(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 114787364 A
(43)申请公布日 2022.07.22

(21)申请号 202080082810.X
(22)申请日 2020.11.20
(30)优先权数据
62/941,568 2019.11.27 US
63/020,197 2020.05.05 US
(85)PCT国际申请进入国家阶段日
2022.05.27
(86)PCT国际申请的申请数据
PCT/US2020/061473 2020.11.20
(87)PCT国际申请的公布数据
WO2021/108248 EN 2021.06.03
(71)申请人 卡里克特公司
地址 美国明尼苏达州
(72)发明人 扎卡里·德莫雷斯特
(74)专利代理机构 北京安信方达知识产权代理有限公司 11262
专利代理师 李敏春 郑霞
(51)Int.Cl.
C12N 15/82(2006.01)

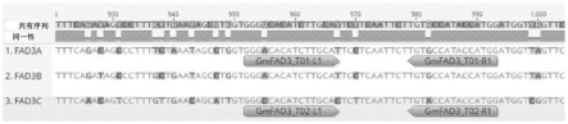
权利要求书4页 说明书18页
序列表25页 附图7页

(54)发明名称

用于基因编辑的TAL效应物核酸酶

(57)摘要

公开了TALEN组合物和使用方法，其包括使用多重组合物同时在若干基因诸如FAD3 A/B/C基因中产生靶突变，使用组合物在单个基因诸如编码FAD2蛋白的基因中产生靶突变，及其组合。所述组合物和方法可以提供与相应的未经改变的植物、植物部分或植物细胞相比具有改进的特性的经基因编辑的植物、植物部分和植物细胞。例如，还提供了大豆植物、植物部分和植物细胞，其能够产生含有油的种子，其与不存在靶突变的相应种子相比具有相对较高水平的油酸和较低水平的亚油酸和亚麻酸。



CN 114787364 A

1.一种组合物,包括:

第一核酸,其编码能够与第一靶基因的第一半位点序列结合的第一转录激活因子样(TAL)效应物核酸酶单体,以及

第二核酸组,每种第二核酸编码能够与所述第一靶基因或一组第二靶基因的第二半位点序列结合的第二TAL效应物核酸酶单体,

其中:

所述第一半位点序列是保守序列;

所述第一半位点序列和每种第二半位点序列不同,并且由间隔序列隔开;以及

所述第一TAL效应物核酸酶单体能够与每种所述第二TAL效应物核酸酶单体形成二聚体。

2.根据权利要求1所述的组合物,其中,当所述第一TAL效应物核酸酶单体与所述第一半位点序列结合并且所述第二TAL效应物核酸酶单体与所述第二半位点序列结合时,所述二聚体能够在活细胞内切割所述靶基因。

3.根据权利要求1或2所述的组合物,其中,所述第一半位点序列是100%保守序列。

4.根据权利要求1-3中任一项所述的组合物,其中,所述间隔序列的长度为约15个至约18个核苷酸。

5.根据权利要求1-4中任一项所述的组合物,其中,所述第一核酸包括FokI核酸内切酶结构域。

6.根据权利要求1-5中任一项所述的组合物,其中,每种第二核酸包括FokI核酸内切酶结构域。

7.根据权利要求1-6中任一项所述的组合物,其中,所述第一核酸在载体中。

8.根据权利要求1-7中任一项所述的组合物,其中,每种第二核酸在载体中。

9.根据权利要求1-8中任一项所述的组合物,其中,所述第一核酸和所述第二核酸组在单个载体中。

10.根据权利要求1-9中任一项所述的组合物,其中,所述第一核酸是质粒中的mRNA。

11.根据权利要求1-10中任一项所述的组合物,其中,每种第二核酸是质粒中的mRNA。

12.根据权利要求1-11中任一项所述的组合物,其中,所述第一核酸和所述第二核酸组是质粒中的mRNA。

13.根据权利要求1-12中任一项所述的组合物,其中,所述第二核酸组包括三种或更多种第二核酸。

14.根据权利要求1-13中任一项所述的组合物,其中,所述第二核酸组包括四种或更多种第二核酸。

15.根据权利要求1-14中任一项所述的组合物,其中,所述第一靶基因是大豆基因的FAD3家族的基因。

16.根据权利要求1-15中任一项所述的组合物,其中,每种第二靶基因是大豆基因的FAD3家族的基因。

17.根据权利要求1-16中任一项所述的组合物,其中,所述第一靶基因是大豆FAD3A基因的等位基因。

18.根据权利要求1-16中任一项所述的组合物,其中,所述第一靶基因是大豆FAD3B基

因的等位基因。

19.根据权利要求1-16中任一项所述的组合物,其中,所述第一靶基因是大豆FAD3C基因的等位基因。

20.根据权利要求17-19中任一项所述的组合物,其中,所述第二靶基因是大豆FAD3A基因的等位基因。

21.根据权利要求17-19中任一项所述的组合物,其中,所述第二靶基因是大豆FAD3B基因的等位基因。

22.根据权利要求17-19中任一项所述的组合物,其中,所述第二靶基因是大豆FAD3C基因的等位基因。

23.根据权利要求1-22中任一项所述的组合物,其中,所述第一半位点序列是SEQ ID NO:18。

24.根据权利要求1-23中任一项所述的组合物,还包括靶向大豆基因的FAD2-1家族的等位基因的一种或多种稀有切割核酸内切酶。

25.根据权利要求24所述的组合物,其中,所述一种或多种稀有切割核酸内切酶是靶向FAD2-1A或FAD2-1B的TAL效应物核酸酶。

26.根据权利要求25所述的组合物,其中,所述TAL效应物核酸酶包括与如SEQ ID NO:27-34中任一个所示的序列结合的单体。

27.根据权利要求25所述的组合物,其中,所述TAL效应物核酸酶包括与选自由以下组成的组的序列结合的单体对:SEQ ID NO:27和28;29和30;31和32;以及33和34。

28.根据权利要求1-15中任一项所述的组合物,其中,每种第二靶基因是大豆基因的FAD2家族的基因。

29.根据权利要求1-15中任一项所述的组合物,其中,所述第一靶基因是大豆FAD2-1A基因的等位基因。

30.根据权利要求1-15中任一项所述的组合物,其中,所述第一靶基因是大豆FAD2-1B基因的等位基因。

31.根据权利要求28-30中任一项所述的组合物,其中,所述第二靶基因是大豆FAD2-1A基因的等位基因。

32.根据权利要求28-30中任一项所述的组合物,其中,所述第二靶基因是大豆FAD2-1B基因的等位基因。

33.根据权利要求1-15中任一项所述的组合物,其中,所述第一半位点序列在SEQ ID NO:25或26内。

34.根据权利要求1-33中任一项所述的组合物,其中,所述组合物包括两种第二核酸,每种所述第二核酸编码能够结合第二半位点序列的第二转录激活因子样(TAL)效应物核酸酶单体,其中所述第二半位点序列是SEQ ID NO:17或SEQ ID NO:19。

35.一种将突变同时引入两种或更多种基因的方法,包括使包括所述两种或更多种基因的细胞群与权利要求1-34中任一项所述的组合物接触。

36.一种植物、植物部分或植物细胞,通过权利要求35所述的方法获得。

37.根据权利要求36所述的植物、植物部分或植物细胞,其中,所述植物、植物部分或植物细胞是大豆植物、植物部分或植物细胞。

38.根据权利要求37所述的植物、植物部分或植物细胞，其中，所述大豆植物部分或植物细胞选自由子叶细胞、种子、胚、胚性愈伤组织细胞和花粉细胞组成的组。

39.一种大豆油组合物，包括由权利要求37所述的大豆植物、植物部分或植物细胞产生的大豆油，其中，与由不存在两种或更多种基因中的突变的相应大豆植物、植物部分或植物细胞产生的油相比，所述大豆油具有增加的油酸含量、减少的亚油酸含量和减少的亚麻酸含量中的一个或多个。

40.一种大豆植物、植物部分或植物细胞，包括降低FAD2-1A基因和FAD2-1B基因中的至少一种的表达的一种或多种突变，其中，所述植物、植物部分或植物细胞产生与由不存在所述一种或多种突变的相应大豆植物、植物部分或植物细胞产生的油相比油酸含量增加的油，并且其中至少一种突变由稀有切割核酸内切酶诱导，所述稀有切割核酸内切酶能够结合来自SEQ ID NO:27-34所示的组的核酸序列，或其功能变体。

41.一种用于生成包括降低FAD2-1A基因和FAD2-1B基因中至少一种的表达的突变的大豆植物的方法，包括：

(a)使来自具有功能性FAD2-1A基因和FAD2-1B基因的大豆植物的大豆植物细胞群与编码稀有切割核酸内切酶的一个或多个核酸序列接触，所述稀有切割核酸内切酶能够结合来自SEQ ID NO:27-34所示的组的核酸序列，或其功能变体；

(b)从所述群中选择其中所述FAD2-1A基因或FAD2-1B基因的表达降低的细胞，以及

(c)将所选择的植物细胞再生成为大豆植物。

42.一种大豆油组合物，包括由大豆植物、植物部分或植物细胞产生的大豆油，所述大豆植物、植物部分或植物细胞包括降低FAD2-1A基因和FAD2-1B基因中的至少一种的表达的一种或多种突变，其中，与由不存在所述一种或多种突变的相应大豆植物、植物部分或植物细胞产生的油相比，所述大豆油具有增加的油酸含量、减少的亚油酸含量和减少的亚麻酸含量中的一个或多个；并且其中所述一种或多种突变包括由稀有切割核酸内切酶诱导的靶突变，所述稀有切割核酸内切酶能够结合来自SEQ ID NO:27-34所示的组的核酸序列，或其功能变体。

43.一种大豆植物、植物部分或植物细胞，包括以下中的第一组突变：

一种或多种FAD3A等位基因和一种或多种FAD3B等位基因，

一种或多种FAD3A等位基因和一种或多种FAD3C等位基因，

一种或多种FAD3B等位基因和一种或多种FAD3C等位基因，或者

一种或多种FAD3A等位基因、一种或多种FAD3B等位基因和一种或多种FAD3C等位基因，

其中，所述第一组突变由以下的表达诱导：

第一核酸，其编码能够与第一靶基因的第一半位点序列结合的第一转录激活因子样(TAL)效应物核酸酶单体，以及

第二核酸组，每种第二核酸编码能够与所述第一靶基因和至少一种第二靶基因的第二半位点序列结合的第二TAL效应物核酸酶单体，

其中：

所述第一半位点序列是保守序列；

所述第一半位点序列和每种第二半位点序列不同，并且由间隔序列隔开；以及

所述第一TAL效应物核酸酶单体能够与每种所述第二TAL效应物核酸酶单体形成二聚

体;以及

以下中的突变:

一种或多种FAD2-1A等位基因,

一种或多种FAD2-1B等位基因,或者

一种或多种FAD2-1A等位基因和一种或多种FAD2-1B等位基因,

其中所述植物、植物部分或植物细胞产生与由不存在所述突变的相应大豆植物、植物部分或植物细胞产生的油相比具有减少的亚麻酸含量、增加的油酸含量和减少的亚油酸含量的油。

44.根据权利要求43所述的大豆植物、植物部分或植物细胞,其中,所述植物、植物部分或植物细胞不包含转基因。

45.根据权利要求43所述的大豆植物、植物部分或植物细胞,其中,所述植物部分是种子。

46.根据权利要求43所述的大豆植物、植物部分或植物细胞,其中,对所述一种或多种FAD2-1A等位基因和所述一种或多种FAD2-1B等位基因的突变是由稀有切割核酸内切酶诱导的。

47.根据权利要求46所述的大豆植物、植物部分或植物细胞,其中,所述稀有切割核酸内切酶是TAL效应物核酸酶,并且其中所述TAL效应物核酸酶与SEQ ID NO:27-34中任一个所示的序列结合。

48.根据权利要求43所述的大豆植物、植物部分或植物细胞,其中,一种或多种FAD3A等位基因、一种或多种FAD3B等位基因和一种或多种FAD3C等位基因、一种或多种FAD2-1A等位基因和一种或多种FAD2-1B等位基因是突变的。

49.一种用于生成包括降低FAD3A基因、FAD3B基因和FAD3C基因中至少两种的表达的突变的大豆植物的方法,包括:

(a)使来自具有功能性FAD3A基因、FAD3B基因和FAD3C基因的大豆植物的大豆植物细胞群与包括第一核酸和第二核酸组的组合物接触,所述第一核酸编码能够与第一靶基因的第一半位点序列结合的第一转录激活因子样(TAL)效应物核酸酶单体,每种第二核酸编码能够与所述第一靶基因和至少一种第二靶基因的第二半位点序列结合的第二TAL效应物核酸酶单体,其中,所述第一半位点序列是保守序列;所述第一半位点序列和每种第二半位点序列不同,并且由间隔序列隔开;并且所述第一TAL效应物核酸酶单体能够与每种所述第二TAL效应物核酸酶单体形成二聚体,其中所述第一靶基因是FAD3A基因,所述至少一种第二靶基因选自FAD3B基因和FAD3C基因;

(b)从所述群中选择其中所述第一靶基因和所述至少一种第二靶基因的表达降低的细胞,以及

(c)将所选择的植物细胞再生成为大豆植物。

50.根据权利要求49所述的大豆植物、植物部分或植物细胞,其中,所述第一单体或第二单体能够结合来自SEQ ID NO:4-19所示的组的核酸序列,或其功能变体。

用于基因编辑的TAL效应物核酸酶

[0001] 序列表

[0002] 本申请包含已以ASCII格式电子提交的序列表，该序列表通过援引整体并入本文。所述ASCII副本创建于2020年11月19日，命名为1702_029PCT1_SL.txt，大小为47,037字节。

背景技术

[0003] 利用稀有切割核酸内切酶(rare-cutting endonuclease)进行的基因编辑可用于产生缺失、插入和启动同源重组。转录激活因子样效应物核酸酶(TALEN)是一种稀有切割核酸内切酶，其可以在DNA中靶向特定序列并产生精准断裂。一个TALEN设计过程的基础是识别约15个碱基对(bp)的左半TALEN(HT)，接着约15-18bp的间隔区和约15bp的右HT识别序列(参见，例如美国专利8,450,471)。该方法针对预期靶标提供高水平的特异性；然而，多重基因编辑能力将是有利的。

发明内容

[0004] 本公开的方面涉及使用稀有切割核酸内切酶例如TALEN来编辑基因(例如，通过引入突变)的组合物和方法。

[0005] 在一个方面，本文提供了一种组合物，其包含：第一核酸，其编码能够与第一靶基因的第一半位点序列结合的第一转录激活因子样(TAL)效应物核酸酶单体；和第二核酸组。每种第二核酸编码能够与所述第一靶基因或一组第二靶基因的第二半位点序列结合的第二转录激活因子样(TAL)效应物核酸酶单体。在该组合物中，第一半位点序列是保守序列，所述第一半位点序列和每种第二半位点序列不同并且由间隔序列隔开，以及所述第一TAL效应物核酸酶单体能够与每种第二TAL效应物核酸酶形成二聚体，从而提供TALEN的组。

[0006] 在一些实施方式中，当所述第一TAL效应物核酸酶单体与所述第一半位点序列结合并且所述第二TAL效应物核酸酶单体与所述第二半位点序列结合时，所述二聚体可以在活细胞内切割所述靶基因。

[0007] 在一些实施方式中，第一半位点序列是100％保守序列。在一些实施方式中，间隔序列的长度为约15至约18个核苷酸。

[0008] 在一些实施方式中，第一核酸包含FokI核酸内切酶结构域。在一些实施方式中，每种第二核酸包含FokI核酸内切酶结构域。

[0009] 在一些实施方式中，第一核酸在载体中。在一些实施方式中，每种第二核酸在载体中。在一些实施方式中，第一核酸和编码第二TAL效应物核酸酶单体的第二核酸组在单个载体中。

[0010] 在一些实施方式中，第一核酸是质粒中的mRNA。在一些实施方式中，每种第二核酸是质粒中的mRNA。在一些实施方式中，第一核酸和第二核酸组是质粒中的mRNA。

[0011] 在一些实施方式中，第二核酸组包括三种或更多种第二核酸。在一些实施方式中，第二核酸组包括四种或更多种第二核酸。

[0012] 在一些实施方式中，第一靶基因是大豆(Glycine max)基因的FAD3家族的基因。在

一些实施方式中，每种第二靶基因是大豆基因的FAD3家族的基因。

[0013] 在一些实施方式中，第一靶基因是大豆FAD3A基因的等位基因。在一些实施方式中，第一靶基因是大豆FAD3B基因的等位基因。在一些实施方式中，第一靶基因是大豆FAD3C基因的等位基因。在一些实施方式中，第二靶基因是大豆FAD3A基因的等位基因。在一些实施方式中，第二靶基因是大豆FAD3B基因的等位基因。在一些实施方式中，第二靶基因是大豆FAD3C基因的等位基因。

[0014] 在一些实施方式中，第一半位点序列是SEQ ID NO:18。在一些实施方式中，其中该组合物包含两种第二核酸，每种第二核酸编码能够结合第二半位点序列的第二转录激活因子样(TAL)效应物核酸酶单体，其中第二半位点序列是SEQ ID NO:SEQ ID NO:17或SEQ ID NO:19。

[0015] 在一些实施方式中，该组合物进一步包括靶向大豆基因FAD2-1家族的等位基因的一种或多种稀有切割核酸内切酶。一种或多种稀有切割核酸内切酶可以是靶向FAD2-1A或FAD2-1B的TAL效应物核酸酶。TAL效应物核酸酶可以包括与如SEQ ID NO:27-34中任一个所示的序列结合的单体。TAL效应物核酸酶可以包括选自由以下组成的组的单体对:SEQ ID NO:27和28;29和30;31和32;以及33和34。

[0016] 在一些实施方式中，每种第二靶基因是大豆基因的FAD2家族的基因。在一些实施方式中，第一靶基因是大豆FAD2-1A基因的等位基因。在一些实施方式中，第一靶基因是大豆FAD2-1B基因的等位基因。在一些实施方式中，第二靶基因是大豆FAD2-1A基因的等位基因。在一些实施方式中，第二靶基因是大豆FAD2-1B基因的等位基因。

[0017] 在一些实施方式中，第一半位点序列在SEQ ID NO:25或26内。

[0018] 在一些实施方式中，该组合物包含两种第二核酸，每种第二核酸编码能够结合第二半位点序列的第二转录激活因子样(TAL)效应物核酸酶单体，其中第二半位点序列是SEQ ID NO:17或SEQ ID NO:19。

[0019] 在另一方面，本公开的特征在于将突变同时引入两种或更多种基因的方法，包括使具有所述两种或更多种基因的细胞群与根据上述一种或多种实施方式的组合物接触。

[0020] 在另一方面，本公开的特征在于通过上述方法获得的植物、植物部分或植物细胞。所述植物、植物部分或植物细胞可以是大豆(soybean)植物、植物部分或植物细胞。大豆植物部分或植物细胞可选自由子叶细胞、种子、胚、胚性愈伤组织细胞和花粉细胞组成的组。

[0021] 在另一方面，本公开的特征在于大豆油组合物，其包含由上述实施方式中描述的大豆植物、植物部分或植物细胞产生的大豆油，其中与由不存在两种或更多种基因中的突变的相应大豆植物、植物部分或植物细胞产生的油相比，所述大豆油具有增加的油酸含量、减少的亚油酸含量和减少的亚麻酸含量中的一个或多个。

[0022] 在另一方面，本公开的特征在于大豆植物、植物部分或植物细胞，其具有降低FAD2-1A基因和FAD2-1B基因中至少一种的表达的一种或多种突变，其中所述植物、植物部分或植物细胞产生与由不存在所述一种或多种突变的相应大豆植物、植物部分或植物细胞产生的油相比油酸含量增加的油，并且其中至少一种突变由稀有切割核酸内切酶诱导，所述稀有切割核酸内切酶能够结合来自SEQ ID NO:27-34中所示的组的核酸序列，或其功能变体。

[0023] 在另一方面，本公开的特征在于用于产生具有降低FAD2-1A基因和FAD2-1B基因中

至少一种的表达的突变的大豆植物的方法，该方法包括：(a)使来自具有功能性FAD2-1A基因和FAD2-1B基因的大豆植物的大豆植物细胞群与编码稀有切割核酸内切酶的一个或多个核酸序列接触，所述稀有切割核酸内切酶能够结合来自SEQ ID NO:27-34中所示的组的核酸序列，或其功能变体；(b)从所述群中选择其中所述FAD2-1A基因或FAD2-1B基因的表达降低的细胞，和(c)将所选择的植物细胞再生为大豆植物。

[0024] 在另一方面，本公开的特征在于大豆油组合物，其包括由大豆植物、植物部分或植物细胞产生的大豆油，所述大豆植物、植物部分或植物细胞包含降低FAD2-1A基因和FAD2-1B基因中的至少一种的表达的一种或多种突变，其中与由不存在所述一种或多种突变的相应大豆植物、植物部分或植物细胞产生的油相比，所述大豆油具有增加的油酸含量、减少的亚油酸含量和减少的亚麻酸含量中的一个或多个；并且其中所述一种或多种突变包括由稀有切割核酸内切酶诱导的靶突变，所述稀有切割核酸内切酶能够结合来自SEQ ID NO:27-34中所示的组的核酸序列，或其功能变体。

[0025] 在另一方面，本公开的特征在于大豆植物、植物部分或植物细胞，其包含以下中的第一组突变：一种或多种FAD3A等位基因和一种或多种FAD3B等位基因、一种或多种FAD3A等位基因和一种或多种FAD3C等位基因、一种或多种FAD3B等位基因和一种或多种FAD3C等位基因或者一种或多种FAD3A等位基因、一种或多种FAD3B等位基因和一种或多种FAD3C等位基因，其中所述第一组突变由以下的表达诱导：第一核酸，其编码能够与第一靶基因的第一半位点序列结合的第一转录激活因子样(TAL)效应物核酸酶单体；和第二核酸组，每种第二核酸编码能够与所述第一靶基因和至少一种第二靶基因的第二半位点序列结合的第二转录激活因子样(TAL)效应物核酸酶单体，其中所述第一半位点序列是保守序列；所述第一半位点序列和每种第二半位点序列不同，并且由间隔序列隔开；并且所述第一TAL效应物核酸酶单体能够与每种第二TAL效应物核酸酶形成二聚体；和以下中的突变：一种或多种FAD2-1A等位基因、一种或多种FAD2-1B等位基因或者一种或多种FAD2-1A等位基因和一种或多种FAD2-1B等位基因，其中植物、植物部分或植物细胞产生与由不存在突变的相应大豆植物、植物部分或植物细胞产生的油相比具有减少的亚麻酸含量、增加的油酸含量和减少的亚油酸含量的油。所述植物、植物部分或植物细胞可以不存在转基因。所述植物部分可以是种子。对所述一种或多种FAD2-1A等位基因和所述一种或多种FAD2-1B等位基因的突变可以是由稀有切割核酸内切酶诱导的。所述稀有切割核酸内切酶可以是TAL效应物核酸酶。所述TAL效应物核酸酶可以与SEQ ID NO:27-34中任一个所示的序列结合。所述一种或多种FAD3A等位基因、一种或多种FAD3B等位基因和一种或多种FAD3C等位基因、一种或多种FAD2-1A等位基因和一种或多种FAD2-1B等位基因可以是突变的。

[0026] 在另一方面，本公开的特征在于用于产生包含降低FAD3A基因、FAD3B基因和FAD3C基因中至少两种的表达的突变的大豆植物的方法，该方法包括：(a)使来自具有功能性FAD3A基因、FAD3B基因和FAD3C基因的大豆植物的大豆植物细胞群与包含第一核酸和第二核酸组的组合物接触，所述第一核酸编码能够与第一靶基因的第一半位点序列结合的第一转录激活因子样(TAL)效应物核酸酶单体，每种第二核酸编码能够与所述第一靶基因和至少一种第二靶基因的第二半位点序列结合的第二转录激活因子样(TAL)效应物核酸酶单体，其中所述第一半位点序列是保守序列；所述第一半位点序列和每种第二半位点序列不同，并且由间隔序列隔开；以及所述第一TAL效应物核酸酶单体能够与每种第二TAL效应物

核酸酶形成二聚体，其中所述第一靶基因是FAD3A基因，所述至少一种第二靶基因选自FAD3B基因和FAD3C基因；(b)从所述群中选择其中所述第一靶基因和所述至少一种第二靶基因的表达降低的细胞，和(c)将所选择的植物细胞再生为大豆植物。第一或第二单体能够结合来自SEQ ID NO:4-19中所示的组的核酸序列，或其功能变体。

附图说明

[0027] 本发明的上述方面和许多伴随的优点将通过参考结合附图进行的如下详细描述而变得更好理解，在附图中：

[0028] 图1A-D示出根据本公开的一种或多种实施方式靶向大豆基因的FAD3家族的示例性多重TALEN的对齐：(A)GmFAD3_T01-L1、GmFAD3_T01-R1、GmFAD3_T02-L1和GmFAD3_T02-R1(分别为SEQ ID NO:45-48，按出现顺序)的对齐；(B)GmFAD3_T03-L1、GmFAD3_T03-R1、GmFAD3_T04-L1和GmFAD3_T04-R1(分别为SEQ ID NO:49-52，按出现顺序)的对齐；(C)GmFAD3_T05-L1、GmFAD3_T05-R1、GmFAD3_T06-L1和GmFAD3_T07-L1(分别为SEQ ID NO:53-56，按出现顺序)的对齐；以及(D)GmFAD3_T08-L1、GmFAD3_T08-R1、GmFAD3_T09-L1和GmFAD3_T09-R1(分别为SEQ ID NO:57-60，按出现顺序)的对齐。

[0029] 图2A、B、B-1和B-2示出根据本公开的一种或多种实施方式靶向大豆基因的FAD2家族的示例性TALEN的对齐：(A)GmFAD2_T01-L1和GmFAD2_T01-R1(分别为SEQ ID NO:61-62，按出现顺序)的对齐和(B)GmFAD2_T02-L1、GmFAD2_T02-R1、GmFAD2_T03-L1、GmFAD2_T03-R1、GmFAD2_T04-L1、GmFAD2_T04-R1、GmFAD2_T05-L1和GmFAD2_T05-R1(分别为SEQ ID NO:63-65，按出现顺序)的对齐。视图B的放大部分显示为视图B-1和B-2，用于序列的放大。图2B-1公开了SEQ ID NO:63-65的残基1-44，以及图2B-2公开了SEQ ID NO:63-65的残基45-103，全部分别按出现顺序。

具体实施方式

[0030] 本文提供了用于使用专门改造的稀有切割核酸内切酶技术靶向单个基因序列或同时靶向多个相似的基因序列的组合物和方法，例如，作为基因家族诸如植物的FAD3基因家族的一部分的基因。

[0031] 稀有切割核酸内切酶可以是TAL效应物核酸酶、兆核酸酶、经改造的归巢核酸内切酶、锌指核酸酶(ZFN)或聚簇有规律间隔的短回文重复序列(CRISPR)和CRISPR相关(Cas)系统(CRISPR/Cas9)。稀有切割核酸内切酶可以是天然的或经改造的蛋白质，其具有针对具有长度为约12-40个碱基对(bp)或更长(例如长度为14-40、15-36或16-32；参见，例如，Baker，Nature Methods 9:23-26,2012)的识别序列(靶序列)的核酸序列的核酸内切酶活性。典型的稀有切割核酸内切酶在其识别位点内引起切割，留下4个带有3'OH或5'OH突出端的核苷酸(nt)交错切割。在一些实施方式中，稀有切割核酸内切酶是兆核酸酶，诸如野生型或变体归巢核酸内切酶(例如，属于十二肽家族的归巢核酸内切酶(参见WO 2004/067736)。

[0032] 稀有切割核酸内切酶可以基于来自II型原核CRISPR(聚簇有规律间隔的短回文重复序列)适应性免疫系统的RNA引导的Cas9核酸酶。该系统允许切割侧接短序列基序的DNA序列，所述短序列基序称为原间隔相邻基序(PAM)。切割是通过改造与靶序列互补的特定crRNA实现的。Cas9核酸内切酶将与CRISPR RNA(crRNA)形成复合物。双重反式激活crRNA

(tracrRNA):crRNA结构充当引导RNA(gRNA),其将Cas9核酸内切酶引导至靶序列。在FAD2-1基因家族或FAD3基因家族中发现的序列中的PAM基序允许设计在异源表达Cas9核酸内切酶并用crRNA转染的植物细胞中特异性引入突变的或使一个或多个靶基因失活的crRNA。

[0033] 在一些实施方式中,稀有切割核酸内切酶是包含DNA结合结构域和具有切割活性的催化结构域的融合蛋白。TALE-核酸酶和ZFN是DNA结合结构域与核酸内切酶FokI的催化结构域融合的实例。定制的TALE-核酸酶以商品名TALEN™(Cellectis,巴黎,法国)市售。转录激活因子样(TAL)效应物的特异性取决于效应物可变重复序列。多态性主要存在于称为重复可变双残基(RVD)的重复位置12和13。TAL效应物的RVD以直接、线性的方式对应于其靶位点中的核苷酸,一个RVD对应一个核苷酸,具有一定的简并性,没有明显的环境依赖性。该蛋白质-DNA识别机制使得能够进行针对新靶标特异性TAL效应物的靶位点预测,以及对所选择位点具有结合特异性的新TAL效应物的靶位点选择和改造。

[0034] TAL效应物DNA结合结构域可以与其他序列诸如核酸内切酶序列融合,使得嵌合核酸内切酶靶向特定的所选择DNA序列,并导致随后在靶序列处或附近切割DNA。例如,在DNA中这样的切割(双链断裂)可以通过NHEJ或同源重组诱导突变成野生型DNA序列。TALE-核酸酶可用于促进复杂基因组中的定点诱变,以高精确度和高效率敲除或以其他方式改变基因功能。

[0035] 用于选择内源靶序列和产生靶向这样的序列的TALE-核酸酶的方法可以如本领域中所述进行。参见例如PCT公开No.WO 2011/072246,其通过援引并入。在一些实施方式中,可以使用特异性识别TALE-核酸酶识别位点的软件。

[0036] 一些核酸内切酶(例如,FokI)作为二聚体起作用,从而增强靶向特异性。当两个TALE核酸酶识别位点接近时,失活的单体可以聚在一起产生切割DNA的功能性酶。通过需要DNA结合来激活核酸酶,可以产生高度位点特异性的限制酶。

[0037] 本公开的TALEN包括能够靶向多种基因的TALEN,所述基因包含共有的序列例如保守序列。本公开的TALEN包含共有的第一TAL效应物核酸酶单体(例如,左半TALEN(左-HT)或右半TALEN或(右-HT)),和第二TAL效应物核酸酶单体组(例如,两种、三种、四种或更多种第二HT),其中共有HT能够与任意第二HT形成二聚体以使得能够靶向所有预期的序列。如本文所用,术语“TAL效应物核酸酶单体”、“半TALEN”和“HT”可互换使用。这些第一TAL效应物核酸酶单体和第二TAL效应物核酸酶单体组可与改造为靶向单个序列的TALEN组合使用,所述单个序列诸如FAD2-1基因(例如,FAD2-1A和FAD2-1B中的至少一种)的序列。

[0038] 所公开的组合物和方法的使用涉及目标基因序列的对齐以确定序列之间接近保守同一性的区域。在一些实施方式中,共有的第一TAL效应物核酸酶单体(例如,左-HT或右-HT)靶向在本公开的组合物和方法靶向的所有基因中约100%保守的、约95%保守的、约90%保守的、约85%保守的或约80%保守的序列。

[0039] 例如,在一种实施方式中,针对左-HT结合结构域(半位点序列)选择100%保守的靶区域。针对右-HT结合结构域,将每个非保守序列的唯一HT设计为使得能够实现靶向。在该非限制性实施例中,左-HT#1和右-HT#1靶向基因1,左-HT#1和右-HT#2靶向基因2,左-HT#和右-HT#3靶向基因3,依此类推。

[0040] 同样,例如,在另一实施方式中,针对右-HT结合结构域,选择100%保守的靶区域。针对左-HT结合结构域,将每个非保守序列的唯一HT设计为使得能够实现靶向。在该非限制

性实例中，左-HT#1和右-HT#1靶向基因1，左-HT#2和右-HT#1靶向基因2，左-HT#3和右-HT#1靶向基因3，依此类推。

[0041] 在一些实施方式中，本公开的组合物和方法靶向的基因可以编码相同的蛋白质或不同的蛋白质(例如，共有的HT靶向在所有基因中保守的或相同的序列)。在一些实施方式中，本公开的组合物和方法靶向的基因是基因的等位基因。

[0042] 在一些实施方式中，靶基因是FAD3A、FAD3B或FAD3C基因。在一些实施方式中，靶基因是SEQ ID NO:1-3。在一些实施方式中，本文公开的方法和组合物可以靶向FAD3A、FAD3B和FAD3C基因中的任意两个，例如引入突变。在一些实施方式中，本文公开的方法和组合物可以靶向FAD3A、FAD3B和FAD3C基因中的全部三个，例如引入突变。一种或多种突变可以存在于FAD3A、FAD3B或FAD3C基因的编码或非编码序列中。天然大豆基因的基因组序列可见于Soybase数据库(www.soybase.org)。在一些实施方式中，本文公开的组合物和方法包括能够结合选自SEQ ID NO:4-19的序列或SEQ ID NO:1-3中任意一个内的序列的转录激活因子样(TAL)效应物核酸酶单体。

[0043] 在一些实施方式中，本文公开的方法和组合物可以靶向FAD2-1A或FAD2-1B基因中的任意一个，例如引入突变。在一些实施方式中，靶基因是FAD2-1A或FAD2-1B基因。一种或多种突变可以存在于FAD2-1A或FAD2-1B基因的编码或非编码序列中。大豆FAD2-1A和FAD2-IB基因的内源示例性基因组序列分别示于SEQ ID NO:23和24中。在一些实施方式中，靶基因是SEQ ID NO:23或24，或其功能变体，所述功能变体与SEQ ID NO:23或24具有至少约80％、诸如至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％或至少约99％的序列同一性。有时，靶序列在FAD2-1A或FAD2-IB基因的编码序列内，诸如SEQ ID NO:25或26中所示的编码序列内。靶序列可以是SEQ ID NO:25或26内的序列，或者其功能变体，所述功能变体与SEQ ID:25或26具有至少约80％、诸如至少约85％、至少约90％、至少约95％、至少约96％、至少约97％、至少约98％或至少约99％的序列同一性。

[0044] 任何核酸序列与序列识别号(SEQ ID NO:)所引用的序列之间的序列同一性百分比可以通过常规方法来测定。在一个实施例中，使用来自包含BLASTN版本2.0.14的独立版本的BLASTZ的BLAST 2Sequences(B12seq)程序将核酸序列与序列识别号所示的序列进行比较。如果所比较的两条序列具有同源性，则指定的输出文件会将那些同源的区域显示为对齐的序列。如果所比较的两条序列不具有同源性，则指定的输出文件将不显示对齐的序列。一旦对齐，则通过对两条序列中显示的相同核苷酸残基的位置的数目进行计数来确定匹配的数目。通过如下所述确定序列同一性百分比：将匹配数目除以所识别的序列(例如，SEQ ID NO:1)中示出的序列的长度，或除以铰接长度(例如，来自所识别的序列中示出的序列的100个连续核苷酸)，然后将结果值乘以100。将序列同一性百分比值四舍五入到最接近的十分之一。例如，靶基因可以具有与SEQ ID NO:20、21、22、25或26具有至少80％、至少85％、至少90％、至少91％、至少92％、至少93％、至少94％、至少95％、至少96％、至少97％、至少98％或至少99％的序列同一性的编码序列。

[0045] 在一些实施方式中，本文公开的组合物和方法包括能够结合FAD2-1A和FAD2-1B基因中至少一种的基因组或编码序列内的序列的转录激活因子样(TAL)效应物核酸酶单体。例如，TALEN单体可以选自SEQ ID NO:27-34中之一，选自由以下组成的组的单体对：SEQ ID NO:27和28；29和30；31和32；33和34；以及35和36，或者可以包括这些序列的功能变体。功能

变体包括例如具有一个或多个核苷酸替换、缺失或插入但保留所需活性的序列。功能变体可以通过本领域技术人员可用的多种方法中的任意一种来产生，诸如通过定点诱变、诱导突变、鉴定为等位基因变体、通过使用限制酶进行切割等。

[0046] 本文公开的组合物包括本公开的TALEN和编码本公开的TALEN的核酸。因此，在一些实施方式中，本文提供了一种组合物，该组合物包含第一核酸和第二核酸组，该第一核酸编码靶向由该组合物靶向的所有基因中保守的共有序列的HT，该第二核酸组编码多个靶向第二序列的第二HT，其中靶向共有序列的HT能够与每种第二HT形成二聚体。

[0047] TALEN基因编辑基于对含有DNA识别结构域和在质粒中编码的FokI核酸内切酶结构域的两个半TALEN蛋白的基因特异性靶向，所述质粒包括适当的启动子、终止子和在受体细胞中表达所必需的其他非编码序列。FokI仅在二聚体组合中具有活性，并且需要两个半TALEN才能实现TALEN基因编辑所需的核酸内切酶DNA切割。

[0048] 本公开的组合物可以被递送至真核细胞、哺乳动物细胞、植物细胞或原核细胞。可以以本领域技术人员已知的多种方式完成向细胞中引入。引入基因编辑TALEN的常用方法包括用携带包括TALEN基因盒的质粒的农杆菌属(Agrobacterium spp.)转化、用编码TALEN序列盒的质粒或mRNA进行基因枪轰击以及用编码TALEN序列盒的质粒或mRNA进行PEG介导的转化。在一些实施方式中，本公开的组合物可用于产生由组合物编码的半TALEN蛋白组。可以通过多种方法分离半TALEN蛋白并随后将其引入靶细胞诸如植物细胞中，所述方法例如包括PEG介导的转染，诸如在例如Luo S.等人Non-transgenic Plant Genome Editing Using Purified Sequence-Specific Nucleases,Mol Plant.2015Sep;8(9):1425-7中描述的技术。

[0049] 植物细胞可以来自大豆植物。术语“大豆植物”或“植物部分”广泛用于包括处于任何发育阶段的大豆植物，或大豆植物的一部分，包括植物插条、植物细胞、植物细胞培养物、植物器官、植物种子和秧苗。植物细胞是植物的结构和生理单元，包括原生质体和细胞壁。植物细胞可以是分离的单个细胞或细胞聚集体的形式，例如脆弱的愈伤组织或培养细胞，或者可以是更高组织单元例如植物组织、植物器官或植物的一部分。因此，植物细胞可以是原生质体、产生配子的细胞，或者可以再生成整株植物的细胞或细胞集合。因此，包含多个植物细胞并且能够再生成完整植物的种子被认为是用于本公开目的的植物细胞。植物组织或植物器官可以是种子、原生质体、愈伤组织或组织成结构或功能单元的任何其他植物细胞的组。植物的特别有用的部分包括可收获的部分和可用于后代植物繁殖的部分。植物的可收获部分可以是植物的任何有用部分，例如花、花粉、幼苗、叶、茎、种子荚、种子、根、结节等。可用于繁殖的植物部分包括例如种子、种子荚、插条、幼苗、根茎等。“种子”是指植物胚珠在其开花的正常成熟点后继续分化而形成的任何植物结构，无论其是否在受精情况下形成，也无论种子结构是否受精。

[0050] 稀有切割核酸内切酶(例如，TALEN)可以由核酸序列编码用于在植物细胞中表达或作为蛋白质引入。可以通过如下所述将编码序列特异性核酸酶的核酸引入大豆植物：农杆菌介导的用编码稀有切割核酸内切酶的T-DNA转化植物部分或植物细胞(例如，叶、茎、叶柄、中间外植体、愈伤组织或原生质体)、用编码稀有切割核酸内切酶的一种或多种核酸对植物部分或植物细胞进行基因枪转化、细胞穿透肽介导的转化和/或聚乙二醇(PEG)介导的转化。例如，可以从表面灭菌的叶子中分离原生质体，并在PEG存在下用编码一种或多种稀

有切割核酸内切酶的质粒转化。可以通过递送可检测的标志物(例如YFP质粒)来监测转化效率,可以使用荧光显微镜或流式细胞仪对可检测的标志物进行可视化。在PEG介导的转化后,可以使用原生质体培养领域的普通技术人员已知的方法和培养基培养原生质体。经过适当长的培养时间后,鉴定为突变体的源自原生质体的愈伤组织可以生长,转移到芽诱导培养基中,然后(一旦形成根)转移到土壤中并生长至成熟用于种子生产。

[0051] 可以使用瞬时表达方法或者通过使用稳定整合到宿主基因组中的方法将稀有切割核酸内切酶递送至大豆植物。为了瞬时递送序列特异性核酸酶,可以将转化的大豆植物部分或植物细胞(例如,使用上述方法)置于不含选择剂的再生培养基上,并且可以再生大豆植物。可以筛选再生植物以鉴定含有诱导突变的那些。为了将基因组改造试剂稳定地整合到宿主基因组中,可以将编码稀有切割核酸内切酶的核酸与编码植物可选择标志物的核酸共同递送。可选择标志物可以与稀有切割核酸内切酶位于同一载体上,或者可以作为单独的载体递送。转化后,可以将大豆植物部分或植物细胞置于含有适当的可选择剂的再生培养基上。经转化的植物细胞可以再生为转基因大豆植物。在一些情况下,大豆植物不包括转基因或任何外源DNA(例如T-DNA)。没有外源DNA(例如序列特异性核酸酶序列)的后代可以通过分离产生。分离可以稳定诱导的突变以及满足生物安全问题。

[0052] 在上述方法中,稀有切割核酸内切酶可以与编码一种或多种核酸外切酶蛋白的质粒共同递送至植物细胞,以提高序列特异性核酸酶诱导的诱变效率。这种核酸外切酶包括但不限于核酸外切酶的TREX(三原修复核酸外切酶)家族的成员,例如TREX2。

[0053] 本公开提供了用于使用稀有切割核酸内切酶(例如,TALE-核酸酶)编辑编码FAD2或FAD3的一种或多种基因以产生大豆植物及相关产物(例如,种子和植物部分)的材料和方法,所述大豆植物及相关产物特别适用于提供高油酸脂肪酸和/或低亚油酸或亚麻酸脂肪酸油。使用本公开的组合物的方法可以包括使表达FAD2-1A基因、FAD2-1B基因、FAD3A基因、FAD3B基因和FAD3C基因中的一种或多种的大豆植物细胞(例如,原生质体)群与被改造为在一种或多种基因中诱导靶突变的稀有切割核酸内切酶接触,从所述群中选择具有降低一种或多种基因表达的突变的细胞。此外,一种方法可以包括从植物细胞中分离包含一种或多种靶基因座(例如,FAD2-1A、FAD2-1B、FAD3A、FAD3B或FAD3C基因座)的至少一部分的基因组DNA的步骤。在一些情况下,该方法包括培养含有突变的植物细胞以产生一种或多种植物品系的步骤。大豆植物的部分诸如种子合成和积累可以识别和量化的脂肪酸。例如,如实施例中所述的植物部分的脂肪酸组成分析可以确认,与不含突变的大豆植物相比,在所产生的植物品系中,突变导致油酸水平提高、亚油酸水平降低、亚麻酸水平降低或这些特征的组合。

[0054] 在特定实施方式中,靶基因的表达因诱导的突变而降低。降低植物、植物部分或植物细胞中基因的表达包括抑制、中断、敲除或敲低基因,使得与相应的其中基因或多肽的表达未被抑制、中断、敲除或敲低的对照植物、植物细胞或者植物群或植物细胞群相比,基因的转录和/或编码的多肽的翻译降低。降低涵盖与相应的对照植物、植物细胞或者植物群或植物细胞群相比表达水平的任何减少(例如,减少10%或更多、20%或更多、30%或更多、40%或更多、50%或更多、60%或更多、70%或更多、80%或更多、90%或更多或者甚至100%)。在一些实施方式中,将表达降低50%或更多可能特别有用。可以使用方法诸如逆转录-聚合酶链式反应(RT-PCR)、Northern印迹、斑点杂交、原位杂交、核连缀(nuclear run-

on)和/或核失控(nuclear run-off)、RNase保护或免疫学和酶促方法诸如ELISA、放射免疫测定和蛋白质印迹来测量表达水平。

[0055] 稀有切割核酸内切酶可以通过编码序列特异性核酸酶的核酸(例如,载体或mRNA)或作为蛋白质引入细胞群中。例如,编码靶向存在于一种或多种FAD2和FAD3基因内的保守核苷酸序列的TALE核酸酶的核酸可用于转化大豆植物细胞或植物部分(例如原生质体),它们可以在大豆植物细胞或植物部分中表达。在一些情况下,可以将TALE核酸酶蛋白引入大豆植物细胞或植物部分(例如原生质体)中。可以使用下一代测序技术(例如,454焦磷酸测序或Illumina测序)对这些细胞或植物部分或从这些细胞产生的植物细胞系或植物部分进行分析以确定是否在一个或多个靶位点处引入了突变。用于测序的模板可以是使用与保守核苷酸序列同源的引物通过PCR扩增的靶基因。还可以分析细胞的任何脱靶突变。

[0056] 在一些实施方式中,编码多种TALE核酸酶的一种或多种核酸可用于转化这些核酸将在其中表达的大豆植物细胞或植物部分(例如原生质体)。这些TALE核酸酶可以作为蛋白质引入植物细胞或部分中。表达或引入的TALE核酸酶可在同一染色体上产生双链断裂(DSB),从而导致间插序列的缺失。可以针对靶基因的损失来分析这些细胞或植物部分或由这些细胞产生的植物细胞系或植物部分。可以通过定性或定量的手段分析含有FAD2-1基因家族或FAD3基因家族的基因座的缺失。例如,可以将第一引物设计为与第一TALE核酸酶靶位点上游的序列同源,可以将第二引物设计为与第二TALE核酸酶靶位点下游的序列互补。如果发生靶缺失,则获得PCR产物。还可以通过qPCR分析包含靶基因的基因座的缺失。相对于未经修饰的对照,靶基因的较低拷贝数可能表明存在预期的缺失。在T7E1测定中,可以从合并的愈伤组织中分离基因组DNA,并且可以对靶基因的TALE核酸酶识别位点侧翼的序列进行PCR扩增。然后可以对扩增产物进行变性和重新退火。如果重新退火的片段形成异源双链,则T7核酸内切酶I在错配位点切割。可以通过凝胶电泳可视化消化的产物以量化TALE核酸酶的诱变活性。

[0057] 用于接触大豆植物细胞群以递送序列特异性核酸酶的方法可以包括用编码序列特异性核酸酶的T-DNA对植物部分或植物细胞(例如叶、茎、叶柄、节间外植体、愈伤组织或原生质体)进行农杆菌介导的转化、用编码序列特异性核酸酶的一种或多种核酸对植物部分或植物细胞进行基因枪转化和/或用纯化的序列特异性核酸酶或编码序列特异性核酸酶的核酸(RNA或DNA)对植物部分或植物细胞进行细胞穿透肽介导的转化。

[0058] 聚乙二醇(PEG)介导的转化可用于递送序列特异性核酸酶。例如,可以从表面经灭菌的叶子中分离原生质体,并在PEG存在下用编码一种或多种序列特异性核酸酶的质粒转化。可以通过递送可检测的标志物(例如YFP质粒)来监测转化效率,可以使用荧光显微镜或流式细胞仪对该可检测的标志物进行可视化。在PEG介导的转化后,可以使用原生质体培养领域普通技术人员已知的方法和培养基培养原生质体。在适当长的培养时间后,被鉴定为突变体的源自原生质体的愈伤组织可以生长,转移到芽诱导培养基中,然后(一旦形成根)转移到土壤中并生长至成熟用于种子生产。

[0059] 在一些情况下,大豆植物包括多种突变,这些突变涉及改变由植物、植物部分或植物细胞产生的油的脂肪酸组成。以具有调节FAD2和FAD3蛋白表达的突变的大豆植物、植物部分或植物细胞为特征的实施方式落在本公开的范围内,其中与由不存在一种或多种突变的相应大豆植物、植物部分或植物细胞产生的油相比,所述植物、植物部分或植物细胞产生

具有增加的油酸和减少的亚油酸和亚麻酸脂肪酸含量的油。靶向FAD2-1A、FAD2-1B或FAD3A/B/C表达突变体可以在任何目标大豆的菌株、种属或栽培品种上实施，但不限于此。在一些实施方式中，可以在已经具有较高水平的油酸和较低水平的亚油酸或亚麻酸特征(例如遗传预先倾向性)的种质或其他植物组织中实施靶基因编辑或其他遗传修饰。

[0060] 突变大豆品系可以包括提供或改变除FAD2和FAD3之外的蛋白质表达的突变。组合效应可以涉及使用多个单独的核酸构建体或转化事件。例如，可以通过相同或不同的方法将如上所述的多个构建体引入植物细胞，包括通过在单个转化载体中包含两个转录盒来引入这样的性状，同时转化两个表达构建体，用第二表达构建体重新转化，或通过传统植物育种方法杂交转基因植物，只要所得产物是具有整合到其基因组中的两种特征的植物。

[0061] 可以将常规育种技术与上述靶向方法结合。在一些情况下，接受构建体的大豆植物是具有一种或多种某些重要农艺性状的优良品系，诸如导致增加的生物质产量、增加的食品产量、改善的食品质量、抗虫性、活力、发育时间(到收获的时间)、增强的营养成分、新颖的生长模式、风味或颜色、盐、热、干旱和寒冷的耐受性等的那些。

[0062] 本公开的转化的大豆植物可用于植物育种程序中以产生新颖且有用的品系和品种。可以通过已知程序进行育种。可以将DNA指纹、SNP或类似技术用于标志物辅助选择(MAS)育种程序，以将调节FAD2和FAD3基因中的一种或多种的表达的突变转移或培育到其他大豆植物中。在一些实施方式中，这样的方法包括在大豆突变体与第二大豆植物之间进行杂交以产生F1植物，或者与可与该突变体互相杂交的种属杂交。该方法可以进一步包括将F1植物与第二大豆植物回交并重复该回交步骤以产生近等基因系，其中突变被整合到所述第二大豆植物的基因组中；其中源自具有整合突变的第二植物的近等基因系具有改变的脂肪酸特性(例如，较高水平的油酸、较低水平的亚油酸和/或较低水平的亚麻酸)。可以通过分子标志物或TILLING®来促进这种方法。因此，本公开的突变大豆植物可用于产生新的和有用的品系和品种。

[0063] 在一些实施方式中，本文所述的方法可用于产生具有部分来自油的脂肪酸特性的优异稳定性和性能的油的大豆品种。商品大豆油主要由五种脂肪酸构成：棕榈酸(10％)、硬脂酸(4％)、油酸(18％)、亚油酸(55％)和亚麻酸(13％)。亚麻酸含量较低的大豆油可增加其氧化和油炸稳定性。这样的油可能更健康，特别是因为它不需要多不饱和脂肪酸的部分氢化来稳定。部分氢化产生反式脂肪酸，其消耗与心脏病风险增加有关。与不存在上述突变的大豆品种相比，本公开的大豆品种可以具有亚麻酸含量减少、油酸含量增加或亚油酸含量减少的油。例如，本公开的在FAD2-1和FAD3基因家族内具有堆叠突变的大豆品种可以具有亚麻酸和亚油酸水平低于3％以及油酸水平高于65％的油。

[0064] 如本文所用，术语“基因”是指包括与基因产物的表达相关的启动子区域的核酸序列。“基因”还涵盖与基因产物表达相关的内含子和外显子区域，以及与基因产物表达相关的5’或3’非翻译区。

[0065] “内源基因”是指核酸分子，其包含存在于野生型植物中的野生型序列的序列，或具有允许其保留编码产物的功能的同一性百分比的序列，例如具有至少90％同一性的序列，并且可以从植物或植物部分的细胞获得，或者可以合成产生。进一步的实施方式提供具有至少91％、92％、93％、94％、95％、96％、97％、98％、99％同一性的序列。

[0066] “FAD2”是指脂肪酸脱饱和酶2(FAD2)蛋白，其由大豆基因组中的三个FAD2脱饱和

酶基因编码,然而,FAD2-1A(Glyma10g42470)和FAD2-1B(Glyma20g24530)在油合成高峰期高表达,并且是大豆种子中油酸和亚油酸水平的主要遗传决定因素。FAD2催化油酸转化为亚油酸。

[0067] “FAD3”是指脂肪酸脱饱和酶3(FAD3)酶,其由FAD3A(Glyma14g37350)、FAD3B(Glyma02g39230)和FAD3C(Glyma18g06950)组成的基因家族产生,并催化亚油酸转化为亚麻酸。

[0068] 术语“功能变体”是指所列举的蛋白质或蛋白质结构域的催化活性突变体、编码执行与所列举的SEQ ID NO相同的功能的蛋白质或蛋白质结构域的突变体的核苷酸序列、或介导与所列举的SEQ ID NO相同的活性(例如,TALEN-介导的切割)的核苷酸序列变体。

[0069] 如本文所用,术语“约”表示相关值可以被修改正负百分之五(+/-5%)并且仍在所公开的实施方式的范围内,另有说明除外。

[0070] 本文公开的本发明的替代要素或实施方式的分组不应被解释为限制。每个组成员可以单独地或与组的其他成员或本文中发现的其他要素以任何组合的方式被提及和要求保护。出于方便和/或可专利性的原因,可以预期组的一个或多个成员可被包括在组中或从组中删除。当发生任何这样的包括或删除时,本说明书在此视为包含修改后的组,从而满足在所附权利要求书中使用的全部马库什组的书面描述。

[0071] 除非另有说明,否则核酸或寡核苷酸以5’至3’方向从左至右书写。

[0072] 除非另有说明,否则当在本申请中使用时,词语“本文”、“上文”和“下文”以及类似含义的词语是指本申请的整体而不是本申请的任何特定部分。如本文所用,词语“约”和“大约”包括所述值周围的微小变化,通常在标准误差范围内,例如在所述值的10%以内。

[0073] 在此提及的所引用的专利、专利申请和科学文献通过援引整体并入本文,就如同每个单独的出版物、专利或专利申请被具体地和单独地指明通过援引并入一样。

[0074] 将在下文的实施例中对本发明进行进一步描述,这些实施例并不限制权利要求书中描述的本发明的范围。

[0075] 实施例

[0076] 1.针对大豆脂肪酸脱饱和酶3(FAD3)基因家族的多重TALEN设计。

[0077] 使用软件诸如Geneious(Biomatters Ltd)对齐大豆FAD3A、FAD3B和FAD3C(SEQ ID:1-3)的基因组序列,并鉴定基因之间的同源区域。将GmFAD3_T01-L1和GmFAD3_T01-R1(SEQ ID:4-5)设计为具有对FAD3A和FAD3B(SEQ ID NO:1-2)的靶向特异性。将GmFAD3_T02-L1和GmFAD3_T02-R1(SEQ ID NO:6-7)设计为对FAD3C(SEQ ID:3)具有特异性(图1A)。将GmFAD3_T03-L1和GmFAD3_T03-R1(SEQ ID NO:8-9)设计为对FAD3A和FAD3B(SEQ ID NO:1-2)具有靶向特异性。将GmFAD3_T04-L1和GmFAD3_T04-R1(SEQ ID NO:10-11)设计为对FAD3C(SEQ ID NO:3)具有特异性(图1B)。将GmFAD3_T05-L1(SEQ ID NO:12)设计为对FAD3A(SEQ ID NO:1)具有特异性,将GmFAD3_T05-R1(SEQ ID NO:13)设计为对FAD3A、FAD3B和FAD3C(SEQ ID NO:1-3)具有特异性,将GmFAD3_T06-L1(SEQ ID NO:14)设计为对FAD3B(SEQ ID NO:2)具有特异性,将GmFAD3_T07-L1(SEQ ID NO:15)设计为对FAD3C(SEQ ID NO:3)具有特异性(图1C)。将GmFAD3_T08-L1和GmFAD3_T08-R1(SEQ ID NO:16-17)设计为对FAD3A和FAD3B(SEQ ID NO:1-2)具有靶向特异性。将GmFAD3_T09-L1和GmFAD3_T09-R1(SEQ ID NO:18-19)设计为对FAD3C(SEQ ID NO:3)具有靶向特异性(图1D)。

[0078] 2.证明大豆原生质体中HT对的核酸酶活性

[0079] 合成半-TAL效应物核酸酶(半-TALEN或HT)，克隆到植物表达载体中，并使用PEG介导的方案转化到大豆原生质体中以评价半-TALEN对的各个组合的活性。转化后48小时从原生质体群中提取DNA，使用PCR扩增并使用Illumina Miseq技术对靶基因FAD3A、FAD3B和FAD3C(SEQ ID:1-3)进行测序。进行分析以确定非同源末端连接(NHEJ)频率作为TALEN核酸酶活性的测量值。对于每个TALEN组合，确定三个靶基因中每一个的核酸酶活性。数据总结于表1中。使用两种HT效应物核酸酶，使用GmFAD3_T08-L1和GmFAD3_T08-R1(22.87％和14.57％)的组合实现FAD3A和FAD3B的最高NHEJ频率。对于FAD3C，使用GmFAD3_T09-L1和GmFAD3_T09-R1(11.32％)观察到使用两种HT组合的最高活性。

[0080] 3.证明大豆原生质体中多重HT的核酸酶活性

[0081] 选择对于FAD3A、FAD3B和FAD3C具有最高活性的三种半TAL效应物核酸酶(半TALEN或HT)的组合，并克隆到单个转化载体中，以使用PEG介导的方案递送到大豆原生质体中。转化后48小时从原生质体群中提取DNA，使用PCR扩增靶基因FAD3A、FAD3B和FAD3C(SEQ ID NO:1-3)并对其进行测序，以如上确定NHEJ活性。GmFAD3_T09-L1、GmFAD3_T08-R1和GmFAD3_T09-R1的组合的递送在全部三种基因靶标FAD3A(16.28％)、FAD3B(14.82％)和FAD3C(21.42％)成功地实现了高效核酸酶活性。数据总结于表2中。

[0082] 表1:原生质体中TALEN对组合的NHEJ诱变频率。

[0083]

左-HT	右-HT	靶标	% NHEJ
GmFAD3_T08-L1	GmFAD3_T08-R1	FAD3A	**22.87**
		FAD3B	**14.57**
		FAD3C	**3.59**
GmFAD3_T08-L1	GmFAD3_T09-R1	FAD3A	**6.56**
		FAD3B	**1.80**
		FAD3C	**7.60**
GmFAD3_T09-L1	GmFAD3_T08-R1	FAD3A	**12.37**
		FAD3B	**11.17**
		FAD3C	**5.98**
GmFAD3_T09-L1	GmFAD3_T09-R1	FAD3A	**6.46**
		FAD3B	**3.61**
		FAD3C	**11.32**

[0084] 表2:原生质体中多重TALEN组合的NHEJ诱变频率

[0085]

左-HT	右-HT	靶标	% NHEJ
GmFAD3_T09-L1	GmFAD3_T08-R1 GmFAD3_T09-R1	FAD3A	16.28
		FAD3B	14.82
		FAD3C	21.42

[0086] 表3:FAD3的TALEN靶序列

[0087]

SEQ ID	识别因子	序列
4	GmFAD3_T01-L1	TGGGACACATCTTGCAT
5	GmFAD3_T01-R1	TTGTGCCATACCATGGA
6	GmFAD3_T02-L1	TGGGCCACATCTTGCAC
7	GmFAD3_T02-R1	TTGTACCATACCATGGA
8	GmFAD3_T03-L1	TTGCTGGGTCAAGAATC

[0088]

9	GmFAD3_T03-R1	AGTTATGTTCTCAGGGA
10	GmFAD3_T04-L1	TTGCTGGGAGAAGAACA
11	GmFAD3_T04-R1	AGTTATGTTCTGAGGGA
12	GmFAD3_T05-L1	TTGGCCCATTCAAGGAA
13	GmFAD3_T05-R1	TTTGTTCTTGGACATGA
14	GmFAD3_T06-L1	TTGCCCCATTCAAGGCA
15	GmFAD3_T07-L1	TTGGCCTGCACAAGGCA
16	GmFAD3_T08-L1	TTAGCCACAGAACTCAC
17	GmFAD3_T08-R1	CACATTGAGAAGGATGA
18	GmFAD3_T09-L1	TTAGCCACAGGACTCAC
19	GmFAD3_T09-R1	CATGTTGAGAAGGATGA

[0089] 在验证TAL效应物核酸酶对在内源靶位点产生靶修饰后，进行实验以产生在两种或更多种FAD3A/B/C中具有突变的大豆植物。示例性大豆植物的特征性突变等位基因示于SEQ ID NO:40-44。

[0090] 分析源自FAD3A/B/C中的两种或更多种中为纯合突变体的大豆品系的种子的脂肪酸组成。简而言之，将单个大豆种子单独粉碎。从一部分基本组织制备DNA，并对其进行分析以确认每个种子的基因型。合并来自双重和三重纯合敲除种子的粉碎组织。然后使用脂肪酸甲酯(FAME)气相色谱法(Beuselinck等人,Crop Sci.47:747-750,2006)测定脂肪酸组成，以评价相对于野生型种子而言，具有各种FAD3A/B/C突变的种子的亚油酸、亚油酸和油酸的比例是否发生改变。

[0091] 4.诱变大豆FAD2-1A和FAD2-1B基因的序列特异性核酸酶

[0092] 为了完全灭活或敲除大豆中FAD2-1A和FAD2-1B基因的等位基因，将序列特异性核酸酶设计为靶向起始密码子附近的蛋白质编码区。设计了八个靶向FAD2-1基因家族的TAL效应物核酸酶对(表4)。GmFAD2_T01-L1和GmFAD2_T01-R1(SEQ ID NO:27和28)、GmFAD2_T02-L1和GmFAD2_T02-R1(SEQ ID NO:29和30)、GmFAD2_T03-L1和GmFAD2_T03-R1(SEQ ID NO:31和32)以及GmFAD2_T04-L1和GmFAD2_T04-R1(SEQ ID NO:33和34)对在FAD2-1A和FAD2-1B基因靶标处实现高效核酸酶活性(表5)。

[0093] 表4:FAD2的TALEN靶序列

[0094]

SEQ ID	识别因子	序列

[0095]

27	GmFAD2_T01-L1	TCTCAAGGGTTCCAAAC
28	GmFAD2_T01-R1	TTGAGTTGGCCAACAGT
29	GmFAD2_T02-L1	TATGTTGTTTATGACCT
30	GmFAD2_T02-R1	TGGTGGCAATGTAGAAA
31	GmFAD2_T03-L1	TCCTATGTTGTTTATGA
32	GmFAD2_T03-R1	TGGCAATGTAGAAAATG
33	GmFAD2_T04-L1	TTGTTTATGACCTTTCA
34	GmFAD2_T04-R1	TAGGTGGTGGCAATGTA
35	GmFAD2_T05-L1	TGCCACCACCTACTTCC
36	GmFAD2_T05-R1	TGCAATGAGGGAAAAGG

[0096] 表5:FAD2靶基因的原生质体中TALEN对组合的NHEJ诱变频率。

[0097]

左-HT	右-HT	靶标	% NHEJ
GmFAD2_T01-L1	GmFAD2_T01-R1	FAD2-1A	3.95
		FAD2-1B	3.41
GmFAD2_T02-L1	GmFAD2_T02-R1	FAD2-1A	0.91
		FAD2-1B	1.19
GmFAD2_T03-L1	GmFAD2_T03-R1	FAD2-1A	4.91
		FAD2-1B	3.94
GmFAD2_T04-L1	GmFAD2_T04-R1	FAD2-1A	4.67
		FAD2-1B	3.54
GmFAD2_T05-L1	GmFAD2_T05-R1	FAD2-1A	0.00
		FAD2-1B	0.00

[0098] 在验证TAL效应物核酸酶对在内源靶位点产生靶向修饰后,进行实验以产生在FAD2-1A和FAD2-1B之一或两者中具有突变的大豆植物。示例性大豆植物的特征性突变等位基因示于SEQ ID NO:37-39中。

[0099] 分析源自FAD2-1A或FAD2-1B中纯合突变体或FAD2-1A和FAD2-1B两者纯合的大豆品系的种子的脂肪酸组成。结果示于表6中。简而言之,将单个大豆种子单独粉碎。从一部分基本组织中制备DNA,并对其进行分析以确认每个种子的基因型。合并来自FAD2-1A纯合、FAD2-1B纯合或FAD2-1A/FAD2-1B双纯合敲除种子的粉碎组织。然后使用脂肪酸甲酯(FAME)气相色谱法(如上所述)测定脂肪酸组成,以评价相对于野生型种子,具有各种FAD2-1突变的种子的亚油酸和油酸比例是否发生改变。

[0100] 5.FAD2和FAD3突变的组合

[0101] 通过使用TAL效应物核酸内切酶靶向一种基因或多种基因,或通过将具有FAD2基因和/或FAD3基因中突变的植物杂交,产生包含一种或多种FAD2和FAD3基因或者蛋白质的敲除活性的突变组合的植物。产生包含一系列突变组合的植物,包括具有2、3、4或5种基因敲除的植物。通过FAME评价温室生长植物的种子油组成。结果示于表6中。

[0102]

表 6：脂肪酸组成

样品	豆蔻酸 C14:0	棕榈酸 C16:0	硬脂酸 C18:0	油酸 C18:1	亚油酸 C18:2	亚麻酸 C18:3	花生酸 C20:0	二十碳烯酸 C20:1	二十二烷酸 C22:0	芥子酸 C22:1	木蜡酸 C24:0
野生型											
Soy 1	0.10	12.16	3.41	29.70	48.55	5.07	0.30	0.24	0.25	0.00	0.22
Soy 2	0.10	12.63	3.19	20.19	57.07	5.97	0.25	0.19	0.25	0.00	0.15
Soy 3	0.10	12.68	3.05	16.19	59.28	7.86	0.24	0.19	0.25	0.00	0.15
Soy 4	0.11	12.83	3.87	17.17	57.78	7.24	0.31	0.19	0.31	0.00	0.19
Soy 5	0.11	12.82	3.88	17.19	57.76	7.24	0.32	0.19	0.31	0.00	0.19
Soy 6	0.00	12.28	3.90	22.74	54.28	6.22	0.32	0.00	0.26	0.00	0.00
Soy 7	0.10	12.34	3.95	24.06	53.52	4.92	0.33	0.25	0.32	0.00	0.22
Soy 8	0.00	13.43	3.24	14.82	58.87	9.06	0.29	0.00	0.30	0.00	0.00
Soy 9	0.10	12.75	3.18	16.59	58.90	7.60	0.26	0.19	0.27	0.00	0.17
Soy 10	0.10	12.64	3.74	16.49	59.03	7.09	0.28	0.20	0.27	0.00	0.15
平均数	**0.08**	**12.66**	**3.54**	**19.51**	**56.50**	**6.83**	**0.29**	**0.16**	**0.28**	**0.00**	**0.14**
低	0.00	12.16	3.05	14.82	48.55	4.92	0.24	0.00	0.25	0.00	0.00
高	0.11	13.43	3.95	29.70	59.28	9.06	0.33	0.25	0.32	0.00	0.22
五种基因敲除（FAD2A/B, FAD3A/B/C）											
Soy 11	0.06	7.84	3.81	81.25	4.79	1.02	0.37	0.39	0.30	0.00	0.17
Soy 12	0.00	8.71	3.79	79.35	5.98	1.04	0.38	0.43	0.33	0.00	0.00

[0103]

Soy 13	0.07	8.07	4.02	79.81	5.65	1.08	0.39	0.41	0.32	0.00	0.18		
Soy 14	0.07	8.77	3.57	79.56	5.74	1.14	0.34	0.38	0.27	0.00	0.15		
Soy 15	0.06	8.28	3.51	80.43	5.48	1.01	0.35	0.41	0.29	0.00	0.16		
Soy 16	0.00	9.08	3.30	78.97	6.49	1.13	0.33	0.38	0.33	0.00	0.00		
Soy 17	0.05	7.65	3.91	81.69	4.76	0.81	0.36	0.33	0.31	0.00	0.14		
Soy 18	0.00	7.69	3.93	80.96	5.26	0.92	0.38	0.40	0.32	0.00	0.15		
Soy 19	0.06	8.93	3.65	79.06	6.01	1.02	0.37	0.40	0.34	0.00	0.15		
Soy 20	0.00	9.17	3.43	78.04	7.09	1.26	0.36	0.34	0.31	0.00	0.00		
平均数	**0.04**	**8.42**	**3.69**	**79.91**	**5.73**	**1.04**	**0.36**	**0.39**	**0.31**	**0.00**	**0.11**		
低	0.00	7.65	3.30	78.04	4.76	0.81	0.33	0.33	0.27	0.00	0.00		
高	0.07	9.17	4.02	81.69	7.09	1.26	0.39	0.43	0.34	0.00	0.18		
四种基因敲除 (FAD2A/B, FAD3A/C)													
Soy 31	0.07	8.47	3.22	81.09	4.43	1.54	0.32	0.36	0.28	0.06	0.16		
Soy 32	0.07	7.99	3.23	81.53	2.89	3.11	0.32	0.36	0.27	0.07	0.16		
Soy 33	0.06	7.89	3.26	81.84	3.38	2.39	0.32	0.36	0.29	0.06	0.16		
Soy 34	0.07	8.82	3.13	79.96	3.48	3.37	0.31	0.37	0.27	0.05	0.16		
Soy 35	0.07	7.82	3.16	82.75	2.00	2.97	0.32	0.37	0.29	0.08	0.16		
Soy 36	0.06	8.45	3.26	81.14	3.44	2.38	0.34	0.39	0.31	0.06	0.16		
Soy 37	0.06	8.08	3.24	82.25	2.30	2.91	0.32	0.36	0.28	0.05	0.15		
Soy 39	0.06	8.25	3.23	81.53	2.52	3.24	0.31	0.36	0.28	0.06	0.14		
Soy 40	0.06	7.93	3.20	81.92	3.35	2.41	0.31	0.36	0.27	0.06	0.15		

[0104]

平均数	**0.06**	**8.19**	**3.22**	**81.56**	**3.09**	**2.70**	**0.32**	**0.37**	**0.28**	**0.06**	**0.15**
低	0.06	7.82	3.13	79.96	2.00	1.54	0.31	0.36	0.27	0.05	0.14
高	0.07	8.82	3.26	82.75	4.43	3.37	0.34	0.39	0.31	0.08	0.16
两种基因敲除 (FAD2A/B)											
Soy 41	0.06	7.55	3.34	79.93	3.30	4.68	0.33	0.29	0.34	0.05	0.13
Soy 42	0.07	7.54	3.49	78.84	3.21	5.64	0.35	0.30	0.35	0.07	0.15
Soy 43	0.06	7.11	3.93	81.30	2.04	4.34	0.37	0.29	0.37	0.06	0.13
Soy 44	0.07	8.22	3.26	77.19	4.51	5.47	0.34	0.33	0.36	0.09	0.17
Soy 45	0.07	8.47	3.22	81.09	4.43	1.54	0.32	0.36	0.28	0.06	0.16
Soy 46	0.07	7.67	3.28	78.96	3.74	5.09	0.32	0.29	0.35	0.09	0.15
Soy 48	0.08	7.50	3.65	80.34	2.70	4.67	0.31	0.27	0.29	0.05	0.13
Soy 50	0.07	8.02	3.55	75.61	4.57	6.89	0.36	0.29	0.39	0.09	0.16
平均数	**0.07**	**7.76**	**3.47**	**79.16**	**3.56**	**4.79**	**0.34**	**0.30**	**0.34**	**0.07**	**0.15**
低	0.06	7.11	3.22	75.61	2.04	1.54	0.31	0.27	0.28	0.05	0.13
高	0.08	8.47	3.93	81.30	4.57	6.89	0.37	0.36	0.39	0.09	0.17
五种/四种基因敲除											
Soy 71[a]	0.08	8.53	2.78	80.56	5.69	1.06	0.30	0.43	0.29	0.08	0.21
Soy 72[b]	0.07	8.06	2.79	82.09	2.78	3.22	0.26	0.31	0.22	0.05	0.16
Soy 73[b]	0.06	8.27	3.32	80.64	3.51	3.01	0.31	0.33	0.27	0.06	0.23
Soy 74[b]	0.08	8.25	3.69	78.86	5.28	2.65	0.33	0.29	0.26	0.08	0.23

[0105]

Soy 75[a]	0.07	8.64	3.29	79.25	6.53	0.99	0.32	0.33	0.31	0.07	0.19
平均数	**0.07**	**8.35**	**3.17**	**80.28**	**4.76**	**2.18**	**0.30**	**0.34**	**0.27**	**0.07**	**0.20**
低	0.06	8.06	2.78	78.86	2.78	0.99	0.26	0.29	0.22	0.05	0.16
高	0.08	8.64	3.69	82.09	6.53	3.22	0.33	0.43	0.31	0.08	0.23

[a]5 种基因敲除 [b]4 种基因敲除

[0106] 相对于来自野生型植物的油，脂肪酸特性示出来自突变植物的油中的若干脂肪酸水平显著变化。亚麻酸、亚油酸和棕榈酸水平减少，而二十碳烯酸、油酸和硬脂酸水平增加。在一些突变体(例如，样品Soy31-50和71-75)的油中，芥子酸水平相对于来自野生型植物的油而言增加。

[0107] 虽然已对说明性实施方式进行了说明和描述，但是应当理解，可以对其做出各种改变而不脱离本发明的精神和范围。

序列表

<110> 卡里克特公司(CALYXT, INC.)

<120> 用于基因编辑的TAL效应物核酸酶

<130> 1702.029PCT1

<140>

<141>

<150> 63/030,197

<151> 2020-05-05

<150> 62/941,568

<151> 2019-11-27

<160> 65

<170> PatentIn版本 3.5

<210> 1

<211> 3966

<212> DNA

<213> 大豆(Glycine max)

<400> 1

caccacgtat ccctgaaaag agagaaaaac acactaagcc aaagccaaag cagcaatggt 60
taaagacaca aagcctttag cctatgctgc taataatgga taccaaaagg aagcttttga 120
tcccagtgct cctccaccgt ttaagattgc agaaatcaga gttgcaatac caaaacattg 180
ctgggtcaag aatccatgga gatccctcag ttatgttctc agggatgtgc ttgtaattgc 240
tgcattgatg gctgctgcaa gtcacttcaa caactggctt ctctggctaa tctattggcc 300
cattcaagga acaatgttct gggctctgtt tgttcttgga catgattggt aattaattaa 360
tttgttgtta ctttttttgtt ataatatgaa tctcacacac tgctttgtta tgcctacctc 420
atttcatttg gctttagaca acttaaattt gagatcttta ttatgttttt tgcttatatg 480
gtaaagtgat tcattcttca cattgaattg aacagtggcc atggaagctt ttcagacagc 540
ccttttctaa atagcctggt gggacacatc ttgcattcct caattcttgt gccataccat 600
ggatggttag ttcatcccgg ctttttttgtt tgtcattgga agttctttta ttgattcaat 660
ttttatagcg tgttcggaaa cgcgtttcag aaaataatga aatacatctt gaatctgaaa 720
gttataactt ttagcttcat tgtcattgaa agttctttta ttaattatat ttttattgcg 780
tgtttggaat cccatttgag aaataagaaa tcacgtttaa aatgtgaaag ttataactat 840
taacttttga ctaaacttga aaaaatcaca tttttgatgt ggaaccaaat ctgatttgag 900
aaccaagttg attttgatgg attttgcagg agaattagcc acagaactca ccatcaaaat 960
catggacaca ttgagaagga tgaatcctgg gttccagtat gtgattaact acttcctcta 1020
tagttatttt tgattcaatt aaatttattt atttaataag ttcaagaaaa aaggaatctt 1080
tatacttcag ataaagctgt tcttgaacat tttttttgt cattatctta gttaaccgag 1140
aagatttaca agaatctaga caacatgaca agacttgtta gattcactgt gccatttcca 1200
ttgtttgtgt atccaattta tttggtgagt gctttttttt ttttacttgg aagactacaa 1260

cacattatta ttattataat atggttcaaa tcaatgactt ttaatttctt tgtgatgtgc 1320
actccatttt cagttctcaa gaagccccgg aaaggaaggt tctcacttca atccctacag 1380
caatctgttc ccacccagtg agagaaaggg aatagcaata tcaacactgt gttgggttac 1440
catgttttct atgcttatct atctctcctt cataactagt ccagttctat tgctcaagct 1500
ctatggaatt ccatattggg taattaaatt actcttacat tacttttcc tctttttttt 1560
tatgggtctt aactagtatc acaaaaatat tggttaaaaa attttaaaaa aatatttatt 1620
atgtaaatca taaaagaaca taaaaaaaat gatgaataac ataattttcg tctcttatta 1680
aaaatatttt tattttaaat ttcttaatca atatatttag aatctggtta acatttttg 1740
aatatttcaa ttctccaatt aaaaatttga aatagtcacc attaattatg taattgtttg 1800
aacacgtgca gatatttgtt atgtggctgg actttgtcac atacttgcat caccatggtc 1860
atcatcagaa actgccttgg tatcgcggca aggtaacaaa aataaataga aaatagtgag 1920
tgaacactta aatgttagat actaccttct tcttcttttt ttttttttt tgaggttaat 1980
gctagataat agctagaaag agaaagaaag acaaatatag gtaaaaataa ataatataac 2040
ctgggaagaa gaaaacataa aaaaagaaat aatagagtct acgtaatgtt tggatttttg 2100
agtgaaatgg tgttcaccta ccattactca aagattctgt tgtctacgta gtgtttggac 2160
tttggagtga aatggtgttc acctaccatt actcagattc tgttgtgtcc cttagttact 2220
gtcttatatt cttagggtat attctttatt ttacatcctt ttcacatctt acttgaaaag 2280
atttttaatt attcattgaa atattaacgt gacagttaaa ttaaaataat aaaaaattcg 2340
ttaaaacttc aaataaataa gagtgaaagg atcatcattt ttcttctttc ttttattgcg 2400
ttattaatca tgcttctctt ctttttttc ttcgctttcc acccatatca aattcatgtg 2460
aagtatgaga aaatcacgat tcaatggaaa gctacaggaa ctttttttgt tttgttttta 2520
taatcggaat taatttatac tccatttttt cacaataaat gttacttagt gccttaaaga 2580
taatatttga aaaattaaaa aaattattaa tacactgtac tactatataa tatttgacat 2640
atatttaaca tgattttcta ttgaaaattt gtatttatta ttttttaatc aaaacccata 2700
aggcattaat ttacaagacc catttttcat ttatagcttt acctgtgatc atttatagct 2760
ttaagggact tagatgttac aatcttaatt acaagtaaat atttatgaaa aacatgtgtc 2820
ttacccctta accttacctc aacaaagaaa gtgtgataag tggcaacaca cgtgttgctt 2880
ttttggccca gcaataacac gtgtttttgt ggtgtacaaa aatggacagg aatggagtta 2940
tttaagaggt ggtctcacaa ctgtggatcg tgactatggt tggatcaata acattcacca 3000
tgacattggc acccatgtta ttcaccatct tttccctcaa attcctcatt atcacctcgt 3060
tgaagcggta tattttacta ttattactca cctaaaaaga atgcaattag tacatttgtt 3120
ttatctcttg gaagttagtc attttcagtt gcatgattgt aatgttctct ctatttttaa 3180
accatgtttt cacacctact tcgtttaaaa taagaatgtg gatactattc taatttctat 3240
taacttcttt taaaaaataa tgtaaaacta gtattaaaaa agaggaaata gattacactc 3300
tactaatact aatagtataa aaaaaattac attgttattt tatcacaaat aattatatat 3360
aattaatttt tacaatcatt atcttaaaag tcatgtatga tatacagttt ttacatgctt 3420
tggtacttat tgtaaagtta gtgatttatt cattatttat gttatataat tggcataaat 3480
atcatgtaac cagctcacta tactataatg ggaacttggt ggtgaaaggg gtttacaacc 3540
ctcttttcta ggtgtaggtg ctttgatact tctggtccct ttttatatca atataaatta 3600

tattttgctg ataaaaaaaa cattattaat atataatcat taacttcttt aaaaaccgta 3660
cctaaaactt tatattatta aaaagaagat tgagatcagc aaaagaaaaa aaaattaaca 3720
gtcatttgaa ttcactgcag acacaagcag caaaatcagt tcttggagag tattaccgtg 3780
agccagaaag atctgcacca ttaccatttc atctaataaa gtatttaatt cagagtatga 3840
gacaagacca cttcgtaagt gacactggag atgtggttta ttatcagact gattctctgc 3900
accttcactc gcaccgagac tgagtttcaa tttttgggtt atttattgga ttctagctac 3960
tcaaat 3966

<210> 2
<211> 4013
<212> DNA
<213> 大豆(Glycine max)
<400> 2

caccacgtat ccctgagaaa agagaggaac cacactctct aagccaaagc aaaagcagca 60
gcagcagcaa tggttaaaga cacaaagcct ttagcctatg ctgccaataa tggataccaa 120
caaaagggtt cttcttttga ttttgatcct agcgctcctc caccgtttaa gattgcagaa 180
atcagagctt caataccaaa acattgctgg gtcaagaatc catggagatc cctcagttat 240
gttctcaggg atgtgcttgt aattgctgca ttggtggctg cagcaattca cttcgacaac 300
tggcttctct ggctaatcta ttgccccatt caaggcacaa tgttctgggc tctctttgtt 360
cttggacatg attggtaata atttttgtgt ttcttactct ttttttttt tttttgttta 420
tgatatgaat ctcacacatt gttctgttat gtcatttctt cttcatttgg ctttagacaa 480
cttaaatttg agatctttat tatgtttttg cttatatggt aaagtgattc ttcattattt 540
cattcttcat tgattgaatt gaacagtggc catggaagct tttcagatag ccctttgctg 600
aatagcctgg tgggacacat cttgcattcc tcaattcttg tgccatacca tggatggtta 660
gttcatactg gctttttttgt ttgttcattt gtcattgaaa aaaaatcttt tgttgattca 720
attattttta tagtgtgttt ggaagcccgt ttgagaaaat aagaaatcgc atctggaatg 780
tgaaagttat aactatttag cttcatctgt cgttgcaagt tcttttattg gttaaatttt 840
tatagcgtgc taggaaaccc attcgagaaa ataagaaatc acatctggaa tgtgaaagtt 900
ataactgtta gcttctgagt aaacgtggaa aaaccacatt ttggatttgg aaccaaattt 960
tatttgataa atgacaacca aattgatttt gatggatttt gcaggagaat tagccacaga 1020
actcaccatc aaaaccatgg acacattgag aaggatgagt catgggttcc agtatgtgat 1080
taattgcttc tcctatagtt gttcttgatt caattacatt ttatttattt ggtaggtcca 1140
agaaaaaagg gaatctttat gcttcctgag gctgttcttg aacatggctc ttttttatgt 1200
gtcattatct tagttaacag agaagattta caagaatcta gacagcatga caagactcat 1260
tagattcact gtgccatttc cattgtttgt gtatccaatt tatttggtga gtgatttttt 1320
gacttggaag acaacaacac attattatta taatatggtt caaaacaatg acttttttctt 1380
tatgatgtga actccatttt ttagttttca agaagccccg gaaaggaagg ctctcacttc 1440
aatccctaca gcaatctgtt tccacccagt gagagaaaag gaatagcaat atcaacactg 1500
tgttgggcta ccatgttttc tctgcttatc tatctctcat tcataactag tccacttcta 1560
gtgctcaagc tctatggaat tccatattgg gtaactaaat tactcctaca ttgttacttt 1620

ttcctccttt tttttattat ttcaattctc caattggaaa tttgaaatag ttaccataat 1680
tatgtaattg tttgatcatg tgcagatatt tgttatgtgg ctggactttg tcacatactt 1740
gcatcaccat ggtcaccacc agaaactgcc ttggtaccgc ggcaaggtaa caaaaataaa 1800
tagaaaatag tgggtgaaca cttaaatgcg agatagtaat acctaaaaaa agaaaaaaat 1860
ataggtataa taaataatat aactttcaaa ataaaaagaa atcatagagt ctagcgtagt 1920
gtttggagtg aaatgatgtt cacctaccat tactcaaaga ttttgttgtg tcccttagtt 1980
cattcttatt attttacata tcttacttga aaagactttt taattattca ttgagatctt 2040
aaagtgactg ttaaattaaa ataaaaaaca agtttgttaa aacttcaaat aaataagagt 2100
gaagggagtg tcatttgtct tctttctttt attgcgttat taatcacgtt tctcttctct 2160
ttttttttt ttcttctctg ctttccaccc attatcaagt tcatgtgaag cagtggcgga 2220
tctatgtaaa tgagtggggg gcaattgcac ccacaagatt ttatttttta tttgtacagg 2280
aataataaaa taaaactttg cccccataaa aaataaatat tttttcttaa aataatgcaa 2340
aataaatata agaaataaaa agagaataaa ttattattaa ttttattatt ttgtactttt 2400
tatttagttt ttttagcggt tagatttttt tttcatgaca ttatgtaatc ttttaaaagc 2460
atgtaatatt tttattttgt gaaaataaat ataaatgatc atattagtct cagaatgtat 2520
aaactaataa taattttatc actaaaagaa attctaattt agtccataaa taagtaaaac 2580
aagtgacaat tatattttat atttacttaa tgtgaaataa tacttgaaca ttataataaa 2640
acttaatgac aggagatatt acatagtgcc ataaagatat tttaaaaaat aaaatcatta 2700
atacactgta ctactatata atattcgata tatattttta acatgattct caatagaaaa 2760
attgtattga ttatatttta ttagacatga atttacaagc cccgtttttc atttatagct 2820
cttacctgtg atctattgtt ttgcttcgct gtttttgttg gtcaagggac ttagatgtca 2880
caatattaat actagaagta aatatttatg aaaacatgta ccttacctca acaaagaaag 2940
tgtggtaagt ggcaacacac gtgttgcatt tttggcccag caataacacg tgtttttgtg 3000
gtgtactaaa atggacagga atggagttat ttaagaggtg gcctcaccac tgtggatcgt 3060
gactatggtt ggatcaataa cattcaccat gacattggca cccatgttat ccaccatctt 3120
ttcccccaaa ttcctcatta tcacctcgtt gaagcggtac attttattgc ttattcacct 3180
aaaaacaata caattagtac atttgtttta tctcttggaa gttagtcatt ttcagttgca 3240
tgattctaat gctctctcca ttcttaaatc atgttttcac acccacttca tttaaaataa 3300
gaacgtgggt gttattttaa tttctattca ctaacatgag aaattaactt atttcaagta 3360
ataattttaa aatattttta tgctattatt ttattacaaa taattatgta tattaagttt 3420
attgatttta taataattat attaaaatta tatcgatatt aatttttgat tcactgatag 3480
tgttttatat tgttagtact gtgcatttat tttaaaattg gcataaataa tatatgtaac 3540
cagctcacta tactatactg ggagcttggt ggtgaaaggg gttcccaacc ctcctttcta 3600
ggtgtacatg ctttgatact tctggtacct tcttatatca atataaatta tattttgctg 3660
ataaaaaaac atggttaacc attaaattct ttttttaaaa aaaaaactgt atctaaactt 3720
tgtattatta aaaagaagtc tgagattaac aataaactaa cactcatttg aattcactgc 3780
agacacaagc agcaaaacca gttcttggag attactaccg tgagccagaa agatctgcgc 3840
cattaccatt tcatctaata aagtatttaa ttcagagtat gagacaagac cacttcgtaa 3900
gtgacactgg agatgttgtt tattatcaga ctgattctct gctcctccac tcgcaacgag 3960

actgagtttc aaacttttg ggttattatt tattggattc tagctactca aat 4013

<210> 3

<211> 2534

<212> DNA

<213> 大豆(Glycine max)

<400> 3

atggttcaag cacagcctct acaacatgtt ggtaatggtg cagggaaaga agatcaagct 60
tattttgatc caagtgctcc accacccttc aagattgcaa atatcagagc agcaattcca 120
aaacattgct gggagaagaa cacattgaga tctctgagtt atgttctgag ggatgtgttg 180
gtagtgactg cattggtagc tgcagcaatc ggcttcaata gctggttctt ctggccactc 240
tattggcctg cacaaggcac aatgttttgg gcacttttg ttcttggaca tgattggtaa 300
ctaattatta ttacaaattg ttatgttatg ttatgttatg ttgttgtgcc tttttctcag 360
tgatgcttta gtcatttcat ttcacttggt tatgcatgat tgttcgttca tatgttctgt 420
catggtgagt tctaatttga ttgatgcatg gaacagtggt catggaagtt tttcaaacag 480
tcctttgttg aacagcattg tgggccacat cttgcactct tcaattcttg taccatacca 540
tggatggtcg gttcctttta gcaacttttc atgttcactt tgtccttaaa tttttttta 600
tgtttgttaa aaaatctttg gtctgattta acaacctaac cattttaca actcatggat 660
tttttgcagg agaattagcc acaggactca ccatcagaac catggccatg ttgagaagga 720
tgaatcatgg gttccggtat tactatgagt ttgcttgatt aatttccaca ttttttcttt 780
cttcttaatt ttaatcagtg gttagatttg gttgtgttcc aatagaagaa aaggggtat 840
ctagagagat gtgaatttca tgaagtggtt catgattatg tgtctttatg cctttatgtc 900
agcttacaga gaaagtttac aagaatctag acaacatgac aagaatgatg agattcactc 960
ttcctttccc catctttgca tacccctttt atttggtgag accctctttt tccagaatga 1020
cagcattatt ttactatata gtacctcaat ttttatattt ctaaaatttt gaattcttga 1080
aattgaaagg aaaggacttt attgggtcta gcatctcact ctctctttgt gatatgaacc 1140
atatatttca gtggagcaga agccctggaa aagaaggctc tcatttcaac ccttacagca 1200
acttgttctc tcctggtgag agaagagatg tgctaacttc aactctgtgt tggggcatca 1260
tgctttctgt gcttctctat ctttccctca caatgggtcc actttttatg ctcaagctct 1320
atggggttcc ctatttggta atctcactct cacactttct ttatacatcg cacaccagtg 1380
tgggttattt gcaacctaca ccgaagtaat gccctataat taatggggtt aacacatgtc 1440
caagtccaat attttgttca cttatttgaa cttgaacatg tgtagatctt cgtcatgtgg 1500
ctggatttcg tcacgtactt gcatcatcat ggttacaagc agaaactgcc ttggtaccgt 1560
ggccaggtat cccatttaac acaatttgtt tcattaacat tttaagagaa ttttttttc 1620
aaaatagttt tcgaaattaa gcaaatacca agcaaattgt tagatctacg cttgtacttg 1680
ttttaaagtc aaattcatga ccaaattgtc ctcacaagtc caaaccgtcc actattttat 1740
tttcacctac tttatagccc aatttgtcat ttggttactt cagaaaagag aaccccattt 1800
gtagtaaata tattatttat gaattatggt agtttcaaca taaaacatat ttatgtgcag 1860
ttttgccatc cttcaaaaga agatagaaac ttactccatg ttactctgtc tatatgtaat 1920
ttcacaggaa tggagttatc taaggggtgg tcttacaaca gtagatcgcg actatggttg 1980

gatcaacaac attcaccatg acattggcac ccatgttatc catcaccttt tccctcaaat 2040
tccacattat catttaatcg aagcggtatt aattctctat ttcacaagaa attattgtat 2100
gtctgcctat gtgatctaag tcaattttca cataacacat gatcaaactt tcttaattct 2160
ttcttctaaa ttgaaaaagt ggattatatg tcaattgaaa attggtcaag accacaaaca 2220
tgtgatgatc tcccacctta catataataa tttctcctat tctacaatca ataatccttc 2280
tatggtcctg aattgttcct ttctttttc attttcttat tcttttgtt gtcccacaat 2340
agactaaagc agcaaaggca gtgctaggaa agtattatcg tgagcctcag aaatctgggc 2400
cattgccact tcatctaata aagtacttgc tccacagcat aagtcaggat cacttcgtta 2460
gcgactctgg cgacattgtg tactaccaga ctgattccca gctccacaaa gattcttgga 2520
cccagtccaa ctaa 2534

<210> 4
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 4
tgggacacat cttgcat 17

<210> 5
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 5
ttgtgccata ccatgga 17

<210> 6
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 6
tgggccacat cttgcac 17

<210> 7
<211> 17

<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 7
ttgtaccata ccatgga 17
<210> 8
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 8
ttgctgggtc aagaatc 17
<210> 9
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 9
agttatgttc tcaggga 17
<210> 10
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 10
ttgctgggag aagaaca 17
<210> 11
<211> 17
<212> DNA
<213> 人工序列
<220>

<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 11
agttatgttc tgaggga 17
<210> 12
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 12
ttggcccatt caaggaa 17
<210> 13
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 13
tttgttcttg gacatga 17
<210> 14
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 14
ttgccccatt caaggca 17
<210> 15
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 15

ttggcctgca caaggca 17

<210> 16

<211> 17

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释="人工序列:合成寡核苷酸的描述"

<400> 16

ttagccacag aactcac 17

<210> 17

<211> 17

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释="人工序列:合成寡核苷酸的描述"

<400> 17

cacattgaga aggatga 17

<210> 18

<211> 17

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释="人工序列:合成寡核苷酸的描述"

<400> 18

ttagccacag gactcac 17

<210> 19

<211> 17

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释="人工序列:合成寡核苷酸的描述"

<400> 19

catgttgaga aggatga 17

<210> 20

<211> 1131

<212> DNA
<213> 大豆(Glycine max)
<400> 20
atggttaaag acacaaagcc tttagcctat gctgctaata atggatacca aaaggaagct 60
tttgatccca gtgctcctcc accgtttaag attgcagaaa tcagagttgc aataccaaaa 120
cattgctggg tcaagaatcc atggagatcc ctcagttatg ttctcaggga tgtgcttgta 180
attgctgcat tgatggctgc tgcaagtcac ttcaacaact ggcttctctg gctaatctat 240
tggcccattc aaggaacaat gttctgggct ctgtttgttc ttggacatga ttgtggccat 300
ggaagctttt cagacagccc ttttctaaat agcctggtgg gacacatctt gcattcctca 360
attcttgtgc cataccatgg atggagaatt agccacagaa ctcaccatca aaatcatgga 420
cacattgaga aggatgaatc ctgggttcca ttaaccgaga agatttacaa gaatctagac 480
aacatgacaa gacttgttag attcactgtg ccatttccat tgtttgtgta tccaatttat 540
ttgttctcaa gaagccccgg aaaggaaggt tctcacttca atccctacag caatctgttc 600
ccacccagtg agagaaaggg aatagcaata tcaacactgt gttgggttac catgttttct 660
atgcttatct atctctcctt cataactagt ccagttctat tgctcaagct ctatggaatt 720
ccatattgga tatttgttat gtggctggac tttgtcacat acttgcatca ccatggtcat 780
catcagaaac tgccttggta tcgcggcaag gaatggagtt atttaagagg tggtctcaca 840
actgtggatc gtgactatgg ttggatcaat aacattcacc atgacattgg cacccatgtt 900
attcaccatc ttttccctca aattcctcat tatcacctcg ttgaagcgac acaagcagca 960
aaatcagttc ttggagagta ttaccgtgag ccagaaagat ctgcaccatt accatttcat 1020
ctaataaagt atttaattca gagtatgaga caagaccact tcgtaagtga cactggagat 1080
gtggtttatt atcagactga ttctctgcac cttcactcgc accgagactg a 1131
<210> 21
<211> 1143
<212> DNA
<213> 大豆(Glycine max)
<400> 21
atggttaaag acacaaagcc tttagcctat gctgccaata atggatacca acaaaagggt 60
tcttcttttg attttgatcc tagcgctcct ccaccgttta agattgcaga aatcagagct 120
tcaataccaa aacattgctg ggtcaagaat ccatggagat ccctcagtta tgttctcagg 180
gatgtgcttg taattgctgc attggtggct gcagcaattc acttcgacaa ctggcttctc 240
tggctaatct attgccccat tcaaggcaca atgttctggg ctctctttgt tcttggacat 300
gattgtggcc atggaagctt ttcagatagc cctttgctga atagcctggt gggacacatc 360
ttgcattcct caattcttgt gccataccat ggatggagaa ttagccacag aactcaccat 420
caaaaccatg gacacattga gaaggatgag tcatgggttc cattaacaga gaagatttac 480
aagaatctag acagcatgac aagactcatt agattcactg tgccatttcc attgtttgtg 540
tatccaattt atttgttttc aagaagcccc ggaaaggaag gctctcactt caatccctac 600
agcaatctgt ttccacccag tgagagaaaa ggaatagcaa tatcaacact gtgttgggct 660
accatgtttt ctctgcttat ctatctctca ttcataacta gtccacttct agtgctcaag 720

ctctatggaa ttccatattg gatatttgtt atgtggctgg actttgtcac atacttgcat 780
caccatggtc accaccagaa actgccttgg taccgcggca aggaatggag ttatttaaga 840
ggtggcctca ccactgtgga tcgtgactat ggttggatca ataacattca ccatgacatt 900
ggcacccatg ttatccacca tcttttcccc caaattcctc attatcacct cgttgaagcg 960
acacaagcag caaaaccagt tcttggagat tactaccgtg agccagaaag atctgcgcca 1020
ttaccatttc atctaataaa gtatttaatt cagagtatga gacaagacca cttcgtaagt 1080
gacactggag atgttgttta ttatcagact gattctctgc tcctccactc gcaacgagac 1140
tga 1143

<210> 22
<211> 1143
<212> DNA
<213> 大豆(Glycine max)
<400> 22

atggttcaag cacagcctct acaacatgtt ggtaatggtg cagggaaaga agatcaagct 60
tattttgatc caagtgctcc accacccttc aagattgcaa atatcagagc agcaattcca 120
aaacattgct gggagaagaa cacattgaga tctctgagtt atgttctgag ggatgtgttg 180
gtagtgactg cattggtagc tgcagcaatc ggcttcaata gctggttctt ctggccactc 240
tattggcctg cacaaggcac aatgttttgg gcacttttg ttcttggaca tgattgtggt 300
catggaagtt tttcaaacag tcctttgttg aacagcattg tgggccacat cttgcactct 360
tcaattcttg taccatacca tggatggaga attagccaca ggactcacca tcagaaccat 420
ggccatgttg agaaggatga atcatgggtt ccgcttacag agaaagttta caagaatcta 480
gacaacatga caagaatgat gagattcact cttcctttcc ccatctttgc ataccccttt 540
tatttgtgga gcagaagccc tggaaaagaa ggctctcatt tcaaccctta cagcaacttg 600
ttctctcctg gtgagagaag agatgtgcta acttcaactc tgtgttgggg catcatgctt 660
tctgtgcttc tctatctttc cctcacaatg ggtccacttt ttatgctcaa gctctatggg 720
gttccctatt tgatcttcgt catgtggctg gatttcgtca cgtacttgca tcatcatggt 780
tacaagcaga aactgccttg gtaccgtggc caggaatgga gttatctaag gggtggtctt 840
acaacagtag atcgcgacta tggttggatc aacaacattc accatgacat tggcacccat 900
gttatccatc acctttttccc tcaaattcca cattatcatt taatcgaagc gactaaagca 960
gcaaaggcag tgctaggaaa gtattatcgt gagcctcaga aatctgggcc attgccactt 1020
catctaataa agtacttgct ccacagcata agtcaggatc acttcgttag cgactctggc 1080
gacattgtgt actaccagac tgattcccag ctccacaaag attcttggac ccagtccaac 1140
taa 1143

<210> 23
<211> 1873
<212> DNA
<213> 大豆(Glycine max)
<400> 23

ccatatacta atatttgctt gtattgatag cccctccgtt cccaagagta taaaactgca 60

tcgaataata caagccacta ggcatggtaa attaaattgt gcctgcacct cgggatattt 120
catgtggggt tcatcatatt tgttgaggaa aagaaactcc cgaaattgaa ttatgcattt 180
atatatcctt tttcatttct agatttcctg aaggcttagg tgtaggcacc tagctagtag 240
ctacaatatc agcacttctc tctattgata aacaattggc tgtaatgccg cagtagagga 300
cgatcacaac atttcgtgct ggttactttt tgttttatgg tcatgatttc actctctcta 360
atctctccat tcattttgta gttgtcatta tctttagatt tttcactacc tggtttaaaa 420
ttgagggatt gtagttctgt tggtacatat tacacattca gcaaaacaac tgaaactcaa 480
ctgaacttgt ttatactttg acacagggtc tagcaaagga aacaacaatg ggaggtagag 540
gtcgtgtggc caaagtggaa gttcaaggga agaagcctct ctcaagggtt ccaaacacaa 600
agccaccatt cactgttggc caactcaaga aagcaattcc accacactgc tttcagcgct 660
ccctcctcac ttcattctcc tatgttgttt atgacctttc atttgccttc attttctaca 720
ttgccaccac ctacttccac ctccttcctc aacccttttc cctcattgca tggccaatct 780
attgggttct ccaaggttgc cttctcactg gtgtgtgggt gattgctcac gagtgtggtc 840
accatgcctt cagcaagtac caatgggttg atgatgttgt gggtttgacc cttcactcaa 900
cacttttagt cccttatttc tcatggaaaa taagccatcg ccgccatcac tccaacacag 960
gttcccttga ccgtgatgaa gtgtttgtcc caaaaccaaa atccaaagtt gcatggtttt 1020
ccaagtactt aaacaaccct ctaggaaggg ctgtttctct tctcgtcaca ctcacaatag 1080
ggtggcctat gtatttagcc ttcaatgtct ctggtagacc ctatgatagt tttgcaagcc 1140
actaccaccc ttatgctccc atatattcta accgtgagag gcttctgatc tatgtctctg 1200
atgttgcttt gttttctgtg acttactctc tctaccgtgt tgcaaccctg aaagggttgg 1260
tttggctgct atgtgtttat ggggtgcctt tgctcattgt gaacggtttt cttgtgacta 1320
tcacatattt gcagcacaca cactttgcct tgcctcatta cgattcatca gaatgggact 1380
ggctgaaggg agctttggca actatggaca gagattatgg gattctgaac aaggtgtttc 1440
atcacataac tgatactcat gtggctcacc atctcttctc tacaatgcca cattaccatg 1500
caatggaggc aaccaatgca atcaagccaa tattgggtga gtactaccaa tttgatgaca 1560
caccatttta caaggcactg tggagagaag cgagagagtg cctctatgtg gagccagatg 1620
aaggaacatc cgagaagggc gtgtattggt acaggaacaa gtattgatgg agcaaccaat 1680
gggccatagt gggagttatg gaagttttgt catgtattag tacataatta gtagaatgtt 1740
ataaataagt ggatttgccg cgtaatgact ttgtgtgtat tgtgaaacag cttgttgcga 1800
tcatggttat aatgtaaaaa taattctggt attaattaca tgtggaaagt gttctgctta 1860
tagctttctg cct 1873

<210> 24
<211> 3401
<212> DNA
<213> 大豆(Glycine max)
<400> 24

tttttctct gttgaaattg aaagtgtttt ccgggaaata ataaaaataa attaaaatct 60
tacacactct aggtaggtac ttctaattta atccacactt tgactctata tatgttttaa 120
aaataattat aatgcgtact tactttctca ttatactaaa tttaacatcg atgattttat 180

tttctgtttc tcttctttcc acctacatac atcccaaaat ttagggtgca attttaagtt 240
tattaacaca tgtttttagc tgcatgctgc ctttgtgtgt gctcaccaaa ttgcattctt 300
ctctttatat gttgtatttg aattttcaca ccatatgtaa acaagattac gtacgtgtcc 360
atgatcaaat acaaatgctg tcttatactg gcaatttgat aaacagccgt ccatttttc 420
tttttctctt taactatata tgctctagaa tctctgaaga ttcctctgcc atcgaatttc 480
tttcttggta acaacgtcgt cgttatgtta ttattttatt ctatttttat tttatcatat 540
atatttctta ttttgttcga agtatgtcat attttgatcg tgacaattag attgtcatgt 600
aggagtagga atatcacttt aaaacattga ttagtctgta ggcaatattg tcttcttttt 660
cctcctttat taatatattt tgtcgaagtt ttaccacaag gttgattcgc tttttttgtc 720
cctttctctt gttcttttta cctcaggtat tttagtcttt catggattat aagatcactg 780
agaagtgtat gcatgtaata ctaagcacca tagctgttct gcttgaattt atttgtgtgt 840
aaattgtaat gtttcagcgt tggctttccc tgtagctgct acaatggtac tgtatatcta 900
ttttttgcat tgttttcatt ttttctttta cttaatcttc attgctttga aattaataaa 960
acaatataat atagtttgaa ctttgaacta ttgcctattc atgtaattaa cttattcact 1020
gactcttatt gtttttctgg tagaattcat tttaaattga aggataaatt aagaggcaat 1080
acttgtaaat tgacctgtca taattacaca ggaccctgtt ttgtgccttt ttgtctctgt 1140
ctttggtttt gcatgttagc ctcacacaga tatttagtag ttgttctgca tacaagcctc 1200
acacgtatac taaaccagtg gacctcaaag tcatggcctt acacctattg catgcgagtc 1260
tgtgacacaa cccctggttt ccatattgca atgtgctacg ccgtcgtcct tgtttgtttc 1320
catatgtata ttgataccat caaattatta tatcatttat atggtctgga ccattacgtg 1380
tactctttat gacatgtaat tgagtttttt aattaaaaaa atcaatgaaa tttaactacg 1440
tagcatcata tagagataat tgactagaaa tttgatgact tattctttcc taatcatatt 1500
ttcttgtatt gatagccccg ctgtcccttt taaactcccg agagagtata aaactgcatc 1560
gaatattaca agatgcactc ttgtcaaatg aaggggggga aatgatacta caagccacta 1620
ggcatggtat gatgctaaat taaattgtgc ctgcacccca ggatatttca tgtgggattc 1680
atcatttatt gaggaaaact ctccaaattg aatcgtgcat ttatattttt ttttccattt 1740
ctagatttct tgaaggctta tggtataggc acctacaatt atcagcactt ctctctattg 1800
ataaacaatt ggctgtaata ccacagtaga gaacgatcac aacattttgt gctggttact 1860
ttttgtttta tggtcatgat ttcactctct ctaatctgtc acttccctcc attcattttg 1920
tacttctcat atttttcact tcctggttga aaattgtagt tctcttggta catactagta 1980
ttagacattc agcaacaaca actgaactga acttctttat actttgacac agggtctagc 2040
aaaggaaaca ataatgggag gtggaggccg tgtggccaaa gttgaaattc agcagaagaa 2100
gcctctctca agggttccaa acacaaagcc accattcact gttggccaac tcaagaaagc 2160
cattccaccg cactgctttc agcgttccct cctcacttca ttgtcctatg ttgtttatga 2220
cctttcattg gctttcattt tctacattgc caccacctac ttccacctcc tccctcaccc 2280
cttttccctc attgcatggc caatctattg ggttctccaa ggttgcattc ttactggcgt 2340
gtgggtgatt gctcacgagt gtggtcacca tgccttcagc aagtacccat gggttgatga 2400
tgttatgggt ttgaccgttc actcagcact tttagtccct tatttctcat ggaaaataag 2460
ccatcgccgc caccactcca acacgggttc ccttgaccgt gatgaagtgt ttgtcccaaa 2520

accaaaatcc aaagttgcat ggtacaccaa gtacctgaac aaccctctag gaagggctgc 2580
ttctcttctc atcacactca caatagggtg gcctttgtat ttagccttca atgtctctgg 2640
cagaccctat gatggttttg ctagccacta ccacccttat gctcccatat attcaaatcg 2700
tgagaggctt ttgatctatg tctctgatgt tgctttgttt tctgtgactt acttgctcta 2760
ccgtgttgca actatgaaag ggttggtttg gctgctatgt gtttatgggg tgccattgct 2820
cattgtgaac ggttttcttg tgaccatcac atatctgcag cacacacact atgccttgcc 2880
tcactatgat tcatcagaat gggattggct gaggggtgct ttggcaacta tggacagaga 2940
ttatggaatt ctgaacaagg tgtttcacca cataactgat actcatgtgg ctcaccatct 3000
tttctctaca atgccacatt accatgcaac ggaggcaacc aatgcaatga agccaatatt 3060
gggtgagtac taccgatttg atgacacacc attttacaag gcactgtgga gagaagcaag 3120
agagtgcctc tatgtggagc cagatgaagg aacatccgag aagggcgtgt attggtacag 3180
gaacaagtat tgatgaacca agcaatgggc catagtggga gttatggaag ttttgtcact 3240
tatcacttaa ttagtagaat gttataaata agtggatttg ccgcgtaatg acttgtgtgc 3300
attgtgaaac agcttgtagc gatccatggc tataatgtaa aaatatgtgg aaagtgttct 3360
gcttataact ttcttcctta aatgcacact ccatagaaca a 3401

<210> 25
<211> 1164
<212> DNA
<213> 大豆(Glycine max)
<400> 25

atgggtctag caaaggaaac aacaatggga ggtagaggtc gtgtggccaa agtggaagtt 60
caagggaaga agcctctctc aagggttcca aacacaaagc caccattcac tgttggccaa 120
ctcaagaaag caattccacc acactgcttt cagcgctccc tcctcacttc attctcctat 180
gttgtttatg acctttcatt tgccttcatt ttctacattg ccaccaccta cttccacctc 240
cttcctcaac ccttttccct cattgcatgg ccaatctatt gggttctcca aggttgcctt 300
ctcactggtg tgtgggtgat tgctcacgag tgtggtcacc atgccttcag caagtaccaa 360
tgggttgatg atgttgtggg tttgaccctt cactcaacac ttttagtccc ttatttctca 420
tggaaaataa gccatcgccg ccatcactcc aacacaggtt cccttgaccg tgatgaagtg 480
tttgtcccaa aaccaaaatc caaagttgca tggttttcca agtacttaaa caaccctcta 540
ggaagggctg tttctcttct cgtcacactc acaatagggt ggcctatgta tttagccttc 600
aatgtctctg gtagacccta tgatagtttt gcaagccact accaccctta tgctcccata 660
tattctaacc gtgagaggct tctgatctat gtctctgatg ttgctttgtt ttctgtgact 720
tactctctct accgtgttgc aaccctgaaa gggttggttt ggctgctatg tgtttatggg 780
gtgcctttgc tcattgtgaa cggttttctt gtgactatca catatttgca gcacacacac 840
tttgccttgc ctcattacga ttcatcagaa tgggactggc tgaagggagc tttggcaact 900
atggacagag attatgggat tctgaacaag gtgtttcatc acataactga tactcatgtg 960
gctcaccatc tcttctctac aatgccacat taccatgcaa tggaggcaac caatgcaatc 1020
aagccaatat tgggtgagta ctaccaattt gatgacacac catttacaa ggcactgtgg 1080
agagaagcga gagagtgcct ctatgtggag ccagatgaag gaacatccga gaagggcgtg 1140

tattggtaca ggaacaagta ttga 1164

<210> 26

<211> 1164

<212> DNA

<213> 大豆(Glycine max)

<400> 26

atgggtctag caaaggaaac aataatggga ggtggaggcc gtgtggccaa agttgaaatt 60
cagcagaaga agcctctctc aagggttcca aacacaaagc caccattcac tgttggccaa 120
ctcaagaaag ccattccacc gcactgcttt cagcgttccc tcctcacttc attgtcctat 180
gttgtttatg acctttcatt ggctttcatt ttctacattg ccaccaccta cttccacctc 240
ctccctcacc ccttttccct cattgcatgg ccaatctatt gggttctcca aggttgcatt 300
cttactggcg tgtgggtgat tgctcacgag tgtggtcacc atgccttcag caagtaccca 360
tgggttgatg atgttatggg tttgaccgtt cactcagcac ttttagtccc ttatttctca 420
tggaaaataa gccatcgccg ccaccactcc aacacgggtt cccttgaccg tgatgaagtg 480
tttgtcccaa aaccaaaatc caaagttgca tggtacacca agtacctgaa caaccctcta 540
ggaagggctg cttctcttct catcacactc acaatagggt ggcctttgta tttagccttc 600
aatgtctctg gcagacccta tgatggtttt gctagccact accaccctta tgctcccata 660
tattcaaatc gtgagaggct tttgatctat gtctctgatg ttgctttgtt ttctgtgact 720
tacttgctct accgtgttgc aactatgaaa gggttggttt ggctgctatg tgtttatggg 780
gtgccattgc tcattgtgaa cggttttctt gtgaccatca catatctgca gcacacacac 840
tatgccttgc ctcactatga ttcatcagaa tgggattggc tgaggggtgc tttggcaact 900
atggacagag attatggaat tctgaacaag gtgtttcacc acataactga tactcatgtg 960
gctcaccatc ttttctctac aatgccacat taccatgcaa cggaggcaac caatgcaatg 1020
aagccaatat tgggtgagta ctaccgattt gatgacacac cattttacaa ggcactgtgg 1080
agagaagcaa gagagtgcct ctatgtggag ccagatgaag gaacatccga gaagggcgtg 1140
tattggtaca ggaacaagta ttga 1164

<210> 27

<211> 17

<212> DNA

<213> 人工序列

<220>

<221> 来源

<223> /注释="人工序列:合成寡核苷酸的描述"

<400> 27

tctcaagggt tccaaac 17

<210> 28

<211> 17

<212> DNA

<213> 人工序列

<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 28
ttgagttggc caacagt 17
<210> 29
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 29
tatgttgttt atgacct 17
<210> 30
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 30
tggtggcaat gtagaaa 17
<210> 31
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 31
tcctatgttg tttatga 17
<210> 32
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"

<400> 32
tggcaatgta gaaaatg 17
<210> 33
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 33
ttgtttatga cctttca 17
<210> 34
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 34
taggtggtgg caatgta 17
<210> 35
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 35
tgccaccacc tacttcc 17
<210> 36
<211> 17
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 36
tgcaatgagg gaaaagg 17
<210> 37

<211> 366
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列: 合成多核苷酸的描述"
<400> 37
atgggtctag caaaggaaac aacaatggga ggtagaggtc gtgtggccaa agtggaagtt 60
caagggaaga agcctctctc aagggttcca aacacaaagc caccattcac tgttggccaa 120
ctcaagaaag caattccacc acactgcttt cagcgctccc tcctcacttc attctcctat 180
gttgtttatg acctttcatt tgcttttcta cattgccacc acctacttcc acctccttcc 240
tcaacccttt tccctcattg catggccaat ctattgggtt ctccaaggtt gccttctcac 300
tggtgtgtgg gtgattgctc acgagtgtgg tcaccatgcc ttcagcaagt accaatgggt 360
tgatga 366
<210> 38
<211> 360
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列: 合成多核苷酸的描述"
<400> 38
atgggtctag caaaggaaac aacaatggga ggtagaggtc gtgtggccaa agtggaagtt 60
caagggaaga agcctctctc aagggttcca aacacaaagc caccattcac tgttggccaa 120
ctcaagaaag caattccacc acactgcttt cagcgctccc tcctcacttc attctcctat 180
gttgtttatg acctttcatt tctacattgc caccacctac ttccacctcc ttcctcaacc 240
cttttccctc attgcatggc caatctattg ggttctccaa ggttgccttc tcactggtgt 300
gtgggtgatt gctcacgagt gtggtcacca tgccttcagc aagtaccaat gggttgatga 360
<210> 39
<211> 360
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列: 合成多核苷酸的描述"
<400> 39
atgggtctag caaaggaaac aataatggga ggtggaggcc gtgtggccaa agttgaaatt 60
cagcagaaga agcctctctc aagggttcca aacacaaagc caccattcac tgttggccaa 120
ctcaagaaag ccattccacc gcactgcttt cagcgttccc tcctcacttc attgtcctat 180

gttgtttatg acctttcatt tctacattgc caccacctac ttccacctcc tccctcaccc 240
cttttccctc attgcatggc caatctattg ggttctccaa ggttgcattc ttactggcgt 300
gtgggtgatt gctcacgagt gtggtcacca tgccttcagc aagtacccat gggttgatga 360
<210> 40
<211> 423
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列: 合成多核苷酸的描述"
<400> 40
atggttaaag acacaaagcc tttagcctat gctgctaata atggatacca aaaggaagct 60
tttgatccca gtgctcctcc accgtttaag attgcagaaa tcagagttgc aataccaaaa 120
cattgctggg tcaagaatcc atggagatcc ctcagttatg ttctcaggga tgtgcttgta 180
attgctgcat tgatggctgc tgcaagtcac ttcaacaact ggcttctctg gctaatctat 240
tggcccattc aaggaacaat gttctgggct ctgtttgttc ttggacatga ttgtggccat 300
ggaagctttt cagacagccc ttttctaaat agcctggtgg gacacatctt gcattcctca 360
attcttgtgc cataccatgg atggagaatt agccacagaa ctcaccatca atggacacat 420
tga 423
<210> 41
<211> 444
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列: 合成多核苷酸的描述"
<400> 41
atggttaaag acacaaagcc tttagcctat gctgctaata atggatacca aaaggaagct 60
tttgatccca gtgctcctcc accgtttaag attgcagaaa tcagagttgc aataccaaaa 120
cattgctggg tcaagaatcc atggagatcc ctcagttatg ttctcaggga tgtgcttgta 180
attgctgcat tgatggctgc tgcaagtcac ttcaacaact ggcttctctg gctaatctat 240
tggcccattc aaggaacaat gttctgggct ctgtttgttc ttggacatga ttgtggccat 300
ggaagctttt cagacagccc ttttctaaat agcctggtgg gacacatctt gcattcctca 360
attcttgtgc cataccatgg atggagaatt agccacagaa ctcacctgga cacattgaga 420
aggatgaatc ctgggttcca ttaa 444
<210> 42
<211> 435
<212> DNA
<213> 人工序列

<220>
<221> 来源
<223> /注释="人工序列: 合成多核苷酸的描述"
<400> 42
atggttaaag acacaaagcc tttagcctat gctgccaata atggatacca acaaaagggt 60
tcttcttttg attttgatcc tagcgctcct ccaccgttta agattgcaga aatcagagct 120
tcaataccaa aacattgctg ggtcaagaat ccatggagat ccctcagtta tgttctcagg 180
gatgtgcttg taattgctgc attggtggct gcagcaattc acttcgacaa ctggcttctc 240
tggctaatct attgccccat tcaaggcaca atgttctggg ctctctttgt tcttggacat 300
gattgtggcc atggaagctt ttcagatagc cctttgctga atagcctggt gggacacatc 360
ttgcattcct caattcttgt gccataccat ggatggagaa ttagccacag aactcaccat 420
caatggacac attga 435
<210> 43
<211> 432
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成多核苷酸的描述"
<400> 43
atggttaaag acacaaagcc tttagcctat gctgccaata atggatacca acaaaagggt 60
tcttcttttg attttgatcc tagcgctcct ccaccgttta agattgcaga aatcagagct 120
tcaataccaa aacattgctg ggtcaagaat ccatggagat ccctcagtta tgttctcagg 180
gatgtgcttg taattgctgc attggtggct gcagcaattc acttcgacaa ctggcttctc 240
tggctaatct attgccccat tcaaggcaca atgttctggg ctctctttgt tcttggacat 300
gattgtggcc atggaagctt ttcagatagc cctttgctga atagcctggt gggacacatc 360
ttgcattcct caattcttgt gccataccat ggatggagaa ttagccacag aactcaccat 420
caaacacatt ga 432
<210> 44
<211> 468
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列: 合成多核苷酸的描述"
<400> 44
atggttcaag cacagcctct acaacatgtt ggtaatggtg cagggaaaga agatcaagct 60
tattttgatc caagtgctcc accacccttc aagattgcaa atatcagagc agcaattcca 120
aaacattgct gggagaagaa cacattgaga tctctgagtt atgttctgag ggatgtgttg 180

gtagtgactg cattggtagc tgcagcaatc ggcttcaata gctggttctt ctggccactc 240
tattggcctg cacaaggcac aatgttttgg gcacttttg ttcttggaca tgattgtggt 300
catggaagtt tttcaaacag tcctttgttg aacagcattg tgggccacat cttgcactct 360
tcaattcttg taccatacca tggatggaga attagccaca ggactatggc catgttgaga 420
aggatgaatc atgggttccg cttacagaga aagtttacaa gaatctag 468

<210> 45
<211> 84
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 45
tttcarayag ycctttkytr aayagcmtkg tgggmcacat cttgcaytcy tcaattcttg 60
trccatacca tggatggtyr gttc 84

<210> 46
<211> 84
<212> DNA
<213> 大豆(Glycine max)
<400> 46
tttcagacag ccctttttcta aatagcctgg tgggacacat cttgcattcc tcaattcttg 60
tgccatacca tggatggtta gttc 84

<210> 47
<211> 84
<212> DNA
<213> 大豆(Glycine max)
<400> 47
tttcagatag ccctttgctg aatagcctgg tgggacacat cttgcattcc tcaattcttg 60
tgccatacca tggatggtta gttc 84

<210> 48
<211> 84
<212> DNA
<213> 大豆(Glycine max)
<400> 48
tttcaaacag tcctttgttg aacagcattg tgggccacat cttgcactct tcaattcttg 60
taccatacca tggatggtcg gttc 84

<210> 49
<211> 70
<212> DNA

<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 49
gywkcaatwc caaaacattg ctgggwsaag aaymcatkga gatcyctsag ttatgttcts 60
agggatgtgy 70
<210> 50
<211> 70
<212> DNA
<213> 大豆(Glycine max)
<400> 50
gttgcaatac caaaacattg ctgggtcaag aatccatgga gatccctcag ttatgttctc 60
agggatgtgc 70
<210> 51
<211> 70
<212> DNA
<213> 大豆(Glycine max)
<400> 51
gcttcaatac caaaacattg ctgggtcaag aatccatgga gatccctcag ttatgttctc 60
agggatgtgc 70
<210> 52
<211> 70
<212> DNA
<213> 大豆(Glycine max)
<400> 52
gcagcaattc caaaacattg ctgggagaag aacacattga gatctctgag ttatgttctg 60
agggatgtgt 70
<210> 53
<211> 60
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 53
tctattgscc yrywcaaggm acaatgttyt gggcwctbtt tgttcttgga catgattggt 60
<210> 54
<211> 60

<212> DNA
<213> 大豆(Glycine max)
<400> 54
tctattggcc cattcaagga acaatgttct gggctctgtt tgttcttgga catgattggt 60
<210> 55
<211> 60
<212> DNA
<213> 大豆(Glycine max)
<400> 55
tctattgccc cattcaaggc acaatgttct gggctctctt tgttcttgga catgattggt 60
<210> 56
<211> 60
<212> DNA
<213> 大豆(Glycine max)
<400> 56
tctattggcc tgcacaaggc acaatgtttt gggcactttt tgttcttgga catgattggt 60
<210> 57
<211> 72
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<220>
<221> 经修饰的_碱基
<222> (1)..(2)
<223> a, c, t, g, 未知或其他
<400> 57
nntttgcag gagaattagc cacagractc accatcaraa ycatggmcay rttgagaagg 60
atgartcmtg gg 72
<210> 58
<211> 72
<212> DNA
<213> 大豆(Glycine max)
<400> 58
gattttgcag gagaattagc cacagaactc accatcaaaa tcatggacac attgagaagg 60
atgaatcctg gg 72
<210> 59
<211> 72

<212> DNA
<213> 大豆(Glycine max)
<400> 59
gattttgcag gagaattagc cacagaactc accatcaaaa ccatggacac attgagaagg 60
atgagtcatg gg 72
<210> 60
<211> 70
<212> DNA
<213> 大豆(Glycine max)
<400> 60
ttttgcagga gaattagcca caggactcac catcagaacc atggccatgt tgagaaggat 60
gaatcatggg 70
<210> 61
<211> 56
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列:合成寡核苷酸的描述"
<400> 61
ctctctcaag ggttccaaac acaaagccac cattcactgt tggccaactc aagaaa 56
<210> 62
<211> 56
<212> DNA
<213> 大豆(Glycine max)
<400> 62
ctctctcaag ggttccaaac acaaagccac cattcactgt tggccaactc aagaaa 56
<210> 63
<211> 103
<212> DNA
<213> 人工序列
<220>
<221> 来源
<223> /注释="人工序列: 合成序列的描述"
<400> 63
ttcattstcc tatgttgttt atgacctttc attkgcyttc attttctaca ttgccaccac 60
ctacttccac ctcctycctc amccctttc cctcattgca tgg 103
<210> 64
<211> 103

<212> DNA
<213> 大豆(Glycine max)
<400> 64
ttcattctcc tatgttgttt atgacctttc atttgccttc attttctaca ttgccaccac 60
ctacttccac ctccttcctc aacccttttc cctcattgca tgg 103
<210> 65
<211> 103
<212> DNA
<213> 大豆(Glycine max)
<400> 65
ttcattgtcc tatgttgttt atgacctttc attggctttc attttctaca ttgccaccac 60
ctacttccac ctcctccctc acccctttc cctcattgca tgg 103

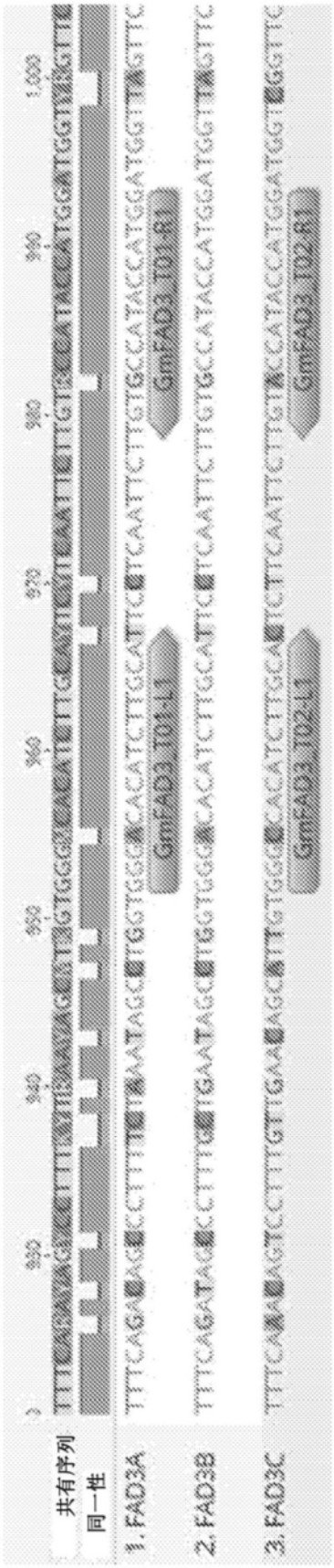

共有序列
同一性
1, FAD3A
2, FAD3B
3, FAD3C
GmFAD3_T01-L1
GmFAD3_T01-R1
GmFAD3_T02-L1
GmFAD3_T02-R1

图1A

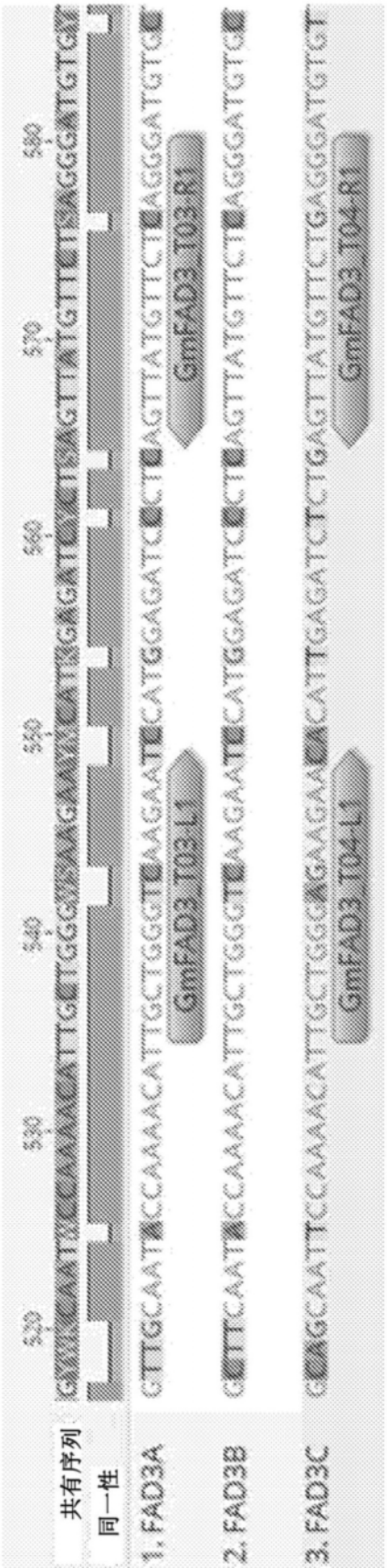

530
540
550
560
570
580
共有序列
同一性
1. FAD3A
2. FAD3B
3. FAD3C
GmFAD3_T03-L1
GmFAD3_T03-R1
GmFAD3_T04-L1
GmFAD3_T04-R1

图1B

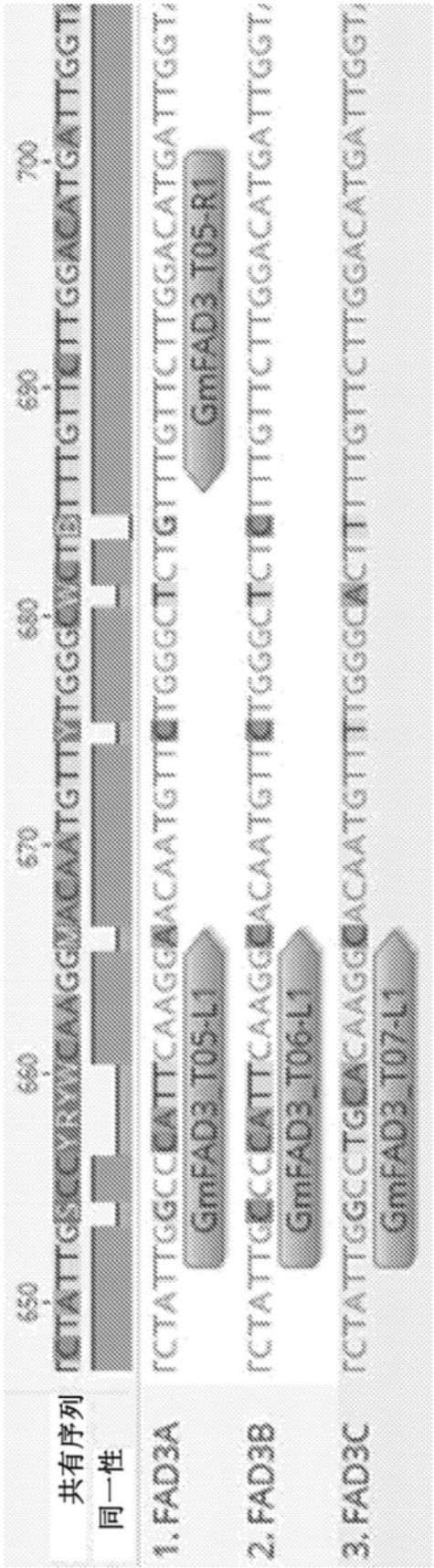

650
660
670
680
690
700
共有序列
同一性
1. FAD3A
2. FAD3B
3. FAD3C
GmFAD3_T05-L1
GmFAD3_T05-R1
GmFAD3_T06-L1
GmFAD3_T07-L1

图1C

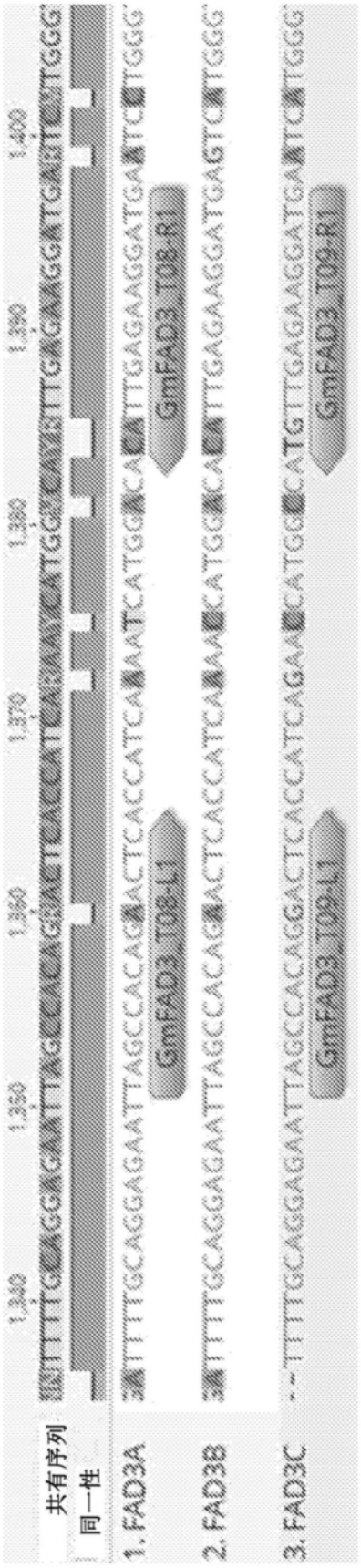

共有序列
同一性
1,340
1,350
1,360
1,370
1,380
1,390
1,400
1. FAD3A
2. FAD3B
3. FAD3C
GmFAD3_T08-L1
GmFAD3_T08-R1
GmFAD3_T09-L1
GmFAD3_T09-R1

图1D

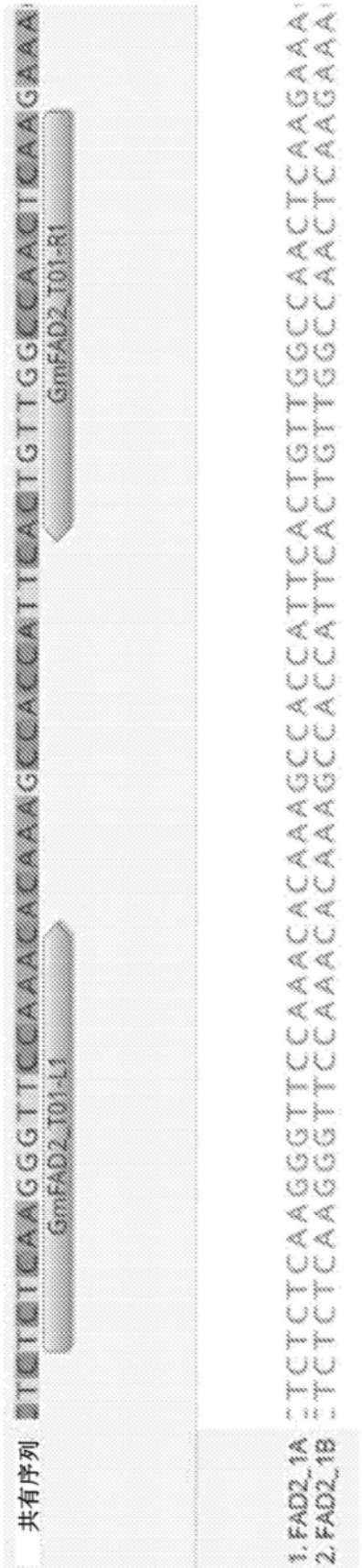


图2A

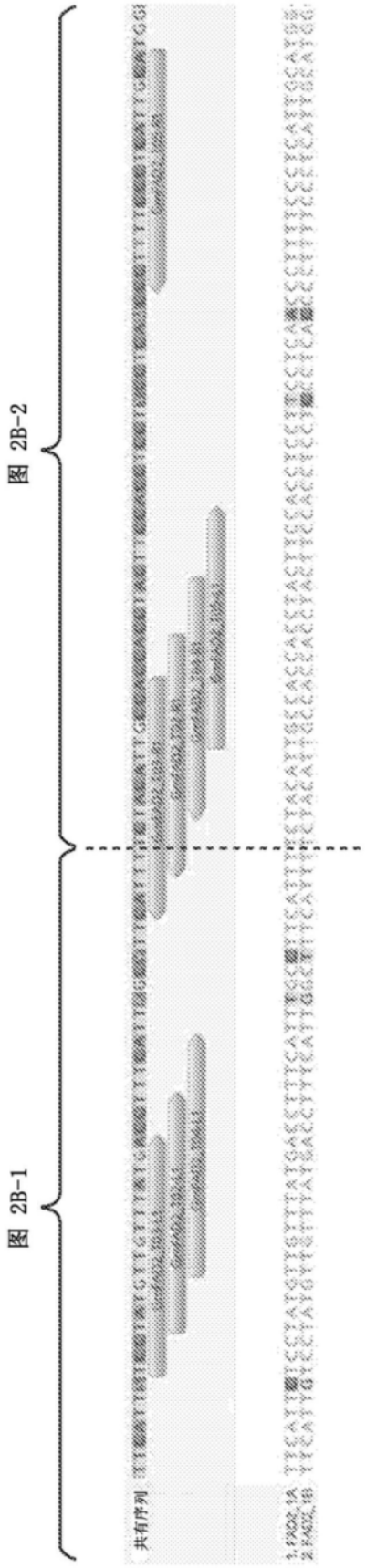


图2B


共有序列
GmFAD2_T03-L1
GmFAD2_T02-L1
GmFAD2_T04-L1
1.FAD2_1A
2.FAD2_1B

图2B-1


GmFAD2_T03-R1
GmFAD2_T02-R1
GmFAD2_T04-R1
GmFAD2_T05-L1
GmFAD2_T05-R1

图2B-2