，植物基因作图的两个遗传需求是多态性遗传标记，和其中所述那些等位基因是分离的一种或多种植物。

[0065] 标记优选以共显性方式遗传，从而可以容易地检测在二倍体基因座上两种等位基因的存在，并且它们没有环境变异，即它们的遗传率为 1。二倍体生物如大豆中标记基因型通常在每个基因座上包含两个标记等位基因。每个基因座的标记等位基因组成可以是纯合或杂合的。纯合性是其中基因座上的两个等位基因的特征是相同的核苷酸序列的情形。杂合性指基因座上该基因的不同的情形。

[0066] 许多不同的标记类型可以用于遗传作图。用于本发明的示例性遗传标记类型包括但不限于限制性片段长度多态性 (RFLP)、简单序列长度多态性 (SSLP)、扩增片段长度多态性 (AFLP)、单核苷酸多态性 (SNP)、核苷酸插入和 / 或缺失 (INDEL) 和同工酶。可以以多种方式测定包含小至单个核苷酸改变的多态性。例如，通过电泳技术可以进行检测，所述技术包括单链构象多态性 (Orita 等, 1989)、变性梯度凝胶电泳 (Myers 等, 1985)，或者切割片段长度多态性 (Life Technologies, Inc., Gathersberg, MD20877)，但是 DNA 测序机器的普遍可用性常常使得仅仅直接测序扩增的产物更容易。一旦多态性序列差异是已知的，就可以设计快速测定用于子代测试，通常包括某种形式的特定等位基因的 PCR 扩增 (PASA, Sommer, 等, 1992)，或者多个特定等位基因的 PCR 扩增 (PAMSA, Dutton 和 Sommer, 1991)。

[0067] 检测 DNA 样品中 SNP 的一种方法是，使用 PCR 与多态性的荧光探针的组合，如在本文中引作参考的 Livak 等, 1995 和美国专利号 5,604,099 所述。简而言之，合成 2 个探针寡核苷酸，其中的一个退火至 SNP 位点，且另一个退火至野生型序列。优选地，SNP 的位点接近探针寡核苷酸的 5' 末端。然后用非荧光猝灭剂和小沟结合部分（它们分别降低背景荧光和降低寡核苷酸的 T_m ）在 3' 末端标记每个探针。用不同荧光染料标记每个探针的 5' 末端，其中荧光依赖于从探针切割下来的染料。这样的染料的有些非限制性实例包括 VICTM 和 6-FAMTM。然后对被怀疑包含给定 SNP 的 DNA 进行 PCR，使用具有 5' -3' 外切核酸酶活性的聚合酶和侧翼引物。在有两种探针寡核苷酸存在下，进行 PCR。如果探针结合试验 DNA 中的互补序列，则聚合酶的外切核酸酶活性释放出荧光标记，激活它的荧光活性。因此，仅含有野生型序列的试验 DNA 将表现出与野生型探针上的标记有关的荧光。另一方面，仅含有 SNP 序列的 DNA 将具有来自 SNP 探针上的标记的荧光活性。但是，在 DNA 来自异质来源的情况下，将观察到两种标记的明显荧光。在已知 SNP 位点处的这类间接基因型分型，使得能高通量地、廉价地筛选 DNA 样品。因而，对于包含突变 Gy 等位基因的子代大豆植物的鉴别，这样的系统是理想的。

[0068] 限制性片段长度多态性 (RFLP) 是 DNA 经限制性内切核酸酶消化后，一般通过琼脂糖凝胶电泳揭示出的 DNA 片段长度可以检测的遗传差异。可以利用多种限制性内切核酸酶，其特征是它们的核苷酸切割位点和它们的来源，例如，EcoRI。RFLP 来自限制位点序列

内的单个碱基对多态性和给定限制片段内的可测量的插入或者缺失。RFLP 是容易和相对廉价产生的（需要克隆的 DNA，但是不需要序列），并且是共显性的。RFLP 的缺点是在分型阶段是劳动密集型的，尽管这可以通过许多任务的复合（complexing）和印迹的再利用在一定程度上减轻。大多数 RFLP 是双等位基因（biallelic）的，并且比微卫星具有更小的多态性含量。由于该原因，在植物遗传学图中 RFLP 的使用已经减少。

[0069] 本领域技术人员将认识到许多类型的分子标记可以用作监控基因遗传的工具，并且不限于 RFLP、SSR 和 SNP，并且技术人员将还理解，可以用多种检测方法来跟踪多种分子标记。本领域技术人员将还认识到，不同类型的标记可以用于作图，特别是随着技术的发展和鉴定出了新类型的标记和鉴定方法。

[0070] 为了方便的目的，可以将遗传的标记基因型转化成数字得分，例如，如果使用特定酶，在特定基因座上有称作 A 和 B 的两种形式的 SNP，或者其它标记，那么可以将二倍体套转化成数字得分，例如，为 AA=2, AB=1, 和 BB=0；或 AA=1, AB=0 和 BB=-1。得分的绝对值不是重要的。重要的是数字标记的加性性质。上面的得分涉及共显性标记。可以给予类似的评分系统，其与显性标记相一致。

[0071] C. 标记辅助的选择

[0072] 本发明提供了具有提高的 β -伴大豆球蛋白含量组合商业上相当大产量和农学上优良特征的大豆植物。此类植物可以根据本发明通过标记辅助的选择方法产生，所述方法包括，测定基因组 DNA 中与非转基因的突变 Gy1、Gy2、Gy3、Gy4、或 Gy5 等位基因包括其所有可能的组合遗传连锁的标记的存在。本发明也提供了具有降低的脂加氧酶含量的大豆植物。这样的植物可以根据本发明通过标记辅助的选择方法产生，所述方法包括，测定基因组 DNA 中与非转基因的突变 lox1、Lox2、或 Lox3 等位基因包括其所有可能的组合遗传连锁的标记的存在。

[0073] 在本发明的某些实施方案中，可能需要得到与 Gy 等位基因连锁的额外标记。这可以例如如下进行：首先通过自交 F_1 杂种制备 F_2 群体，所述 F_1 杂种通过近交品种的杂交产生，所述近交品种中仅一个包含赋予降低的大豆球蛋白含量的突变 Gy 等位基因。然后可以制备重组近交系（RIL）（遗传上相关的系；通常 $>F_5$ ，从朝纯合性连续自交 F_2 系开发），并将其用作作图群体。从显性标记得到的信息可以通过使用 RIL 最大化，因为所有基因座都是纯合的或者接近纯合的。也可能需要得到与 Lox 等位基因连锁的额外标记。这可以例如如下进行：首先通过自交 F_1 杂种制备 F_2 群体，所述 F_1 杂种通过近交品种的杂交产生，所述近交品种中仅一个包含赋予降低的脂加氧酶含量的突变 Lox 等位基因。如上所述，然后可以制备重组近交系，并将其用作作图群体，且从显性标记得到的信息可以通过使用 RIL 最大化，因为所有基因座都是纯合的或者接近纯合的。

[0074] 本发明也提供了具有降低的含量的 KTI、例如 KTI 无效性状的大豆植物，其可以通过标记辅助的选择来得到，并与商业上相当大产量和农学上优良特征组合提供。在本发明的某些实施方案中，可以将这样的植物定义为在至少 10 种环境中具有超过每英亩约 35、37、39、41、43 或 45 蒲式耳的产量。在本发明的某些实施方案中，用于这样的标记辅助的选择的标记可以包括 SNP 或 INDEL。在本发明的具体实施方案中，一个或多个 INDEL 可以存在于编码 Kunitz 胰蛋白酶抑制剂的基因中。在本发明的一个实施方案中，所述标记可以包含在编码 Kunitz 胰蛋白酶抑制剂的基因（SEQ ID NO :167）的位置 622-623 处的缺失和 / 或

在位置 664 处的插入,如图 9 所示,且所述植物可以根据本发明通过标记辅助的选择方法产生,所述方法包括,测定基因组 DNA 中这种标记的存在。

[0075] 携带以前不存在的性状的回交群体(例如从所需的品种(回归亲本)和另一品种(供方亲本)之间的杂交产生)也可以用作作图群体。可以进行与回归亲本的一系列回交以恢复其所需的性状的大多数。从而,产生由与回归亲本相似的个体组成的群体,但是每个个体携带来自供方亲本的不同量的基因组区域。如果回归亲本中的所有基因座都是纯合的并且供方亲本和回归亲本具有对比多态性标记等位基因,那么回交群体可以用于显性标记作图(Reiter 等,1992)。

[0076] 用于作图目的的有用的群体是近等基因系(NIL)。通过多次回交以产生一系列个体来产生 NIL,所述个体除了所需的性状或者基因组区域外在遗传组成上接近相同,其可以用作作图群体。在用 NIL 作图中,预期仅部分多态性基因座作图到所选的区域。也可以在转化的植物系上进行作图。

[0077] D. 植物育种方法

[0078] 本发明的某些方面提供了标记辅助的植物育种的方法,其使得能将非转基因的突变 Gy 等位基因导入异源大豆遗传背景中。本发明的某些方面还提供了标记辅助的植物育种的方法,其使得能将非转基因的突变 lox 等位基因导入异源大豆遗传背景中。一般而言,育种技术利用了植物的传粉方法。有两种一般的传粉方法:自花传粉,其在如果来自一朵花的花粉转移到相同植物的相同或者另一朵花时发生;和异花传粉,其在如果花粉来自不同植物上的花时发生。已经自花传粉并且对许多世代进行类型选择的植物在几乎所有基因座上都纯合,并且产生真实育种子代的均一群体,纯合植物。

[0079] 在合适的品种的开发中,可以使用系谱育种。特定性状的系谱育种方法涉及杂交两种基因型。每种基因型可以具有另一种基因型中缺乏的一种或多种所需的特征;或者每种基因型可以与另一种互补。如果两种最初的亲本基因型不提供所有所需的特征,那么在育种群体中可以包括其它基因型。将作为这些杂交产物的优异植物自交并在每个连续世代中再次自交。每个连续世代由于自花传粉和选择而变得更均质。通常,这种育种方法涉及 5 代或者更多代的自交和选择: $S_1 \rightarrow S_2 ; S_2 \rightarrow S_3 ; S_3 \rightarrow S_4 ; S_4 \rightarrow S_5$ 等等。自交的世代(S)可以被认为是一种类型的杂交后代(F)并且同样命名为 F。至少 5 代后,认为近交植物是遗传上纯的。

[0080] 每个育种程序将包括育种过程效率的周期性的客观评价。评价标准依赖于目标和目的而变。将有希望的较晚育种系在代表商业靶区域的环境中充分测试并与合适的标准一般比较 3 年或更多年。难以鉴定遗传上优异的个体,因为混杂的植物性状或者环境因素可以掩盖基因型值。鉴定优异植物的一种方法是观察其相对于其它实验植物和一种或多种广泛栽培的标准品种的性能。单次观察可能是不确定的,但是重复观察提供了对遗传价值的更好的估计。

[0081] 混合选择和轮回选择可以用于改进自花或者异花传粉的农作物的群体。通过几种不同亲本的互交可以鉴定或者产生杂合个体的遗传变异群体。基于个体优越性、杰出的子代或者卓越的组合能力选择最好的植物。将所选的植物互交以产生新的群体,其中继续其它轮的选择。关于常用于不同性状和农作物的其它育种方法的描述可以见几种参考书籍之一(例如,Allard,1960;Simmonds,1979;Sneep 等,1979;Fehr,1987a,b)。

[0082] 在育种程序中对具有目的性状（例如，高产量、抗病性、脂肪酸谱）的基因型的选择的有效性将取决于：1) 群体中个体植物的目的性状的变异性是遗传因素的结果并且从而传递到所选基因型的子代的程度；和 2) 植物中目的性状中多少变异性是由于不同基因型所生长的环境。性状的遗传范围为通过一种主要基因控制（其表达不受环境的影响（即，质量性状））到受到许多基因的控制（其效果受到环境的很大影响（即数量性状））。对于数量性状如产量的育种的其它特征是如下事实：1) 每种基因的作用导致的差异是小的，从而使得难以或者不可能单独地鉴定它们；2) 有助于一种性状的基因的数目很大，从而很少（如果得到的话）得到不同的分离比；和 3) 基因的作用可以基于环境变异以不同方式表达。因此，具有目的性状的超亲分离物（transgressive segregate）或者优异基因型的准确鉴定极端困难，并且它的成功取决于植物育种者使得影响群体中数量性状的表达的环境变异最小化的能力。

[0083] 随着组合到一种基因型的性状的数目的增加，鉴定超亲分离子的可能性很快减小。例如，如果在三种复杂性状（如产量、 β -伴大豆球蛋白含量和至少第一种农学性状）有差别的栽培种之间进行杂交，那么不用分子工具极其难以通过将三种性状的每一种的最大数目的有利基因重组到一个基因型中同时回收。因此，所有育种者一般所希望的是得到除所选基因外到一个基因型中的第一种复杂性状的基因的有利分配与第二种性状的基因的有利分配组合。

[0084] 回交是转移特定所需性状的有效方法。这可以通过例如首先将优异品种近交系（inbred）(A)（回归亲本）与供方近交系（非回归亲本）杂交来实现，所述供方近交系携带对于研究中的性状合适的基因（Fehr, 1987）。然后将该杂交的子代与优异回归亲本（A）回交，接着在所得子代中选择将从非回归亲本转移的所需性状。这种选择可以基于如下面提到的遗传测定，或者备选地，可以基于子代植物的表型。经过 5 或者更多回交世代选择所需的性状后，子代对于控制正转移的性状的基因座是杂合的，但是对于大多数或者几乎所有其它基因与优异亲本相似。将回交的最后世代自交，或者同胞交配，以得到所转移的基因（例如，提供具有降低的种子大豆球蛋白含量的植物的基因座）的纯育子代。

[0085] 在本发明的一个实施方案中，回交转变的方法可以定义为包括下面步骤的方法：

[0086] (a) 将含有一种或多种所需的基因、DNA 序列或者元件（如与降低的种子大豆球蛋白含量有关的突变 Gy1、Gy2、Gy3、Gy4 和 / 或 Gy5 等位基因）的第一种基因型的植物与缺乏所述所需的基因、DNA 序列或者元件的第二种基因型的植物杂交；

[0087] (b) 选择含有所述所需的基因、DNA 序列或者元件的一种或多种子代植物；

[0088] (c) 将所述子代植物与第二种基因型的植物杂交；并

[0089] (d) 重复步骤 (b) 和 (c)，以用于将来自第一种基因型的植物的所述所需的基因、DNA 序列或者元件转移到第二种基因型的植物中。

[0090] 特定 DNA 元件或者一组元件向植物基因型中的渐渗现象定义为回交转变方法的结果。DNA 序列已经渐渗入其中的植物基因型可以称作回交转变的基因型、系、近交系或者杂种。类似地，缺乏所需的 DNA 序列的植物基因型可以称作未转变的基因型、系、近交系或者杂种。在育种过程中，与降低的大豆球蛋白含量连锁的遗传标记可以用于辅助育种，以用于产生具有降低的大豆球蛋白含量、且优选提高的 β -伴大豆球蛋白含量的大豆植物。回交和标记辅助的选择尤其可以用于本发明，以将根据本发明的降低的大豆球蛋白含量性状

引入任意品种,这通过用有关的非转基因的突变 Gy1、Gy2、Gy3、Gy4 和 / 或 Gy5 等位基因转变该品种来实现。

[0091] 合适的回归亲本的选择是成功的回交过程的重要步骤。回交规程的目的是改变或者代替最初近交系中的性状或者特征。为了完成该目的,将回归近交系的一个或多个基因座用来自非回归亲本的所需基因修饰或者代替,而保持最初近交系的基本上所有剩余的所需遗传的,并且因此所需生理学和形态学的构成。具体非回归亲本的选择将取决于回交的目的,其在本发明的情况下可以为加入赋予降低的大豆球蛋白含量的一种或多种等位基因。确切的回交规程将取决于被改变的特征或者性状以确定合适的测试规程。尽管当被转移的特征是显性等位基因时,回交方法可以简化,但也可以转移隐性等位基因。在该情况下,必须引入子代的测试以确定是否已经成功地转移了所需的特征。对于本发明,可以测试在回交程序期间产生的子代系的大豆球蛋白含量,例如通过 SDS-PAGE / 考马斯染色,以及使用本文描述的标记系统来基于标记而不是基于视觉性状选择系。

[0092] 通过天然或者机械技术可以杂交大豆植物(大豆(Glycine max L.))(见例如, Fehr, 1980)。通过自花传粉或者天然异花传粉在大豆中发生天然传粉,其通常得到传粉生物的帮助。在天然或者人工杂交中,开花和开花时间是重要的考虑因素。大豆是短日照植物,但是对光周期的敏感性有相当大的遗传变异(Hamner, 1969; Criswell 和 Hume, 1972)。开花的临界日长范围为对于适应热带纬度的基因型的约 13 小时到对于在更高纬度生长的光周期不敏感的基因型的 24 小时(Shibles 等, 1975)。大豆在出苗(emergence)后 9 天对日长似乎不敏感。需要 7 到 26 天的短于临界日长的光周期来完成花诱导(Borthwick 和 Parker, 1938; Shanmugasundaram 和 Tsou, 1978)。

[0093] 使用和不使用雌花的去势,可以通过用镊子从雄性亲本的花取出雄蕊和雌蕊并对雌花的柱头轻轻刷花药进行手工传粉。通过除去前萼片和龙骨瓣花瓣,或者用闭合的镊子穿刺龙骨瓣并允许它们打开以推开花瓣可以实现接近雄蕊。在柱头上刷花药导致它们破裂,且当在柱头上清晰可见花粉时,得到最高百分比的成功杂交。可以通过在刷柱头前轻打花药检查花粉的脱落。当条件不利时,可能必须使用几朵雄花以得到合适的花粉脱落,或者可以使用具有良好花粉脱落的相同雄花对几朵花传粉。

[0094] 遗传雄性不育可以在大豆中得到并且可以用于促进本发明背景中的杂交,尤其对于轮回选择程序(Brim 和 Stuber, 1973)。完全分离杂交区组(block)所需的距离还不清楚;然而,当雄性不育植物与外来花粉来源为 12m 或者更远时,异型杂交小于 0.5% (Boerma 和 Moradshahi, 1975)。杂交区组的边界上的植物与外来花粉维持最多的异型杂交并且可以在收获时将其除去以使污染最小化。

[0095] 一旦收获,就通常将豆荚在至多 38°C 的温度下风干,直到种子含有 13% 水分或者更小,然后手工取出种子。如果相对湿度为 50% 或者更低,则可以将种子在约 25°C 下令人满意的保存最长达 1 年。在湿润气候中,萌发百分比快速降低,除非种子被干燥到 7% 水分并且在密封容器中室温下保存。通过将种子干燥到 7% 水分并将其在 10°C 或更低在维持 50% 相对湿度的房间或者在密封容器中保存最好地完成了在任何气候中的长期保存。

[0096] III. 用于修饰和改进大豆品种的性状

[0097] 在某些实施方案中,本发明提供的大豆植物可以包含一种或多种转基因。此类转基因的一个实例赋予除草剂抗性。通常的除草剂抗性基因包括赋予草甘膦抗性的 EPSPS 基

因,赋予对卡那霉素抗性的新霉素磷酸转移酶 II (nptII) 基因 (Fraley 等, 1983), 赋予对抗生素潮霉素抗性的潮霉素磷酸转移酶基因 (Vanden Elzen 等, 1985), 赋予对草铵膦或者 broxynil 的抗性的基因 (Comai 等, 1985 ;Gordon-Kamm 等, 1990 ;Stalker 等, 1988), 如二氢叶酸还原酶和乙酰乳酸合酶 (Eichholtz 等, 1987, Shah 等, 1986, Charest 等, 1990)。其它实例包括赋予对咪唑啉酮类 (imidazolinone) 或者磺酰脲类的抗性的突变 ALS 和 AHAS 酶 (Lee 等, 1988 ;Miki 等, 1990)、赋予膦丝菌素抗性的膦丝菌素 - 乙酰 - 转移酶基因 (欧洲申请 0 242 246)、赋予苯氧基丙酸和环己酮 (cyclohexones) 如稀禾定 (sethoxydim) 和吡氟氯禾灵的抗性的基因 (Marshall 等, 1992) ;和赋予对三嗪的抗性的基因 (psbA 和 gs+ 基因) 和对芸腈抗性的基因 (腈水解酶基因) (Przibilla 等, 1991)。

[0098] 本发明的植物还可以包含赋予对昆虫、害虫、病毒或者细菌攻击的抗性的基因。例如,赋予对害虫如大豆胞囊线虫的抗性的基因在 PCT 申请 W096 / 30517 和 PCT 申请 W093 / 19181 中描述。Jones 等, (1994) 描述了关于对暗黄枝孢 (*Cladosporium fulvum*) 的抗性的番茄 Cf-9 基因的克隆 ;Martin 等, (1993) 描述了关于对丁香假单胞菌致病变种 (*Pseudomonas syringae* pv.) 的抗性的番茄 Pto 基因, 以及 Mindrinos 等, (1994) 描述了关于对丁香假单胞菌 (*Pseudomonas syringae*) 的抗性的鼠耳芥 (*Arabidopsis*) RSP2 基因。苏云金芽孢杆菌 (*Bacillus thuringiensis*) 内毒素也可以用于昆虫抗性 (见, 例如, Geiser 等, (1986))。维生素结合蛋白, 如抗生素蛋白也可以用作杀幼虫剂 (PCT 申请 US93 / 06487)。

[0099] 已知病毒外壳蛋白在转化的植物细胞中的应用赋予对该外壳蛋白基因所来源的病毒以及相关病毒影响的病毒感染和 / 或疾病发展的抗性 (见 Beachy 等, 1990)。已经赋予转化的植物对苜蓿花叶病毒、黄瓜花叶病毒、烟草线条病毒、马铃薯 X 病毒、马铃薯 Y 病毒、烟草蚀斑病毒、烟草脆裂病毒和烟草花叶病毒的外壳蛋白介导的抗性。同前。还可以使用通过病原体或者寄生物自然产生的发育停顿蛋白。例如, Logemann 等, (1992) 已经表明表达大麦核糖体灭活基因的转基因植物对真菌疾病已具有增强的抗性。

[0100] 还可以使用赋予增加的营养价值或者另一种增值性性状的转基因。一个实例是修饰的脂肪酸代谢, 例如, 通过用硬脂酰 -ACP 去饱和酶的反义基因转化植物以增加植物的硬脂酸含量 (见 Knutzon 等, 1992)。还可以引入有义去饱和酶基因以改变脂肪酸含量。通过导入肌醇六磷酸酶编码基因来增强肌醇六磷酸的分解来修饰肌醇六磷酸含量, 从而向转化的植物添加更多的游离磷酸。也可以例如通过用编码改变淀粉的分支模式的酶的基因转化植物来影响改变的糖类组成。(见 Shiroza 等, 1988) (变异链球菌 (*Streptococcus mutans*) 果糖基转移酶基因的核苷酸序列) ;Steinmetz 等, (1985) (枯草芽孢杆菌 (*Bacillus subtilis*) 果聚糖蔗糖酶基因的核苷酸序列) ;Pen 等, (1992) (产生表达地衣芽孢杆菌 (*Bacillus licheniformis*) α - 淀粉酶的转基因植物) ;Elliot 等, (1993) (番茄转化酶基因的核苷酸序列) ;Søgaard 等, (1993) (大麦 α - 淀粉酶基因的定点诱变) ;和 Fisher 等, (1993) (玉米胚乳淀粉分支酶 II))。

[0101] 转基因还可以用于改变蛋白代谢。例如, 美国专利号 5,545,545 描述了赖氨酸不敏感的玉米二氢吡啶二羧酸合酶 (DHPS), 其显著抗要不然抑制天然 DHPS 活性的 L- 赖氨酸的浓度。类似地, EP0640141 描述了编码赖氨酸不敏感的天冬氨酸激酶 (AK) 的序列, 其能够导致高于正常的苏氨酸生产, 以及编码反义赖氨酸 - 酮戊二酸还原酶的亚片段, 其用于

增加赖氨酸。

[0102] 在另一实施方案中,可以使用改变植物糖类代谢的转基因。例如,已知果糖激酶基因用于转基因植物和它们的果实中果糖激酶基因表达的代谢工程改造(见美国专利号6,031,154)。可以使用的转基因的另一实例是改变谷粒产量的基因。例如,美国专利号6,486,383描述了用腺苷二磷酸葡萄糖焦磷酸化酶(“ADPG PPase”)的亚基蛋白修饰植物中的淀粉含量。在EP0797673中,讨论了转基因植物,其中特定DNA分子的导入和表达导致在液泡外形成容易活动化的磷酸库和增强的生物量产生和/或改变的开花行为。还已知改变植物成熟度的基因。美国专利号6,774,284描述了编码植物脂肪酶的DNA和其用于控制植物中衰老的方法。美国专利号6,140,085讨论了改变开花特征,尤其开花时间选择的FCA基因。美国专利号5,637,785讨论了基因修饰植物,其具有受调节的花发育,如具有早期花分生组织发育并且在其基因组中包含编码LEAFY蛋白的结构基因。

[0103] 用于改变植物形态学特征的基因也是已知的并且可以根据本发明使用。美国专利号6,184,440讨论了基因工程改造的植物,其由于表达细胞壁调节转基因而显示出改变的结构或者形态学。细胞壁调节转基因的实例包括纤维素结合结构域,纤维素结合蛋白,或者细胞壁修饰蛋白或者酶,如内切木葡聚糖(endoxylglucan)转移酶、木葡聚糖内切转糖基酶、扩展蛋白、纤维素合酶,或者新分离的内切-1,4- β -葡聚糖酶。

[0104] 导入转基因的方法是本领域公知的并且包括生物学和物理学的植物转化规程。见例如Miki等(1993)。

[0105] 一旦转基因被导入品种中,它就可以容易地通过杂交转移。通过使用回交,除了通过回交技术转移到品种中的基因座外,回收了品种的基本上所有所需的形态学和生理学特征。回交方法可以用于本发明以改善特征或向植物中引入特征(Poehlman等,1995;Fehr,1987a,b)。

[0106] IV. 大豆植物的组织培养物和体外再生

[0107] 本发明的另一方面涉及本发明的大豆品种的组织培养物。本文使用的术语“组织培养物”指包含相同或不同类型的分离的细胞的组合物或者组织成植物的部分的一组这类细胞。组织培养物的示例性类型是在植物或植物部分中完整的原生质体、愈伤组织和植物细胞,如胚、花粉、花、叶、根、根尖、花药等等。在优选实施方案中,组织培养物包含胚、原生质体、分生细胞、花粉、叶或花药。

[0108] 用于制备可再生大豆细胞的组织培养物和从其再生大豆植物的示例性方法公开在美国专利号4,992,375;美国专利号5,015,580;美国专利号5,024,944和美国专利号5,416,011中,将它们每一个的公开内容都完整地特别引入本文作为参考。

[0109] 组织培养物的重要的能力是能够再生能育的植物。这允许例如转化组织培养物细胞后再生转基因植物。为了有效转化和成功转化,必须将DNA导入细胞中,该细胞产生植物或者种系组织。

[0110] 大豆通常通过两种不同的过程再生:枝条形态发生和体细胞胚胎发生(Finer,1996)。枝条形态发生是枝条分生组织组构和发育的过程。枝条从来源组织长出并被切断和生根以得到完整植物。在体细胞胚胎发生中,从体细胞植物组织形成含有枝条和根轴的胚(类似于合子胚)。从体细胞胚的萌发得到完整植物而不是生根的枝条。

[0111] 枝条形态发生和体细胞胚胎发生是不同的过程并且再生的特定途径主要依赖于

用于组织培养操作的外植体来源和培养基。尽管系统是不同的,但是两种系统都显示出品种特异性应答,其中一些系对于组织培养操作比其它系更有应答性。在枝条形态发生中高度应答的系不能产生许多体细胞胚。在“诱导”步骤过程中产生许多胚的系可能不产生快速生长的增殖培养物。因此,需要对每种大豆系最优化组织培养条件。这些最优化可以由组织培养技术领域中技术人员通过小规模培养研究容易地实施。除了系特异性应答外,也可以在枝条形态发生和体细胞胚胎发生两者中观察到增殖性培养物。增殖对于两种系统都是有益的,因为它允许单个被转化的细胞增殖到将有助于种系组织的点。

[0112] Wright 等 (1986) 首次报导枝条形态发生为从大豆幼苗的子叶节从头得到枝条所借助的系统。枝条分生组织在表皮下形成并且形态发生组织可以在含有苄基腺嘌呤 (BA) 的培养基上增殖。如果认识到枝条的表皮下多细胞起源并且利用增殖培养物,那么该系统可以用于转化。想法是靶定将产生新枝条的组织并且增殖分生组织内的那些细胞以减少与嵌合状态有关的问题。如果被转化的细胞不足够增殖并且不产生种系组织,那么在分生组织内仅单个细胞的转化得到的嵌合体的形成是有问题的。一旦充分理解了该系统并且令人满意的重复该系统,那么它就可以用作大豆转化的一种靶组织。

[0113] Christianson 等 (1983) 首次报导大豆中体细胞胚胎发生是其中最初从合子胚轴得到胚胎发生组织的系统。这些胚胎发生培养物是增殖的,但是该系统的重复性低并且没有报导胚的起源。后来对不同增殖性胚胎发生大豆培养物的组织学研究表明增殖性胚是顶端或者表面起源的,其中少数组细胞有助于胚形成。初生胚 (来自最初外植体的第一个胚) 的起源取决于外植体组织和诱导培养基中的生长素水平 (Hartweck 等, 1988)。使用增殖性胚培养物,“较老”体细胞胚的单个的细胞或者小组表面细胞形成“较新”的胚。

[0114] 如果认识到胚的起源并且理解了增殖性胚胎发生培养物的生物学限制,那么胚胎发生培养物还可以成功地用于再生,包括转基因植物的再生。生物学限制包括开发增殖性胚胎发生培养物的困难和与从长期增殖性胚胎发生培养物再生的植物有关的降低的能育性问题 (培养物诱导的变异)。在长时间的培养中这些问题中的一些更明显。使用更近期培养的细胞可以减小或者消除此类问题。

[0115] V. 大豆植物的利用

[0116] 本发明提供的大豆植物可以用于认为有价值的任意目的。通常的用途包括制备用于人类消费的食物、用于非人动物消费的饲料和工业用途。本文使用的“工业用途”或者“工业使用”指大豆或基于大豆的产品的非食物和非饲料用途。

[0117] 通常将大豆加工成两种主要产品:大豆蛋白(粗粉)和粗大豆油。这两种产品通常被进一步精炼以用于特定目的。精炼的油产品可以分解成甘油、脂肪酸和固醇。这些可以用于食物、饲料或工业使用。可食用的食品用途例子包括稀奶油、人造黄油、蛋黄酱、药物、色拉调料、起酥油、烘焙产品和巧克力糖衣。

[0118] 大豆蛋白产品(例如,粗粉)可以分成具有食物/饲料和工业用途的大豆面粉浓缩物和分离物。大豆面粉和渣(grit)通常用于生产肉类补充剂(extender)和类似物、宠物食物、烘焙成分和其它食品。由大豆面粉和分离物制备的食品包括婴儿食物、糖果产品、谷类、食物饮料、面条、酵母、啤酒、上面啤酒等等。大豆粗粉尤其常用作家畜饲养,主要是猪和家禽的蛋白来源。饲料用途从而包括但不限于,水产养殖饲料、蜜蜂饲料、小牛饲料替代品、鱼饲料、家畜饲料、家禽饲料和宠物饲料等等。

[0119] 完整大豆产品也可以用作食物或饲料。通常的食物使用包括这样的产品，诸如种子、豆芽、烘焙的大豆、用于各种烘焙产品的全脂大豆面粉、用作糖食的烘烤大豆、大豆坚果仁酱 (soy nut butter)、大豆咖啡 (soy coffee) 和东方食物的其它大豆衍生物。对于饲料使用，通常从大豆取出壳并将其用作饲料。

[0120] 大豆还具有许多工业用途。大豆的一种常见的工业使用是制备可以用于生产复合材料的粘合剂。例如，使用改性的大豆蛋白、水解的大豆蛋白和 PF 树脂的混合物、含有粉末树脂的大豆面粉和含有泡沫胶水的大豆蛋白可以生产木材复合材料。利用基于大豆的粘合剂生产常见的木材产品，如胶合板已经超过 70 年了。尽管脲甲醛和酚醛树脂的引入已经降低了木材产品中基于大豆的粘合剂的使用，但是环境关心和消费者对从可再生的原料生产的粘合剂的偏爱已经导致重新出现对开发用于木材复合材料工业的新的基于大豆的产品的兴趣。

[0121] 粘合剂的制备代表大豆的另一常见的工业使用。大豆粘合剂的实例包括大豆水解产物粘合剂和大豆面粉粘合剂。大豆水解产物是无色的水溶液，其通过将大豆蛋白分离物在 5% 氢氧化钠溶液中在热 (120 °C) 和压力 (30psig) 下反应而制备。得到的降解的大豆蛋白溶液在室温下是碱性的 (pH11) 和可流动的 (约 500cps)。大豆面粉是从大豆制备的精细研磨的脱脂粗粉。可以从大豆面粉制备各种粘合剂制剂，第一步通常需要将面粉溶解在氢氧化钠溶液中。所得制剂的强度和其它性质将依赖于制剂中的添加剂而变。大豆面粉粘合剂还可能与其它可商购的树脂组合。

[0122] 大豆油还可以应用于多种工业用途。大豆油是世界上最容易得到的和成本最低的植物油之一。大豆油的常见的工业用途包括用作抗静电剂、填隙化合物、消毒剂、杀真菌剂、墨水、油漆、保护涂层、壁板、消泡剂、醇、人造黄油、油漆、墨水、橡胶、起酥油、化妆品等等的组分。大豆油多年来也是醇酸树脂的主要成分，所述醇酸树脂溶解在载体溶剂中以制备基于油的油漆。在热和压力下将植物油转化成为醇酸树脂的基本化学是本领域技术人员充分了解的。

[0123] 处于可商购的未精炼或者精炼的食用级状态的大豆油是相当稳定和干燥缓慢的油。还可以改性大豆油以增强它在环境条件下的反应性，或者输入各种形式的能量，以导致油共聚或者熟化成干燥膜。这些形式的改性中的一些包括环氧化、醇解或者酯交换、直接酯化、复分解、异构化、单体修饰、和各种形式的聚合作用，包括热粘化。对于许多工业用途，大豆油的反应性亚油酸组分与其双键可以比主要的油酸和亚油酸组分更有用。

[0124] 还可以使用基于大豆的成分制备溶剂。例如，豆油脂肪酸甲酯 (methyl soyate) - 一种基于大豆油的甲酯 - 正在市场上被接受为在诸如部件清洁和脱脂、油漆和墨水去除和漏油补救应用中的优秀的溶剂替代备选方案。它还以多种配制的消费者产品上市，所述产品包括洗手剂、汽车蜡和乱画洗净剂。通过大豆油与甲醇的酯交换产生豆油脂肪酸甲酯。它可以从多种生产商和供应商商购得到。作为溶剂，豆油脂肪酸甲酯具有重要的环境和安全性相关的性质，其使得它对于工业应用是有吸引力的。它的毒性比大多数其它溶剂更低，容易生物降解，并且具有非常高的闪点和低水平的挥发性有机化合物 (VOC)。豆油脂肪酸甲酯与金属、塑料、大多数高弹体和其它有机溶剂的相容性是优良的。豆油脂肪酸甲酯的当前用途包括清洁剂、脱漆剂、漏油清除和生物整治 (bioremediation)、杀虫剂佐剂、防腐剂和生物柴油机燃料添加剂。

[0125] VI. 保藏信息

[0126] 本文公开的大豆系B2G2和系谱3的至少2500粒种子已经保藏于美国典型培养物保藏中心(ATCC), 10801 University Boulevard, Manassas, Va. 20110-2209 USA, 保藏日为2005年7月28日。所述保藏物确定的ATCC保藏号分别为PTA-6893和PTA-6892。种子由ATCC保藏。按照布达佩斯条约, 在本申请未决期间, 可以获取该保藏物。所述保藏物将维持在ATCC保藏机构(它是公开的保藏机构), 保藏期为30年或保藏机构最后接到要求之后至少5年, 或保藏到该专利的可实施寿命, 哪一项时间更长就取决于哪一项, 并且在所述时间段内, 如果保藏物变得不可存活, 则要进行更换。

[0127] VII. 试剂盒

[0128] 本文所述的任意组合物可以包含在试剂盒中。在非限制性实施例中, 用于检测本文所述多态性的组合物和/或其它试剂, 可以包含在试剂盒中。试剂盒因而可以在合适的容器中包含用于检测多态性的探针或引物和/或本发明的其它试剂。在具体实施方案中, 试剂盒将允许检测至少一种非转基因的Gy无效等位基因, 且另外可以检测脂加氧酶和/或KTI无效等位基因, 例如, 通过检测这样的等位基因和/或否则含有等位基因的连锁不平衡中的多态性。

[0129] 试剂盒可以包含适当等分试样的本发明试剂组合物, 其可以标记或未标记, 用于检测这样的等位基因的任意测定形式中。试剂盒组分可以包装在含水介质中, 或以冻干的形式。试剂盒的容器通常包括至少一个小瓶、试管、烧瓶、瓶、注射器或其它容器, 其中可以装有组分, 且优选地适当等分试样的。当试剂盒中存在超过一种组分时, 试剂盒通常也将含有第二个、第三个或其它额外容器, 其中可以分别装有其它组分。但是, 组分的各种组合可以包含在小瓶中。本发明的试剂盒通常也将包括容纳检测组合物的工具和商业销售紧密密封的任意其它试剂容器。这样的容器可以包括注射或吹塑塑料容器, 其中保持所需小瓶。

[0130] 当试剂盒的组分提供在一种和/或多种液体溶液中时, 液体溶液可以是含水溶液, 无菌含水溶液是特别优选的。但是, 试剂盒的组分可以作为干燥粉末来提供。当试剂和/或组分作为干燥粉末来提供时, 粉末可以通过加入合适溶剂来重建。预见到, 溶剂也可以在另一个容器中提供。容器通常将包括至少一个小瓶、试管、烧瓶、瓶、注射器和/或其它容器, 其中可以装有用于检测无效等位基因的组合物, 优选地适当分配。试剂盒也可以包含第二个容器, 其装有无菌缓冲液和/或其它稀释剂。

[0131] 本发明的试剂盒也通常将包括容纳商业销售紧密密封的小瓶的工具, 例如注射和/或吹塑塑料容器, 其中保持所需小瓶。无论容器的数目和/或类型, 本发明的试剂盒也可以包含用于辅助使用检测组合物的仪器, 和/或与其一起包装。

[0132] VIII. 定义

[0133] 在下面的描述和表中, 使用了许多术语。为了提供对说明书和权利要求书的清楚和一致的理解, 提供下面的定义:

[0134] 一个/种: 当与词语“包含”或者权利要求书中的其它开放用语结合使用时, 词语“一个/种”指“一个/种或多个/种”。

[0135] 农学上优良的: 如本文使用的, 指基因型, 其具有许多可辨别性状的最佳表现, 如种子产量、突出体、萌发势、营养势、抗病性、种子结实、站立能力(standability)和脱粒能力, 其允许生产者收获具有商业重要性的产品。

[0136] 等位基因 : 基因座的一种或多种可选形式的任一种,所有等位基因都涉及性状或者特征。在二倍体细胞或者生物中,给定基因的两个等位基因占据一对同源染色体上的对应基因座。

[0137] 回交 : 一种方法,其中育种者将杂种子代例如第一代杂种 (F_1) 与杂种子代的亲本之一重复杂交。回交可以用于从一种遗传背景向另一遗传背景引入一个或多个单一基因座转变。

[0138] 商业上相当大产量 : 对栽培者具有商业重要性的谷粒产量,当生长在相同条件下时,由对照系 AG2703 和 DKB23-51 的至少 95% 的实际谷粒产量来表示。

[0139] 杂交 : 两种亲本植物的交配。

[0140] 异花传粉 : 通过来自不同植物的两种配子的联合受精。

[0141] 下调突变 : 为了本申请的目的,将下调突变定义为降低给定基因的蛋白表达水平的突变。因而,下调突变包含无效突变。

[0142] F_1 杂种 : 两种非等基因植物杂交的第一代子代。

[0143] 基因型 : 细胞或生物的基因组成。

[0144] INDEL : 由核苷酸序列插入或缺失引起的遗传突变。

[0145] 工业用途 : 大豆植物的非食物和非饲料用途。术语“大豆植物”包括大豆植物的植物部分和衍生物。

[0146] 连锁 : 一种现象,其中相同染色体上的等位基因比如果等位基因的传递是独立的而偶然预期的更倾向于一起分离。

[0147] 标记 : 容易检测的表型,优选以共显性方式遗传(在二倍体杂合子中基因座上的两个等位基因容易检测),没有环境变异组分,即遗传率为 1。

[0148] 非转基因突变 : 天然发生的或通过常规方法(例如将植物暴露于辐射或诱变化合物)诱导的突变,不包括使用重组 DNA 技术产生的突变。

[0149] 无效表型 : 本文使用的无效表型是指,给定蛋白不以可检测的水平表达。在 Gy 亚基的情况下,通过 SDS-PAGE 和考马斯染色测定表达水平。

[0150] 表型 : 细胞或者生物的可检测的特征,该特征是基因表达的表现。

[0151] 数量性状基因座 (QTL) : 数量性状基因座 (QTL) 指在一定程度上控制通常连续分布的以数值表示的性状的遗传基因座。

[0152] SNP : 当对比 2 个同源序列时,指单核苷酸多态性或单核苷酸突变。

[0153] 严格条件 : 指 5X SSC、50% 甲酰胺和 42℃ 的核酸杂交条件。

[0154] 基本上等同的 : 一种特征,当被比较时,不与平均值显示出统计学显著差异(例如, $p=0.05$)。

[0155] 组织培养物 : 组合物,其包含相同或不同类型的分离的细胞或者一组组织成植物部分的此类细胞。

[0156] 转基因 : 遗传基因座,其包含已经通过转化导入大豆植物的基因组的序列。

[0157] IX. 实施例

[0158] 包括下面的实施例用于证明本发明的优选实施方案。本领域技术人员应明白,下面实施例中公开的技术代表本发明人发现在本发明的实践中良好运行的技术,并且因此可以被认为组成它的实践的优选方式。然而,本领域技术人员根据本公开内容应明白,在所公

开的特定实施方案中可以在不背离本发明的精神和范围的情况下做出许多改变，并且仍然得到同样或者相似的结果。

[0159] 实施例 1

[0160] 在研究中使用的大豆品种

[0161] B2G2 或“11S 无效”大豆品种具有独特的种子组成，包括高水平的 β -伴大豆球蛋白和低量大豆球蛋白。但是，B2G2 品种表现出农学上劣等的特征，例如低产量、过度倒伏和绿种子。开发了许多育种系，其携带 B2G2 系中存在的突变的全部或部分。15 个这样的系以及 B2G2 系用作重复测序组 (resequencing panel) 的突变系。8 个野生型用于本研究中的对比。表 1 列出了在测序组中使用的所有系。

[0162] 表 1

[0163]

系名称 或符号	系谱
JB1	AH_(A3244/(B2G2/A1923:.077.):0001.0097.0015.)/DJW2500C0R:@.0042.0006.@.
JB2	AH_(A3244/(B2G2/A1923:.077.):0001.0064.0001.)/DKB19-51:@.0232.0002.@.
JB3	AH_(A3244/(B2G2/A1923:.077.):0001.0064.0001.)/DKB19-51:@.0228.0015.@.
JB4	AH_(A3244/(B2G2/A1923:.077.):0001.0097.0011.)/DAK2501A0R:@.0314.0009.@.
JB5	AH_(A3244/(B2G2/A1923:.077.):0001.0097.0015.)/DJW2500C0R:@.0042.0012.@.
JB6	AH_(A3244/(B2G2/A1923:.077.):0001.0011.0008.)/AG2402:@.0028.0010.@.
JB7	AH_(A3244/(B2G2/A1923:.077.):0001.0008.0016.)/DBL3201A0X:@.0256.0014.@.
JB8	AH_(A3244/(B2G2/A1923:.077.):0001.0097.0015.)/DJW2500C0R:@.0234.0019.@.
JB9	AH_704416-24/((A3244/(B2G2/A1923:.077.):0001.0097.0015.)/DJW2500C0R:@.0005.):@.0067.0016.
JB10	AH_704416-24/((A3244/(B2G2/A1923:.077.):0001.0097.0015.)/DJW2500C0R:@.0005.):@.0067.0007.
JB11	AH_704416-24/((A3244/(B2G2/A1923:.077.):0001.0097.0015.)/DJW2500C0R:@.0005.):@.0067.0003.
JB12	AH_704416-24/((A3244/(B2G2/A1923:.077.):0001.0097.0015.)/DJW2500C0R:@.0005.):@.0067.0002.
JB13	AH_DAK2301A1R/((A3244/(B2G2/A1923:.077.):0001.0097.0015.)/DJW2500C0R:@.0013.):@.0018.0001.
JB14	AH_DAK2301A1R/((A3244/(B2G2/A1923:.077.):0001.0097.0015.)/DJW2500C0R:@.0013.):@.0114.0011.
JB15	AH_DAK2301A1R/((A3244/(B2G2/A1923:.077.):0001.0097.0015.)/DJW2500C0R:@.0013.):@.0130.0001.
B2G2	
A1923	A1923
A3244	A3244
AG2403	AG2403
AG2703	AG2703
AG3201	AG3201
AG3202	AG3202
DKB17-51	DKB17-51
DKB19-51	DKB19-51

[0164] 实施例 2

[0165] Gy 等位基因的标记的设计

[0166] 所有大豆球蛋白基因的 DNA 序列都可来自 GenBank (NCBI)。这些序列用作针对 Monsanto 序列数据库的 blast 的查询。使用“blastn”程序，得到了许多高得分命中 (hit)。比对从 blast 搜索得到的序列，来提供用于引物设计的高质量共有序列。设计嵌套引物，来完全覆盖在每个基因座处的整个基因，并从不同系产生扩增子。比对这些扩增子的序列，来鉴别与高 β -伴大豆球蛋白表型有关的 SNP 和 INDEL。最初，为 Gy1 设计 10 对引物，为 Gy2 和 Gy3 各设计 7 对，为 Gy4 设计 14 对，为 Gy5 设计 10 对，且为 Gy7 设计 11 对。一旦从该研究中得知它们的序列，就设计其它引物。表 2 列出了在该研究中使用的引物。

[0167] 使用 Qiagen 植物 DNA 试剂盒分离 DNA，并用 KOD 规程 (EMD Biosciences, Inc, Madison, WI) 进行 PCR。反应混合物包括 3.4 μ 15M 甜菜碱，2 μ 110x KOD 缓冲液，2 μ 12mM

dNTPs, 0.8 μl 125mM MgSO₄, 0.2 μl KOD 酶 (1U / μl), 1.6 μl 引物 (5 μM) 和 10 μl DNA 模板 (2ng/ μl)。PCR 循环如下 :94°C 5 分钟 ;8 个 94°C 40 秒、62°C 40 秒、72°C 1 分钟、94°C 40 秒、60°C 40 秒、72°C 1 分钟、94°C 40 秒、58°C 40 秒、72°C 1 分钟、94°C 40 秒、56°C 40 秒、72°C 1 分钟的循环 ;3 个 94°C 40 秒、55°C 40 秒、72°C 1 分钟的循环 ; 在 72°C 保持 7 分钟。在 1% 琼脂糖凝胶中, 通过电泳分析 PCR 产物。为了测序, 将 5 μl PCR 产物转入新管中, 并加入 1 μl 外切核酸酶 I (1 :10 稀释的) 和 1 μl 虾 (Shrimp) 碱性磷酸酶 (1 :100 稀释的)。将混合物在 37°C 温育 20 分钟, 然后在 80°C 温育 20 分钟, 以灭活酶。加入 40ul H₂O, 将 6 μl 用作模板, 具有 1 μl 测序引物。使用 Capillary 测序仪 ABI3730 进行测序。装配序列, 并使用 DNASTar™ 的 SeqMan II 程序 (Laser Gene) 比对。

[0168] 设计终点 SNP/Taqman® 测定, 并通过为它们提供的基于 SNP 序列的 Applied Biosystems 来生产。根据提供的说明书 (Applied Biosystems), 进行 SNP 检测。Taqman® 测定或实时 PCR, 通过扩增反应期间双标记的荧光探针的杂交和切割, 检测特定 PCR 产物的积累。Taqman 测定包括 4 种寡核苷酸, 其中的 2 种用作 PCR 引物, 并产生包括待测多态性的 PCR 产物。另外 2 种是等位基因 - 特异性的荧光共振能转移 (FRET) 探针; 每个探针具有独特的荧光团, 其在 Taq DNA 聚合酶降解探针后释放, 有效地指示样品中存在的每个等位基因的量。

[0169] 表 2

[0170]

引物 名称	正向引物	反向引物
Gy1_1	SEQ ID NO:1 GCCTAAGTACGTACTAAAATGCCAA	SEQ ID NO:2 CTACACCTCATGAAGTCATGGTGTGA
Gy1_2	SEQ ID NO:3 CCATGCATGGTCCCCTCGTCATCACGA	SEQ ID NO:4 CCCTCATTTATCAAACCCCTAAACATATT
Gy1_3	SEQ ID NO:5 GAACITCATGAGGTGTAGCACCCAAGGCTT	SEQ ID NO:6 GATTATGTTACGTATATGGAAGAAATCAA
Gy1_4	SEQ ID NO:7 CCATATGACGTAAACATAATCATATCATTGAT	SEQ ID NO:8 GAATTATAATATCTAATATTGCTATGTGGC
Gy1_5	SEQ ID NO:9 CTCAACAAAGAGGACAAGCAGCAGACCA	SEQ ID NO:10 GATGACCTCCTTGCTCTGCTGATATT
Gy1_6	SEQ ID NO:11 CACCTTGGATTCTTGGAACATGCATTCA	SEQ ID NO:12 GGAGAGATCCAAACTCAGCACTGAGTC
Gy1_7	SEQ ID NO:13 GCGTGGACAAGCAGATAGCGAA	SEQ ID NO:14 CTTGGAGAGATCCAAACTCA
Gy1_8	SEQ ID NO:15 GAGGATGAGAAGCCACAGTGCAAGGG	SEQ ID NO:16 GTATGTTGATCTTGATGAATGATGTACGTA
Gy1_9	SEQ ID NO:17 GGATGAATTGTTGTGACTCTGCATGCA	SEQ ID NO:18 CTGAGACTCCTGAGGTGGAACCAGGAAC
Gy1_10	SEQ ID NO:19 GCAGATAAAAGAACAAACAACCCTTCAAG	SEQ ID NO:20 CAACACTTCTAAAGATATCATCGATCAA
Gy2_1	SEQ ID NO:21 GCAATTGCATGCAATACAAACACACTT	SEQ ID NO:22 GCCACAGTTCAATCAATTAACTAACAA
Gy2_2	SEQ ID NO:23 CCATGAACTTAATGAGGTGTAACACACAA	SEQ ID NO:24 AAGATAGTTGGACGGTTAAGAAGAA
Gy2_3	SEQ ID NO:25 ACCGTCCAACCTATCTTATATATTCAA	SEQ ID NO:26 CACCCCTCTGAAGCGATGTACCTTT
Gy2_4	SEQ ID NO:27 CCTAGCACTTATCAAGAGCCGCAAGAAT	SEQ ID NO:28 CTGCATGTTACGCCAACGCTTCTTT

[0171]

引物 名称	正向引物	反向引物
Gy2_5	SEQ ID NO:29 ATATCAGCAGCAGCAAGGAGGTTCC	SEQ ID NO:30 GCAAGTGCTAAGATAACTTGTGTC
Gy2_6	SEQ ID NO:31 CTTGAAAGAACGCGTCCGGCGTGAACAT	SEQ ID NO:32 GCAAGTGCTAAGATAACTTGTGTC
Gy2_7	SEQ ID NO:33 GGATGAATAACATGTTGATTAAACGTA	SEQ ID NO:34 CAAGGAAGCTGAAAGGGTTGTTCTTC
Gy3_1	SEQ ID NO:35 CACCATTAACITAATAGTGTAAAGACAG	SEQ ID NO:36 CCTTGTTGAATAAAGGTTGTAAGTGGATT
Gy3_2	SEQ ID NO:37 GATTCCGAAGCCACCTTACACCATTAACTTA	SEQ ID NO:38 GGATAAAATGAACCTGTTGAATAAAGGTT
Gy3_3	SEQ ID NO:39 GTCTTAAGCTCAGCACCCCCACTCTGAGT	SEQ ID NO:40 GGATAAAATGAACCTGTTGAATAAAGGTT
Gy3_4	SEQ ID NO:41 CCTCAAGAGTAACGTTAAGGACATCGATA	SEQ ID NO:42 CAGTTATTTAAAGTGATTCACACGAGG
Gy3_5	SEQ ID NO:43 AAGAAATTGGACAACGTTGTAACATGCA	SEQ ID NO:44 CAATGTTGCTTCTCGTCACAATCTGG
Gy3_6	SEQ ID NO:45 AAGAAATTGGACAACGTTGTAACATGCA	SEQ ID NO:46 GCTTTTATAACATGAATTAATGATGTAAGTA
Gy3_7	SEQ ID NO:47 GATTAACGTACACTTGTATGGTGCA	SEQ ID NO:48 GCATAGGTACTTGAGTGACTCATTACACAA
Gy4_1	SEQ ID NO:49 GCACAGTAAAACAGTTCAAATTGAGAA	SEQ ID NO:50 CATTCTCACCTTGCATGGCTATTGTT
Gy4_2	SEQ ID NO:51 GCAAGGTGAAGAATGTCACAAACTCAGCAA	SEQ ID NO:52 GGTGACAAATGGATTAATATAACTGAGAA
Gy4_3	SEQ ID NO:53 GGATGATCATCATCGCCCCAAGGTAAT	SEQ ID NO:54 CTGGTGACTGTCCTGTAGCTGCTGCTT
Gy4_4	SEQ ID NO:55 TCAAGGTCGCAGAACAGCAGCTACAG	SEQ ID NO:56 TGAGAGGGAATTGTCATCTCATCAT
Gy4_5	SEQ ID NO:57 GATGATGAAGATGAACAAATTCCCTCTCAC	SEQ ID NO:58 GTAGAGGACAACATATTGGGCACTGAGTTG
Gy4_6	SEQ ID NO:59 CACCCCTCCCAGGCCCTCCGCCAATTCCA	SEQ ID NO:60 GTCACATAGATCACACTGTTGCATTCA
Gy4_7	SEQ ID NO:61 TTACTCTCCACATTGGAATCTGAATGCA	SEQ ID NO:62 GCCACTCAGATATAAACATAGGCTCGCTG
Gy4_8	SEQ ID NO:63 CATAAATGACAAGCATGATGGTGTGAGGA	SEQ ID NO:64 CCAGTAAACATATAATCAGTATTACTCATTT
Gy4_9	SEQ ID NO:65 AGCCATGCAAGGTGAAGAATGTCACAAA	SEQ ID NO:66 AAGAGTATCACCAGCATTCTCAGTGT
Gy4_10	SEQ ID NO:67 CTTGTTGACATATCAATCACCTAA	SEQ ID NO:68 GTGAAAGAATTAACAAGTAAGGAGAACAA
Gy4_11	SEQ ID NO:69 GTTCTCCTTACTTGTAAATTCTTCACTT	SEQ ID NO:70 GTCCTGGTCTGGTCTTGTACCGCTT
Gy4_12	SEQ ID NO:71 GAAGATCAACCTCGCAAGAGCCCGAAT	SEQ ID NO:72 CAACTACCCCTAGAGAATCACTAAAGAAT
Gy4_13	SEQ ID NO:73 GTCAGTAAGTATGTTAGGGTTGGATT	SEQ ID NO:74 CCATCATGCTTGTCAATTATGCGACTTT
Gy4_14	SEQ ID NO:75 CAAGGCTCACCCCGTGTAAAGTCGCAT	SEQ ID NO:76 GAATAAAGACAAAACGTGAAGACTGACAT
Gy5_1	SEQ ID NO:77 CTCCCTCAAAACTTATTAAACACTTT	SEQ ID NO:78 CCTTGAACGACAATGATCATTT
Gy5_2	SEQ ID NO:79 CTCAAGGTCGCAGCAGCAACTACAA	SEQ ID NO:80 GAGGGAGTTGTTCATATTCTCG
Gy5_3	SEQ ID NO:81 GAAGAATATGAACAAACTCCCTCTT	SEQ ID NO:82 GGACAACATATTGGGCACTGAGTCC

[0172]

引物 名称	正向引物	反向引物
Gy5_4	SEQ ID NO:83 CTCCCAGCCCTCCGCCAATTGG	SEQ ID NO:84 CACATAGATCACACTGTTCGCGTCAAG
Gy5_5	SEQ ID NO:85 TACTCTCCACATTGGAACCTTGAACGCG	SEQ ID NO:86 CACTCAGATATTAAACATAGGCTGGGTC
Gy5_6	SEQ ID NO:87 GGCCCTTGGTCAACCCATAAATAA	SEQ ID NO:88 AAGACTGACATTATTAAAGGCGATT
Gy5_7	SEQ ID NO:89 CATGGAACCTCTAACACCCCTGAGCTGCAA	SEQ ID NO:90 CCTCTAGATATAAGATAGTGTTCTCAA
Gy5_8	SEQ ID NO:91 GTCACTGTTCCAACACGCACCCCAA	SEQ ID NO:92 GGGTTCCCAGCAAGGTAAAATACCTT
Gy5_9	SEQ ID NO:93 GTGTTCCCTACTGGACCTATAACACT	SEQ ID NO:94 GGGTGATCAGGACGAGGTTGATCIT
Gy5_10	SEQ ID NO:95 GTCACTGCATAGTATCATAACACACTT	SEQ ID NO:96 GTCTGACATCCTCTTCCACGTGGTT
Gy7_1	SEQ ID NO:97 CGAGAACAAATAGAAATAGACCATCAGG	SEQ ID NO:98 CCTCTTCACACAATGATCCAAACTC
Gy7_2	SEQ ID NO:99 GTCCAGGGTCCATGTTATCGTCT	SEQ ID NO:100 CATTGTCCTTCCTTACTGATTCTCC
Gy7_3	SEQ ID NO:101 TGACTGCATGTATCATGTATGTGAAAG	SEQ ID NO:102 GATTGATGAGGTTGTCAGGTTTC
Gy7_4	SEQ ID NO:103 CGGAGAACAGTAAGGAAAGGACAA	SEQ ID NO:104 TGAGGTTGTTGGATACCTGGAGTA
Gy7_5	SEQ ID NO:105 ACCATTCTGCGCTCCATTATTATT	SEQ ID NO:106 AGTTTTCGTACAACCAGGAATTACA
Gy7_6	SEQ ID NO:107 GTTAACCAATTCTGCGCTCCATTAT	SEQ ID NO:108 GAATATGTCACCCCTGCTTCAGGTAA
Gy7_7	SEQ ID NO:109 TGTAATTCCCTGGTTGTGACGAAACT	SEQ ID NO:110 GTGGCCTGATAATGCTAAGACCTTT
Gy7_8	SEQ ID NO:111 CGTTACCTGAAGCAGGGTGACATA	SEQ ID NO:112 CACCTCTTCTCTCCTCTCCTTCTT
Gy7_9	SEQ ID NO:113 CTTGGAACACGAAGTTAGAGAACCA	SEQ ID NO:114 CGGTTATTGTTGTAAGTGTGGT
Gy7_10	SEQ ID NO:115 CCCAATGGGTAAACTCTACAAGGT	SEQ ID NO:116 ATTATTCTTGAGCTCGCTCACTTCC
Gy7_11	SEQ ID NO:117 GTGGTGAACCTCCCAGGGAAAGT	SEQ ID NO:118 GAGGAAAGTACAAATAGCAACTGACAA
Gy1_473	SEQ ID NO:119 CGAAGGCCACCTCACACCATGAACCTCAT	SEQ ID NO:120 CAGAAGTAGGGTGCTGAGCTTGAGACATT
Gy1_579	SEQ ID NO:121 GTCCCTCATTCACCTTCTCTTCCCTAT	SEQ ID NO:122 GCTTGGCCATGGTGATGACTGATGAGTGT
Gy1_794	SEQ ID NO:123 CTCAATGCCCTCAAACCGGATAACCGTA	SEQ ID NO:124 CAACACCGGCACACTGGAATGGCTTGT
Gy3_1264	SEQ ID NO:125 GCATGATATTCCCGGGTTGTCCTAGCACAT	SEQ ID NO:126 CCCTCTCTGAAGTGATAGATCTCTGGTGA
Gy3_1356	SEQ ID NO:127 CACTTCAGAGAGGGTGATTGATTGCAGT	SEQ ID NO:128 GTTCTGGAAAGCTGTTGGTGTCAATAAGAGA

[0173] 实施例 3

[0174] Gy 表达分析的方法

[0175] 在分别源自 AAH3504T0C / AH0209439-130 之间和 AAH2104J0C / AH0209439-130 之间的杂交的命名为 JB0305602 和 JB0305605 的 2 个 F2 群体中, 进行 Gy 标记的遗传构造。取样 400 株 F2 单个植物, 用 SNP 标记对 372 株植物进行基因型分型。

[0176] 如下进行蛋白分析: 收集 8 粒大豆种子, 使用 CAT Mega- 研磨机 (SOP Asci-01-0002) 进行研磨。将研磨的样品保藏在 4℃。为了分析, 将约 30mg 来自每种的面粉称量进 96 孔 2ml 微量滴定板的一个孔中。在摇动下, 在含有 0.1M 二硫苏糖醇 (DTT) 作为还原剂的 1.0ml 1XLaemmli SDS 缓冲液 pH6.8 中, 提取蛋白 1 小时。离心后, 在 SDS 缓冲

液中进一步稀释每种提取物的部分,以产生 0.2–0.5 μg / mL 总蛋白,加热至 90–100°C 10 分钟,并冷却。对于每种样品,使用 12 通道移液管将 1–2 μg 总蛋白装载到 26 泳道 15% T 梯度 Tris / HCl Criterion 凝胶上。在每个凝胶中,在 2 个泳道中包括分子量标准品和亲本对照。踪迹染料达到凝胶底部之前,对凝胶电泳约 1.2 小时,然后在胶体考马斯蓝 G-250 中染色过夜,在 DI(去离子)水中脱色,且使用 GS800 校准的光密度计成像。在图 1 中显示了染色凝胶的示例性图像,且在图的左侧指示着与 Gy 等位基因有关的蛋白带。使用 Bio-Rad Quantity One™ 软件,进行定量。使用软件来测定样品泳道中每个带的相对量。将百分比酸性大豆球蛋白和百分比 β-伴大豆球蛋白蛋白亚基带记录为泳道中总蛋白的相对百分比。使用 Master LIMS™,追踪样品特性和重量。

[0177] 实施例 4

[0178] Gy1 和 Gy2 中的突变

[0179] 分析 F2 子代植物中 Gy1 和 Gy2 编码的蛋白的总含量。如图 2 所示,将植物分配入 2 个表型组,一个组具有小于 3% 的 Gy1,2 编码的蛋白,且另一组具有 3.1% 更大的 Gy1,2 编码的蛋白。卡方分析(图 2B)与作为隐性性状的连锁的突变 Gy1,2 等位基因相一致。

[0180] 在 Gy1 基因的大部分中得到了良好的序列覆盖度。由于 2 个引物对的失败,存在 2 个小缺口。将共有序列列为 SEQ ID NO :163。在表 3 中给出了在 Gy1 的一些选定核苷酸位置处的等位基因评分。在 JB7 / JB8 和其它系之间,存在 3 个在位置 643、835 和 839 处的 SNP。因为在这些位置处没有从 B2G2 系重新获得序列,所以没有确定 SNP 是否是从 B2G2 突变系遗传而来。没有从突变系 B2G2 和它的衍生系 JB3、JB4、JB13、JB14 和 JB15 得到扩增子或在 Gy1 基因的 5' 末端的序列,从而表明在突变系中可能存在序列缺失。用前 3 对引物进行另外的 PCR 反应,并得到一致的结果。这证实了突变系中的缺失。缺失跨越上游启动子区、外显子 I 和内含子 I。还没有精确测定缺失的接头。在缺失上设计显性标记,以用作 Gy1 基因座的诊断标记。

[0181] Gy2 基因的测序仅仅指明了 3 个可鉴别的 SNP,且罕见的等位基因都来自 JB7 系。所有其它系都是相同的。因为 Gy1、Gy2 和 Gy3 之间的高序列同源性,甚至用基因座特异性的引物也发生偶然的交叉污染,因而可能存在没有鉴别出的其它变异。

[0182] 表 3

系	位置/等位基因											
	195	450	600	643	800	835	839	1000	1300	2762	27703	
JB1			C	T	G	A	C	T	C	G	G	
JB2		C	C	T	G	A	C	T		G	G	
JB3											G	
JB4												
JB5	A		C	T						G	G	
JB6	A	C	C	T	G	A	C	T	C			
JB7		C	A	G	T	T		T			G	
JB8			A		T	T						
JB9				A	G	A	C	T		A/G	A/G	
JB10	A	C	C	T	G	A	C	T		G	G	
JB11										G	G	
JB12	A		C	T	G	A		T		G	G	
JB13										G	G	
JB14										G	G	
JB15										G	G	
A1923			C	N	G	A	C	T		G	G	
A3244	A	C	C	T	G	A	C	T	C	G		
AG2403	A	C	C	T	G	A	C	T	C	G	G	
AG2703	A	C	C	T	G	A	C	T	C	G	G	
AG3201		N	C	T	G	A	C	T	C			
AG3202	A	C	C	T	G	A	C	T	C	G	G	
DKB17-51		C	C	T	G	A	C	T	C			
DKB19-51		C	C	T	G	A	C	T		G	G	
B2G2									C	G	G	

[0184] 实施例 5

[0185] Gy3 中的突变

[0186] 分析了 F2 子代植物中 Gy3 编码的蛋白的总含量。如图 3 所示, 将植物分配入 2 个表型组, 一个组具有小于 1% 的 Gy3 编码的蛋白, 且另一组具有 1.1% 或更大的 Gy3 编码的蛋白。卡方分析 (图 3B) 与作为隐性性状的减少的 Gy3 表达相一致。

[0187] 当对比 Gy1,2 和 Gy3 亚基的表达水平时 (图 4), 发现 Gy1,2 编码的蛋白的表达与 Gy3 编码的蛋白的表达正相关, 相关系数为 .88, 见图 4B。该数据表明, 低 Gy3 编码的蛋白水平的表达可以基于 Gy1 和 / 或 Gy2 基因型来测定。

[0188] 通过在多个系上对整个基因重新测序, 测定在 Gy3 基因座处的 DNA 序列变异。将得到的共有序列列为 SEQ ID NO :164。在表 4 中给出了在该基因座处多态性的等位基因评分。在突变型和野生型之间检测到 5 个 SNP 和 2 个 INDEL。原始突变系 B2G2 携带在位置 848-851 处的插入 (TGAT), 而所有其它系携带在该位置处的缺失。B2G2 也携带在位置 1083、1120 和 1866 处的 3 个 SNP 上的罕见等位基因, 而所有其它系携带丰富的等位基因。数据表明, 这些 SNP 和 INDEL 与 Gy3 编码的蛋白的低表达处于连锁不平衡 (LD) 中。但是, 在该基因的 3' 末端处, 位置 2504-2505、2574 和 3189 的 3 个序列变异, 似乎没有与 Gy3 编码的蛋白的低表达处于 LD 中, 因为罕见等位基因不仅在 B2G2 上检测到, 而且也在野生型 AG2703 和 DKB19-51 上检测到。在该研究中使用的 B2G2 衍生系都没有遗传在 Gy3 基因座处的 B2G2 等位基因。

[0189] 表 4

系	位置/等位基因									
	234	1318	1600	1866	2200	2504-2505	2574	2850	3189	
JB1	A	A	A	C	A	--	G		C	
JB2	A	A	A	C	A			G	C	
JB3	A	A	A	C		--	G	G		
JB4	A	A	A	C	A			G	C	
JB5	A	A	A	C	A			G	C	
JB6	A	A	A	C	A	--	G	G	C	
JB7	A	A	A	C	A			G	C	
JB8	A	A	A	C	A			G	C	
JB9	A							G	C	
JB10	A	A	A	C						
JB11	A	A	A	C				G		
JB12	A	A	A	C	A			G	C	
JB13	A	A	A	C	A					
JB14	A	A	A	C				G		
JB15	A	A	A	C				G		
A1923	A	A	A	C				G		
A3244	A	A	A	C	A	--	G		C	
AG2403	A	A	A	C				G		
AG2703	A	A	A	C	A	AT	A	G	T	
AG3201	A	A	A	C	A	--	G			
AG3202	A	A	A	C	A			G	C	
DKB17-51	A	A	A	C	A	--	G			
DKB19-51	A	A	A	C	A			G	T	
B2G2	A			A	A			G		

[0191] 实施例 6

[0192] Gy4 中的突变

[0193] 对子代植物进行蛋白分析,以测定表达的 Gy4 编码的多肽的量。如图 5 所示,将植物分配入 2 个表型组,一个组是 Gy4 无效的,且另一组表现出 Gy4 编码的多肽的表达。卡方分析(图 5B)与作为隐性性状的 Gy4 无效等位基因相一致。

[0194] Gy4 等位基因的序列分析揭示了在突变系 B2G2 和它的衍生系中的翻译起始密码子处的突变(表 5,位置 682, SEQ ID NO :165)。突变将 ATG 改变成 ATA。由于丧失翻译起始密码子,由该基因编码的肽亚基最可能不能翻译。该 SNP 可以理想地用作指示 Gy4 丧失的分子标记。使用得到的序列来设计用于检测称作 NS0199003 的该标记的引物(见表 8)。

[0195] 在来自 B2G2 和某些子代系的有些序列处,也观察到在位置 1620 和 1632 处的其它多态性,但是它们与 Gy4 编码的蛋白亚基的丧失无关。因而,这些不一致可能由该区域引物的非特异性扩增造成。

[0196] 表 5

系	位置/等位基因		
	682	1620	1632
JB1	A	C	T
JB2	A	C	T/C
JB3	A	C	T
JB4	G	C	T
JB5	A	C	T
JB6	G	C	T
JB7	G	C	T
JB8	G	C	T
JB9	A	C	T
JB10	A	C	T/C
JB11	A	C	T
JB12	A	C	T
JB13	A	C	T
JB14	A	C	T
JB15	A	C	T
A1923	G	C	T
A3244	G	C	T
AG2403	G	C	T
AG2703	G	C	T
AG3201			
AG3202	G	C	T
DKB17-51			
DKB19-51	G	C	T
B2G2	A	T/C	T/C

[0197]

[0198] 实施例 7

[0199] Gy5 中的突变

[0200] Gy5 等位基因的序列分析表明,与其它亲本品种相比,在 B2G2 植物中存在 2 个 SNP (SEQ ID NO :166 上的位置 363 和 612) 和 2 个 INDEL (位置 447-453, 519-524) (表 6)。另外,在外显子 II 的位置 752 处鉴别出了 SNP, 这将氨基酸残基从丝氨酸变成天冬酰胺。所有 5 个 SNP 或 INDEL 在测试的系中形成 2 个单元型。突变系 B2G2 和它的衍生系 JB1、JB5 和 JB8 共有一个单元型, 而其它系共有另一个单元型。这些 SNP / INDEL 似乎是连锁不平衡的,且与“11s 无效”表型相关。仍然不清楚这些 SNP 或 INDEL 是否确实造成如图 1 所示的 B2G2 中 A3 亚基的丧失。因为它们都是在编码区中检测到的序列变异,所以在启动子区可能存在一些其它变异,它们造成在蛋白凝胶上观察到的 A3 带的丧失。

[0201] 表 6

[0202]

系	位置 / 等位基因				
	363	447-453	519-524	612	752
JB1	C	*****	*****	A	A
JB2	G	TTTTAG	TAATAA	T	G
JB3	G	TTTTAG	TAATAA	T	G
JB4	G	TTTTAG	TAATAA	T	G
JB5	C	*****	*****	A	A
JB6					
JB7	G	TTTTAG	TAATAA	T	G
JB8	C	*****	*****	A	A
JB9	G	TTTTAG	TAATAA	T	G
JB10	G	TTTTAG	TAATAA	T	G
JB11	G	TTTTAG	TAATAA	T	G
JB12					
JB13	G	TTTTAG	TAATAA	T	G
JB14	G	TTTTAG	TAATAA	T	G
JB15	G	TTTTAG	TAATAA	T	G
A1923	G	TTTTAG	TAATAA	T	
A3244	G	TTTTAG	TAATAA	T	G
AG2403	G	TTTTAG	TAATAA	T	G
AG2703					
AG3201	G	TTTTAG	TAATAA	T	G
AG3202	G	TTTTAG	TAATAA	T	G
DKB17-51	G	TTTTAG	TAATAA	T	G
DKB19-51	G	TTTTAG	TAATAA	T	G
B2G2	C	*****	*****	A	A

[0203] 实施例 8

[0204] Gy 突变标记的开发

[0205] 为上面鉴别出的 SNP 或 INDEL, 设计 PCR 分析, 例如 **Taqman®** 测定。表 7 列出了每个测定的引物和探针序列以及为每个标记分配的标记名称。使用在不同位置处的 SNP, 分别为 Gy1 和 Gy3 设计 2 个测定。这些测定首先运行在用于在该研究中重复测序的标准组上, 然后用于分离群体中。

[0206] 表 7

[0207]

基因	标记名称	测定	描述	序列
<i>Gy1</i>	NS0199008	<i>GY1</i> conA-644	正向引物	SEQ ID NO:129 AATAACCACGCCCTCAGGTTCTC
<i>Gy1</i>	NS0199008	<i>GY1</i> conA-644	反向引物	SEQ ID NO:130 GAGTGTAAAGGACCAATGGAGAGA
<i>Gy1</i>	NS0199008	<i>GY1</i> conA-644	Vic 探针	SEQ ID NO:131 CTTCACAACCAAACAT
<i>Gy1</i>	NS0199008	<i>GY1</i> conA-644	FAM 探针	SEQ ID NO:132 TTCACAAACACAAACAT
<i>Gy1</i>	NS0199009	<i>GY1</i> conB-839	正向引物	SEQ ID NO:133 CCCTCAAACCGGATAACCGTATAG
<i>Gy1</i>	NS0199009	<i>GY1</i> conB-839	反向引物	SEQ ID NO:134 CACTGGAATGGCTTGTGTTAGG
<i>Gy1</i>	NS0199009	<i>GY1</i> conB-839	Vic 探针	SEQ ID NO:135 ATGTCTCAATGAGCCC
<i>Gy1</i>	NS0199009	<i>GY1</i> conB-839	FAM 探针	SEQ ID NO:136 ATGTCTCAATGAACCC
<i>Gy2</i>	NS0199002	<i>GY2</i> -102	正向引物	SEQ ID NO:137 CGTACATCATACATGTTATAAATTAAAGCTAACAA
<i>Gy2</i>	NS0199002	<i>GY2</i> -102	反向引物	SEQ ID NO:138 GCATATGCAAGTGCTAAGATAACTTTGT
<i>Gy2</i>	NS0199002	<i>GY2</i> -102	Vic 探针	SEQ ID NO:139 ACACATTTAAATTACTATATATAACT
<i>Gy2</i>	NS0199002	<i>GY2</i> -102	FAM 探针	SEQ ID NO:140 CACATTTAAATTACTATATATAGCT
<i>Gy3</i>	NS0199004	<i>GY3</i> -89	正向引物	SEQ ID NO:141 AGAGCCCTTTGCATGTGCTA
<i>Gy3</i>	NS0199004	<i>GY3</i> -89	反向引物	SEQ ID NO:142 TCGTTCTTAATTATTGCTACGCACACT
<i>Gy3</i>	NS0199004	<i>GY3</i> -89	Vic 探针	SEQ ID NO:143 CAAAAGGACAAAAGTGT
<i>Gy3</i>	NS0199004	<i>GY3</i> -89	FAM 探针	SEQ ID NO:144 AAAAGGACGAAAGTGT
<i>Gy3</i>	NS0199010	<i>GY3</i> conA-1866	正向引物	SEQ ID NO:145 GGAACCAAGAGCAAGAGTTCTACA
<i>Gy3</i>	NS0199010	<i>GY3</i> conA-1866	反向引物	SEQ ID NO:146 CGCTTCCCTTCTGGCTTGAGTA
<i>Gy3</i>	NS0199010	<i>GY3</i> conA-1866	Vic 探针	SEQ ID NO:147 CTCCTTGCTGCTTCT
<i>Gy3</i>	NS0199010	<i>GY3</i> conA-1866	FAM 探针	SEQ ID NO:148 CCTCCTTCTGCTTCT
<i>Gy4</i>	NS0199003	<i>GY4</i> -93	正向引物	SEQ ID NO:149 TCCAATTCAACCAACTCCTCAAAC
<i>Gy4</i>	NS0199003	<i>GY4</i> -93	反向引物	SEQ ID NO:150 CAAAGGAAAGAAAGAGAGAGGTGA
<i>Gy4</i>	NS0199003	<i>GY4</i> -93	Vic 探针	SEQ ID NO:151 CTTCCTTAGTTCAATATAGG
<i>Gy4</i>	NS0199003	<i>GY4</i> -93	FAM 探针	SEQ ID NO:152 TCCTTAGTTCAATATGGG
<i>Gy7</i>	NS0199001	<i>GY7</i> -72	正向引物	SEQ ID NO:153 CATAGGAGAACACGAGGGATGTG
<i>Gy7</i>	NS0199001	<i>GY7</i> -72	反向引物	SEQ ID NO:154 GGCTTCTACTTTGCTTCCCTCTT
<i>Gy7</i>	NS0199001	<i>GY7</i> -72	Vic 探针	SEQ ID NO:155 AATGCGAAGATAAAG
<i>Gy7</i>	NS0199001	<i>GY7</i> -72	FAM 探针	SEQ ID NO:156 ATGCGAAAATAAAG

[0208] 实施例 9

[0209] Gy 标记和蛋白亚基之间的高度一致

[0210] 分析 F1 子代植物来测定大豆球蛋白和 β -伴大豆球蛋白蛋白的总含量。图 6A 中的图显示了每种植物的总大豆球蛋白对总 β -伴大豆球蛋白蛋白。数据表明, 大豆球蛋白的低表达与 β -伴大豆球蛋白 (Cgy) 亚基的高表达相关 (图 6B)。

[0211] 在 2 个 F2 分离群体 JB0305602 和 JB0305605 中, 进行分子标记和它们的对应大豆球蛋白亚基之间的遗传相关。为了确定哪个遗传标记能指示突变 Gy 等位基因, 在 SDS-PAGE 上分析 2 个群体的所有个体的蛋白含量, 并用在研究中开发的 SNP 标记以及许多贯穿基因组中选择的标记, 进行基因型分型。如上面和在图 1 中指出的, Gy1 和 Gy2 的蛋白带聚簇在一起, 且因而它们被测量为一个单位。表 8 表明, 蛋白带的分离 (表达为总蛋白的百分比) 与 SNP 标记高度相关。例如, 具有在 Gy1 处的突变等位基因的植物总是含有更低的大豆球蛋白 A1a、A1b 和 A2 亚基, 在 2 个群体中分别是 2.5 和 2.4%, 而具有“TT”等位基因的那些则含有更高的大豆球蛋白 A1a、A1b 和 A2 亚基, 分别是 7.6 和 8.1%。相关是非常显著的 (P- 值分别是 1.7×10^{-55} 和 4.1×10^{-62}), 且在 Gy2 的情况下观察到类似的相关。在 Gy4 基因座处, 携带“AA”等位基因的个体含有更低的大豆球蛋白亚基 A4B3A5, 在它们各自的群体中为 0.6% 和 0.1%, 而杂合子 (“AG”) 在 2 个群体中含有 1.8%, 且具有“GG”的那些含有最高的大豆球蛋白亚基 A4B3A5, 在它们各自的群体中为 2.9% 和 3.3%。相关是非常显著的, P- 值分别是 2.3×10^{-40} 和 2.3×10^{-69} 。数据表明, NS0199008、NS0199002 和 NS0199003 标记与指定的大豆球蛋白亚基的减少的表达特别相关, 且可以用于预测减少的大豆球蛋白表型。

[0212] 表 8

[0213]

群体	基因	标记名称	等位基因	蛋白平均值 (%)	F2植物的数目	DF	ProbF	Rsquare
JB0305605	Gy1	NS0199008	**	2.5	103	1	1.70E-55	0.49
	Gy1	NS0199008	TT	7.6	269	1		
	Gy2	NS0199002	**	2.8	109	1	1.00E-49	0.45
	Gy2	NS0199002	TT	7.6	263			
JB0305602	Gy1	NS0199008	**	2.4	89	1	4.10E-62	0.49
	Gy1	NS0199008	TT	8.1	283	1		
	Gy2	NS0199002	**	3	102	1	2.70E-54	0.48
	Gy2	NS0199002	TT	8.1	270			
JB0305605	Gy4	NS0199003	AA	0.6	112	2	2.30E-40	0.43
	Gy4	NS0199003	AG	1.8	128			
	Gy4	NS0199003	GG	2.9	92			
JB0305602	Gy4	NS0199003	AA	0.1	93	2	2.30E-69	0.62
	Gy4	NS0199003	AG	1.8	149			
	Gy4	NS0199003	GG	3.3	88			

[0214] 实施例 10

[0215] Gy3 和 Gy4 无效大豆植物的蛋白表达的变化

[0216] 分析包含赋予无效 Gy3 和 Gy4 表型和降低的 Gy1 / Gy2 表型的非转基因突变的农学上优良的大豆植物中各种大豆球蛋白和 β -伴大豆球蛋白亚基的总含量。研究了 3 个不同系, 系谱 1 (AH_DAK2301A1R / ((A3244 / (B2G2 / A1923 :. 077.) :0001. 0097. 0015.) / DJW2500C0R :@. 0013.) :@. 0114. 0008. @); 系谱 2 (AH_DAK2301A1R / ((A3244 / (B2G2 / A1923 :. 077.) :0001. 0097. 0015.) / DJW2500C0R :@. 0013.) :@. 0096. 0007. @); 和 系谱 3 (AH_DAK2301A1R / ((A3244 / (B2G2 / A1923 :. 077.) :0001. 0097. 0015.) / DJW2500C0R :@. 0013.) :@. 0105. 0006. @) (保藏为 ATCC 登记号 PTA-6892), 且各自的数据分别显示在表 9、10 和 11 中。

[0217] 在每个系的情况下,分析了在 2004–2005 季期间生长在各个位置的植物。显示在这些表中的数据指示了 Gy1 / Gy2、Gy5、 α' BC、 α BC 和 β BC 蛋白亚基的总含量,表达为总种子蛋白含量的百分比。也指示了酸性大豆球蛋白亚基的总含量和 β -伴大豆球蛋白总含量。对于 3 个系谱中的每一个,在每个表的最后两行中指示了平均数据 (Avg.) 和平均值的标准差 (St. Dev.)。使用上面实施例 3 所述的方法,计算值。在任何情况下,都没有观察到可检测的 Gy3 或 Gy4 编码的蛋白。

[0218] 在表 9–11 中所示的研究表明,在测试的 3 个植物系(例如来自不同基因组背景的作用)和在不同环境条件(在不同位置)下生长的植物之间各种大豆球蛋白和 β -伴大豆球蛋白蛋白亚基的含量是可变的。这些数据突出了使用多态性标记选择 Gy 突变植物的优点,因为这些标记的检测将不会发生基于基因组背景和环境条件的变异性。

[0219] 表 9

[0220]

系谱	Gy1,2	Gy5	α' BC	α BC	β BC	酸性大豆球蛋白	β -伴大豆球蛋白
1	1.98	4.36	12.300	15.630	7.00	6.34	34.93
1	2.46	6.83	12.276	14.274	9.91	9.29	36.46
1	3.01	6.12	12.375	16.158	7.19	9.13	35.72
1	2.98	6.69	12.63	16.731	6.79	9.67	36.15
1	2.35	4.94	12.693	15.393	4.94	7.29	33.03
1	2.48	4.97	11.906	14.579	9.06	7.45	35.55
Avg. 1	2.54	5.65	12.36	15.46	7.48	8.20	35.31
St. Dev. 1	0.39	1.03	0.28	0.93	1.77	1.35	1.23

[0221] 表 10

[0222]

系谱	Gy1,2	Gy5	α' BC	α BC	β BC	酸性大豆球蛋白	β -伴大豆球蛋白
2	1.69	5.31	13.301	18.618	6.01	7.00	37.93
2	2.33	5.72	13.351	16.495	6.98	8.05	36.83
2	2.21	5.44	11.661	18.332	8.47	7.65	38.46
2	2.32	5.61	13.194	18.020	7.28	7.93	38.49
2	2.22	6.43	14.027	18.156	6.14	8.65	38.32
2	1.93	5.92	12.711	18.036	9.62	7.85	40.37
Avg. 2	2.12	5.74	13.04	17.94	7.42	7.86	38.40
St. Dev.2	0.25	0.40	0.80	0.74	1.40	0.54	1.15

[0223] 表 11

[0224]

系谱	Gy1,2	Gy5	α' BC	α BC	β BC	酸性大豆球蛋白	β -伴大豆球蛋白
3	2.70	5.73	14.090	19.802	6.63	8.43	40.52
3	1.90	4.45	15.315	21.201	6.08	6.35	42.60
3	2.33	5.94	14.608	20.974	8.63	8.27	44.21
3	2.66	5.40	12.277	18.842	8.12	8.06	39.24
3	2.91	6.96	14.863	20.699	5.40	9.87	40.96
3	1.77	5.73	13.303	17.909	8.65	7.50	39.86
Avg. 3	2.38	5.70	14.08	19.90	7.25	8.08	41.23
St. Dev.3	0.46	0.81	1.12	1.31	1.40	1.16	1.85

[0225] 实施例 11

[0226] Gy 基因组标记可以用于选择低大豆球蛋白大豆植物

[0227] 在 2 个 F2 群体上, 分析了 Gy1 (NS0199008)、Gy2 (NS0199002) 和 Gy4 (NS0199003) 标记的分离。如表 9 所示, Gy1 和 Gy2 都是显性标记, 且各自以 3 : 1 比分离。Gy4 是共显性标记。

[0228] 3 个鉴别出的标记可以用于鉴别具有低大豆球蛋白和高 β -伴大豆球蛋白种子含量的植物。图 7 解释了 F2 植物的大豆球蛋白和 β -伴大豆球蛋白含量。基于亚基表达的蛋白分析 (图 7A) 或使用 3 个突变 Gy 标记 (图 7B), 选择植物。数据显示, 基于标记的选择

错误鉴别小于 1% (7 / 754) 的 F2 植物。

[0229] 实施例 12

[0230] 生产的低大豆球蛋白大豆植物的大豆种子的颜色分析

[0231] 大豆农场主和消费者经常认为绿种子颜色是较不合需要的。因此需要消除系的绿种子颜色。对如本文所述生产的低大豆球蛋白大豆，分析绿种子颜色的消除程度。使用 ColorFlex™ 反射分光光度计（型号 45 / 0），对完整大豆进行颜色分析。使用黑色玻璃和白色砖，标准化分光计。使用具有生产商 Hunter Associates Laboratory, Inc. (Reston, VA, USA) 证实过的颜色值的绿色砖，检查标准化。

[0232] 该分光光度计测量 CIE 三色激励色标值 X、Y 和 Z，并从这些值计算 CIELAB 色标值 L*、a* 和 b*。CIELAB 色标允许按照三维空间描述色感觉 (CIE, Colorimetry, Publication 15.2, 第二版, Vienna, 1986)，CIELAB 色空间组织为立方体。L*- 轴是从顶到底，且称作光亮度值，它从 0 (黑色) 延伸至 100 (白色)。a* 和 b* 轴没有具体的数字边界。坐标 a* 在正值时代表红色，在 0 时代表灰色，且在负值时代表绿色。坐标 b* 在正值时代表黄色，在 0 时代表灰色，且在负值时代表蓝色。

[0233] 在 35x10mm 型聚苯乙烯组织培养皿中装满多层完整大豆，将盖子置于培养皿底部，以便保护读数表面。在置于分光光度计口上之前，从组织培养皿的底部去除盖子。将含有完整大豆的组织培养皿置于分光光度计口上，待测的一侧朝向口。将挡光器置于样品上，以便限制任何外部光干扰。

[0234] 在 ColorFlex 上显示 CIE 三色激励色标屏。按下 Read 键，且测量完整大豆样品的 CIE 三色激励色标值。颜色捕获程序将值记录到 Excel 电子制表软件中。将图像转换 (toggle) 到 CIELAB 色标屏上，按下 Read 键，使用分光光度计计算完整大豆样品的 CIELAB 色标值 I*、a* 和 b*。

[0235] 为每个样品测量一式五份多层排列的完整大豆。将相同的完整大豆重新包装 4 次，以得到 5 个颜色测量结果。在下面的表 12 中，总结了颜色分析的结果。

[0236] 表 12

[0237]

		a*	stdev	L*	stdev
商业化对照 (AG1901)	商业化对照 (Ag1901)	7.08	0.3	54.55	1
B2G2	B2G2	1.91	0.28	51.65	1.34
JB5	对颜色的HBC系分离	3.73	0.59	53.65	0.76
AJB2002KOC (A3244/(B2G2/A1923:.077.))	HBC系保留一些不标准颜色	7.36	0.25	52.39	1.69
系谱3	优良HBC品种	7.34	0.2	57.66	0.84
	优良 = a* > 5 和 L* > 54				

[0238] 实施例 13

[0239] 脂加氧酶基因中的序列变异

[0240] 将来自 GenBank 的 Lox2 序列 G1505137 用作针对 Monsanto 序列数据库的 blast 的查询。使用 SeqMan 程序 (DNASTAR, INC, Madison, WI)，下载并装配 60 个高命中。在 Monsanto DNA 序列收集中，鉴别出 2 个不同的转录物。转录物之一对应着 GenBank 中的脂加氧酶 -1 (Lox1) 基因，且因而命名为 1x1 (SEQ ID NO :157)，而另一个对应着 Lox2 (1x2) 基

因 (SEQ ID NO :158)。设计基因 - 特异性的引物，并用于从 8 系组产生扩增子。该组由 6 个突变型和 2 个野生型组成。表 13 列出了在测序组中使用的系。

[0241] 表 13

[0242]

Pos	品种	系谱	等位基因
A	PI86023		1x21x2
B	L2-3(aka PI561405)	Century x PI86023	1x21x2
C	IA2025		1x11x11x21x21x31x3
D	IA2032		1x11x11x21x21x31x3
E	PI408251		1x11x1Lx2Lx2Lx3Lx3
F	PI417458		LX1LX1LX2LX2LX3LX3
G	A3469		LX2LX2
H	A2247		LX2LX2

[0243] 扩增子测序揭示，在 8 个系中在 Lox1 基因座上存在 27 个多态性，包括 21 个 SNP 和 6 个 INDEL (表 14)。在这些多态性中，10 个位于外显子中。基于在所有这些多态性处的等位基因评分，序列组中的 8 个系落入 6 个单元型。野生型 A3469 和 A2247 属于相同的单元型，而其它突变型属于不同单元型。在 IA2025、IA2032 和 PI408251 中检测到 74bp 缺失。该缺失似乎与 1x1 突变型表型有关 (表 13)。

[0244] 表 14

[0245]

品种	Haplo	EX	INT	INT	INT	INT	EX	INT	INT	INT	INT	EX	INT	INT	INT
PI86023	Haplo	178-180	326	363	380	713	1196	1253	1372	1388	1527	1554	2267	3055]	
L2-3	Hap2	GCG	*	C	A	A	A	T	T	T	A	A	G	C	
IA2025	Hap5	***	*	A	T	*	G	A	T	T	T	A	A	C	A
IA2032	Hap6	GCG	*	C	A	A	A	T	T	T	A	A	C	A	
PI408251	Hap4	***	A	T	*	G	T	A	G	C	T	C	C	A	
PI417458	Hap1	GCG	*	C	A	A	T	A	G	C	T	C	C	A	
A3469	Hap3	***	A	T	*	G	T	A	G	C	T	C	C	A	
A2247	Hap3	***	A	T	*	G	T	A	G	C	T	C	C	A	

品种	Haplo	INT	EX	INT	EX	INT	INT	EX							
PI86023	Haplo	3088	3125	3139	3204	3278	3832-3905	4043	4057	4193	4225	4247	4267	4430	4439
L2-3	Hap2	T	*	T	C	C	GATC#	A	*	C	T	A	C	A	G
IA2025	Hap5	G	G	A	G	A	****	C	A	G	C	G	T	C	A
IA2032	Hap6	G	G	A	G	A	****	C	A	G	C	G	T	C	A
PI408251	Hap4	G	G	A	G	A	****	C	A	G	C	G	T	C	A
PI417458	Hap1	G	G	A	G	A	GATC	C	A	G	C	G	T	C	A
A3469	Hap3	G	G	A	G	A	GATC	C	A	G	C	G	T	C	A
A2247	Hap3	G	G	A	G	A	GATC	C	A	G	C	G	T	C	A

[0246] 注 : #GATC 代表在 IA2025、IA2032 和 PI408251 中缺失的 74bp 序列。

[0247] 在表 15 中显示了在 Lox2 基因座处的多态性。检测到 6 个 SNP 和 2bp 的 INDEL, 且在这 8 个系中形成了 2 个不同的单元型。单元型明显与 1x2 表型有关 (表 13)。除了一个在位置 2542 处以外, 所有这些多态性都位于内含子中。在位置 2542 处的 SNP 是错义突变, 从而造成遗传密码子从编码组氨酸的 CAT 变成编码谷氨酰胺的 CAA。

[0248] 表 15

[0249]

品种	单元型	323	439	1390	1431	1458	2486-87	2542
PI86023		C	A	A	C	G	**	A
L2-3	Hap1	C	A	A	C	G	**	A
IA2025	Hap1	C	A	A	C	G	**	A
IA2032	Hap1	C	A	A	C	G	**	A
PI408251	Hap2	T	G	G	T	T	AT	T
PI417458	Hap2	T	G	G	T	T	AT	T
A3469	Hap2	T	G	G	T	T	AT	T
A2247	Hap2	T	G	G	T	T	AT	T

[0250] 为在位置 2542 处的 SNP, 设计 Taqman 测定。测定信息在表 16 中给出。该测定的 allelogram 如图 8 所示。该标记允许清楚区分在 Lox2 处来自突变型的“A”等位基因和来自野生型的“T”等位基因。

[0251] 表 16

[0252]

标记	NS0203296
基因	Lox2
SNP	203296
测定 ID	1915582-19230-203296
正向引物	GCTATCATCAACTCATGAGCCATTG (SEQ ID NO :159)
反向引物	GTGTCGGTTGTTGCTATGATGAAT (SEQ IID NO :160)
VicProbe	CAATCACCGCTTGAGTAT (SEQID NO :161)
VicAllele	A
FamProbe	AATCACCGCATGAGTAT (SEQ ID NO :162)
FamAllele	T

[0253] 实施例 14

[0254] 大豆中与 Kunitz 胰蛋白酶抑制剂无效表型有关的序列变异

[0255] 从编码大豆的 KTI 的 4 个候选序列中, 鉴别出编码大豆的 Kunitz 胰蛋白酶抑制剂蛋白 (KTI) 的候选序列。一个候选序列 KTIA (SEQ ID NO :167) 用作模板, 来设计用于随后 PCR 扩增反应的引物。从该候选序列设计基因座特异性的引物, 且从 5 个 KTI 无效突变系和 7 个野生型系产生 PCR 扩增子 (表 17)。比对这些扩增子的序列, 允许鉴别出与 Kunitz 表型有关的核苷酸变异 (图 9)。在所有 Kunitz 无效突变型中检测到在位置 622-623 处的 2-bp 缺失和在位置 624 处的单碱基突变 (见图 9), 而“GAG”在所有野生型中存在于相同位置中。缺失 / 插入是无义突变, 从而造成蛋白过早终止, 这解释了表现出 Kunitz 无效表型的突变系表型。该 INDEL 可以用作 Kunitz 无效大豆系和品种的标记辅助选择的遗传标记。

[0256] 表 17

[0257]

系	基因型
PI542044	Titi
PI157440	Titi
IS206-17	Titi
PI547656	Titi

PI547816	Titi
PI518671	TiTi
A3935	TiTi
DKB19-15	TiTi
AG2703	TiTi
AG3302	TiTi
AG2403	TiTi

[0258] 采用引物（例如 SEQ ID NO:168-171）的基于 PCR 的 TAQMAN 标记 - 辅助的测定，允许鉴别 KTI 无效性状（表 18），且测定的验证表明，KTI 无效标记与 KTI 无效表型一起分离。

[0259] 表 18

[0260] KTI 测定信息。标记名称：NS0201535

[0261]	引物1	32_1-1F	GAGAACAAAGATGCAATGGATGGTT (SEQ ID NO:168)
	引物2	32_1-1R	GCTGTGGACAGAACACAAGCTTATA (SEQ ID NO:169)
	Vic_探针	32_1-1V2	AGAAAACCTCTCTCAAGTCT (SEQ ID NO:170)
	FAM_探针	32_1-1M2	CATCAGAAACTCTAAGTCT (SEQ ID NO:171)
	KTI_无效		***
	野生型		GAG

[0262] ***

[0263] 根据本说明书，无需过多的实验就可以制备和实施本文所公开的和本文所要求保护的所有组合物和方法。尽管业已按照优选实施方案对本发明的组合物和方法进行了说明，但本领域技术人员将明白，在不背离本发明的概念、精神和范围的情况下，可以对本文所描述的组合物和方法以及所述方法的步骤或步骤的页序进行改变。更具体地，将明白的是，可以用化学上和生理学上相关的某些试剂取代本文所描述的试剂，同时能获得相同或相似的结果。对本领域技术人员来说显而易见的所有这样的类似取代和修饰，都被视为在如所附权利要求书所限定的本发明的精神、范围和概念内。

[0264] 参考文献

[0265] 在本文中具体引入下面的参考文献作为参考，其程度为它们提供补充本文所给出的那些的示例性程序或者其它细节。

[0266] 美国专利 4,992,375

[0267] 美国专利 5,015,580

[0268] 美国专利 5,024,944

[0269] 美国专利 5,416,011

[0270] 美国专利 5,545,545

[0271] 美国专利 5,604,099

- [0272] 美国专利 5,637,785
- [0273] 美国专利 6,031,154
- [0274] 美国专利 6,140,085
- [0275] 美国专利 6,184,440
- [0276] 美国专利 6,486,383
- [0277] 美国专利 6,774,284
- [0278] Adamset al., *J. Nutr.*, 134(3) :511–516, 2004.
- [0279] Allard, In :*Principles of Plant Breeding*, John Wiley&Sons, NY, 50–98, 1960.
- [0280] Baba et al., *J. Nutr. Sci. Vitaminol. (Tokyo)*, 50(1) :26–31, 2004.
- [0281] Beachy et al., *Ann. rev. Phytopathol.* 28 :451(1990)
- [0282] Beilinson et al., *Theor. Appl. Genet.*, 104(6–7) :1132–1140, 2002.
- [0283] Boerma and Moradshahi, *Crop Sci.*, 15 :858–861, 1975.
- [0284] Borthwick and Parker, *Bot. Gaz.*, 100 :374–387, 1938.
- [0285] Brim and Stuber, *Crop Sci.*, 13 :528–530, 1973.
- [0286] Charest et al., *Plant Cell Rep.* 8 :643(1990)
- [0287] Chen and Shoemaker, *J. Hered.*, 89 :211–215, 1998.
- [0288] Chrispeelsetal., *J. Cell Biol.*, 93 :306–313, 1982.
- [0289] Christianson et al., *Science*, 222 :632–634, 1983.
- [0290] Comai et al., *Nature* 317 :741–744(1985)
- [0291] CrisWell and Hume, *Crop Sci.*, 12 :657–660, 1972.
- [0292] de Moraes et al., *Euphytica* 149 :221–226, 2006.
- [0293] Diers et al., *Theor. Appl. Genet.*, 89 :297–304, 1993.
- [0294] Duranti et al., *j. Nutr.*, 134(6) :1334–1339, 2004.
- [0295] Dutton and Sonmer, *Biotechniques*, 11(6) :700–7002., 1991.
- [0296] Eichholtz et al., *Somatic Cell Mol. Genet.* 13 :67(1987)
- [0297] Elliot et al., *Plant Molec. Biol.* 21 :515(1993)
- [0298] European Appln. 0 242 246
- [0299] European Appln. 0640141
- [0300] European Appln. 0797673
- [0301] Fehr, In :*Theory and Technique, and Crop Species Soybean*, Iowa State Univ., Macmillian Pub. Co., NY, (1) (2) :360–376, 1987b.
- [0302] Fehr, In :*Soybeans : Improvement, Production and Uses*, 2nd Editioi, Manograph., 16 :249, 1987a.
- [0303] Fehr, In :*Hybridization of Crop Plants*, Fehr and Hadley (Eds.), Am. Soc. Agron. and Crop Sci. Soc. Am., Madison, WI, 90–599, 1980.
- [0304] Finer et al., In :*Soybean :Genetics, Molecular Biology and Biotechnology*, CAB Intl., Verma and Shoemaker (ed), Wallingford, Oxon, UK, 250–251, 1996.
- [0305] Fisher et al., *Plant physiol.*, 102(3) :1045–1046, 1993.
- [0306] Fraley et al., *Proc. Natl. Acad. Sci. USA*, 80 :4803, 1983.

- [0307] Geiser et al., Gene, 48 :109, 1986.
- [0308] Gordon-Kamm et al., Plant Cell, 2 :603–618, 1990.
- [0309] Hajika et al., Jpn. J. Breed., 42 :787–792, 1992.
- [0310] Hamner, In :The Induction of Flowering :Some Case Histories, Evans (ed), Cornell Univ. Press, Ithaca, NY, 62–89, 1969.
- [0311] Harada et al., Japan J. Breed., 33 :23–30, 1983.
- [0312] Hartweck et al., In Vitro Cell. Develop. Bio., 24 :821–828, 1988.
- [0313] Hildebrand et al., Crop Sci., 22 :851–853, 1982.
- [0314] Jones et al., Science, 266 :789, 1994.
- [0315] Kitamura et al., Jpn. J. Breed., 35 :413–420, 1985.
- [0316] Kitamura, Agric. Biol. Chem., 27 :234–239, 1984.
- [0317] Knutzon et al., Proc. Natl. Acad. Sci. USA, 89 :2624, 1992.
- [0318] Ladin et al., Plant Physiol., 84 :35–41, 1987.
- [0319] Lander and Botstein, Genetics, 121(1) :185–199, 1989.
- [0320] Lee et al., EMBO J., 7 :1241, 1988.
- [0321] Livak et al., Nat. Genet., 9 :341–342, 1995.
- [0322] Logemann et al., Bio / Technology, 10 :305, 1992.
- [0323] Marshall et al., Theor. Appl. Genet., 83 :435, 1992.
- [0324] Martin et al., Science, 262 :1432, 1993.
- [0325] Miki et al., Theor. Appl. Genet., 80 :449, 1990.
- [0326] Mindrinos et al., Cell, 78 :1089, 1994.
- [0327] Moriyama et al., Biosci. Biotechnol. Biochem., 68(2) :352–359, 2004.
- [0328] Myers, EPO 0273085
- [0329] Natarajan et al., J. Pl. Physiol. (in press).
- [0330] Nielsen et al., In :Cellular and molecular biology of plant seed development, Larkins and Vasil IK(Eds)., Kluwer Academic Publishers, Dordrecht, The Netherlands, 151–220, 1997.
- [0331] Nielsen et al., Plant Cell., 1 :313–328, 1989.
- [0332] Nishi et al., J. Nutr., 133(2) :352–357, 2003.
- [0333] Orita et al., Genomics, 8(2) :271–278, 1990.
- [0334] PCT Appln. US93 / 06487
- [0335] PCT Appln. WO93 / 19181
- [0336] PCT Appln. WO96 / 30517
- [0337] Pen et al., Bio / Technology, 10 :292, 1992.
- [0338] Poehlman and Sleper, In :Breeding Field Crops, Iowa State University Press, Ames, 1995.
- [0339] Przibila et al., Plant Cell, 3 :169, 1991.
- [0340] Reiter et al., Proc. Natl. Acad. Sci. USA, 89(4) :1477–1481, 1992.
- [0341] Shah et al., Science, 233 :478, 1986.

- [0342] Shanmugasundaram and Tsou, *Crop Sci.*, 18 :598–601, 1978.
- [0343] Shibata et al., *J. Biol. Chem.*, 262 :10080–10085, 1987.
- [0344] Shibata et al., *J. Biol. Chem.*, 263 :6816–6821, 1988.
- [0345] Shibles et al., In :*Crop Physiology, Some Case Histories*, Evans (ed), Cambridge Univ. Press, Cambridge, England, 51–189, 1975.
- [0346] Shiroza et al., *J. Bacteol.*, 170 :810, 1988.
- [0347] Simmonds, In :*Principles of crop improvement*, Longmai, Inc., NY, 369–399, 1979.
- [0348] Sneep and Hendriksen, In :*Plant breeding perspectives*, Wageningen (ed), Center for Agricultural Publishing and Documentation, 1979.
- [0349] **Søgaard** et al., *J. Biol. Chem.*, 268(30) :22480–22484, 1993.
- [0350] Sommer et al., *Biotechniques*, 12(1) :82–87, 1992.
- [0351] Stalker et al., *Sciernce*, 242 :419–423, 1988.
- [0352] Steinmetz et al., *Mol. Gen. Genet.*, 20 :220, 1985.
- [0353] Utsumi, In :*Advances in Food and Nutrition Research*, Kinsella (Ed.), 36 : 89–208, Academic Press, San Diego, CA, 1992.
- [0354] Vanden Elzen et al., *Plant Mol. Biol.*, 5 :299, 1985.
- [0355] Wright et al., *Plant Cell Reports*, 5 :150–154, 1986.
- [0356] Yenofsky et al., *Mol. Gen. Genet.*, 211 :215–222, 1988.

[0001]

序列表

<110> Wu, Kunsheng
 Horejsi, Thomas
 Byrum, Joe
 Bringe, Neal
 Yang, Julie
 Pei, Donghong
 Reiter, Robert

<120> 农学上优良的具有高 β -伴大豆球蛋白含量的大豆

<130> MSUT:015W0

<140> 11/517, 186
 <141> 2006-09-07

<150> 60/714, 779
 <151> 2005-09-07

<150> 60/722, 493
 <151> 2005-09-30

<160> 171

<210> 1
 <211> 26
 <212> DNA
 <213> 人工序列

<220>

<223> 合成的引物

<400> 1

gcctaagtac gtactcaaaa tgccaa 26

<210> 2

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 2

ctacacctca tgaagttcat ggtgtga 27

[0002]

<210> 3		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 3		
ccatgcattg tccccctcgta atcacga		27
<210> 4		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 4		
ccctcatatca tcaaaccctt aaacatatt		29
<210> 5		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 5		
gaacttcatg aggtgttagca cccaaaggctt		30
<210> 6		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 6		
gattatgtta cgtcatatgg aagaaaatcaa		30

[0003]

<210> 7		
<211> 31		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 7		
ccatatgacg taacataatc atatcattga t		31
<210> 8		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 8		
gaattataat atctaataatt gctatgtggc		30
<210> 9		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 9		
ctcaacaaag aggacaaagc agcagacca		29
<210> 10		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 10		
gatgacctcc ttgctttgc tgatattt		28

[0004]

<210>	11	
<211>	29	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的引物	
<400>	11	
	caccctggaa ttcttggAAC atgcattca	29
<210>	12	
<211>	27	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的引物	
<400>	12	
	ggagagatcc aaactcagca ctgagtc	27
<210>	13	
<211>	22	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的引物	
<400>	13	
	gcgtggacAA gcAGATAGCG AA	22
<210>	14	
<211>	22	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的引物	
<400>	14	
	cttgcgGAGA gATCCAAACT CA	22

[0005]

<210> 15	
<211> 26	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 15	
gaggatgaga agccacagtg caaggg	26
<210> 16	
<211> 31	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 16	
gtatgttcat ctttgatgaa ttagttagt a	31
<210> 17	
<211> 29	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 17	
ggatgaattt gtttgtactc ttgcattca	29
<210> 18	
<211> 29	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 18	
ctgagactcc tgaggtggaa ccaggaact	29

[0006]

<210> 19		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 19		
gcagataaag aacaacaacc ctttcaag		28
<210> 20		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 20		
caacacccatcc taaaagatatc atcgatcaa		29
<210> 21		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 21		
gcaattgcattt gcaatacataaa cacactt		27
<210> 22		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 22		
gccacagttt caatcaattt tactaacaa		29

[0007]

<210> 23	
<211> 29	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 23	
ccatgaactt aatgaggtgt aacacacaaa	29
<210> 24	
<211> 26	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 24	
aagataggtt ggacggtaa gaagaa	26
<210> 25	
<211> 27	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 25	
accgtccaac ctatctata tattcaa	27
<210> 26	
<211> 26	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 26	
caccctctct gaagcgatgt accttt	26

[0008]

<210>	27
<211>	28
<212>	DNA
<213>	人工序列
<220>	
<223>	合成的引物
<400>	27
cctagcactt atcaagagcc gcaagaat	28
<210>	28
<211>	27
<212>	DNA
<213>	人工序列
<220>	
<223>	合成的引物
<400>	28
ctgcatgttc acgccgaacg cttcttt	27
<210>	29
<211>	28
<212>	DNA
<213>	人工序列
<220>	
<223>	合成的引物
<400>	29
atatcagcag cagcagcaag gaggttcc	28
<210>	30
<211>	27
<212>	DNA
<213>	人工序列
<220>	
<223>	合成的引物
<400>	30
gcaaggcgcta agataacttt gtctgtca	27

[0009]

<210> 31		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 31		
cttggaaagaa gcgttcggcg tgaacat		27
<210> 32		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 32		
gcaaggtagcta agataacttt gtcgtca		27
<210> 33		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 33		
ggatgaataaa catgttgtga ttaacgta		28
<210> 34		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 34		
caaggaagct gaaagggttg ttgttcttc		29

[0010]

<210> 35		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 35		
caccatataac ttaatagtgt aagacag		27
<210> 36		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 36		
ccttggaa taaagggtgt aagttggatt		30
<210> 37		
<211> 31		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 37		
gattccgaag ccacccatac ccattaaactt a		31
<210> 38		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 38		
ggataaaaatg aaccccttgt aataaagggtt		30

[0011]

<210> 39		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 39		
gtcttaagct cagcacccca cttctgagt		29
<210> 40		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 40		
ggataaaaatg aacccttggtg aataaaagggtt		30
<210> 41		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 41		
cctcaagagt aacgttaagg acatcgata		29
<210> 42		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 42		
cagttattta aagtgatttc accacgagg		29

[0012]

<210> 43		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 43		
aagaaaattgg acaacgttgtt aacatgca		28
<210> 44		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 44		
caatgtttgtt ctttctcgtc acaatctgg		29
<210> 45		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 45		
aagaaaattgg acaacgttgtt aacatgca		28
<210> 46		
<211> 31		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 46		
gcttttataa catgaattaa tcatgttaagt a		31

[0013]

<210> 47		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 47		
gattaaacgtacacttgatgtatggcgca		28
<210> 48		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 48		
gcataggtaacttggactcattacacaa		30
<210> 49		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 49		
gcacagtaaaaacagttcaaaatggagaa		27
<210> 50		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 50		
cattcttcaccttgcatggctattgtt		27

[0014]

<210> 51		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 51		
gcaaggtaaa gaatgtcaca aactcagcaa		30
<210> 52		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 52		
ggtgacaaaat ggattataat acactgagaa		30
<210> 53		
<211> 26		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 53		
ggatgatcat catgcggcaa ggtaat		26
<210> 54		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 54		
ctgggtgactg tccctgttagct gctgctt		27

[0015]

<210> 55		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 55		
tcaaggtcgc agaaggcagca gctacag		27
<210> 56		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 56		
tgagagggaa tttgttcatc ttcatcat		28
<210> 57		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 57		
gatgatgaag atgaacaaat tccctctcac		30
<210> 58		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 58		
gttagaggaca acatattggg cactgagttg		30

[0016]

<210> 59		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 59		
caccctcccc gccctccgcc aattcca		27
<210> 60		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 60		
gtcacataga tcacactgtt tgcattcaga		30
<210> 61		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 61		
ttaactctcca catttggaaatc tgaatgca		28
<210> 62		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 62		
gccactcaga tataaacata ggctcgctg		29

[0017]

<210> 63		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 63		
cataaatgac aagcatgatg gtgtgagga		29
<210> 64		
<211> 31		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 64		
ccagtaaaaca tataatcagt attactcatt t		31
<210> 65		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 65		
agccatgcaa ggtgaagaat gtcacaaaa		28
<210> 66		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 66		
aagagtatca ccagcatttc tcagtgt		27

[0018]

<210> 67		
<211> 26		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 67		
ctttgttgac atatcaatca ccttaaa		26
<210> 68		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 68		
gtgaaaagaat taacaaggtaa ggagaaca		28
<210> 69		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 69		
gttctcctta cttgttaatt ctttcactt		29
<210> 70		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 70		
gtcctggtcc tggctttgtt cacgctt		27

[0019]

<210> 71		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 71		
gaagatcaac ctcgcaagag ccgcgaat		28
<210> 72		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 72		
caactacccc tagagaatca ctaaagaat		29
<210> 73		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 73		
gtcagtaagt atgttgtagg gttggatt		28
<210> 74		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 74		
ccatcatgct tgtcatttat gcgacttt		28

[0020]

<210> 75	
<211> 28	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 75	
caaggcteac cccgtttaa agtcgcat	28
<210> 76	
<211> 29	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 76	
gaataaagac aaaacgtcaa gactgacat	29
<210> 77	
<211> 24	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 77	
c当地ttcaaa cttatataaca ct当地	24
<210> 78	
<211> 22	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 78	
c当地tgaacga caatgatcat tt	22

[0021]

<210> 79		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 79		
ctcaaggctcg cagcagcaac tacaa		25
<210> 80		
<211> 24		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 80		
gagggaggttt gttcatatcc ttccg		24
<210> 81		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 81		
gaagaatatg aacaaaactcc ctctt		25
<210> 82		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 82		
ggacaacata ttgggcactg agtcc		25
[0022]		

<210>	83	
<211>	23	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的引物	
<400>	83	
	ctccccagccc tecggccaatt cgg	23
<210>	84	
<211>	28	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的引物	
<400>	84	
	cacatagatc acactgttgcg cgttcaag	28
<210>	85	
<211>	27	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的引物	
<400>	85	
	tactctccac attggaaacctt gaacgct	27
<210>	86	
<211>	27	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	合成的引物	
<400>	86	
	cactcagata ttaaacatagg ctgggtc	27

[0023]

<210> 87		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 87		
ggccctttgg tcaaccata aataa		25
<210> 88		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 88		
aagactgaca ttttattaag gcgattc		27
<210> 89		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 89		
catggaaactc tcaacaccct gagctgcaa		29
<210> 90		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 90		
cctcttagata taagatagtg ttcttcaa		28

[0024]

<210> 91	
<211> 26	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 91	
gtcaactgttt ccaaacgcac cctcaa	26
<210> 92	
<211> 26	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 92	
gggttcccaag caaggtaaaa tacctt	26
<210> 93	
<211> 26	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 93	
gtgttcctta ctggacctat aacact	26
<210> 94	
<211> 25	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 94	
gggtgatcat gacgaggttg atctt	25

[0025]

<210> 95	
<211> 26	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 95	
gtcaactgcat agtatacatac acacttt	26
<210> 96	
<211> 25	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 96	
gtcttgacatc ctcttccacg tggtt	25
<210> 97	
<211> 27	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 97	
cgagaacaat agaaaatagac catcagg	27
<210> 98	
<211> 25	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 98	
cctcttcaca caatgtatcca aactc	25

[0026]

<210> 99		
<211> 23		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 99		
gtccagggtc catgttatacg tct		23
<210> 100		
<211> 26		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 100		
cattgtccctt tcctttactga ttctcc		26
<210> 101		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 101		
tgactgcatt tatcatgtat gtgaaag		27
<210> 102		
<211> 24		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 102		
gattcgatga ggttgtcagg tttc		24

[0027]

<210> 103		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 103		
eggagaatca gtaaggaaag gacaa		25
<210> 104		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 104		
tgaggttgtt ggataccttg gagta		25
<210> 105		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 105		
accattctgc gctccattat tattt		25
<210> 106		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 106		
agtttcgtca caaccaggaa ttaca		25

[0028]

<210> 107		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 107		
gtttaaccat tctgctcc attat		25
<210> 108		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 108		
gaatatgtca ccctgcttca ggtaa		25
<210> 109		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 109		
tgttaattcct ggttgtacg aaact		25
<210> 110		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 110		
gtggcctgat aatgctaaga ccttt		25

[0029]

<210> 111		
<211> 24		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 111		
cgttacctga agcagggtga cata		24
<210> 112		
<211> 26		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 112		
cacctttct tctccttctc cttttt		26
<210> 113		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 113		
cttgaaacac gaagtttagag aagca		25
<210> 114		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 114		
cggttattgt ggttgtaagt gtgggt		25

[0030]

<210> 115		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 115		
cccaatgggt taaactctac aaggt		25
<210> 116		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 116		
attattcttg agctcgctca cttecc		25
<210> 117		
<211> 22		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 117		
gtggtaact cccaggaaaa gt		22
<210> 118		
<211> 27		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 118		
gaggaaagta caaatagcaa ctgacaa		27

[0031]

<210> 119		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 119		
cgaaggccacc tcacaccatg aacttcat		28
<210> 120		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 120		
cagaagtagg gtgttgagct tgagacatt		29
<210> 121		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 121		
gtcccttcatt caccttcctc tcttccctat		30
<210> 122		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 122		
gcttggccat ggtgatgact gatgagtgt		29

[0032]

<210> 123		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 123		
ctcaaatgcce tcaaaccgga taaccgta		28
<210> 124		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 124		
caacacccggc acactggaat ggcttggtt		28
<210> 125		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 125		
gcatgatatt cccgggttgt cctagcacat		30
<210> 126		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 126		
ccctctctga agtgatagat cttctggta		30

[0033]

<210> 127		
<211> 29		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 127		
cacttcagag agggtgattt gattgcagt		29
<210> 128		
<211> 30		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 128		
gttctggaag ctgttgtgt caataagaga		30
<210> 129		
<211> 22		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 129		
aataaccacg cctcagggttc tc		22
<210> 130		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 130		
gagtgttaa ggaccaatgg agaga		25

[0034]

<210> 131		
<211> 17		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 131		
cttcacaact caaacat		17
<210> 132		
<211> 16		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 132		
ttcacaaacac aaacat		16
<210> 133		
<211> 24		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 133		
ccctcaaacc ggataaccgt atag		24
<210> 134		
<211> 23		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 134		
cactggaatg gcttgttgg agg		23

[0035]

<210> 135		
<211> 16		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 135		
atgtctcaat gagecc		16
<210> 136		
<211> 16		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 136		
atgtctcaat gaaccc		16
<210> 137		
<211> 35		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 137		
cgtacatcat acatgttata aattaagctc aacaa		35
<210> 138		
<211> 28		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 138		
gcatatgcaa gtgctaagat aactttgt		28

[0036]

<210> 139

<211> 26

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 139

acacatttta attactatat ataaact

26

<210> 140

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 140

cacattttaa ttactatata tagct

25

<210> 141

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 141

agagcccttt ttgcatgtgc ta

22

<210> 142

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 142

tcgttcttat ttattgctac gcacact

27

[0037]

<210> 143		
<211> 17		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 143		
caaaaaggaca aaagtgt		17
<210> 144		
<211> 16		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 144		
aaaaggacga aagtgt		16
<210> 145		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 145		
ggaacccaaga gcaagagttt ctaca		25
<210> 146		
<211> 24		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 146		
cgctttcctt tctggcttg agta		24

[0038]

<210> 147		
<211> 15		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 147		
cttcattgtcg ctttgt		15
<210> 148		
<211> 16		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 148		
cctccatttctt gctttct		16
<210> 149		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 149		
tccaaattcac caactccatc aaact		25
<210> 150		
<211> 25		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 150		
caaaggaaag aaagagagag agtga		25

[0039]

<210> 151		
<211> 20		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 151		
cttccttagt tcaatataagg		20
<210> 152		
<211> 18		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 152		
tccttagttc aatatggg		18
<210> 153		
<211> 23		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 153		
cataggagaa cacgagggat gtg		23
<210> 154		
<211> 24		
<212> DNA		
<213> 人工序列		
<220>		
<223> 合成的引物		
<400> 154		
ggcttctact ttgtcttcc tctt		24
[0040]		

<210> 155	
<211> 15	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 155	
aatgcgaaga taaag	15
<210> 156	
<211> 14	
<212> DNA	
<213> 人工序列	
<220>	
<223> 合成的引物	
<400> 156	
atgcgaaaat aaag	14
<210> 157	
<211> 5495	
<212> DNA	
<213> 大豆 (Glycine max)	
<220>	
<221> misc_feature	
<222> (1)..(5495)	
<223> k 是 g 或 t	
<220>	
<221> misc_feature	
<222> (1)..(5495)	
<223> m 是 a 或 c	
<220>	
<221> misc_feature	
<222> (1)..(5495)	
<223> w 是 a 或 t	
<220>	
<221> misc_feature	
<222> (1)..(5495)	
<223> y 是 c 或 t	

[0041]

<400> 157

aaaaccctaa acctaaaccc ccaagcgctg cattggcat ccattgtgg aacattctct 60
gattaccaac aatgagaccc cggagaggcg gtggacgcgg cggcggattc agaggtggcc 120
gcgcggcgttgg tggtcgcggc agaggtggtt ttggcgcgg cggagggtgtt ttggcgcgg 180
ggggcggcggc cggatatcg gacgaaggac caccctecga agtcgttagt tgittctt 240
tccatttga aacttcagaa gcttcaat ttctatgtt tctgtgggt attcatcaa 300
ccgtttga agattaaaaaaa aaaaaatgtt ctttgggtt ttgtgtgtt atgacatgt 360
ttygtttctg gttttgata atttgcagg aaataacata actttatcta gcttagttaa 420
aaaaagattt ttattttat ttgtgtgtt ttatcttt taatteattt ggtgtatatg 480
cagagggtgtc atctttatg catgcatcg aggagatgc agtgcacaag ctacaaatg 540
agaaagtcc cttttcaat getcttattt atctgaaaaa catgactcag attggaaag 600
ttgatgaaat atttgcccc atcaatgaag ctgtgggtt ctcttattt ttgtttgtt 660
gttttaacat gttgtgtgt tttttgtt gttatggaaac tgggtttga 720
tggtatctgg tgcttcgtt gtgaacatgc agtacttctc aattaagatg atgaaaggaa 780
ttgttgetac ttcttattca tccggcgaca agttttat tttatccagg aaactgttgc 840
ctcttgcag attttttcca caaccaagg gacaatcagc tggtagaggt ggaggtggag 900
gtggcgttgg tggatgtaga ggtggccgtg gaggtgggtt tttcgttgg aggggcgctc 960
caaggggtgg gagaggtggt cttcccgagg gtggtgcagg tttttttt ttcttctt 1020
ttttttttt tatttttca ttttgc当地 cttttttttt tttttttt tttttttt 1080
agatgttttcc agcaggccat aagatcaaag ggacagtgtt gttgatgc aagaatgagt 1140
tggaaagttaa ccctgtatgc tcagcagttt acaaccttaa tgctttttt ggccwgagt 1200
tttttttttca gttttttttt tttttttt tttttttt tttttttt tttttttt 1260
acttttttcc atcatcatca tccacaatca tcactttctt tttttttt tttttttt 1320
tttttttttca gttttttttt tttttttt tttttttt tttttttt tttttttt 1380
gattcaayga tagtatacac attttttca ttttttttca atagagggtt ttttttttca 1440
tttagaaatat atcagttttt aatacataac caaagatatt tatatagtttt aaacttactc 1500
ggtagtacg taacatagta ttattawtat gttgttatga ttttttttgc gcamatggaa 1560
aaggaaaatg tggaaaggat acgttttttca aaggatttttca ttttttttca 1620
gagcaggaga gtttttttca aatatttca ttttttttca ttttttttca 1680
gttgttttttca ctttttttca ttttttttca ttttttttca 1740
aagccatttca aatatttca ttttttttca ttttttttca 1800
aacttttca aatatttca ttttttttca ttttttttca 1860
gttttttttca aatatttca ttttttttca ttttttttca 1920
gcttttttca aatatttca ttttttttca ttttttttca 1980
gcaccacttgc ttttttttca ttttttttca ttttttttca 2040
gagcgcagg aatatttca ttttttttca ttttttttca 2100
gataaggatgtt ttttttttca ttttttttca ttttttttca 2160
ctgttttttca ttttttttca ttttttttca ttttttttca 2220
aatttttttca ttttttttca ttttttttca ttttttttca 2280
ttgttttttca ttttttttca ttttttttca ttttttttca 2340
atatttttttca ttttttttca ttttttttca ttttttttca 2400
agacccgaat acttgcaggaa aaggcgaatg ttttttttca ttttttttca 2460
tcacttgcagg ttttttttca ttttttttca ttttttttca 2520
ggccggcgttgc gtttttttca ttttttttca ttttttttca 2580
agatgttgc ttttttttca ttttttttca ttttttttca 2640
tatttttttca ttttttttca ttttttttca ttttttttca 2700
tccacaaccttca ctttttttca ttttttttca ttttttttca 2760

[0042]

aatcttgaat taatcaatga cataataaat ataataaaacc actgatactt gaaatcttga 2820
 attaatttaat gacattttgg tcatacaaat cgtatcgat tcacgaacaa attcaaaaact 2880
 cattatgttt tttatgaaat ggttattcca ttaataagg attttcattt tattattaat 2940
 ttgggtagtg agtcagtcg catggatgac tgatgaagaa tttgcgttag agatgattgc 3000
 tgggttaaat ccctgcgtaa ttctggtct tgaagtaagt tcaaataattt atttataaaa 3060
 teaaatttga tgcacacgt aaatttaggtt gtatatcgcc agtaataaac taaaatacta 3120
 aatgggttgtt ctctttaat taaatggatt gttaggatgc cctccaaaaa gtaatcttga 3180
 tcctgcaatc tatggtgatc aaagcgttaa gataacagca gattcccttgc atctagatgg 3240
 gtacacaatg gacgaggtaa acatgcgaag cgaaggaaatt gatacaataa tttectctt 3300
 aaagtaatta attaaaacta attgaatttt gtggtagcagg cacttggtag tcgaaggta 3360
 ttatgttag attaccatga tatcttcatg ccatatgtga ggcagataaa tcagctgaat 3420
 tctgccaaga cttatgcac aaggactatc cttttttttga gagaagatgg aactttaaag 3480
 ccagtggcca tgaatttaag tttgccacat tctgtgggg atctgtcagc tgccgtcagt 3540
 caagtcgtct tacctgtttaa ggaaggtgtt gagagcacaa tttggctact agccaaagct 3600
 tatgtcatcg taaatgactt ttgttaccat caactcatga gccattggta taaaatttca 3660
 ttttcatttt cattgttata tatcaacata tatagataaa ttcaatctt aatcccttctc 3720
 tetcaagtgtt atattaatta cttttgattt ctctcataggt taaatactca tggggcgatg 3780
 gagccattcg tcatagcaac acaccgacat cttagctgc ttcacccat ttaacaagctt 3840
 ctgactccctc actatcgtaa caacatgaac atcaacgcac ttgccaggca atctctaatt 3900
 aatgctaattg gcataataga gacaacctt ttgcctcaa agtattctgt ggagatgtct 3960
 tcggcggittt ataagaattt ggttttcaact gatcaagcac tacctgcaga tcttatcaag 4020
 aggttaattaa ttaagtaat ctcctatat cactagaaaaaaa aaaatgtgac ataaagggtt 4080
 gaaaatataa ggtatacatt ttgttacaaa gcataatgat ggtatgtt atatattata 4140
 ttttggtttga cttagaggag tggcaattaa ggatccatca acccccacatg gaggctgtct 4200
 tctgatagag gactatccctt atgcccgttga tggactggag atatggctg caattaagac 4260
 atggatttcaaa gaatatgtgc ctttgtacta tgcaagagat gatgtatca aaaaatgattc 4320
 tgaactccaa cattgggttga aagaagctgtt agagaaaggc catggtgatt tggaaagacaa 4380
 gccatgggttgc cctaagttgc agacacttga agaccttgtt gaagtttgcc tcattatcat 4440
 atggatttgc tcaagcttcc atgcagccgt taattttgtt cagttatctt atggagggtt 4500
 gataatgaac cgcccaactt cttctagaag gttgtttctt gagaaggca ccccaagaata 4560
 tgaagaaatg attaataacc atgaaaaggc ttatgttata gacaattatcat caaagtggcc 4620
 gactctcattt agcctttcag tgatagagat ctgtcgaca catgttcttgc atgagggteta 4680
 cttggccag agggacaacc cacattggac ctctgttca aaagcattac aagccttca 4740
 aaagtttggaa aacaagctca aggaaatttga gaaaaactt gtgaggagga acaatgttcc 4800
 gagtcgtcag ggcaatcgac ttggcccggt tcaactgcca tacacttgc tttatcttag 4860
 cagtggaaatggtaactt tttagggaaat tccaaatagtc atctctatct aaggggcct 4920
 gtggtttact ctaattactt tataatagt gcatgtgtac ctccaaataaa aaaaatgtca 4980
 agactagaga tccgataaat cttgcattttt atctaatgtt tcaattatct tgggtttaa 5040
 ttaatgttgtt aattgagtc acagtttggt tttgttttgc aaataataag agcagtggc 5100
 actattatgtt tttgttttgc ttctgttcatc atgtttggta ttatacattt tgcagattac 5160
 gtaacacaaa tttatgttat tagtaaaaaaa actgttattgtt acgaatattt ttttattataa 5220
 gtatataaga aataaaaattt attagtatta gtttttcagg ggaagggggaa gatcatagag 5280
 caattttatgtt taatgtttt cacattgttgc cagttttttt cctttggaga taacatagat 5340
 gttgaatttgg ctttttatgtt ctatgcattc gagttcatgt tggttgatta tttgtttatc 5400
 attatggggaa agtcgattgtt tttgtttagt ctctatagaa attgtgttgc atttctatccc 5460
 tctcaaatcg aaatggaaag tggaaacttgg aaaaaa 5495

[0043]

<210> 158
 <211> 3856
 <212> DNA
 <213> 大豆 (Glycine max)

<400> 158

```

atttctgtat tttagcaa at taaaatagta gtgttattgg taggttttgaa caaagatgtt 60
ttcagtccca ggggtgtcgga gaatccgtaaa cagaggagga gggcataaaga taaaaggac 120
agtgggtttg atgaggaaga atgtgttggaa cttcaacagc gtggctgate ttactaaagg 180
aaatgttggg ggactcatag gcacccgcct caacgttggt ggctcaacac ttgataaacct 240
cactgtttc ttggccgaa gtgtccctt acacgttaccat agtgcatacca aacctcttgg 300
ttcattttt cttcccttcca cacaatcaat aacttctata ttcaaaatata agtgtttaaat 360
ctctataactc teatttcattt cattcaatga aaaaaaaaaatc ataagactttt taactaaaaat 420
taacctatgt aaagaatcac aaacaaaaaaa ctatataata ttaagtttat ttactttttt 480
tataatgaca aaaaaaaaaat ttttatatgg tgcacaaaattt tttgtactct taaaatata 540
tcactttata catagccaaa catatttattt ttgtatagta ttaacttattt tgggtacgta 600
ctttaataat attattatgt gtgtatgtat ggtctgtttt tagcaaaatgg aaaaggaaaa 660
gttggaaagg ataegettctt ggaagggattt attgtgttgt taccaactttt gggagcagg 720
gagtctgcat tcaatattca gtttgaatgg gacgaaagca tggaaatccc cggtgcgttt 780
tacataaaga actacatgca agttgagttt tacctcaaga gtttgaetct tgaagacg 840
ccaaaccaag gaaccattcg ctttggggc aactcatggg tttataaacac taaaactttac 900
aaaagcggtc gcattttttt tgccaaaccat gtaagctattt tattttacgt acttagctag 960
tagtttataa ttaaaattca gtatgtat atatataat ttcattttgtc cgtctcta 1020
ggcaagatct ttaatacgtt aatttaatga attttagaca tatgttccaa gtggagacacc 1080
agcagactt gtgggttaca gagaagaaga attgaagaat ttaagaggag atggaaaagg 1140
agagcgc当地 gaacatgata ggatctatga ttatgtatc tacaatgattt tggcaatcc 1200
agatcacggtaaaaatttttgc tctgggggattt ttttttttttccatc atcccttaccc 1260
tcttaggggaa agaactggc ttttttttttccatc atcccttaccc 1320
tatttttttta agtgaatttga tattttatgt gtgaagctaa tttaattttt tttttttttt 1380
tttttttttacg atgttatgttgc ctttttttttccatc atcccttaccc 1440
tttttttttatttttttatttttttatttttttatttttttatttttttatttttttattttttt 1500
attttttttca atcaatttca acagatcaga atttctgagaa gcccggccaa gttttttttt 1560
caagagatga aaatttttttgc ttttttttttccatc atcccttaccc 1620
ctttatctca atatgtctta ccggccgttgc aatcttttttccatc atcccttaccc 1680
atgagtttgc tagtttccaa gatgtgcgttgc accttccacgaa aggtggattt aagcttccaa 1740
cagaagtaat tagcacaattt atgccccatc cgggttttttccatc atcccttaccc 1800
gtgaacaatgtt ctttttttttccatc atcccttaccc 1860
tgaccgttgc agaatttttttgc agagagatgg ttgggggttgc aaatccatgc gtaattttttt 1920
gttttttttca gtttttttttccatc aaaaggccatc ttttttttttccatc atcccttaccc 1980
gtaagataac agcagatgcc ttttttttttccatc atcccttaccc 2040
tcaagtttca taatttttacc acagggaaatgg ggaactttaa gtaatttttttgc gttttttttt 2100
atgatctgtat tacaatataat ttaatttttttgc gcacaggeac tagcaagttgc gaggttattttt 2160
atgttagattt accatgtatgtt atccatgc ttttttttttccatc atcccttaccc 2220
aaggcttgc ttttttttttccatc atcccttaccc 2280
ggccatcgaat taagtttgc acatcttgcgtt gggggccatc ttttttttttccatc atcccttaccc 2340
atcttacccgc caaaagaagg ttttttttttccatc atcccttaccc 2400
gttttttttgc ttttttttttccatc atcccttaccc 2460
tcaatcttccatc ttttttttttccatc atcccttaccc 2520

```

[0044]

tgttcatagg taaaatactc aagcggtgat tgagccattc atcatagcaa caaaccgaca 2580
 ccttagtgct ctteacccaa tttataagct tctaacttct eactaccgtg acaccatgaa 2640
 catcaacgca cttgcttaggc aatctctcat taatgtctgat ggcataatag agaaatctt 2700
 tttgccctca aagcattccg ttgagatgtc ttcagcggtt tataagaatt gggtttcac 2760
 tgcataagca ctacctgcag atcttatcaa gaggttaatta atctctctaa actctaaaca 2820
 taaaactatg aaatataggt acatagttt ttattgattt atatittgtt tgattcagag 2880
 gagtggcaat taaggatcca tctgccccac atggacttgc acttctgata gaggactacc 2940
 ctatgttgt tgcattggata gagatatggg ctgcaattaa gacaagggtt caagaatatg 3000
 tgcattgtt ctatgcaaga gatgatgtat tc当地acactga ttctgaactc caacagtgg 3060
 gaaaaaaggc tgcataagaaa ggtcatgggt atctgaaaga caagccatgg tggcctaagt 3120
 tgc当地acaat tgaagagctt gttgaaattt gcaccattat catatggact gcttcagccc 3180
 tccatgcagc cgttaactt ggtcaatattc catatggagg ttcttctg aatgc当地aa 3240
 ctatgttagt aagggtgtt cctgagaaag gc当地cccaaa atatgaaatggtggaaaa 3300
 gtc当地aaaaa ggcttattt agaacttattt catcaagtt tcaacttca gttgacctt 3360
 cagtgtataga gatcttgtca aggcatgtt ctgatgagg ctacccggc当地aggac 3420
 acccacattt gacccctgtac tcaaaaagcat tacaaggctt tcaaaaattt gaaacaaggc 3480
 tcaaaaat tgaggaaaaa ct当地caagga agaacaatga tcaacttca tccaatgcac 3540
 ttggccgggt tcaactgcca facactttgc tccatctaa cagtggggaa ggggttactt 3600
 gcaggggggat tccataatgc atcttctatctt aaggggcat gtggctact ttaatttacag 3660
 tactgtaccc accttccaaat aaaaagatg caaggctaga gaiccaataa atcttgcattc 3720
 ctatctaattt ttcaattt ct当地gtttt aatgctgtt当地 tggagccctac aatttgc当地 3780
 ttaatttacatg gacccggacat ctatgtttt gctttagct tctgtcatgt tt当地atgat 3840
 atatttgc当地 gatgtt 3856

<210> 159

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 159

gcttatcatca actcatgaggc cattt

25

<210> 160

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 160

gtgtcggtt gttgttatga tgaat

25

[0045]

<210> 161

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 161

caatcacccgc ttgagttat

18

<210> 162

<211> 17

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 162

aatcacccgca tgagtat

17

<210> 163

<211> 3524

<212> DNA

<213> 大豆 (Glycine max)

<220>

<221> misc_feature

<222> (1)..(3524)

<223> k 是 g 或 t

<220>

<221> misc_feature

<222> (1)..(3524)

<223> r 是 a 或 g

<220>

<221> misc_feature

<222> (1)..(3524)

<223> w 是 a 或 t

<220>

<221> misc_feature

<222> (1)..(3524)

[0046]

〈223〉 y 是 e 或 t

〈400〉 163

[0047]

ccacactaca acctgaacge gaacagcata atatacgat tgaatggacg ggcattgata 2760
 caagtggta attgcaacgg tgagagagt tttgatggag agctgcaaga gggacgggtg 2820
 ctgatcgac cacaaaaactt tgtggtggt gcaagatcac agagtgacaa cttecgatgtat 2880
 gtgtcattca agaccaatga tacacccatg atcgccactt ttgcaggggc aaactcattt 2940
 ttgaacgcac taccagagga agtggatcag cacacttca acctaaaaag ccagcaggcc 3000
 aggccatcaa agaacaacaa cccttcacag ttcttgcgtt caccctcaggaa gtctcagaag 3060
 agacgtgtgg cttagageccc tttttgtatg tgcttaccca cttttgtctt ttggcaata 3120
 gtgtcgtacaa ccaataaaata ataataataaa taatgaataa gaaaacaaag gcttttagctt 3180
 gccttttgtt cactgtaaaa taataatgtt agtactctt ataatgagtc acgaaacttt 3240
 tgcggaaata aaaggagaaa ttccatgtt ttttctgtca aatcttctt tgcgtcgtc 3300
 tctctctt ttttttttctt ttcttctgtt cttcttgcac aacaaaaggc aaacaataaac 3360
 gattggtcca atgatagttt gcttgcgtca tgatatctt aggaagtgtt ggcaggacag 3420
 gacatgtgtt agaagactaa aattgaaagt attgcagacc caatgttga agattaactt 3480
 taagaatgaa gacgttctt caggttctt atgacttggaa gctc 3524

<210> 164

<211> 3640

<212> DNA

<213> 大豆 (Glycine max)

<220>

<221> misc_feature

<222> (1)..(3640)

<223> d 是 a 或 g 或 t

<220>

<221> misc_feature

<222> (1)..(3640)

<223> r 是 a 或 g

<220>

<221> misc_feature

<222> (1)..(3640)

<223> w 是 a 或 t

<400> 164

tgaaaggaaa gaaagaagca gaggaagaaa agaaatgaaa ccatgcattt tccccacccc 60
 agacatcat gggtttctgc catttgcattt acaaacttgg aaacacctttt ctctttgtca 120
 cgtaatcgat attccgaaacg cacccttacac cattaacttta atatgttgcg acagaagggt 180
 tccatagccca tgcataacttgc agaatgttgc aagctcggca ccccaacttgc gagacgtgtc 240
 ctttcatcactt ctttcttctt tccctataaa taaccacgc tcaatgttgc tgcgttccaa 300
 cacaacattt ctctccattt tgccttgcattt ataatacttca gcatggctaa gcttgcgtt 360
 tccctttgtt ttctgtttt cagttgcgtt tgcgttgcgtt tcaatgttgcg agacgtgtc 420
 cagaaacatc agtgcgttgc ccaacgcgtt aatgcgttgc aaccggatca ccttgcgtt 480
 tcagaagggtt gcttgcgttgc gacatggaaac ccttgcgttgc aaccggatca gcttgcgtt 540
 gttgcgttgc ttcgttgcgtt ctttgcgttgc aacgcgttgc gcaatgttgc tcaatgttgc 600
 gtcgttgcgttgc gtcgttgcgttgc aacgcgttgc gcaatgttgc tcaatgttgc 660

[0048]

ttagtttat taaaaatatt aatcgaaaaa aatccaactt acaaccctta ttcaacaagg 720
 ttcatttat ccacactta taatacatac atacatatat atatatatat atatatatat 780
 atatatatat atatatatat atatatatat atatatatat atatatatgtat 840
 gtatgtatga ttgatttact aacgtatata cctatatttt ttagcagag gactgttatt 900
 tagtacaaca aatcaatgc acaaattgcaaa gaaaagaaaac atcaacggct gaacttllgt 960
 tcaaaaacaaa tacttaattg gaaaataata aaataaaaaac tggaaattga cttaatgcac 1020
 tcttcatgct ttttttaact aaaactgttt cctacattaa caaatacttg cggttat 1080
 caacttacat aticttatct ttttagtata aatacatttg caaattatfat cattgtat 1140
 tttaaagatat tccttatcta taattcaactg gtaaaaaaaac catattagca aaaccctcaa 1200
 gagtaacgtt aaggacatcg ataatgtttt gtttgtatca gttttcagaa taaaacaa 1260
 ttgaaacgaa aatggttt atagtaatgt ttttgtattt gtttaatgt aataggtagt 1320
 ggtatttttt gcatgatatt cccgggttgt cctagcacat ttgaagagcc tcaacaaaaa 1380
 ggacaaagca gcaggccccca agaccgtcac cagaagatct atcacifcag agagggtgat 1440
 ttgatttgcag tgccaaacccg ttttgcatac tggatgtaca acaatgaaga cactccgtt 1500
 gttgcgtttt ctcttattga caccaacage tccagaacc agctcgacca gatgcctagg 1560
 gtacgtgagc cacatatata gcattagata tttagattct taaaagattt aaatatctt 1620
 ttggtttgtt tgcaatttgc ttttttaattt tagtatactt agattaaatg tcaattttat 1680
 cttaattttt gttttttt atgcaatttgc ttccctgtgg tgaaatcact tttaaataact 1740
 ggaaaaaaattt gtagctgata aaaaaaaagaa attggacaac gttgtacat gcattaagat 1800
 tagaactttaat aagtcttaat actggttaca gaattaaacaa gtatttttt tgataaaaata 1860
 tacacttggc agagatttca tcttgcgtgg aaccaagagc aagagtttct acagttatcag 1920
 ccacagaage agcaaggagg tacicaaage cagaaaggaa agcgtcagca agaagaagaa 1980
 aacgaaggag gcagcatatt gagtggcttcc gcccggaaat tcttggaaaca tgcgttcgtc 2040
 gtggacagggc agatagttag aagactacaa ggtgagaacg aagaggaaga gaagggtgcc 2100
 attgtgacag tgaaaggagg tctcagcgtg ataagccac ccacggaaaga gcagcaacaa 2160
 agacccgagg aagaggagaa gccagattgt gacgagaaag acaaacattt ccaaagccaa 2220
 agcagaaatg geattgacga gaccatttgc acaatgagac ttgcacccaa cattggccag 2280
 acttcatcac ctgacatattt caaccctcaa gctggtagca tcacaaccgc taccagccct 2340
 gacttcccttccatcactt gctcaactt agtgccttgc ttggatcact ccgcaggta 2400
 ettacatcat taattcatgt taaaatgtt caaaaaatgt taaaatgata ctaatagttt 2460
 tatatagttta taaaatgtt tgacaaaatc atcttaaacac ctgcattgtat atgcattatt 2520
 ttttgcattttt ttttggatcata aatatgcgtt gtatgtctat atatatatat gttatggaaa 2580
 atgtttaaccat gtttttttgg atatggatcata aggaatgaga aataattttt ctaagattra 2640
 gtttttaatat gatcttcata ataaacattt tttttttttt ttttaattat gtacttagta 2700
 cactaattaa caaaatcata tataatgtat tttcataggtt ttgcaaaagaa aagagatatg 2760
 attgagagat ggaataattt ctgtataat aaaaagatattt gttatgatca acgtacactt 2820
 gatgtatggt gcaaatgtt atgttgcgtc cacaactacaa ectgaacgc aacagcataa 2880
 tatacgcattttaat gatggacgg gcattggatc aagtggatc ttgcaatgtt gagagatgtt 2940
 ttgatggaga gctgcaagag ggacagggtt taattgtgcc acaaaaactttt ggggtggctg 3000
 caagatcaca gacgcacaac ttgcgtatg tttcatttcaaa gaccaatgtt agaccctcga 3060
 tcggcaacctt tgcagggtca aacattttgt tgaacgcattt gccggaggaa gtgatttcagc 3120
 aaacttttaat cctaaggagg cagcaggccaa ggcagggtcaaa gaacaacaa cctttagt 3180
 tcctggttcc acctaaggagg tctcagaggaa gatgttgcgtgg ttagagccctt ttttgcattt 3240
 gctaccacac tttcgatctt ttggcacta gttgttgcgtc gtagcaatataa ataaacacgt 3300
 ataaaaacaaac aaaggctttt ctgccttttgc ttaagtataa aataacgggtt aatgtat 3360
 tacccttgcgtt taatggatcata ctcaagtacc tttgcgggaa aatcataaaat taaaagagaa 3420
 atttcaatgtt attttgcattt ttttttttattt ttttttttgc tttttttatcc 3480
 ttcttgcaga acaaaaataat gaaaatagia ttgaaaacaaat ttttttttttttgc 3540

[0049]

ctttgacaat gatatataat agtttttatt gtagaatttc gtgacatgtt gtgegcteta	3600
atttatttgtt aagaaccttc attggttcg tatgcgttd	3640

<210> 165
 <211> 3544
 <212> DNA
 <213> 大豆 (Glycine max)

<220>
 <221> misc_feature
 <222> (1)..(3544)
 <223> r 是 a 或 g

<220>
 <221> misc_feature
 <222> (1)..(3544)
 <223> s 是 c 或 g

<400> 165
 ctataacaata taagatcata gtactgacaa aatgcacagt aaaacagttc aaattgagaa 60
 ggattsttaa cacaccatag tatttaaat atatctttac agagacaatt atgctggagg 120
 attcaggcaa agatttatata ttgtggattt gtttttaat aattaacgca tcataatgaaa 180
 gatecatgtat atatactaatt ggttataaga aaaatatita acagtttctta taacctttt 240
 ctttatctt ttactgtat attatttattt ttatttcaca ttttaatca gcttatctca 300
 tttataaacg aaattgtata aaaatataca tgatgaactg aatagaacaa tattgatctg 360
 atattctcat attgtataag aggatagact ttgagacgag gagaatctgt aggaggggac 420
 cattcagagt gcctccaattt tgggttgtt tcattgtacc attgcaaaata taaaacgaaac 480
 atgcatgttt atgtatgagg tgtaacaaaa ttggaaacaa tagccatgca aggtgaagaa 540
 tgcacaaac tcagcaaccc ttatttcattt acgtgtccct cagtcactct cctctcatac 600
 ctataaaatca ccactcctca tgggttgtt tcattgtacc attgcaaaata taaaacgaaac 660
 aacacttctt tagtcaataa trggaaagcc cttcacttctc ttcattttttt ccctttgttt 720
 gtcacatctt tgggttgtt tcattgtacc ttacttcattt ttcattttttt ccctttgttt 780
 caacaaccc aacgcgttgg aaccgcacca ccgcgttgg tccgaagggtg gtttggattca 840
 aacatggAAC tctcaacacc ctgagctgaa atgcgcgggt gtcactgttt ccaaactcac 900
 cctcaaccgc aatggcctcc acttgccttcc ttacttcattt ttcattttttt ccctttgttt 960
 cggccaaagggt aatcatatata aaggagtgct ttcattttttt ccctttgttt ccctttgttt 1020
 agcaattctc agtgtatattt aatccattttt tcattttttt ccctttgttt ccctttgttt 1080
 taccatagat catttactaa agataataat gatTTAGTA aatagtatct ctatagtaaa 1140
 ttttacatga ttatTTAact acaaattttt attattttat atagaatgac ttgttgttgc 1200
 tatcaatcac cttaaaaatgtt ttatTTAactt atatataatca actaagatat ctgattaaat 1260
 aaaaatgtga ttgtttttt tgggttgtt tgatgtacag ggaaaggagc acttggagtt 1320
 gcaattccag gatgtcctga gacgttttggag gagccacaaag aacaatcaaa cagaagaggc 1380
 tcaaggctgc agaaggcageca gtcacaggac agtccaccaaa cttcaatggaa 1440
 ggagacgtac tgggttgtt ccattttttt ccctttttt ccctttgttt ccctttgttt 1500
 ccagggttttccatcattt ttcattttttt ccctttttt ccctttgttt ccctttgttt 1560
 ccttagggtaa ttatcaattt aatttttttccatcattt ttcattttttt ccctttttt ccctttgttt 1620
 tgtaatTTTtttccatcattt ggttttttttccatcattt ttcattttttt ccctttttt ccctttgttt 1680

[0050]

gagaccatgc aacaacaaca acagcagaaa agtcatggc gacgcaagca gggcaacac 1740
 cagcaggagg aagaggaaga aggtggcagc gtgcctactg gcttcagcaa acacttttg 1800
 geacaatccct tcaacaccaa cgaggacata gctgagaac ttctacttcc agacgacgaa 1860
 aggaagcaga tcgtgacagt ggaaggaggt ctcagegtta tcagccccaa gtggcaagaa 1920
 caacaagatg aagatgaaga tgaagacgaa gatgatgaag atgaacaaat tccctctcac 1980
 ccctctcgcc gaccaagcca tggaaagcgt gaacaagacg aggacgagga cgaagatgaa 2040
 gataaaacctc gtccttagtc accaagccaa ggaaagcgt aacaagacca ggaccaggac 2100
 gaggacgaag atgaagatga agatcaaccc cgcaagagcc gccaatggag atcgaaaaag 2160
 acacaacccca gaagacctag acaagaagaa ccacgtaaa gaggatgcga gacaagaaac 2220
 ggggttgggg aaaatatctg caccctgaa ctctacgaga acatgtctcg cccttcacgc 2280
 gctgacttct acaaccctaa agctggcgc attagtaccc tcaacagect caccctccaa 2340
 gcccctcgcc aattccaact cagtgcctaa tatgttgtec tetacaaggt atgttaattca 2400
 cctcattcat attactaagt aatcaacatg aaactaat acgtacatac ttacacatct 2460
 accagtaatt ttccgtgaa tattcaattt tcaatttagt tatcttgaga aaattaagaa 2520
 ataaaaaagaa agcacaaaaag gaaaaaatct ttatgtcata aatcatatga tataataatt 2580
 tagaagacat ataaaaatgt cagtaagtat gttgttaggt tggattccct taaatgtcat 2640
 taaaatatca ttgatatgg gtaattctt agtgatttc tagggtagt tgaactgtaa 2700
 tgtatataa ttgtgcattt attttatgtat gttactttaa cagtcaatg aagacttatt 2760
 tgataataat tatagttact tttttttt actactttt ataaaaaaat aataaaaaata 2820
 ttgggtttaaa tatataatata ataataataa tttatgtatg acgtacacaca ttgttattata 2880
 tccatgcaga atgaaatttta ctctccacat tggaaatctgatgcaaacag tttgtatctat 2940
 gtgactcgag gacaaggaaa gtttagagtt gtgaactgcc aagggatgc agtgttcgac 3000
 ggtgagettta ggaggggaca atttgtgggtt gtaccacaga acttgcgtt ggcggagc 3060
 gcccggagaac aaggattcga atacatagta ttcaagacac accacaacgc agtcaactage 3120
 tacttgaagg atgtgttttag ggcaattttcc tcagaggattt ttgtccattt ttacaaccc 3180
 egacagagtc aagtgtctga gtttaagtat gaaggaaattt ggggttccctt ggtcaaccc 3240
 gtttgcataac aaggctcacc ccgtgttaaa gtcgcataaa tgacaagcat gatgggtgtga 3300
 ggttgcataatc attttatgaa ataataacaa ataaataaaat tttttatgtat aataaaaaagt 3360
 atggcccatg taccatccca gcgacccat gtttataatct gttttttttt gttttttttt 3420
 atcgccctaa taaaatgtca gtcttcacgt ttgttcttta ttctgtgtttt atttttttt 3480
 ttgtgggcaaa gttttttt atctacittt aatgtgatgaa tactgattt atgtttactg 3540
 gasg 3544

<210> 166
 <211> 3355
 <212> DNA
 <213> 大豆 (Glycine max)

<220>
 <221> misc_feature
 <222> (1)..(3355)
 <223> r 是 a 或 g

<400> 166
 ccaacttattt aacactttcc ttgttcaat atggggaaac cttttttcac 60
 tctctctttt tctttttt gtttgcactt cttgtcgagt gcatgtttt ctattaccc 120
 cagcaagtcc aacgagtgcg aactcaacaa cctcaacgcg ttggaaacccg accaccgcgt 180

[0051]

[0052]

tgtttccat egctcaagtt gtgtgcctta attgettata actcgagatg tctgtctatt 3120
 aacgaaaatc aatattaaca actgaacttc tgctccatt aaaggtgaa aaaaaatgct 3180
 ttatattgtt aattaatttt atatataagt aactttttt agtaatcaaataatgatta 3240
 tataaattaa tcttttaag taagatacaa atttgatatt ttattaattg aatttgaatt 3300
 tattacgata gaagtttatga ttataatgt aattaaattha attattataa aatta 3355

<210> 167

<211> 1011

<212> DNA

<213> 大豆 (Glycine max)

<400> 167

gtacttcctt tattctgac gttttatata caagtggaca tacgtgaaga tttaattat 60
 cagtcfaaat atttcattag cacttaatac ttttctgtt tattcttata ctataaggtag 120
 tcccgattct cccaacattt cttattcaca caactaacta agaaagtctt ccatagcccc 180
 ccaaaaatga agagcacat ttctttgtt ctctttctt tttgtgcctt caccaccta 240
 tacctaccctt cagccatcgcc tgatttcgtt ctcgataatg aaggtaaccc tcttggaaat 300
 ggtggcacat attatatctt gtcagacata acagcattt gtggaaataag ageagcccc 360
 acggaaatg aaagatgcc tctcacatgt gtgcataatc gcaatgatct cgacaaagg 420
 attggaaacaa tcatactcgcc cccatatcgat atccgtttt tcgcggagg ccattttt 480
 agcattaaatg tcgatttcatt tgcaattata atgcgtgtt ttggaaattcc taccgatgg 540
 tctgttgttgg aggatctacc agaaggacat gctgttaaaa ttggtgagaa caaagatgca 600
 atggatgggtt ggtagact tgagagagtt tctgtatgt aattcaataa ctataagctt 660
 gtgttcgttc cacagcaagc tgaggatgac aaatgtggg atattggat tagtattgtat 720
 catgtatgtt gaaccaggcg tttgggtgggt tctaaagaca aaccgttagt ggtagttt 780
 caaaaacttg ataaagaatc actggccaag aaaaatcatg gccttctcg cagtgtgt 840
 gacacaatgt tgagatgtt aaataaatgc ttgggtgtt cggaaatcattt acactaaata 900
 aaataatcaa agcttatata tgccttccgc taaggccaa tgcaaaagaaa ttggttttt 960
 ctctttatctt ttggccactt ttactatgtac gtattaaatc ctacttaatc a 1011

<210> 168

<211> 25

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 168

gagaacaaag atgcaatgga tggtt

25

<210> 169

<211> 25

<212> DNA

<213> 人工序列

[0053]

<220>

<223> 合成的引物

<400> 169

gctgtggaca gaacacaagg ttata

25

<210> 170

<211> 18

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 170

agaaaactctc tcaagtct

18

<210> 171

<211> 19

<212> DNA

<213> 人工序列

<220>

<223> 合成的引物

<400> 171

catcagaaac tctaagtct

19

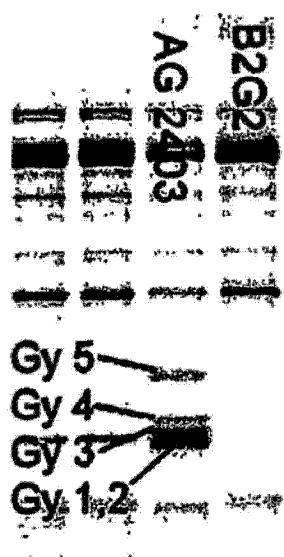


图 1

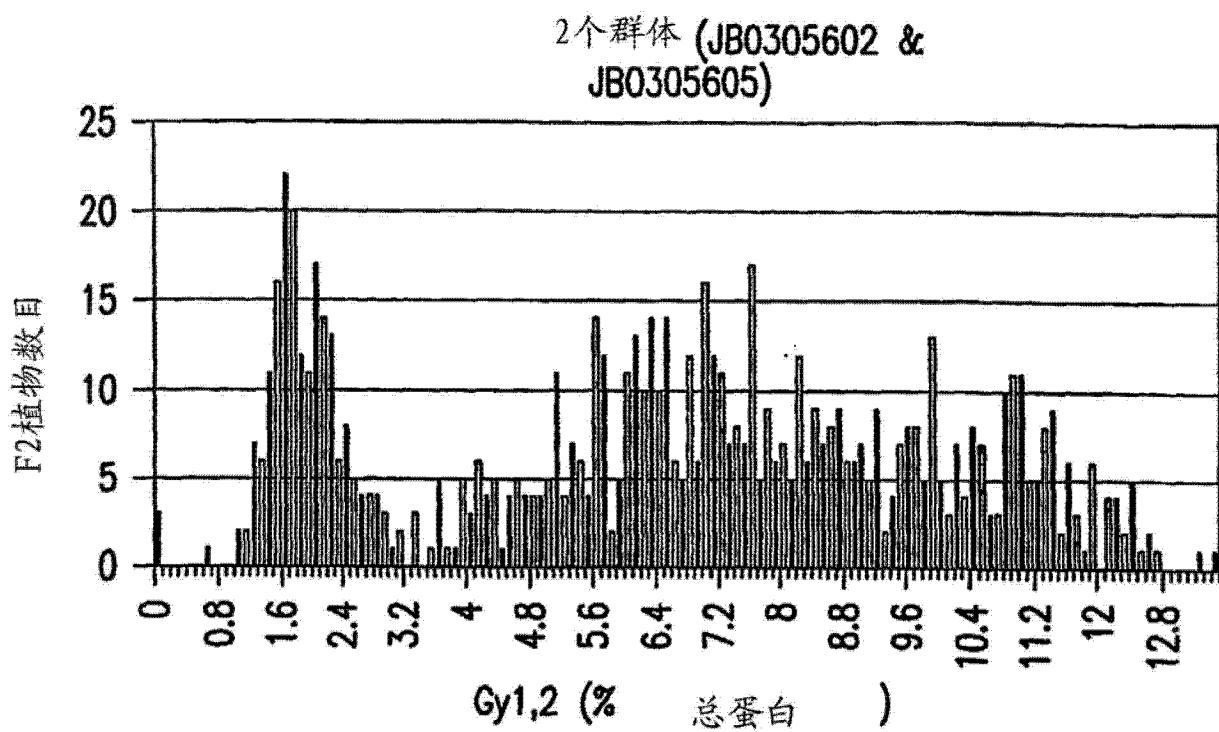


图 2A

	观察的植物数目		卡方	P 值
Gy1,2 水平	$\leq 3.0\%$	$\geq 3.1\%$		
JB0305602	92	307	0.854	0.355
JB0305605	97	301	0.120	0.942
总计	189	608	0.670	0.715

图 2B

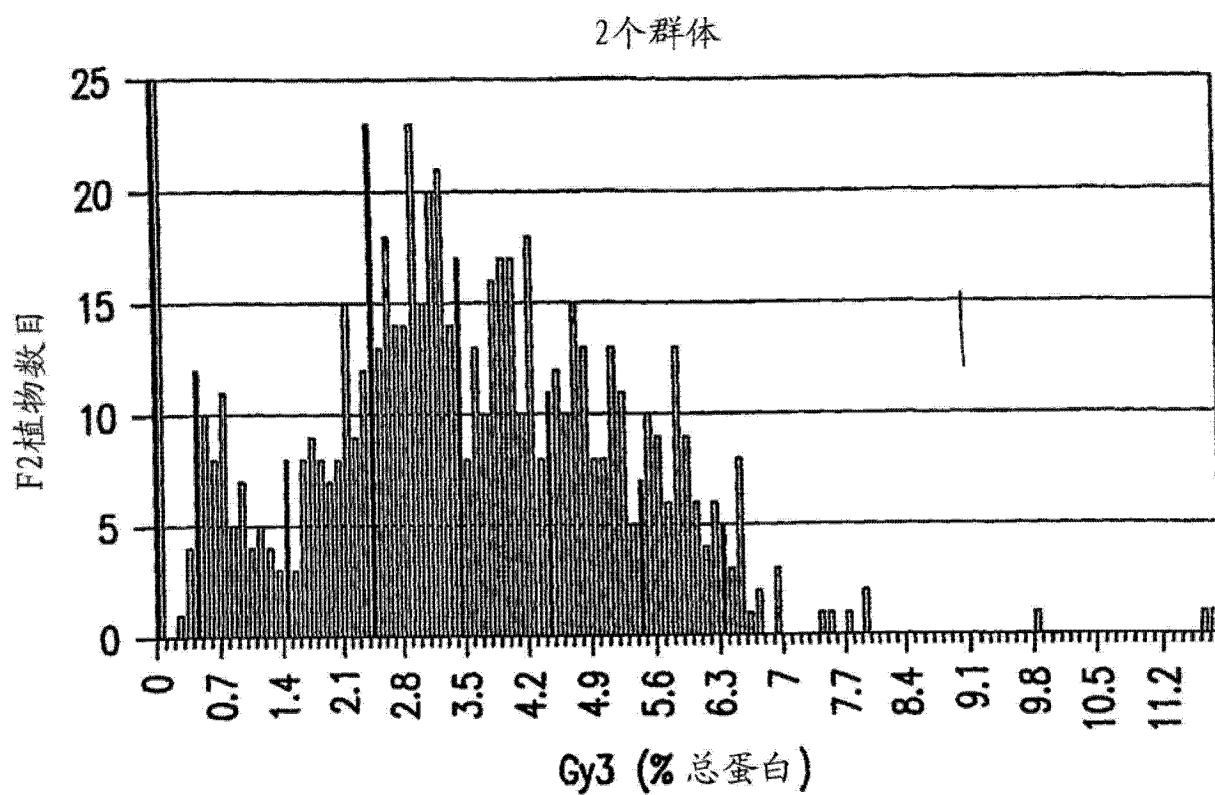


图 3A

	观察的植物数目			
Gy3 水平	≤ 1.0	≥ 1.1	卡方	P值
JB0305602	95	305	0.333	0.564
JB0305605	91	309	1.080	0.583
总计	186	614	1.307	0.520

图 3B

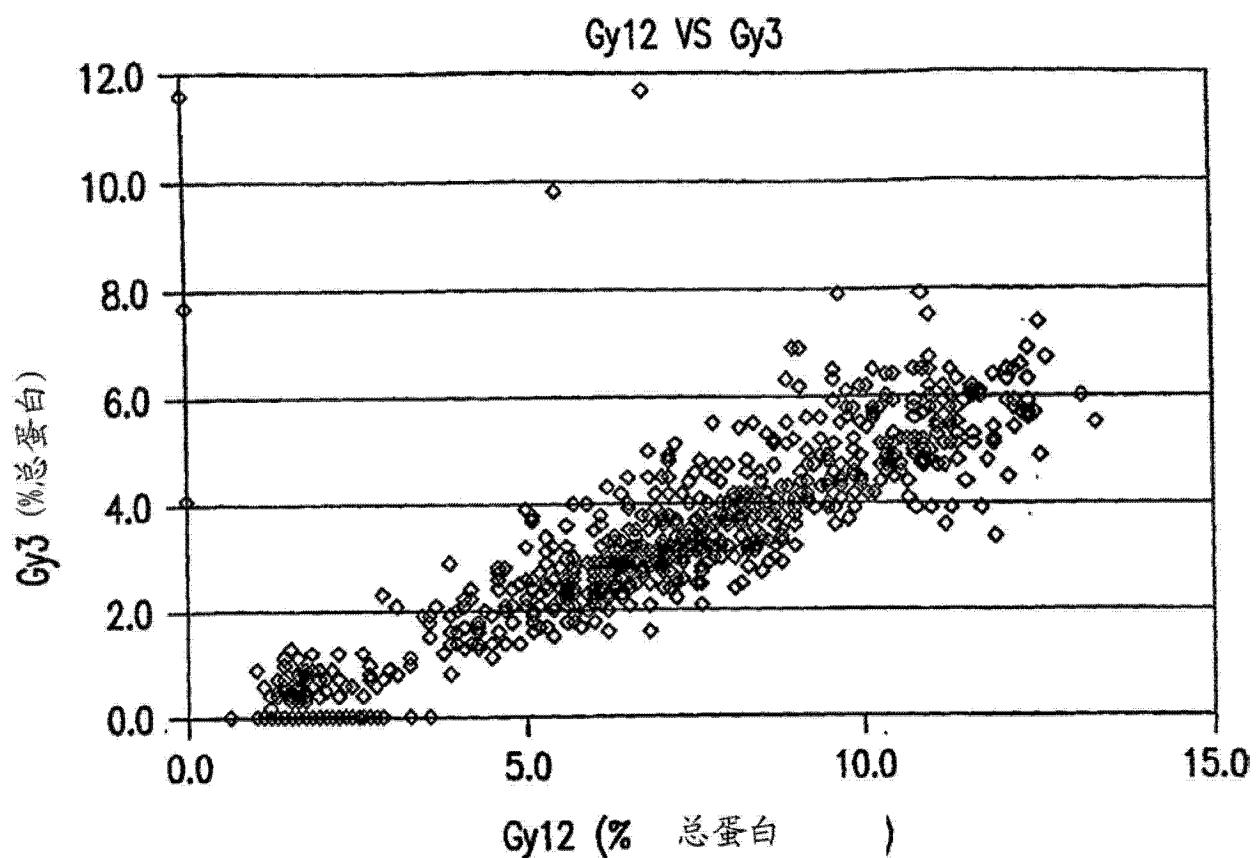


图 4A

表型相关系数 (r)

	Gy1,2	Gy3	Gy4	Gy5
Gy1,2	1			
Gy3	0.88	1		
Gy4	-0.23	-0.16	1	
Gy5	-0.59	-0.62	0.04	1

图 4B

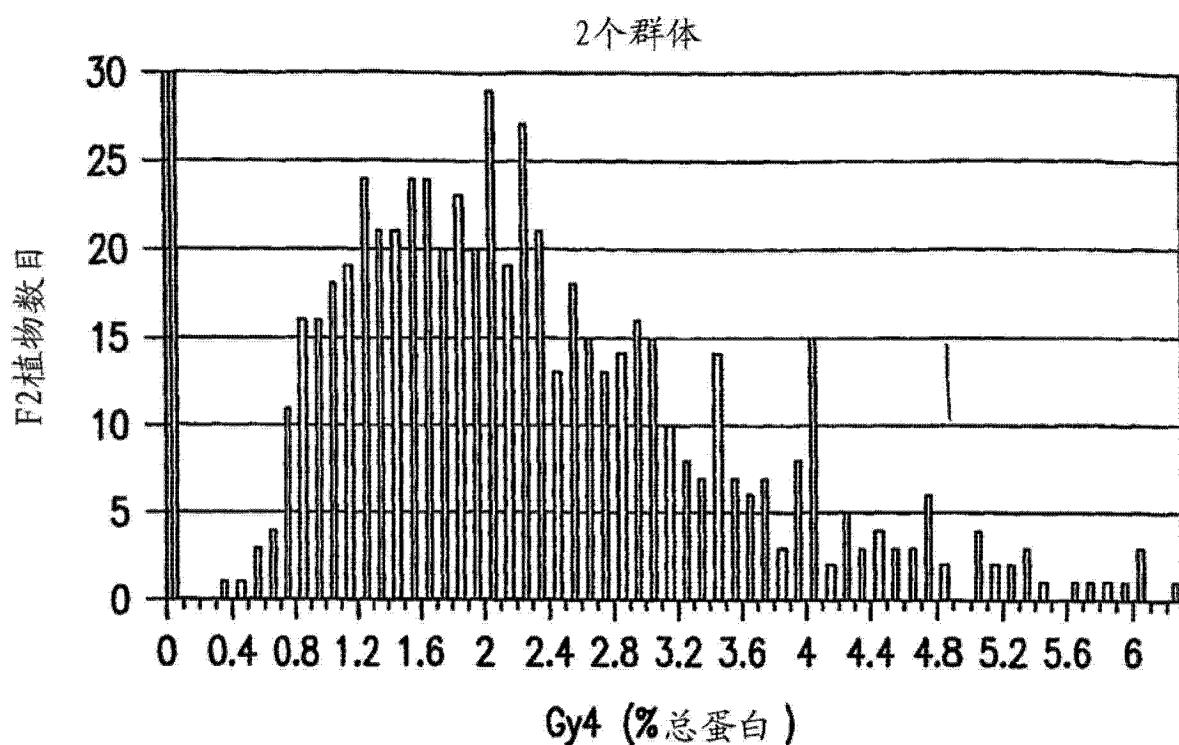


图 5A

观察的植物数目			
Gy4 水平	0	>0	卡方 P值
JB0305602	106	294	0.480 0.488
JB0305605	95	305	0.333 0.846
总计	201	599	0.007 0.997

图 5B

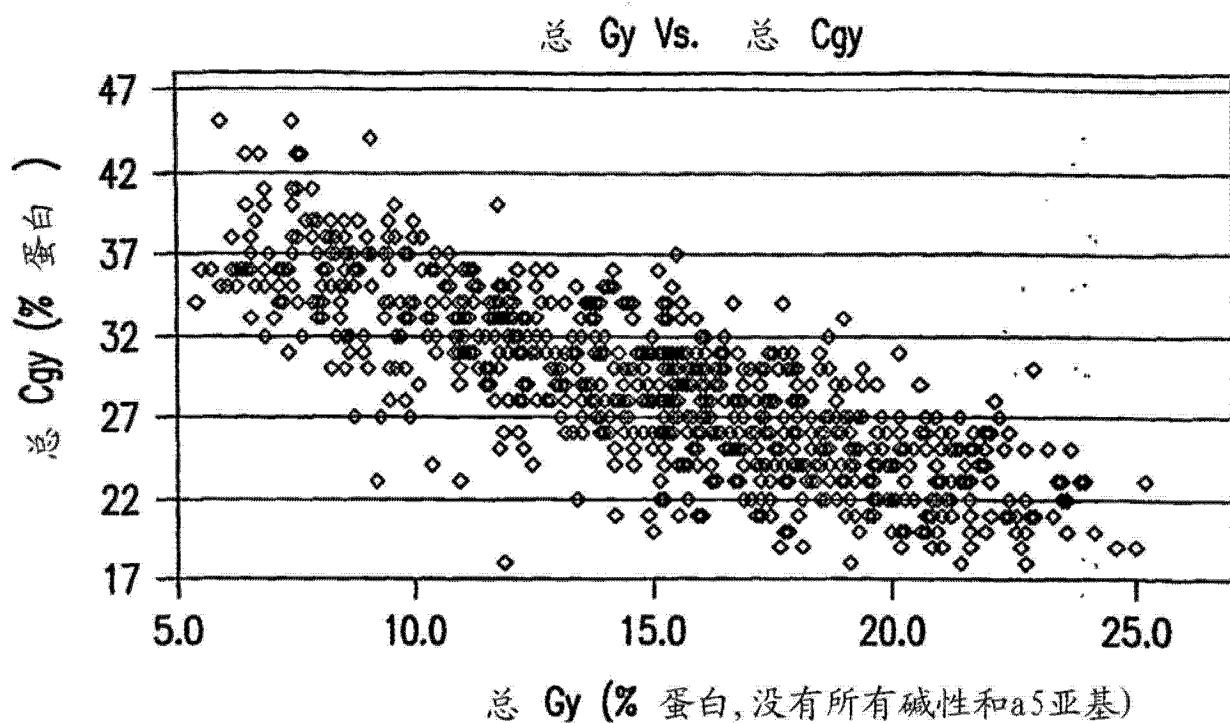


图 6A

	Gy 总计
Cgy1	-0.83
Cgy2,3	-0.79
Cgy4	-0.36
Cgy 总计	-0.78

图 6B

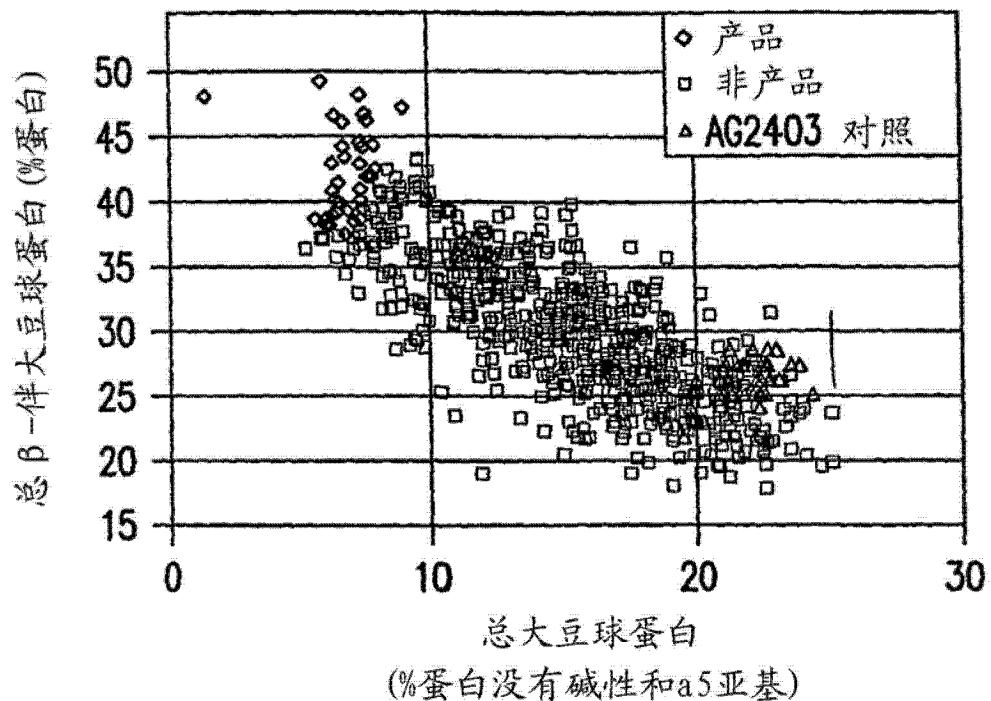
通过表型比较大豆球蛋白对 β -伴大豆球蛋白

图 7A

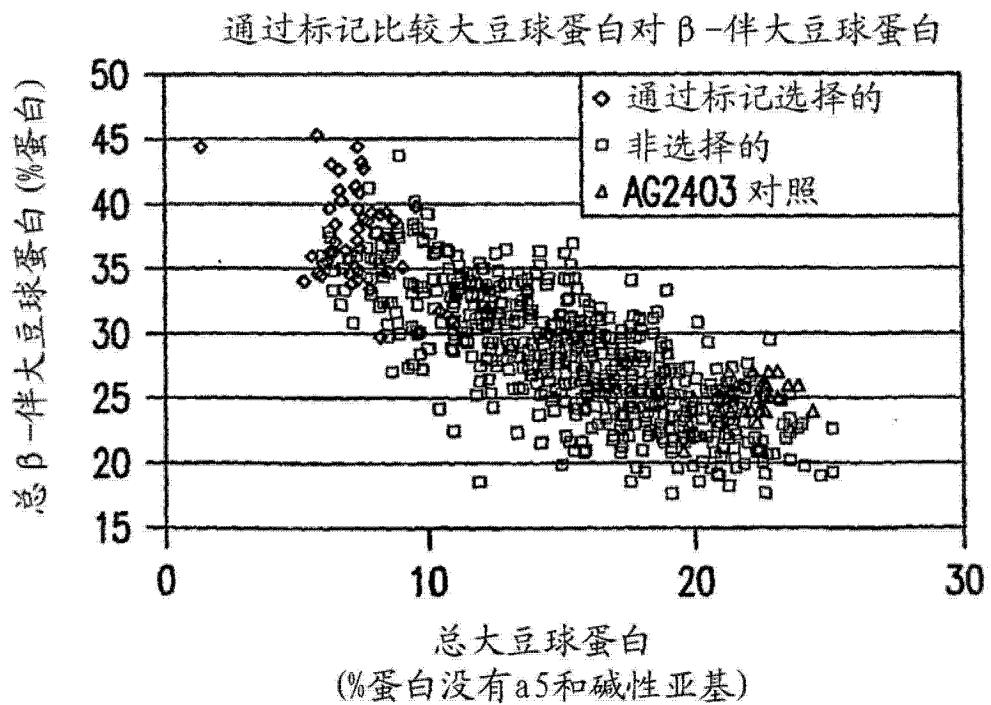


图 7B

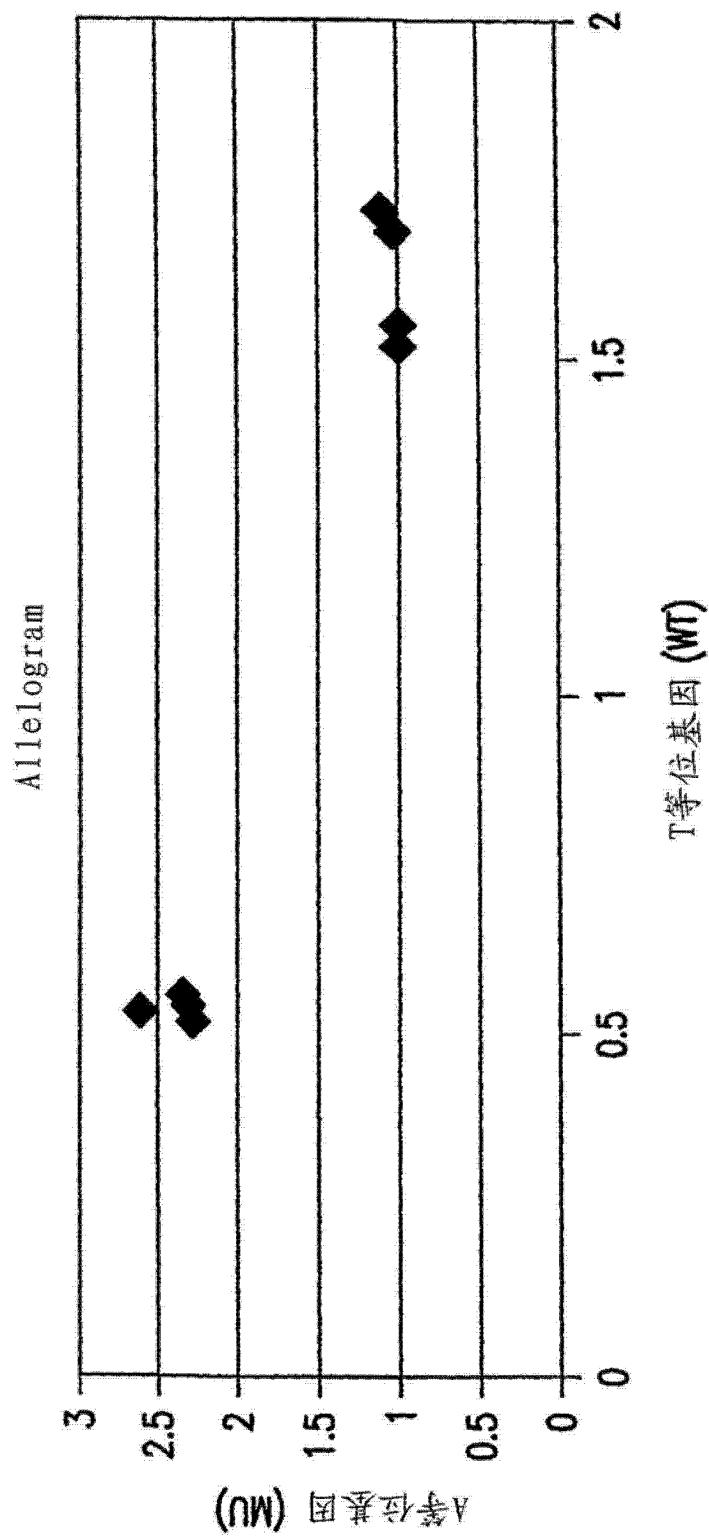


图 8

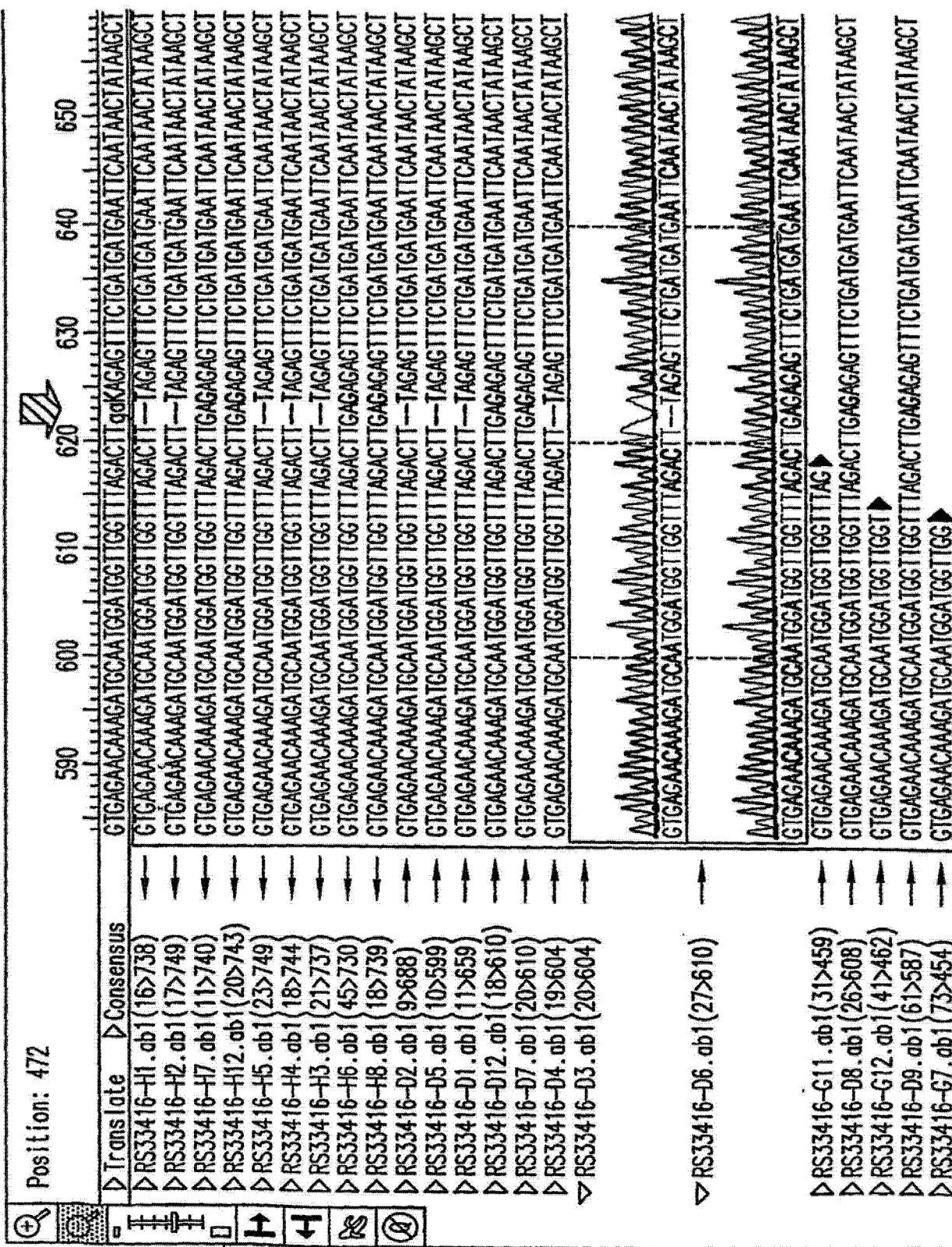


图 9