

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200580048074.1

[51] Int. Cl.

C12N 15/82 (2006.01)

C12N 9/02 (2006.01)

A01H 5/00 (2006.01)

[43] 公开日 2008 年 2 月 6 日

[11] 公开号 CN 101120091A

[22] 申请日 2005.12.19

[21] 申请号 200580048074.1

[30] 优先权

[32] 2004.12.20 [33] US [31] 60/637,531

[86] 国际申请 PCT/US2005/046027 2005.12.19

[87] 国际公布 WO2006/073787 英 2006.7.13

[85] 进入国家阶段日期 2007.8.13

[71] 申请人 巴斯福植物科学有限公司

地址 德国路德维希港

[72] 发明人 H·黑特尔 J·吉布森 P·任

[74] 专利代理机构 北京市中咨律师事务所

代理人 黄革生 隗永良

权利要求书 4 页 说明书 94 页 序列表 114 页
附图 17 页

[54] 发明名称

编码植物脂肪酸去饱和酶基因的核酸分子及其使用方法

[57] 摘要

本发明一般性涉及与植物中种子贮存化合物的存在相关的蛋白质的编码核酸序列。更具体而言，本发明涉及脂质代谢调节蛋白质的 FAD2 样编码核酸序列，及这些序列在转基因植物中的用途。具体而言，本发明涉及脂质代谢相关化合物的操作方法，以及在植物和种子中提高油水平和改变脂肪酸组成的方法。本发明还涉及应用这些新的植物多肽刺激植物生长和/或提高产率和/或种子贮存化合物组成的方法。

1. 分离的多肽, 包含选自如下的氨基酸序列:

$X^1X^2X^3LX^4X^5X^6X^7LX^8X^9PX^{10}YL$, 其中 X^1 不为 M, 且 X^3 不为 T, 且 X^6 不为 F, 且 X^7 不为 V,

$GX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}X^{17}X^{18}HX^{19}X^{20}PX^{21}X^{22}X^{23}X^{24}X^{25}X^{26}X^{27}X^{28}ER$, 其中 X^{15} 不为 G, 且 X^{20} 不为 F, 且 X^{21} 不为 N, 且 X^{22} 不为 A,

$HX^{29}X^{30}PX^{31}X^{32}X^{33}X^{34}X^{35}X^{36}X^{37}X^{38}ER$, 其中 X^{30} 不为 F, 且 X^{31} 不为 N, 且 X^{32} 不为 A,

$LX^{39}X^{40}X^{41}X^{42}X^{43}X^{44}X^{45}GX^{46}X^{47}X^{48}X^{49}X^{50}X^{51}X^{52}YX^{53}X^{54}P$, 其中 X^{41} 不为 Y, 且 X^{45} 不为 Q, 且 X^{48} 不为 S, 且 X^{49} 不为 M, 且 X^{50} 不为 I,

$TX^{55}X^{56}X^{57}X^{58}HX^{59}X^{60}X^{61}X^{62}X^{63}X^{64}X^{65}X^{66}X^{67}X^{68}X^{69}X^{70}X^{71}T$, 其中 X^{67} 不为 N, 以及

$PX^{72}X^{73}X^{74}X^{75}X^{76}X^{77}X^{78}X^{79}X^{80}X^{81}X^{82}X^{83}X^{84}X^{85}X^{86}$, 其中 X^{84} 不为 W, 且 X^{85} 不为 Y, 且 X^{86} 不为 V,

且其中若无上述其他限定, X 代表任何氨基酸。

2. 分离的多肽, 包含选自如下的氨基酸序列:

AWYT

$PYX^{87}YNNPX^{88}GRLVHIX^{89}VQLTLGWPLYLAX^{90}NX^{91}SGRPYPREACHFD$
 $CHFD$
 $DPYGPIYNDRER$,

FISDVG

$ALX^{92}KLX^{93}SX^{94}FGFWVVRVYGVP$,

$ILGEYYQFDX^{95}TPVAKAT$,

且其中 X 代表任意氨基酸。

3. 分离的核酸序列, 编码包含权利要求 1 或 2 的氨基酸序列的蛋白质。

4. 分离的多肽, 由权利要求 3 的核酸序列编码。

5. 分离的核酸, 包含选自如下的多核苷酸序列:

如 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、

SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 所示的多核苷酸序列；

编码如 SEQ ID NO: 6、SEQ ID NO: 8、SEQ ID NO: 10、SEQ ID NO: 12、SEQ ID NO: 14、SEQ ID NO: 16、SEQ ID NO: 18、SEQ ID NO: 20、SEQ ID NO: 22、SEQ ID NO: 24、SEQ ID NO: 26、SEQ ID NO: 28、SEQ ID NO: 30、SEQ ID NO: 32、SEQ ID NO: 34 或 SEQ ID NO: 36 所示多肽的多核苷酸序列；

与上述 a) 或 b) 中核酸具有至少 70% 序列同一性的多核苷酸序列；

与上述 a) 或 b) 中核酸互补的多核苷酸序列；以及

在严紧条件下与上述 a) 或 b) 中核酸杂交的多核苷酸序列。

6. 分离的多肽，由权利要求 5 的多核苷酸序列编码。

7. 权利要求 5 的分离的核酸，其中所述分离的核酸编码作为微生物或植物中种子贮存化合物调节物起作用的多肽。

8. 权利要求 6 的分离的核酸，其中分离的 LMP 多肽序列作为微生物或植物中种子贮存化合物调节物。

9. 表达载体，含有权利要求 3 或 5 的核酸，其中所述核酸有效连接于选自种子特异性启动子、根特异性启动子及非组织特异性启动子的启动子。

10. 产生转基因植物的方法，所述转基因植物的种子贮存化合物重量百分比的水平与野生型相比发生改变，所述方法包括：

a. 第一步将含有核酸的表达载体引入植物细胞，和

b. 第二步从该植物细胞产生转基因植物，

其中所述核酸编码作为植物中种子贮存化合物调节物起作用的多肽，且其中所述核酸包含选自如下的多核苷酸序列：

a. 如 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO:

29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 所示的多核苷酸序列；

b. 编码如 SEQ ID NO: 6、SEQ ID NO: 8、SEQ ID NO: 10、SEQ ID NO: 12、SEQ ID NO: 14、SEQ ID NO: 16、SEQ ID NO: 18、SEQ ID NO: 20、SEQ ID NO: 22、SEQ ID NO: 24、SEQ ID NO: 26、SEQ ID NO: 28、SEQ ID NO: 30、SEQ ID NO: 32、SEQ ID NO: 34 或 SEQ ID NO: 36 所示多肽的多核苷酸序列；

c. 与上述 a)或 b)中核酸具有至少 70%序列同一性的多核苷酸序列；

d. 与上述 a)或 b)中核酸互补的多核苷酸序列；以及

e. 在严紧条件下与上述 a)或 b)中核酸杂交的多核苷酸序列。

11. 权利要求 10 的方法，其中的核酸包含与权利要求 5 中 a)或 b)的多核苷酸序列具有至少 90%序列同一性的多核苷酸序列。

12. 权利要求 10 的方法，其中转基因植物的油酸水平与该植物野生型品种相比提高。

13. 产生转基因植物的方法，所述转基因植物的油酸重量百分比水平与野生型相比提高，所述方法包括：

第一步用 RNA 前体构建体转化植物细胞，和

第二步从该植物细胞产生转基因植物，

其中所述构建体含有在植物细胞中驱动表达的启动子，该启动子有效连于编码前体微 RNA 序列的核苷酸序列，其中编码所述微 RNA 前体序列的核苷酸序列选自：

a. 如 SEQ ID NO: 47 所示的核苷酸序列

b. 与上述 a)中核酸具有至少 70%序列同一性的多核苷酸序列；

c. 与上述 a)中核酸互补的多核苷酸序列；以及

d. 在严紧条件下与上述 a)中核酸杂交的多核苷酸序列。

14. 权利要求 13 的方法，其中设计编码前体微 RNA 序列的核苷酸序列，从而使编码 SEQ ID NO: 37 所示微 RNA 的核苷酸序列被编码 SEQ ID NO: 40 所示微 RNA 的核苷酸序列所取代。

15. 分离的核酸，包含选自如下的多核苷酸序列：

如 SEQ ID NO: 47 所示的核苷酸序列；

与上述 a) 中核酸具有至少 70% 序列同一性的多核苷酸序列；

与上述 a) 中核酸互补的多核苷酸序列；以及

在严紧条件下与上述 a) 中核酸杂交的多核苷酸序列。

16. 调节植物中种子贮存化合物重量百分比水平的方法，所述方法包括：

a. 第一步将含有核酸的表达载体引入植物细胞，和

b. 第二步从该植物细胞产生转基因植物，

其中所述核酸编码作为植物中种子贮存化合物调节物起作用的多肽，且其中所述核酸包含权利要求 5 的多核苷酸序列。

17. 权利要求 16 的方法，其中油酸重量百分比的水平改变。

18. 通过权利要求 10、13 或 16 的方法制备的转基因植物。

19. 权利要求 18 的转基因植物，其中转基因植物的油酸重量百分比水平与该植物野生型品种相比提高。

20. 权利要求 18 的转基因植物，其中植物选自油菜籽、芸苔、亚麻籽、大豆、向日葵、玉米、燕麦、裸麦、大麦、小麦、胡椒、万寿菊、棉花、油椰、椰树、亚麻、蓖麻、甜菜、稻和花生。

21. 由权利要求 18 的转基因植物所产生的种子，其中所述植物表达作为种子贮存化合物调节物起作用的多肽，且其中所述植物为种子贮存化合物重量百分比水平与该植物野生型品种相比改变的真实遗传。

编码植物脂肪酸去饱和酶基因的核酸分子及其使用方法

本发明要求于2004年12月20日递交的美国临时申请60/637531的优先权，此处整体引用作为参考。

发明领域

此处描述了植物遗传工程领域的发明，包括编码脂肪酸去饱和酶2样（*FAD2*样）多肽的分离的核酸分子，以提高其农艺、园艺和品质性状。本发明一般性涉及与植物中种子贮存化合物的存在相关的蛋白质的编码核酸序列。更具体而言，本发明涉及脂质代谢调节蛋白质的*FAD2*样编码核酸序列，及这些序列在转基因植物中的用途。具体而言，本发明涉及脂质代谢相关化合物的操作方法，以及在植物和种子中提高油水平和改变脂肪酸组成的方法。本发明还涉及应用这些新植物多肽刺激植物生长和/或提高产率和/或种子贮存化合物组成的方法。

发明背景

对植物进行研究和遗传操作的历史已久，甚至始于格里戈·孟德尔（Gregor Mendel）的著名研究之前。从提高土豆块茎的淀粉含量到提高或改变油料种子植物如芸苔（canola）和向日葵的脂肪酸含量，在该学科的完善中，科学家们已经实现了对植物特定性状的改良。随着植物油的消费和使用不断增加，对于种子油含量和种子油水平的改良也日益广泛（如 Töpfer 等，1995, *Science* 268:681-686）。对转基因植物中生物合成路径的操作作为分子生物学家及植物生物化学家们提供了影响植物代谢从而产生特定的更高价值产品的许多机会。已改变了许多传统油料种子植物和非传统油料种子植物的种子油的产量或组成，传统油料种子植物如大豆（美国专利 No.

5,955,650)、芸苔(美国专利 No. 5,955,650)、向日葵(美国专利 No. 6,084,164)和油菜籽(Töpfer 等人, 1995, Science 268:681-686), 非传统油料种子植物如烟草(Cahoon 等, 1992, Proc. Natl. Acad. Sci. USA 89:11184-11188)。

植物种子油同时含有中性及极性脂质(见表1)。中性脂质主要包括甘油三酯, 为种子油体中累积的主要贮存脂质。极性脂质则主要见于种子细胞的多种膜内, 如内质网、微粒体膜和细胞膜。中性和极性脂质含有几种常见的脂肪酸(见表2)以及一系列较少见的脂肪酸。膜脂质中脂肪酸的组成受到高度调节, 其中仅存在有限几种脂肪酸。另一方面, 许多植物物种种子的中性贮存脂质中含有大量的少见脂肪酸(Van de Loo F.J.等人, 1993, Unusual Fatty Acids in Lipid Metabolism in Plants, 第91-126页, 编辑 TS Moore Jr. CRC Press; Millar 等, 2000, Trends Plant Sci. 5: 95-101)。

脂质由脂肪酸合成且其合成可分为两部分: 原核路径和真核路径(Browse 等人, 1986, Biochemical J. 235:25-31; Ohlrogge & Browse, 1995, Plant Cell 7: 957-970)。原核路径位于质体中, 后者为脂肪酸生物合成的初始位点。通过乙酰辅酶 A 羧化酶(ACCase)将乙酰辅酶 A 转化为丙二酰辅酶 A, 起始脂肪酸的生物合成。丙二酰辅酶 A 随后由丙二酰辅酶 A:ACP 酰基转移酶转移到丙二酰酰基载体蛋白(ACP)上。 β -酮酰基 ACP 合酶 III (KAS III)催化缩合反应, 其中将来自乙酰辅酶 A 的酰基基团转移到丙二酰 ACP 上形成 3-酮丁酰 ACP。在随后系列的缩合、还原以及脱水反应中, ACP 辅因子上新生的脂肪酸链通过一步步地添加(缩合)由丙二酰 ACP 所提供的两个碳原子而得以延长, 直至形成 16 或 18 碳的饱和脂肪酸链。质体 Δ^9 酰基 ACP 去饱和酶将第一个非饱和的双键引入脂肪酸。硫酯酶将脂肪酸从 ACP 辅因子上切下, 而游离脂肪酸被输出到胞质中, 在那里作为脂肪酰基辅酶 A 酯参与真核路径。在该路径中, 脂肪酸分别通过甘油-3-磷酸乙酰转移酶和溶血磷脂酸酰基转移酶酯化甘油-3-磷酸的 sn-1 和 sn-2 位, 产生磷脂酸(PA)。PA 是其他极性和中性脂质的前体, 其中后者生成于 Kennedy 路径(Voelker, 1996, Genetic Engineering, Setlow 编: 18:111-113;

Shanklin & Cahoon, 1998, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49:611-641; Frentzen, 1998, *Lipids* 100:161-166; Millar 等人, 2000, *Trends Plant Sci.* 5:95-101)。

种子中的贮存脂质由糖类衍生的前体合成。植物胞质溶胶中具有完整的糖酵解路径(Plaxton, 1996, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47:185-214), 且业已表明油菜籽质体中也存在完整的路径 (Kang & Rawsthorne, 1994, *Plant J.* 6:795-805)。蔗糖是碳和能源的主要来源, 其从叶转移至发育的种子中。在种子贮存期间, 蔗糖转化进胞质溶胶, 提供代谢前体葡萄糖-6-磷酸和丙酮酸。它们被转运进质体并转化为乙酰辅酶 A, 作为脂肪酸合成的主要前体。质体中的乙酰辅酶 A 是脂质生物合成的中心前体。乙酰辅酶 A 可通过不同的反应在质体中形成, 且各反应的确切作用仍在讨论中(Ohlrogge & Browse, 1995, *Plant Cell* 7:957-970)。然而人们认可大部分乙酰辅酶 A 来自从细胞质输入质体的葡萄糖-6-磷酸和丙酮酸。蔗糖产生于源器官(叶或光合作用发生的任何地方)并被转运至发育的种子, 后者也称为库器官。在发育的种子中, 蔗糖是所有贮存化合物即淀粉、脂质和部分种子贮存蛋白质的前体。因此, 显然蔗糖在其中起重要作用的糖类代谢对于种子贮存化合物的积累非常重要。

贮存化合物如三酰基甘油(种子油)是发芽和幼苗生长时使用的碳和能量储备。种子(植物)油也是人类饮食的重要成分和为化学工业提供原料的有价值的商品。

虽然传统植物育种方法可以改变种子油的脂质和脂肪酸含量和/或组成, 重组 DNA 技术的来临使得对植物种子油含量的操作更为简单, 某些情况下能通过仅靠育种法不能完成的方法改变种子油(见如 Töpfer 等人, 1995, *Science* 268:681-686)。例如, 向转基因烟草中引入 Δ^{12} 羟化酶核酸序列, 导致烟草种子油中引入新脂肪酸蓖麻油酸(Van de Loo 等人, 1995, *Proc. Natl. Acad. Sci USA* 92:6743-6747)。也通过引入和表达来自芜菁的酰基 ACP 去饱和酶, 遗传改造烟草植物, 使之产生低水平的岩芹酸(Cahoon 等人, 1992, *Proc. Natl. Acad. Sci USA* 89:11184-11188)。

改变植物的种子油含量具有重大的医学、营养和经济价值。就医学价值而言,可见于许多种子油的长链脂肪酸(C18及更长)与高胆固醇血症和其他冠心病相关临床疾病的减少相关(Brenner, 1976, Adv. Exp. Med. Biol. 83:85-101)。因此,消耗这些类型脂肪酸水平增高的植物可能会降低心脏疾病的风险。种子油含量的水平增高也增加了种子油的大规模生产,从而降低这些油的成本。

为了提高或改变植物中化合物如种子油的水平,必须鉴定调节脂质和脂肪酸代谢的核酸序列和蛋白质。如早先所述,已经克隆了若干去饱和酶核酸如 Δ^6 去饱和酶核酸、 Δ^{12} 去饱和酶核酸和酰基 ACP 去饱和酶核酸,并证实其编码多种植物物种中脂肪酸合成必需的酶((Miquel&Browse, in Seed Development and Germination, Galili 等人编, Marcel Dekker, New York, 169-193 页, 1994; Ohlrogge & Browse, 1995, Plant Cell 7:957-970)。还克隆了来自不同物种如芸苔、大豆、胡萝卜、松树和拟南芥(*Arabidopsis thaliana*)的油体蛋白(oleosin)核酸序列,并确定其编码与这些植物中油体磷脂单层膜相关的蛋白质。

尽管已知若干广泛影响植物和种子发育的化合物,仍明确需要特异地鉴定对于贮存化合物积累的发育调节更为特异的因子以及鉴定能够赋予其宿主植物和其他植物物种油产量改变或提高的基因。

本发明涉及的另一问题在于提供使脂肪酸去饱和酶沉默的更有效方法。本发明涉及的另一问题在于特异性修饰种子油的脂肪酸含量。本发明涉及的另一问题在于提高种子油的油酸含量。本发明涉及的另一问题在于降低种子油的亚油酸含量。

本发明公开了来自拟南芥、大豆(*Glycine max*)、稻(*Oryza sativa*)、玉米(*Zea mays*)、亚麻(*Linum usitatissimum*)、大麦(*Hordeum vulgare*)或小麦(*Triticum aestivum*)的核酸序列。这些核酸序列可用于改变或提高植物中种子贮存化合物如蛋白质、糖和油的水平,所述植物包括转基因植物,例如芸苔、亚麻子、大豆、向日葵、玉米、燕麦、裸麦、大麦、小麦、稻、胡椒、万寿菊、棉花、油椰、椰树、亚麻、蓖麻和花生,它们是含有

大量脂质化合物的油料种子植物。

发明概述

本发明提供了与植物种子贮存化合物代谢相关的新的分离的核酸和氨基酸序列，具体而言涉及 *FAD2* 样序列。

本发明的另一主题为包含选自如下的氨基酸序列的分离的多肽

a. $X^1X^2X^3X^4X^5X^6X^7LX^8X^9PX^{10}YL$ ，其中 X^1 不为 M，且 X^3 不为 T，且 X^6X^7 不为 FV，

b. $GX^{11}X^{12}X^{13}X^{14}X^{15}X^{16}X^{17}X^{18}HX^{19}X^{20}PX^{21}X^{22}X^{23}X^{24}X^{25}X^{26}X^{27}X^{28}ER$ ，其中 X^{15} 不为 G，且 X^{20} 不为 F，且 $X^{21}X^{22}$ 不为 NA，

c. $HX^{29}X^{30}PX^{31}X^{32}X^{33}X^{34}X^{35}X^{36}X^{37}X^{38}ER$ ，其中 X^{30} 不为 F，且 $X^{31}X^{32}$ 不为 NA，

d. $LX^{39}X^{40}X^{41}X^{42}X^{43}X^{44}X^{45}GX^{46}X^{47}X^{48}X^{49}X^{50}X^{51}X^{52}YX^{53}X^{54}P$ ，其中 X^{41} 不为 Y，且 X^{45} 不为 Q，且 X^{48} 不为 S，且 X^{49} 不为 M，且 X^{50} 不为 I，

e. $TX^{55}X^{56}X^{57}X^{58}HX^{59}X^{60}X^{61}X^{62}X^{63}X^{64}X^{65}X^{66}X^{67}X^{68}X^{69}X^{70}X^{71}T$ ，其中 X^{67} 不为 N，

f. $PX^{72}X^{73}X^{74}X^{75}X^{76}X^{77}X^{78}X^{79}X^{80}X^{81}X^{82}X^{83}X^{84}X^{85}X^{86}$ ，其中 $X^{84}X^{85}X^{86}$ 不为 WYV，

且其中若无上述其他限定，X 可代表任何氨基酸。

优选利用 Vector NTI Suite 9.0 的 Align X (2003 年 8 月 22 日)生成确定权利要求 1 中共有肽序列 a 至 f 的序列比对。用于多重比对的参数优选如下：空位开放罚分：10；空位衍生罚分：0.05；空位分离(separation)罚分范围：8；比对延迟(delay)的同一性百分比：40。

在本发明优选实施方案中，要求包含选自如下的氨基酸序列的分离的多肽

a. AWYPYX⁸⁷YX⁸⁸NPX⁸⁹GRLVHIX⁹⁰VQLTLGWPLYLAX⁹¹NX⁹²SGR
PYPRFACHFDPIYGPIYNDREER，

b. FISDVGIV，

c. ALX⁹³KLX⁹⁴SX⁹⁵FGFWVVRVYGVP,

d. ILGEYYQFDX⁹⁶TPVAKAT,

且其中 X 代表任意氨基酸。

优选利用 Vector NTI Suite 9.0 的 Align X (2003 年 8 月 22 日)生成确定权利要求 2 中共有肽序列 a 至 d 的序列比对。用于多重比对的参数优选如下：空位开放罚分：10；空位延伸罚分：0.05；空位分离罚分范围：8；比对延迟的同一性百分比：40。

本发明分离的多肽可包括 1 个、2 个、3 个、4 个、5 个或 6 个权利要求 1 的氨基酸序列。本发明分离的多肽可包括 1 个、2 个、3 个或 4 个权利要求 2 的氨基酸序列。若权利要求 1 中未另行指定，X 可代表任意氨基酸，特别是 G、A、V、L、I、F、Y、W、P、D、E、N、Q、S、T、C、M、K、R、H。

X¹ 不为 M。在优选实施方案中 X¹ 为选自 G、A、V、L、I、F、Y、W、P、D、E、N、Q、S、T、C、K、R 和 H 的氨基酸，在更优选的实施方案中选自 F、T 和 H，而在甚至更优选的实施方案中为 H。

X³ 不为 T。在优选实施方案中 X³ 为选自 G、A、V、L、I、F、Y、W、P、D、E、N、Q、S、C、M、K、R 和 H 的氨基酸，在更优选的实施方案中选自 L、A、V、F，而在甚至更优选的实施方案中为 V 或 F。

X⁶ 不为 F 且 X⁷ 不为 V。在优选实施方案中 X⁶ 和 X⁷ 为分别独立选自 G、A、L、I、Y、W、P、D、E、N、Q、S、T、C、M、K、R 和 H 的氨基酸，在更优选的实施方案中分别独立选自 P、L 和 T，而在甚至更优选的实施方案中 X⁶ 为 P 或 L，而 X⁷ 为 L 或 T，而在更优选的实施方案中 X⁶ 为 L 而 X⁷ 为 T。

X¹⁵ 不为 G。在优选实施方案中 X¹⁵ 为空白或选自 A、V、L、I、F、Y、W、P、D、E、N、Q、S、T、C、M、K、R 和 H 的氨基酸，在更优选的实施方案中 X¹⁵ 为 R 或空白，更优选的 X⁵ 为 R。空白是指该位点上没有氨基酸。

X²⁰ 不为 F。在优选实施方案中 X²⁰ 为选自 G、A、V、L、I、Y、W、

P、D、E、N、Q、S、T、C、M、K、R和H的氨基酸，在更优选的实施方案中为N或D，而在甚至更优选的实施方案中为D。

X^{21} 不为N且 X^{22} 不为A。在优选实施方案中 X^{21} 和 X^{22} 为分别独立选自G、V、L、I、F、Y、W、P、D、E、Q、S、T、C、M、K、R和H的氨基酸。在更优选的实施方案中分别独立选自D、Y、H、S、G、I。在甚至更优选的实施方案中 X^{21} 为D、Y或H，更优选Y，而 X^{22} 为S或G，更优选G。

X^{30} 不为F。在优选实施方案中 X^{30} 为选自G、A、V、L、I、Y、W、P、D、E、N、Q、S、T、C、M、K、R和H的氨基酸，在更优选的实施方案中 X^{30} 为N或D，而在甚至更优选的实施方案中 X^{30} 为D。

X^{31} 不为N且 X^{32} 不为A。在优选实施方案中 X^{31} 和 X^{32} 为分别独立选自G、V、L、I、F、Y、W、P、D、E、Q、S、T、C、M、K、R和H的氨基酸，在更优选的实施方案中分别独立选自D、Y、H、S、G。在甚至更优选的实施方案中 X^{31} 为D、Y或H，更优选Y，而 X^{32} 为S或G，更优选G。

X^{41} 不为Y。在优选实施方案中 X^{41} 为选自G、A、V、L、I、F、W、P、D、E、N、Q、S、T、C、M、K、R和H的氨基酸，在更优选的实施方案中 X^{41} 为L。

X^{45} 不为Q。在优选实施方案中 X^{45} 为选自G、A、V、L、I、F、Y、W、P、D、E、N、S、T、C、M、K、R和H的氨基酸，在更优选的实施方案中选自M、K和F，而在甚至更优选的实施方案中 X^{45} 为F。

X^{48} 不为S，且 X^{49} 不为M且 X^{50} 不为I。在优选实施方案中 X^{48} 、 X^{49} 和 X^{50} 为分别独立选自G、A、V、L、F、Y、W、P、D、E、N、Q、T、C、K、R和H的氨基酸，在更优选的实施方案中分别独立选自Q、W、L和V。在甚至更优选的实施方案中 X^{48} 为Q或W，更优选 X^{48} 为W，而 X^{49} 、 X^{50} 分别独立选自L或V，更优选 X^{49} 、 X^{50} 分别独立选自V。

X^{67} 不为N。在优选实施方案中 X^{67} 为选自G、A、V、L、I、F、Y、W、P、D、E、Q、S、T、C、M、K、R和H的氨基酸，在更优选的实

施方案中 X^{67} 为 H 或 R, 而在甚至更优选的实施方案中 X^{67} 为 H。

X^{84} 不为 W, 且 X^{85} 不为 Y 且 X^{86} 不为 V。在优选实施方案中 X^{84} 、 X^{85} 和 X^{86} 为分别独立选自 G、A、L、I、F、P、D、E、N、Q、S、T、C、M、K、R 和 H 的氨基酸, 在更优选的实施方案中 $X^{84}X^{85}X^{86}$ 分别独立选自 S、F、V、P、M、A、L、K 和 G, 而在甚至更优选的实施方案中 $X^{84}X^{85}X^{86}$ 为 VAK。

本发明也提供了编码含有上述(权利要求 1 或权利要求 2 中)氨基酸序列的蛋白质的分离的核酸序列, 以及由(权利要求 3 中)此类核酸序列编码的分离的多肽。

在本发明的另一实施方案中, 上述(权利要求 1 或权利要求 2 中)分离的多肽作为种子贮存化合物的调节物在微生物或植物中发挥作用。

在本发明的另一实施方案中, 上述(权利要求 1 或权利要求 2 中)分离的多肽用于提高转基因植物中的油酸水平, 相对该植物野生型品种提高 1 重量%、2.5 重量%、5 重量%、7.5 重量%、10 重量%、12.5 重量%、15 重量%、17.5 重量%、20 重量%、22.5 重量%、25 重量%或更多。

在优选实施方案中, 上述(权利要求 1 或权利要求 2 中)分离的多肽具有如 SEQ ID NO: 6、SEQ ID NO: 8、SEQ ID NO: 10、SEQ ID NO: 12、SEQ ID NO: 14、SEQ ID NO: 16、SEQ ID NO: 18、SEQ ID NO: 20、SEQ ID NO: 22、SEQ ID NO: 24、SEQ ID NO: 26、SEQ ID NO: 28、SEQ ID NO: 30、SEQ ID NO: 32、SEQ ID NO: 34 或 SEQ ID NO: 36 中所公开的多肽序列。

在更优选的实施方案中, 上述(权利要求 1 或权利要求 2 中)分离的多肽选自:

a. 如 SEQ ID NO: 6、SEQ ID NO: 8、SEQ ID NO: 10、SEQ ID NO: 12、SEQ ID NO: 14、SEQ ID NO: 16、SEQ ID NO: 18、SEQ ID NO: 20、SEQ ID NO: 22、SEQ ID NO: 24、SEQ ID NO: 26、SEQ ID NO: 28、SEQ ID NO: 30、SEQ ID NO: 32、SEQ ID NO: 34 或 SEQ ID NO: 36 中所公开的多肽序列

b. 由如 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 中公开的多核苷酸序列所编码的多肽序列;

c. 与上述 a)或 b)中多肽具有至少 70%序列同一性的多肽序列。

本发明还提供包含选自如下的多核苷酸序列的分离的核酸:

a. 如 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 中所公开的多核苷酸序列;

b. 编码如 SEQ ID NO: 6、SEQ ID NO: 8、SEQ ID NO: 10、SEQ ID NO: 12、SEQ ID NO: 14、SEQ ID NO: 16、SEQ ID NO: 18、SEQ ID NO: 20、SEQ ID NO: 22、SEQ ID NO: 24、SEQ ID NO: 26、SEQ ID NO: 28、SEQ ID NO: 30、SEQ ID NO: 32、SEQ ID NO: 34 或 SEQ ID NO: 36 中所公开的多肽的多核苷酸序列;

c. 与上述 a)或 b)中核酸具有至少 70%序列同一性的多核苷酸序列;

d. 与上述 a)或 b)中核酸互补的多核苷酸序列; 以及

e. 在严紧条件下与上述 a)或 b)中核酸杂交的多核苷酸。

本发明还提供选自如下的分离的多肽:

a. 如 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 中公开的多核苷酸序列所编码的多肽序列;

b. 如 SEQ ID NO: 6、SEQ ID NO: 8、SEQ ID NO: 10、SEQ ID NO: 12、SEQ ID NO: 14、SEQ ID NO: 16、SEQ ID NO: 18、SEQ ID NO: 20、

SEQ ID NO: 22、SEQ ID NO: 24、SEQ ID NO: 26、SEQ ID NO: 28、SEQ ID NO: 30、SEQ ID NO: 32、SEQ ID NO: 34 或 SEQ ID NO: 36 中所公开的多肽序列;

c. 与上述 a)或 b)中多肽具有至少 70%序列同一性的多肽序列。

本发明也提供了来自拟南芥、大豆、稻、玉米、亚麻、大麦或小麦的分离的编码脂质代谢蛋白(LMP)或其部分的核酸。这些序列可用于改变或提高微生物及植物中脂质和脂肪酸、辅因子和酶,例如提高油酸水平 1 重量%、2.5 重量%、5 重量%、7.5 重量%、10 重量%、12.5 重量%、15 重量%、17.5 重量%、20 重量%、22.5 重量%、25 重量%或更高。

已知拟南芥植物中产生大量的脂肪酸如亚油酸和亚麻酸(参见例如表 2),并且它们与芸苔属(*Brassica*)油料作物植物在许多方面(基因同源性等)非常近似。因此,来自植物如拟南芥、欧洲油菜(*Brassica napus*)、大豆、稻、玉米、亚麻、大麦或小麦或相关生物的核酸分子特别适合于改变宿主特别是微生物和植物中的脂质和脂肪酸代谢。另外,来自植物拟南芥、欧洲油菜、大豆、稻、玉米、亚麻、大麦或小麦或相关生物的核酸可用于鉴定其他物种中的那些 DNA 序列和酶,这些 DNA 序列和酶可用于改变相应生物中脂肪酸前体分子的生物合成。

本发明还提供了分离的核酸,其包含来自植物(拟南芥、大豆、稻、玉米、亚麻、大麦或小麦)的编码 LMP 或其部分的核酸的至少 15 个核苷酸的片段。

本发明还提供了所述核酸编码的多肽,包含由所述核酸编码的多肽的异源多肽以及这些多肽的抗体。

另外,本发明涉及并提供了 LMP 核酸产生转基因植物的用途,所述转基因植物种子贮存化合物的水平或组成发生改变。至于改变的组成,本发明可用于如提高油酸相对于其他植物油(如亚麻酸或亚油酸)的百分比,如提高 1 重量%、2.5 重量%、5 重量%、7.5 重量%、10 重量%、12.5 重量%、15 重量%、17.5 重量%、20 重量%、22.5 重量%、25 重量%或更高。产生种子贮存化合物水平或组成发生改变的转基因植物的方法包括用

含有 LMP 核酸的表达载体转化植物细胞，并从植物细胞产生种子贮存化合物水平或组成发生改变的植物的步骤。在优选的实施方案中，植物为产油物种，例如选自芸苔、亚麻子、大豆、向日葵、玉米、燕麦、裸麦、大麦、小麦、稻、胡椒、万寿菊、棉花、油椰、椰树、亚麻、蓖麻和花生。

根据本发明，本文描述的组合物和方法可用于改变转基因植物的 LMP 组成，以及提高或降低转基因植物的 LMP 水平，包括增加或减少该植物中 LMP 核酸的表达。可以通过转基因过表达、共抑制方法、反义抑制方法或 LMP 核酸的体内诱变来增加或减少 LMP 核酸的表达。本发明也可用于提高或降低种子油中的脂质水平（如 1 重量%、2.5 重量%、5 重量%、7.5 重量%、10 重量%、12.5 重量%、15 重量%、17.5 重量%、20 重量%、22.5 重量%、25 重量%或更多）、提高或降低种子油中的脂肪酸水平（如 1 重量%、2.5 重量%、5 重量%、7.5 重量%、10 重量%、12.5 重量%、15 重量%、17.5 重量%、20 重量%、22.5 重量%、25 重量%或更多），或提高或降低种子或植物中的淀粉水平（如 1 重量%、2.5 重量%、5 重量%、7.5 重量%、10 重量%、12.5 重量%、15 重量%、17.5 重量%、20 重量%、22.5 重量%、25 重量%或更多）。

微 RNA (miRNA) 是进化保守的植物和动物基因表达的基于 RNA 的调节物。MiRNA (~ 21 至 25 nt) 来自较大的具有茎环结构的前体，所述前体由非蛋白质编码基因转录而来。miRNA 寻靶特定的 mRNA，在转录后水平（即降解 mRNA）或翻译水平（即抑制蛋白质合成）上抑制基因表达 (Bartel D 2004, Cell 116, 281-297)。

可以设计微 RNA 前体（前 miRNA），使由前 miRNA 编码的内源性 miRNA 能被某 miRNA 取代，以寻靶目的基因如 dsRed 报告基因。

本发明还提供了产生转基因植物的方法，所述转基因植物的油酸水平与其野生型相比提高，所述方法包括，

- a. 第一步用 RNA 前体构建体转化植物细胞，且
- b. 第二步从该植物细胞产生转基因植物，

其中所述构建体含有在植物细胞中驱动表达的启动子，该启动子有效

连于编码前体微 RNA 序列的核苷酸序列,其中编码所述微 RNA 前体序列的核苷酸序列选自:

- a. 如 SEQ ID NO: 47 所示的核苷酸序列
- b. 与上述 a)中核酸具有至少 70%序列同一性的多核苷酸序列;
- c. 与上述 a)中核酸互补的多核苷酸序列; 以及
- d. 在严紧条件下与上述 a)中核酸杂交的多核苷酸序列。

编码脂肪酸去饱和酶的玉米基因表达于许多组织包括种子中。与玉米去饱和酶 mRNA 的编码区或 5'UTR 和 3'UTR 互补的 19 至 21 nt (如 SEQ ID NO: 40 所述的 ACCAGACCCCGAACGCCGC), 可用于取代如 SEQ ID NO: 37 描述的 Zm miR166 (5' tcggaccaggettcattcccc 3'), 以及 SEQ ID NO: 38 的 Zm miR166 前体。该转基因随后可转化入玉米。遗传改造的 Zm miR166 基因的表达可受玉米种子特异性启动子 (例如胚乳特异性 10 KD 玉米醇溶蛋白启动子或 Glob1 胚特异性启动子) 的控制。

微 RNA (例如 SEQ ID NO: 40 所述的 ACCAGACCCCGAACGCCGC) 一般在加工遗传改造的 Zm miR166 前体时在种子中产生。该 miRNA 可特异性结合于玉米脂肪酸去饱和酶 mRNA 中与该 miRNA 互补的区域, 从而通过基因沉默机器在转录或翻译水平导致种子中所靶向的玉米去饱和酶的表达下降。因此, 转基因植物优选玉米能够具有期望的脂肪酸水平和组成, 例如种子中低亚油酸和/或中或高油酸重量百分比。

本发明还提供了通过产生油酸水平相对于野生型提高的转基因植物, 改变、优选降低脂肪酸去饱和酶表达的方法, 所述脂肪酸去饱和酶尤其是由 FAD2 直向同源物编码, 更优选由附录 A 中核酸编码, 在优选实施方案中如 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 所述, 该方法包括:

- a. 第一步用 RNA 前体构建体转化植物细胞, 且
- b. 第二步从该植物细胞产生转基因植物,

其中所述构建体含有在植物细胞中驱动表达的启动子，该启动子有效连于编码前体微 RNA 序列的核苷酸序列，其中编码所述微 RNA 前体序列的核苷酸序列选自：

- a. 如 SEQ ID NO: 47 所示的核苷酸序列
- b. 与上述 a) 中核酸具有至少 70% 序列同一性的多核苷酸序列；
- c. 与上述 a) 中核酸互补的多核苷酸序列；以及
- d. 在严紧条件下与上述 a) 中核酸杂交的多核苷酸序列。

在优选实施方案中，设计编码前体微 RNA 序列的核苷酸序列，使编码 SEQ ID NO: 37 所示微 RNA 的核苷酸序列被编码 SEQ ID NO: 40 所示微 RNA 的核苷酸序列所取代。

设计的微 RNA 前体和微 RNA 在调节基因表达中的用途众所周知，叙述于如 US 2004/0268441 中，后者在此处整体引用。设计的微 RNA 前体可用于调节一个或数个靶基因的表达，例如 1 个、2 个、3 个、4 个或 5 个如 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 所示的核苷酸序列。设计的微 RNA 前体和微 RNA 在调节基因表达中的用途可与其他本领域技术人员公知的基因工程方法联合使用。

启动子可为普遍表达的或组织特异性的，例如种子特异性和胚乳特异性的。启动子优选为种子特异性启动子。该方法可用于有效提高种子中的油酸水平，如提高 1 重量%、2.5 重量%、5 重量%、7.5 重量%、10 重量%、12.5 重量%、15 重量%、17.5 重量%、20 重量%、22.5 重量%、25 重量%或更多。

设计的微 RNA 前体和微 RNA 在调节基因表达中的用途可用于各种植物，特别是用于此处所述的植物，在优选实施方案中用于单子叶植物，而在更优选实施方案中用于玉米。

本发明的另一目的在于分离的核酸，其含有选自如下的多核苷酸序列：

- a. 如 SEQ ID NO: 47 所示的核苷酸序列
- b. 与上述 a) 中核酸具有至少 70% 序列同一性的多核苷酸序列;
- c. 与上述 a) 中核酸互补的多核苷酸序列; 以及
- e. 在严紧条件下与上述 a) 中核酸杂交的多核苷酸序列。

该核苷酸序列可用于调节目的基因的表达, 特别是用于下调靶基因的表达, 尤其是用于上述核苷酸序列。

本发明的另一目的在于微 RNA 前体, 由选自如下的核苷酸序列编码:

- a. 如 SEQ ID NO: 47 所示的核苷酸序列
- b. 与上述 a) 中核酸具有至少 70% 序列同一性的多核苷酸序列;
- c. 与上述 a) 中核酸互补的多核苷酸序列; 以及
- d. 在严紧条件下与上述 a) 中核酸杂交的多核苷酸序列。

本发明的另一目的在于如 SEQ ID NO: 40 所示的微 RNA。

更具体而言, 本发明包括并提供了提高种子中总油含量的方法, 包括: 用核酸构建体转化植物, 所述构建体中含有作为有效连接组分的启动子和能调节 *FAD2* 样 mRNA 或 *FAD2* 样蛋白质水平的核酸序列, 并种植该植物。此外, 本发明包括并提供了提高种子中油酸水平的方法, 包括: 用核酸构建体转化植物, 所述构建体中含有作为有效连接组分的启动子、能提高油酸水平的结构核酸序列, 并种植该植物。

本文也包括由 LMP DNA 序列转化的转基因植物所产生的种子, 其中所述种子含有 LMP DNA 序列且该植物是用于改变种子贮存化合物水平的真实育种。本发明还包括上述种子产生的种子油。

本发明还提供了含有所述核酸的载体、含有所述载体的宿主细胞, 以及用所述核酸和/或载体转化植物细胞所产生的子代植物材料。

根据本发明, 本文所述的化合物、组合物以及方法可用于提高或降低种子油中的脂质相对百分比, 提高或降低种子油中的脂质水平, 或提高或降低种子油中的脂肪酸水平, 或提高或降低种子或植物中的淀粉或其他糖类水平, 或提高或降低种子或植物中的蛋白质水平, 例如提高或降低 1 重量%、2.5 重量%、5 重量%、7.5 重量%、10 重量%、12.5 重量%、15 重

量%、17.5重量%、20重量%、22.5重量%、25重量%或更多。本文所述的操作方法也可用于促进种子发芽和幼苗及植物的生长，以及提高植物中种子贮存化合物的产率。

还提供了在转基因植物中产生高于或低于正常或一般水平的贮存化合物的方法，所述转基因植物表达来自拟南芥、欧洲油菜、大豆、稻、玉米、亚麻、大麦或小麦的 LMP 核酸，其中转基因植物为拟南芥、欧洲油菜、大豆、稻、玉米、亚麻、大麦、小麦、向日葵 (*Helianthus annuus*) 或甜菜 (*Beta vulgaris*)，或与拟南芥、欧洲油菜、大豆、稻、玉米、亚麻、大麦或小麦不同的物种。本文也包括改变种子贮存化合物生产效率的组合物和方法。本文所用短语拟南芥、欧洲油菜、大豆、稻、玉米、亚麻、大麦、小麦、向日葵或甜菜时，也是指拟南芥和/或欧洲油菜和/或大豆和/或稻和/或玉米和/或亚麻和/或大麦和/或小麦和/或向日葵和/或甜菜。

因此，本发明的一个目的在于提供来自大豆、稻、玉米、亚麻、大麦或小麦的新的分离的 LMP 核酸及分离的 LMP 氨基酸，以及其活性片段、类似物和直系同源物。这些活性片段、类似物和直系同源物可来自不同的植物物种，因为本领域技术人员应当理解其他植物物种也含有这些或相关核酸。

本发明的另一目的在于提供种子贮存化合物水平改变的转基因植物，特别是脂质、脂肪酸或糖水平的改变。

本发明的多核苷酸及多肽，包括其激动剂和/或其片段，也具有如下用途，包括调节植物生长，及潜在地调节植物产率，优选提高在逆境条件（干旱、寒冷、光、紫外线）下的植物生长。此外，本发明的拮抗剂可能具有包括调节植物生长和/或产率的用途，优选通过提高植物生长和产率来实现。而在另一实施方案中，利用组成型启动子过表达的本发明多肽可用于通过调节光利用效率，提高植物在应激条件（干旱、光、寒冷、紫外线）下的产量。此外，本发明的多核苷酸和多肽能促进种子萌发和种子休眠，从而改善植物生长和/或种子贮存化合物的产率。

本发明的分离的核酸分子还含有有效连接的启动子或部分启动子区

域。该启动子可为组成型启动子、诱导型启动子或组织特异性启动子。组成型启动子的实例如超级启动子(Ni 等人, Plant J. 7:661-676, 1995; US5955646)。组织特异性启动子可以在营养组织或繁殖组织中具有活性。在繁殖组织中具有活性的组织特异性启动子可为种子特异性启动子。在营养组织中具有活性的组织特异性启动子可为根特异性、芽特异性、分生组织特异性或叶特异性启动子。本发明分离的核酸分子还可含有 5'非翻译序列、3'非翻译序列、内含子或它们的组合。

本发明也提供了提高植物中一种或多种植物器官的数量和/或大小的方法, 所述表达来自拟南芥、欧洲油菜、大豆、稻、玉米、亚麻、大麦或小麦的编码 LMP 或其部分的分离的核酸。更具体而言, 可操作种子大小和/或种子数量和/或重量。

本发明的另一目的在于提供生产这类前述转基因植物的方法。

本发明的另一目的在于提供这类前述转基因植物的种子和种子油。

本领域普通技术人员在阅读以下详述的公开实施方案和附加权利要求书后, 可清楚了解本发明的这些及其他目的、特征和优势。

附图简述

根据以下详述以及构成本申请一部分的所附图表和序列表, 可更充分地理解本发明。

图 1A-D. SEQ ID NO: 1-4 – 拟南芥基因 *AtFAD-01* 的核酸序列、核酸的开放读框和氨基酸序列。

图 2A-C. SEQ ID NO: 5-8 – 大豆基因 *GmFAD-01* 的核酸序列、核酸的开放读框和氨基酸序列。

图 3A-C. SEQ ID NO: 9-12 – 大豆基因 *GmFAD-02* 的核酸序列、核酸的开放读框和氨基酸序列。

图 4A-C. SEQ ID NO: 13-16 – 大豆基因 *GmFAD-03* 的核酸序列、核酸的开放读框和氨基酸序列。

图 5A-C. SEQ ID NO: 17-20 – 玉米基因 *ZmFAD-01* 的核酸序列、核酸

的开放读框和氨基酸序列。

图 6A-C. SEQ ID NO: 21-24 – 稻基因 *OsFAD-01* 的核酸序列、核酸的开放读框和氨基酸序列。

图 7A-C. SEQ ID NO: 25-28 – 亚麻基因 *LuFAD-01* 的核酸序列、核酸的开放读框和氨基酸序列。

图 8A-C. SEQ ID NO: 29-32 – 大麦基因 *HvFAD-01* 的核酸序列、核酸的开放读框和氨基酸序列。

图 9A-C. SEQ ID NO: 33-36 – 小麦基因 *TaFAD-01* 的核酸序列、核酸的开放读框和氨基酸序列。

图 10. 由 USP 启动子驱动且转化入 *fad2* 拟南芥突变株（该转化株的遗传背景为 Columbia-2，每一栏代表用 5 mg 个体植物的群体种子得到的脂肪酸数据）的 *OsFAD-01* 得到的 T2 种子脂肪酸数据。

图 11. 由 USP 启动子驱动且转化入 *fad2* 拟南芥突变株（该转化株的遗传背景为 Columbia-2，每一栏代表用 5 mg 个体植物的群体种子得到的脂肪酸数据）的 *HvFAD-01* 得到的 T2 种子脂肪酸数据。

图 12. 图示了公开的 *AtFAD-01*、*GmFAD-01*、*GmFAD-02*、*GmFAD-03*、*LuFAD-01*、*HvFAD-01*、*TaFAD-01*、*OsFAD-01* 及 *ZmFAD-01* 氨基酸序列间的相对同源性。该图表是利用 Vector NTI Suite 9.0 的 Align X (2003 年 8 月 22 日) 所产生。用于该多重比对的参数如下：空位开放罚分：10；空位延伸罚分：0.05；空位分离罚分范围：8；比对延迟的同一性百分比：40。

图 13. 表格显示了 *AtFAD-01*、*GmFAD-01*、*GmFAD-02*、*GmFAD-03*、*LuFAD-01*、*HvFAD-01*、*TaFAD-01*、*OsFAD-01* 及 *ZmFAD-01* 氨基酸序列间的相似性。该表格是利用 Vector NTI Suite 9.0 的 Align X (2003 年 8 月 22 日) 所产生。其他参数参见图 12 的图例。

图 14. 图示了公开的 *AtFAD-01*、*GmFAD-01*、*GmFAD-02*、*GmFAD-03*、*LuFAD-01*、*HvFAD-01*、*TaFAD-01*、*OsFAD-01* 及 *ZmFAD-01* 核酸序列间的相对同源性。该图表是利用 Vector NTI Suite 9.0 的 Align X

(2003年8月22日)所产生。用于该多重比对的参数如下：空位开放罚分：15；空位延伸罚分：6.66；空位分离罚分范围：8；比对延迟的同一性百分比：40。

图 15. 表格显示了 AtFAD-01、GmFAD-01、GmFAD-02、GmFAD-03、LuFAD-01、HvFAD-01、TaFAD-01、OsFAD-01 及 ZmFAD-01 核酸序列间的相似性。该表格是利用 Vector NTI Suite 9.0 的 Align X (2003年8月22日)所产生。其他参数参见图 14 的图例。

图 16. AtFAD-01、GmFAD-01、GmFAD-02、GmFAD-03、LuFAD-01、HvFAD-01、TaFAD-01、OsFAD-01 及 ZmFAD-01 氨基酸序列的序列比对。该比对是利用 Vector NTI Suite 9.0 的 Align X (2003年8月22日)所产生。用于该多重比对的参数如下：空位开放罚分：15；空位延伸罚分：6.66；空位分离罚分范围：8；比对延迟的同一性百分比：40。

一般定义

应当认识到本发明并不仅限于此处所述的具体方法、方案、细胞系、植物物种或属、构建体及试剂。也应当认识到本文所用的术语仅为描述具体的实施方案所需，并不在于限制本发明的范围，后者只受附加权利要求书的限制。应当注意到此处及附加权利要求书中所用的单数形式包括复数指代，除非上下文中另有清楚说明。因此，例如提及“载体”是指一种或多种载体，并包括本领域技术人员已知的其等效物，等等。

本文所用的术语“约”，是指大约、粗略、左右或在此范围内。当术语“约”与数字范围联合使用时，在所列数值边界上下延伸而修饰其范围。一般而言，本文使用术语“约”来修饰数值为所列值的上下波动 20%、优选 10%、更优选上下（高低）5%。

本文所用的词“或”，是指具体列单内的任一成员，并且也包括该列单中成员的任意组合。

本文所用的术语“氨基酸序列”是指代表氨基酸残基的缩写、字母、字或词语的序列。此处可用氨基酸的通用 3 个字母符号或由 IUPAC-IUB

生物化学命名委员会推荐的单字母符号来指代氨基酸。与之类似，用通用的单字母代码指代核苷酸。本文所用的缩写为常规的氨基酸单字母代码：A，丙氨酸；B，天冬酰胺或天冬氨酸；C，半胱氨酸；D 天冬氨酸；E，谷氨酸；F，苯丙氨酸；G，甘氨酸；H 组氨酸；I 异亮氨酸；K，赖氨酸；L，亮氨酸；M，甲硫氨酸；N，天冬酰胺；P，脯氨酸；Q，谷氨酰胺；R，精氨酸；S，丝氨酸；T，苏氨酸；V，缬氨酸；W，色氨酸；Y，酪氨酸；Z，谷氨酰胺或谷氨酸(见 L. Stryer, Biochemistry, 1988, W. H. Freeman and Company, New York)。此处用于氨基酸序列中的字母“x”可以代表任一氨基酸残基。

术语“核酸”指单链或双链、正义或反义形式的脱氧核糖核苷酸或核糖核苷酸及其聚合物或杂交物。

此次所用的短语“核酸序列”是指代表核苷酸的缩写、字母、字或词语的连续序列。在一个实施方案中，核酸可为相对短核酸的“探针”，通常长度少于 100 个核苷酸。通常核酸探针的长度为约 50 个核苷酸至约 100 个核苷酸。核酸的“靶区”为核酸中鉴定为目标的部分。核酸的“编码区”为核酸的一部分，所述部分当处于合适调控序列的控制下时可通过序列特异性方式转录并翻译，以产生特定的多肽或蛋白质。此编码区被表述为编码此多肽或蛋白质。除非另有说明，除了明确显示的序列之外，特定的核酸序列也隐含地包括其经保守性修饰的变体(例如，简并密码子取代)及互补序列。此处术语“核酸”可与“基因”、“cDNA”、“mRNA”、“寡核苷酸”和“多核苷酸”交换使用。

本文所用的术语“互补”或“互补性”是指通过碱基配对原则相关的核苷酸序列。例如，序列 5'-AGT-3'与序列 5'-ACT-3'互补。互补可为“部分”或“全部”互补。“部分”互补是指一个或多个核酸碱基并未根据碱基配对原则配对。核酸之间的“全部”或“完全”互补是指每一个核酸碱基都根据碱基配对原则与另一碱基配对。核酸链间的互补程度对核酸链杂交的效率和强度都有显著影响。本文所用的核酸序列的“互补物”是指其核酸与本发明核酸序列的核酸完全互补的核苷酸序列。

术语“基因组”或“基因组 DNA”是指宿主生物的可遗传遗传信息。所述基因组 DNA 包括核 DNA（也称为染色体 DNA），也包括质体（例如叶绿体）及其他细胞器（如线粒体）DNA。优选术语基因组或基因组 DNA 指核染色体 DNA。

术语“染色体 DNA”或“染色体 DNA 序列”应理解为独立于细胞周期状态的细胞核基因组 DNA。因此，染色体 DNA 可被组织为染色体或染色单体，它们可为压缩或未解旋的形式。可以通过多种本领域已知的方法证实并分析染色体 DNA 中的插入，例如聚合酶链式反应（PCR）分析、Southern 印迹分析、原位荧光杂交（FISH）及原位 PCR。

术语“野生型”、“天然”或“天然来源”的是指生物、多肽或核酸序列，其中所述生物天然存在，或可获自至少一种天然存在的未经人改变、突变或其他操作的生物。

术语“异源核酸序列”或“异源 DNA”可交换使用，是指连接或经操作后连接于其天然并不连接的核酸序列的核苷酸序列，或连接于核酸序列上的位置不同于其天然连接位点。异源 DNA 并非其引入细胞内源产生的，而是来自另一细胞。一般而言，尽管并非必需，此类异源 DNA 编码表达其的细胞通常不产生的 RNA 和蛋白质。如果启动子、转录调控序列或其他遗传元件与另一序列（例如编码标记序列或农艺相关性状）在其天然环境中并非组合或有效连接，则所述启动子、转录调控序列或其他遗传元件相对于另一序列为“异源”。优选所述序列在其天然环境中并非有效连接（即来自不同基因）。最优选所述调控序列共价连接并邻接其天然环境中并不邻接的核酸。

本文所用的术语“转基因”是指任何引入细胞基因组或经人试验操作的核酸序列。优选所述序列产生于与天然存在生物不同的基因组中（例如，若所述序列为所述生物的内源序列，则将其引入与其天然位置不同的位置，或提高或减少其拷贝数目）。转基因可为“内源 DNA 序列”、“外源 DNA 序列”（如外源基因）或“异源 DNA 序列”。术语“内源 DNA 序列”指天然地见于其所引入的细胞的核苷酸序列，只要其相对于天然存在序列并不

含有某些改变（如点突变、存在选择标记基因等）。

当术语“转基因”或“重组”用于细胞或生物（例如对于大麦植物或植物细胞而言）时，指含有转基因或其基因组已被引入的转基因改变的细胞或生物。转基因生物或组织可以含有一个或多个转基因细胞。优选该生物或组织主要由转基因细胞组成（即所述生物或组织中多于 80%，优选 90%，更优选 95%，最优选 99%的细胞为转基因细胞）。

“重组多肽”为非天然存在的多肽，其序列与天然存在的多肽相差至少一个氨基酸残基。产生此类重组多肽和/或核酸的优选方法包括直接或非直接诱变、DNA 改组或其他递归重组(recursive recombination)的方法。

相对于目的杂交条件而言的术语“等同”杂交条件指该杂交条件及目的杂交条件下，具有相同范围的同源性百分比(%)的核酸序列能够杂交。例如，如果目的杂交条件下，第一核酸序列能与与第一核酸序列同源性为 80% - 90%的其他核酸序列杂交，则另一杂交条件与目的杂交条件等同是指在该另一杂交条件下，第一核酸序列也能与与第一核酸序列同源性为 80% - 90%的其他核酸序列杂交。

在本发明优选实施方案中，除非另有说明，两个核酸或多肽序列间的序列同一性百分比是利用 Vector NTI 7.0 (PC)软件包(InforMax, 7600 Wisconsin Ave., Bethesda, MD 20814)确定的。优选使用空位开放罚分 15 及空位延伸罚分 6.66 来确定两个核酸间的同一性百分比。优选使用空位开放罚分 10 及空位延伸罚分 0.1 来确定两个多肽间的同一性百分比。所有其他参数优选设为默认设置。为进行多重比对 (Clustal W 算法)，在优选实施方案中，blosum62 矩阵的空位开放罚分为 10 而空位延伸罚分为 0.05。应当理解为了确定 DNA 序列和 RNA 序列比较时的序列同一性，胸腺嘧啶核苷等同于尿嘧啶核苷。

本领域已充分了解核酸杂交中可使用多种等同条件以包含低或高严谨条件；需考虑的因素如探针长度和性质（DNA、RNA、碱基组成）、靶标性质（DNA、RNA、碱基组成、存在于溶液之中或固定化等），以及盐和其他成分的浓度（例如甲酰胺、葡聚糖硫酸盐、聚乙二醇的存在与否），且

杂交溶液可变动，以产生与上述条件有别但等同的低或高严紧性杂交的条件。本领域技术人员已知优选使用较高严紧性来减少或消除非特异性结合，而优选较低严紧性来检测同源性不同的较大数目的核酸序列。

术语“基因”指编码区，其有效连接于能以某种方式调控多肽表达的合适调控序列。基因包括处于编码区（开放读框 ORF）前（上游）或后（下游）的 DNA 非翻译调控区（例如启动子、增强子、阻遏物等），以及适用的情况下位于单个编码区（即外显子）之间的间隔序列（即内含子）。本文所用的术语“结构基因”是指可转录为 mRNA 随后翻译为特定多肽特征性氨基酸序列的 DNA 序列。

本文所用的术语“编码区”当用于结构基因时，是指编码氨基酸的核苷酸序列，所述氨基酸见于 mRNA 分子翻译所得到的新生多肽。在真核细胞中，编码区 5'端为编码起始甲硫氨酸的三联核苷酸“ATG”，3'端为定义终止密码子的 3 种三联核苷酸（即 TAA、TAG、TGA）之一。除了含有内含子，基因的基因组形式还包括 RNA 转录物序列中位于 5'-和 3'-末端的序列。这些序列被称为“侧翼”序列或“侧翼”区域（这些侧翼序列位于 mRNA 转录物上非翻译序列的 5'-或 3'-端）。5'侧翼区可含有调控序列如启动子和增强子，控制或影响基因的转录。3'侧翼区可含有指导转录终止、转录后切割和多聚腺苷化的序列。

此处术语“多肽”、“肽”、“寡肽”、“多肽”、“基因产物”、“表达产物”和“蛋白质”可交换使用，指连续氨基酸残基的多聚物或寡聚物。

本文所用的术语“分离的”是指材料从其初始环境中分离出来。例如，活体动物中天然存在的多核苷酸或多肽并非分离的，但与某些或所有其天然系统中共存材料分离开的同一多核苷酸或多肽是分离的。此类多核苷酸可为载体的部分，和/或此类多核苷酸或多肽可为组合物的一部分，若该载体或组合物不为其原始环境的一部分，则所述多核苷酸或多肽为分离的。

术语“遗传修饰生物”或“GMO”是指含有转基因 DNA 的任何生物。示例性生物包括植物、动物和微生物。

本文所用的单数形式的术语“细胞”或“植物细胞”指单个细胞。复

数形式的术语“细胞”指细胞群。细胞群可为含有一种细胞类型的纯细胞群。同样，细胞群也可含有一种以上的细胞类型。在本发明中，对细胞群中可包括的细胞类型数目并无限制。细胞可为同步化或非同步化。本发明意义上的植物细胞可为分离的（例如处于悬浮培养物中）或含于任意发育阶段的植物组织、植物器官或植物中。

对于植物而言的术语“器官”（或“植物器官”）是指植物的一部分，可包括（但不限于）例如根、果实、芽、茎、叶、花药、萼片、花瓣、花粉、种子等。

对于植物而言的术语“组织”（或“植物组织”）是多个植物细胞的排布，包括植物的分化和未分化组织。植物组织可构成植物器官的部分（例如植物叶的上皮），也可构成肿瘤组织（例如愈伤组织）及多种类型的培养细胞（例如单个细胞、原生质体、胚、愈伤组织、原始球茎样体等）。植物组织可为植物原位、器官培养物、组织培养物或细胞培养物。

本文所用的术语“植物”是指高度分化为植物发育中任何阶段存在的结构的一群植物细胞。此类结构包括一个或多个植物器官包括（但不仅限于）果实、芽、茎、叶、花瓣等。

术语“染色体 DNA”或“染色体 DNA 序列”指独立于细胞周期状态的细胞核基因组 DNA。因此，染色体 DNA 可被组织为染色体或染色单体，它们可为压缩或未解旋的。可以通过多种本领域已知的方法证实并分析染色体 DNA 中的插入，例如 PCR 分析、Southern 印迹分析、原位荧光杂交（FISH）及原位 PCR。

本文所用的术语“结构基因”是指可转录为 mRNA 随后翻译为特定多肽特征性氨基酸序列的 DNA 序列。

术语“表达”指基因产物的生物合成。例如，在结构基因的情况下，表达包括将结构基因转录为 mRNA，并任选随后将 mRNA 翻译为一个或多个多肽。

本文所用的术语“表达盒”或“表达构建体”是指将任意待表达核酸序列有效连接于启动子序列和（任选）有助于所述核酸序列表达的其他元

件（例如像终止子和/或多聚腺苷化序列）的组合。

本文所用的术语“启动子”、“启动子元件”或“启动子序列”是指位于核苷酸序列 5'端的指导转录起始的核苷酸序列（即能控制该核苷酸序列转录为 mRNA）。启动子一般（但不必须）位于核苷酸序列的 5'端（即上游，例如邻近结构基因的转录起始位点），其调控所述核苷酸序列向 mRNA 的转录，并提供 RNA 聚合酶和其他转录因子特异结合的位点以起始转录。启动子序列为驱动下游基因表达所必需（但并不充分）。一般而言，真核启动子包括位于转录起始（帽子）位点 5'端 10-30 bp 处、与共有序列 5'-TATAAT-3'(TATA)盒同源的特征性 DNA 序列，方便起见将所述转录起始位点编号为 +1。位于帽子位点 3'端的碱基用正数表示，而位于帽子位点 5'端的碱基用负数表示，分别反映其远离帽子位点的距离。另一启动子成分 CAAT 盒，常见于 TATA 盒 5'端 30 至 70 bp 处，与经典形式的 5'-CCAAT-3' (Breathnach 1981)具有同源性。植物中 CAAT 盒有时被称为 AGGA 盒的序列所取代，后者为三联 G(或 T)NG 侧翼具有对称性分布腺苷残基的区域(Messing 1983)。其他对转录有调控作用的序列可见于启动子区域内，并延伸至帽子位点 5'端以外多达 1000 bp 处或更远。当术语“组成型”用于启动子时，表示该启动子能够在缺乏刺激（例如热休克、化学物质、光等）的情况下指导有效连接的核酸序列进行转录。一般而言，组成型启动子能够指导转基因在基本任何细胞和任何组织中表达。

调节控制指通过主要（但不全）位于转录起始位点上游（5'端）的 DNA 序列元件来调节基因表达。调节可能会造成对环境刺激的全有或全无应答，或可能造成基因表达水平的多样化。本发明中，热休克调控元件能在应答突发温度上升时发挥作用，瞬时增强下游基因表达的水平。

多聚腺苷化信号是指任何能影响 mRNA 加工的核酸序列，通常以向 mRNA 前体的 3'端添加多聚腺苷酸为特征。多聚腺苷化信号 DNA 片段自身可为若干来源（天然存在或合成的）的片段组合物，可来自基因组 DNA 或 RNA 衍生的 cDNA。多聚腺苷化信号通常因典型形式 5'-AATAA-3'同源物的存在而被识别，尽管距离的差异、部分“通读”和多重串联典型序列

并不少见 (Messing 1983)。应当认识到典型的“多聚腺苷化信号”可能实际上导致转录的终止而并非仅仅多聚腺苷化本身 (Montell 1983)。

热休克元件是指调控应答突发温度升高应激的基因表达的 DNA 序列。在该应答中可见到下游基因表达水平突发但短暂的增强。关于热休克基因的初步工作是在果蝇中完成的, 但许多其他物种包括植物 (Barnett 1980) 对应激也表现出类似的应答。已描述果蝇热休克元件的关键主要成分具有共有序列 5'-CTGGAATNTTCTAGA-3' (其中 N=A、T、C 或 G), 并定位于转录起始位点上游-66 至-47 bp 的残基之间 (Pelham 1982)。化学合成的该共有序列的寡核苷酸拷贝可以取代天然序列, 赋予热休克诱导性。

前导序列指包括位于转录起始位点和翻译起始位点间的约 100 个核苷酸的 DNA 序列。前导序列中包括的是指定核糖体结合位点的区域。

本文的内含子或间隔序列指与编码序列 (外显子) 一同翻译、但形成成熟 mRNA 时被移除的那些 DNA 序列区域。内含子可位于转录序列中的任何部位——位于相同或不同基因的编码序列之间, 位于基因的编码序列中 (中断并分开其氨基酸序列), 以及位于启动子区 (翻译起始位点 5' 端)。初级转录物中的内含子被切割出来, 而编码序列被同步且精确连接以形成成熟 mRNA。内含子和外显子的连接处构成剪接位点。内含子的碱基序列以 GU 开始, 以 AG 结束。相同的剪接信号见于许多高等真核细胞中。

术语“有效连接”或“有效连接的”应理解为, 例如调控元件 (如启动子) 与待表达的核酸序列的有序排列 (若合适, 还包括调控元件 (如终止子)), 从而使每一调控元件都能完成其预期的功能, 即允许、修饰、促进或以其他方式影响所述核酸序列的表达。造成的表达可能取决于核酸序列相对正义或反义 RNA 的排列。此处, 并不必需化学意义上的直接连接。遗传控制序列如增强子序列也可从远处的位点或实际上在其他 DNA 分子上对靶序列施加影响。优选的排列为将待重组表达的核酸序列置于作为启动子起作用的序列之后, 从而这两个序列彼此共价连接。启动子序列和待重组表达的核酸序列之间的距离优选少于 200 个碱基对, 尤其优选少于 100 个碱基对, 特别优选少于 50 个碱基对。可以通过 (如见于 Maniatis 1989;

Silhavy 1984; Ausubel 1987; Gelvin 1990)所述的常规重组和克隆技术产生有效连接及表达盒。然而,作为具有限制性酶特异切割位点的接头序列或信号肽起作用的其他序列,也可置于两个序列之间。序列的插入也可导致表达融合蛋白质。优选,由启动子和待表达核酸连接组成的表达盒,可以载体整合形式存在,通过如转化插入植物基因组中。

本文所用的术语“转化”是指将遗传物质(例如转基因)引入细胞。细胞的转化可为稳定的或瞬时性的。术语“瞬时转化”或“瞬时转化的”是指将一个或多个转基因引入细胞,但转基因并不整合于宿主细胞基因组中。可通过例如酶联免疫吸附试验(ELISA)检测瞬时转化,ELISA 试验可检测是否存在由所述一个或多个转基因编码的多肽。或者,通过如此处所示检测由转基因(如 uid A 基因)编码的蛋白质(如 β -葡萄糖苷酸酶)活性[例如通过用 X-gluc 染色组化测定 GUS 酶活性,当存在 GUS 酶时 X-gluc 可表现出蓝色沉淀;利用 GUS-Light 试剂盒(Tropix)对 GUS 酶活性进行化学发光测定]来检测瞬时转化。术语“瞬时转化株”指瞬时掺入一个或多个转基因的细胞。与之相对,术语“稳定转化”或“稳定转化的”指一个或多个转基因引入并整合于细胞的基因组中,优选形成染色体整合和减数分裂下的稳定遗传力。可以通过将细胞基因组 DNA 与能结合所述一个或多个转基因的核酸序列进行 Southern 印迹杂交,检测细胞的稳定转化。或者也可通过细胞基因组 DNA 的聚合酶链式反应扩增转基因序列来检测稳定的细胞转化。术语“稳定转化株”指将一个或多个转基因稳定整合于基因组 DNA(包括质体 DNA 和核 DNA),优选整合入核内染色体 DNA 的细胞。因此,稳定转化株与瞬时转化株的不同之处在于,稳定转化株的基因组 DNA 含有一个或多个转基因,而瞬时转化株的基因组 DNA 中不含转基因。转化也包括以植物病毒载体的形式将遗传物质引入植物细胞,包括染色体外复制以及减数分裂稳定性性质不同的基因表达。转化也包括以植物病毒载体的形式将遗传物质引入植物细胞,涉及减数分裂稳定性不同的染色体外复制及基因表达。优选术语“转化”包括将遗传物质引入植物细胞,从而导致染色体整合和减数分裂所致的稳定遗传力。

术语细菌“感染”和“侵染”指在一定条件下将靶生物样品（例如细胞、组织等）与细菌共孵育，从而使细菌中所含的核酸序列被引入靶生物样品的一个或多个细胞中。

术语“农杆菌”指存于土壤中可导致冠瘿病的革兰氏阴性杆状植物致病细菌。术语“农杆菌”包括但不限于根瘤农杆菌（*Agrobacterium tumefaciens*）菌株（一般可导致感染植物的冠瘿病），以及发根农杆菌（*Agrobacterium rhizogenes*）菌株（导致感染宿主植物的发根病）。农杆菌感染植物细胞一般会导致感染细胞产生冠瘿碱（例如胭脂碱、农杆菌碱、章鱼碱等）。因此，导致产生胭脂碱的农杆菌菌株（例如菌株 LBA4301、C58、A208）被称为“胭脂碱型”农杆菌；导致产生章鱼碱的农杆菌菌株（例如菌株 LBA4404、Ach5、B6）被称为“章鱼碱型”农杆菌；而导致产生农杆菌碱的农杆菌菌株（例如菌株 EHA105、EHA101、A281）被称为“农杆菌碱型”农杆菌。

术语“轰击”和“基因枪轰击”指向靶生物样品（如细胞、组织等）加速引入颗粒的方法，从而导致靶生物样品中细胞的细胞膜损伤，和/或颗粒进入该靶生物样品中。基因枪轰击的方法为本领域已知（例如 US 5,584,807，其内容在此次引用作为参考），并可购得（例如氦气驱动显微投射加速器(PDS-1000/He)(BioRad)）。

本文所用的术语“杂交”包括“使一条核酸链通过碱基配对与互补链结合的任意方法”(Coombs 1994)。杂交及杂交强度（即核酸间结合的强度）受以下因素影响，如核酸间的互补程度、涉及条件的严紧性、所形成杂合体的 T_m ，以及核酸内的 G:C 比率。

本文所用的术语“ T_m ”用于指“解链温度”。解链温度为双链核酸分子群半解离形成单链的温度。计算核酸 T_m 的方程式为本领域所公知。如标准参考文献所提示，可通过方程式 $T_m = 81.5 + 0.41(\% G + C)$ 简单估算 T_m 值，其中核酸处于 1 M NaCl 的水溶液中[见如 Anderson 和 Young, Quantitative Filter Hybridization, Nucleic Acid Hybridization (1985)]。其他参考包括在 T_m 计算中考虑结构和序列特点的更复杂的计算。

除非另有说明,核酸杂交的低严紧性条件包括等同于如下的条件:于68℃在含5×SSPE(43.8 g/L NaCl、6.9 g/L NaH₂PO₄·H₂O和1.85 g/L EDTA,用NaOH调节pH为7.4)、1% SDS、5×Denhardt's试剂[50×Denhardt's,其中每500 mL含有:5 g Ficoll (Type 400, Pharmacia)、5 g BSA (Fraction V; Sigma)],以及100 μg/mL变性鲑精DNA的溶液中进行结合或杂交,然后于室温在含0.2×SSPE和0.1% SDS的溶液中洗涤,其中使用长约100至1000个核苷酸的DNA探针。

除非另有说明,核酸杂交的高严紧性条件包括等同于如下的条件:于68℃在含5×SSPE、1% SDS、5×Denhardt's试剂和100 μg/mL变性鲑精DNA的溶液中进行结合或杂交,然后于68℃在含0.1×SSPE和0.1% SDS的溶液中洗涤,其中使用长约100至1000个核苷酸的DNA探针。

相对于目的杂交条件而言的术语“等同”杂交条件指该杂交条件及目的杂交条件下,具有相同范围同源性百分比(%)的核酸序列能够杂交。例如,如果目的杂交条件下,第一核酸序列能与与第一核酸序列同源性为80%–90%的其他核酸序列杂交,则另一杂交条件与目的杂交条件等同是指在该另一杂交条件下,第一核酸序列也能与与第一核酸序列同源性为80%–90%的其他核酸序列杂交。

本领域已充分了解核酸杂交中可使用多种等同条件以包含低或高严紧条件;需考虑的因素如探针长度和性质(DNA、RNA、碱基组成)、靶标性质(DNA、RNA、碱基组成、存在于溶液之中或固定化等),以及盐和其他成分的浓度(例如甲酰胺、葡聚糖硫酸盐、聚乙二醇存在与否),且杂交溶液可变动,以产生与上述条件有别但等同的低或高严紧性杂交的条件。本领域技术人员已知优选使用较高严紧性来减少或消除非特异性结合,而优选较低严紧性来检测同源性不同的较大数目的核酸序列。

发明详述

参考以下对本发明优选实施方案的详细描述及所附的实施例,可更容易地理解本发明。

在公开并描述本发明的化合物、组合物及方法之前，应当理解本发明不仅限于特定核酸、特定多肽、特定细胞类型、特定宿主细胞、特定条件或特定方法等，因为这些理所当然可以有所变化，其多种修饰和变化对于本领域技术人员而言是显而易见的。还应当理解本文所用的术语仅出于描述具体实施方案的目的而并非旨在限制。如说明书和权利要求书中所用，

“一个”或“一种”可以表示一个/一种或多个/多种，视上下文而定。从而，例如，述及“一个/一种细胞”可以表示能够利用至少一个/一种细胞。

本发明部分基于对编码 FAD2 样 LMP 的核酸分子的分离和表征，所述核酸分子来自植物包括拟南芥、大豆、稻、玉米、亚麻、大麦和小麦及其他相关作物物种如玉米、大麦、亚麻、甜菜或向日葵。

根据本文所体现和详细描述的本发明目的，本发明一方面提供了来自植物（拟南芥、大豆、玉米、稻、亚麻、大麦或小麦）的编码脂质代谢蛋白(LMP)或其部分的分离的核酸。

本发明的一方面涉及编码 LMP 多肽或其生物学活性部分的分离的核酸分子，也涉及足以作为鉴定或扩增 LMP 编码核酸（如 LMP DNA）的杂交探针或引物的核酸片段。如本文所用的术语“核酸分子”旨在包括 DNA 分子（如 cDNA 或基因组 DNA）和 RNA 分子（如 mRNA），以及利用核苷酸类似物产生的 DNA 或 RNA 类似物。该术语也包括基因编码区 3'和 5'端的非翻译序列：基因编码区 5'端上游序列至少约 1000 个核苷酸以及编码区 3'端下游序列至少约 200 个核苷酸。核酸分子可以为单链或双链，但优选双链 DNA。“分离”的核酸分子为与所述核酸天然来源中存在的其他核酸分子充分分离的核酸分子。优选“分离”的核酸中基本不含在其来源生物基因组 DNA 中天然位于所述核酸两侧的序列（即位于所述核酸 5'和 3'端的序列）。例如，在多个实施方案中，分离的 LMP 核酸分子可含有少于约 5 kb、4kb、3kb、2kb、1 kb、0.5 kb 或 0.1 kb 的核苷酸序列，所述核苷酸序列在该核酸来源的细胞（如拟南芥、大豆、稻、玉米、亚麻、大麦或小麦细胞）基因组 DNA 中与之天然侧接。此外，“分离”的核酸分子如 cDNA 分子中基本不含有使用重组技术生产时的其他细胞材料或培养

基，或采用化学合成时的化学前体或其他化学物。

可以使用标准分子生物学技术及本文提供的序列信息分离本发明的核酸分子，例如具有附录 A 所示核苷酸序列，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 或其部分的核酸分子。例如，可以使用附录 A 所示序列之一，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 的全部或部分作为杂交探针，利用标准杂交技术（如 Sambrook 等所述，1989, *Molecular Cloning: A Laboratory Manual* 第二版, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY），从拟南芥、大豆、稻、玉米、亚麻、大麦或小麦文库中分离拟南芥、大豆、玉米、稻、亚麻、大麦或小麦的 LMP cDNA。此外，可以使用根据相同核酸序列设计的寡核苷酸引物进行聚合酶链式反应分离包含附录 A 所示序列之一，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 的全部或一部分的核酸分子（例如可以利用根据附录 A 所示相同序列，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 设计的寡核苷酸引物，通过聚合酶链式反应分离包含附录 A 所示序列之一，如优选实施方案中公

开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 的全部或一部分的核酸分子)。例如,可以从植物细胞中分离 mRNA(如通过 Chirgwin 等, 1979, *Biochemistry* 18:5294-5299 所述的胍盐-硫氰酸盐萃取法), 并且可以利用逆转录酶制备 cDNA(如获自 Gibco/BRL, Bethesda, MD 的 Moloney MLV 逆转录酶; 或获自 Seikagaku America, Inc., St. Petersburg, FL 的 AMV 逆转录酶)。可以根据附录 A 所示核苷酸序列, 如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一设计用于聚合酶链式反应扩增的合成寡核苷酸引物。可根据标准 PCR 扩增技术选择合适的寡核苷酸引物, 以 cDNA 或基因组 DNA 为模板扩增本发明的核酸。将如此扩增的核酸克隆至合适的载体中, 并通过 DNA 序列分析进行表征。此外, 可以通过标准合成技术, 如使用自动 DNA 合成仪, 制备与 LMP 核苷酸序列相应的寡核苷酸。

在优选实施方案中, 本发明的分离的核酸包含附录 A 所示核苷酸序列之一, 如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 9、SEQ ID NO: 13、SEQ ID NO: 17、SEQ ID NO: 21、SEQ ID NO: 25、SEQ ID NO: 29 或 SEQ ID NO: 33。附录 A 所示序列, 如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 9、SEQ ID NO: 13、SEQ ID NO: 17、SEQ ID NO: 21、SEQ ID NO: 25、SEQ ID NO: 29 或 SEQ ID NO: 33 对应于本发明的拟南芥、大豆、稻、玉米、亚麻、大麦或小麦 LMP cDNA。这些 cDNA 中含有编码 LMP 的序列(即附录 A 所示的“编码区”)以及 5' 非翻译序列及 3' 非翻译序列。可选地, 核酸分子可只包含附录 A 所示任何序列, 如优选实施方案中公开的 SEQ ID NO: 7、SEQ ID NO: 11、SEQ ID NO: 15、SEQ ID

NO: 19、SEQ ID NO: 23、SEQ ID NO: 27、SEQ ID NO: 31 或 SEQ ID NO: 35 的编码区，或者可包含分离自基因组 DNA 的完整基因组片段。

就本申请的目的而言，应该理解附录 A 中公开的每个序列具有标识记录号（例如 *TaFAD-01*）。这些序列中每一个通常各包含三个部分：5' 上游区、编码区和下游区。这些序列的编码区标记为“ORF 位置”（表 3）。

在另一优选实施方案中，本发明的分离的核酸分子含有这样的核酸分子，所述核酸分子为附录 A 所示核苷酸序列，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一或其部分的互补序列。与附录 A 所示核苷酸序列，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一互补的核酸分子是这样的核酸分子，它们与附录 A 所示核苷酸序列充分互补，从而能够与附录 A 所示核苷酸序列之一杂交，由此形成稳定的双链体。

又在另一个优选实施方案中，本发明的分离的核酸分子含有核苷酸序列，所述核苷酸序列与附录 A 所示核苷酸序列，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 或其部分之间具有至少约 50-60%，优选至少约 60-70%，更优选至少约 70-80%、80-90% 或 90-95%，也优选至少约 85%、86%、87%、88%、89%、90%、91%、92%、93% 或 94%，甚至更优选至少约 95%、96%、97%、98%、99% 或更高的同源性。优选利用 Vector NTI Suite 9.0 的 Align X (2003 年 8 月 22

日)确定核苷酸序列的同源性,设定参数如下:空位开放罚分:15;空位延伸罚分:6.66;空位分离罚分范围:8;比对延迟的同一性百分比:40。

在另一优选的实施方案中,本发明的分离的核酸分子含有核苷酸序列,所述核苷酸序列可与附录A所示核苷酸序列之一或其部分杂交,例如在严谨条件下杂交。这些杂交条件包括用盐浓度为约0.02 M的溶液于pH 7约60℃洗涤。

此外,本发明的核酸分子可仅含有附录A所示序列,如优选实施方案中公开的SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33或SEQ ID NO: 35之一的编码区的一部分,例如可用作探针或引物的片断,或编码LMP生物学活性部分的片断。通过克隆拟南芥、大豆、稻、玉米、亚麻、大麦或小麦的LMP基因而确定的核苷酸序列能够用于产生为在其他细胞类型及生物中鉴定和/或克隆LMP同源物以及来自其他植物或相关物种的LMP同源物所设计的探针和引物。因此本发明也提供含有本文公开的核酸或其片断的化合物。这些化合物包含附着于其他部分的核酸。其他部分包括但不限于检测部分、杂交部分、纯化部分、递送部分、反应部分、结合部分等。探针/引物一般含有基本上纯的寡核苷酸。所述寡核苷酸一般含有可在严格条件下与附录A所公开序列,如优选实施方案中公开的SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33或SEQ ID NO: 35之一的正义链、附录所A公开序列,如优选实施方案中公开的SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33或SEQ ID NO: 35之一的反

义序列或其天然突变体中至少约 12 个, 优选约 25 个, 更优选约 40、50 或 75 个连续核苷酸杂交的核苷酸序列区。基于附录 A 所示核苷酸序列, 如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 的引物可用于 PCR 反应以克隆 LMP 同源物。基于 LMP 核苷酸序列的探针可用于检测转录物或编码相同或同源蛋白质的基因组序列。在优选的实施方案中, 探针还含有与之相连的标记基团, 例如, 标记基团可为放射性同位素、荧光化合物、酶或酶辅因子。此类探针可作为基因组标记物测试试剂盒的一部分, 用于如通过测量细胞样品中 LMP 编码核酸的水平 (如检测 LMP mRNA 水平) 鉴定表达 LMP 的细胞, 或确定基因组 LMP 基因是否发生了突变或缺失。

在一个实施方案中, 本发明的核酸分子编码蛋白质或其部分, 所述蛋白质或其部分含有与附录 A 所示序列, 如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一所编码的氨基酸序列基本同源的氨基酸序列, 从而使蛋白质或其部分能保持与野生型蛋白质相同或相似的功能。本文所用的“基本同源”, 是指蛋白质或其部分具有的氨基酸序列中包括最小数目的与某氨基酸序列同一或等价的氨基酸残基 (例如, 与附录 A 所示序列, 如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 中 ORF 之一的氨基酸残基具有相似侧链的氨基酸残基), 从而使蛋白质或其部分能够参与植物中种子贮存化合物的产生、微生

物或植物中细胞膜的构建、或分子的跨膜转运所必需的化合物代谢。调节蛋白如 DNA 结合蛋白、转录因子、激酶、磷酸酶，或代谢路径如脂质、淀粉和蛋白质生物合成路径的蛋白质成员、或膜转运系统可能在种子贮存化合物的生物合成中起作用。本文描述了此类活性的实例（见表 3 的推定注释）。LMP 编码核酸序列的实例在附录中 A 公开，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35。

由于在多种植物如玉米、小麦、裸麦、燕麦、黑小麦、稻、大麦、大豆、花生、棉花、芸苔、油菜籽、木薯、胡椒、向日葵、甜菜和万寿菊、茄属植物如马铃薯、烟草、茄子及番茄、蚕豆属物种、豌豆、紫花苜蓿、灌木植物（咖啡、可可树、茶树）、柳树属物种、树（油棕、椰树）以及多年生草本和饲料作物中，改变或提高糖和/或脂肪酸的产量是希望得以遗传的主要性状，这些作物植物也是本发明另一实施方案中用于基因工程的优选靶植物。

本发明的 LMP 核酸分子编码的蛋白质部分优选为 LMP 之一的生物学活性部分。本文所用的术语“LMP 的生物学活性部分”旨在包括这样的部分，如参与种子贮存脂质的生物合成、微生物或植物中细胞膜的构建或分子的跨膜转运所必需的化合物代谢、或具有表 3 所列活性的 LMP 结构域/基序。为确定 LMP 或其生物学活性部分是否能够参与种子贮存化合物的产生及细胞膜所需的化合物代谢，可进行酶活测定。此类测定法为本领域技术人员所熟知，并描述于实施例 14。

LMP 的生物学活性部分包括含有氨基酸序列的肽，所述氨基酸序列来源于 LMP 氨基酸序列（如由附录 A 所示核酸序列，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、

SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 编码的氨基酸序列或与 LMP 同源的蛋白质的氨基酸序列,其包括比全长 LMP 或与 LMP 同源的全长蛋白质更少的氨基酸),且表现出至少一种 LMP 活性。通常,生物学活性部分(如长度为如 5、10、15、20、30、35、36、37、38、39、40、50、100 或更多个氨基酸的肽)含有具有至少一种 LMP 活性的结构域或基序。此外,可以通过重组技术制备蛋白质其他区域缺失的其他生物学活性部分,并评估本文所述的一种或多种活性。优选地,LMP 生物学活性部分包括一个或多个选择的结构域/基序或其具有生物学活性的部分。

可以制备编码 LMP 生物活性部分的其他核酸片断,分离这些序列之一的部分、表达 LMP 或肽的编码部分(如通过体外重组表达)并评估 LMP 或肽编码部分的活性。

本发明还包括由于遗传密码简并而有别于附录 A 所示核苷酸序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一(及其部分)的核酸分子,因而也编码与附录 A 所示核苷酸序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 编码的 LMP 相同的 LMP。在另一实施方案中,本发明的核酸分子编码全长蛋白质,其与附录 A 所示开放读码框编码的多肽的氨基酸序列基本同源。在一个实施方案中,全长核酸或蛋白质或者核酸或蛋白质的片断来自拟南芥、大豆、稻、玉米、亚麻、大麦或小麦。

除本文附录 A 所示的拟南芥、大豆、稻、玉米、亚麻、大麦或小麦 LMP 核苷酸序列外,本领域技术人员会理解种群(如拟南芥、欧洲油菜、大豆、稻、玉米、亚麻、大麦或小麦种群)中可能存在导致 LMP 氨基酸

序列改变的 DNA 序列多态性。由于天然变异,种群个体间可能存在此类 LMP 基因的遗传多态性。本文所用的术语“基因”和“重组基因”是指含有编码 LMP 的开放读框的核酸分子,其中 LMP 优选为拟南芥、大豆、稻、玉米、亚麻、大麦或小麦 LMP。此类天然变异一般会造成 LMP 基因的核苷酸序列 1-40% 的变异。任何及所有此类天然变异导致的且不会改变 LMP 功能性活性的核苷酸变异及其所产生的 LMP 氨基酸多态性,均在本发明的范围内。

对应于天然变异体和本发明拟南芥、大豆、稻、玉米、亚麻、大麦或小麦 LMP cDNA 的非拟南芥、大豆、稻、玉米、亚麻、大麦或小麦直向同源物的核酸分子,可以根据其与本文公开的拟南芥、大豆、稻、玉米、亚麻、大麦或小麦 LMP 核酸的同源性,以拟南芥、大豆、稻、玉米、亚麻、大麦或小麦 cDNA 或其部分为杂交探针,根据严格杂交条件下的标准杂交技术进行分离。本文所用的术语“直向同源物”是指来自不同物种的两种核酸,但其系自共同祖先基因通过物种形成进化而来。通常直向同源物编码具有相同或相似功能的蛋白质。因此,在另一实施方案中,本发明分离的核酸分子长度至少为 15 个核苷酸,且可在严紧条件下与含有附录 A 所示核苷酸序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 的核酸分子杂交。在其他实施方案中,核酸长度至少为 30、50、100、250 或更多个核苷酸。本文所用的术语“在严紧条件下杂交”用以描述杂交和洗涤条件,在所述条件下彼此间至少有 60% 同源性的核苷酸序列通常可仍然彼此杂交。优选的条件可使彼此具有至少约 65%,更优选至少约 70%,甚至更优选至少约 75% 或更高同源性的序列通常仍然彼此杂交。此类严紧条件为本领域技术人员所公知,可见于《Current Protocols in Molecular Biology》, John Wiley & Sons, N.Y., 1989, 6.3.1-6.3.6。优选的严紧杂交条件的非限制性实例为在 $6\times$ 氯化钠/柠檬酸钠(SSC)中于约 45℃

杂交,随后在 $0.2 \times \text{SSC}$ 、 $0.1\% \text{SDS}$ 中于 $50-65^{\circ}\text{C}$ 洗涤一次或多次。优选地,可在严紧条件下与附录 A 所示序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 杂交的本发明的分离的核酸分子对应于天然存在的核酸分子。本文所用的术语“天然存在的”核酸分子是指具有天然存在(如编码天然蛋白质)的核苷酸序列的 RNA 或 DNA 分子。在一个实施方案中,核酸编码天然存在的拟南芥、大豆、稻、玉米、亚麻、大麦或小麦 LMP。

除了种群中可能存在的 LMP 序列天然变体之外,技术人员还理解可以通过突变向附录 A 所示核苷酸序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一引入变化,由此改变所编码 LMP 的氨基酸序列,而不改变 LMP 的功能活性。例如,可以对附录 A 所示序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 进行核苷酸取代,从而引起“非必需”氨基酸残基的氨基酸取代。

“非必需”氨基酸残基是指 LMP (附录 A,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35)之一的野生型序列中可以改变而不改变所述 LMP 活性的残基,而“必需”氨基酸残基则是 LMP 活

性所必需的。然而，其他氨基酸残基（如那些 LMP 活性结构域中非保守或仅为半保守的残基）对于其活性可能是非必需的，因此这些残基可能适于被改变而不改变 LMP 活性。

因此，本发明的另一方面涉及编码 LMP 的核酸分子，所述 LMP 含有 LMP 活性非必需的氨基酸残基改变。此类 LMP 在氨基酸序列上与还保留本文所述的至少一种 LMP 活性的序列有所不同。在一个实施方案中，分离的核酸分子包含编码蛋白质的核苷酸序列，其中所述蛋白质含有与附录 A，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 所示核酸序列编码的氨基酸序列具有至少约 50%同源性的氨基酸序列，且能够参与拟南芥、大豆、稻、玉米、亚麻、大麦或小麦中种子贮存化合物的产生或细胞膜所必需的化合物代谢，或具有表 3 列出的一种或多种活性。优选地，核酸分子编码的蛋白质与附录 A 所示的核酸，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一所编码的序列具有至少约 50-60%的同源性，更优选与附录 A 所示的核酸，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一所编码的序列具有至少约 60-70%的同源性，甚至更优选与附录 A 所示的核酸，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、

SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一所编码的序列具有至少约 70-80%、80-90%、90-95%的同源性,还优选与附录 A 所示的核酸,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一所编码的序列具有至少约 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%或 95%的同源性,并且最优选与附录 A 所示的核酸,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一所编码的序列具有至少约 96%、97%、98%或 99%的同源性。优选利用 Vector NTI Suite 9.0 的 Align X (2003 年 8 月 22 日)确定多肽序列的同源性,设定参数如下:空位开放罚分: 10; 空位延伸罚分: 0.05; 空位分离罚分范围: 8; 比对延迟的%一致性: 40。

为确定两个氨基酸序列(如附录 A 所示核酸,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一所编码的序列与其突变形式)之间或两个核酸之间的同源性百分比,将所述序列按最佳比较目进行比对(例如,可在一蛋白质或核酸序列中引入空位以便与另一蛋白质或核酸进行最佳比对)。然后比较相应氨基酸位置或核苷酸位置的氨基酸残基或核苷酸。如果其中一个序列(如由附录 A 所示核酸,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、

SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 所编码的序列之一)中某一位置被与另一序列(如选自附录 A 所示核酸序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 所编码多肽的序列的突变形式)相应位置上相同的氨基酸残基或核苷酸所占据,则两个分子在该位置上同源(即本文所用的氨基酸或核酸“同源性”与氨基酸或核酸“同一性”的意义等同)。两个序列之间的同源性百分比为这两个序列所共有的相同位置数的函数(即,同源性% = 相同位置数/总位置数 × 100)。

可以通过向附录 A 所示核苷酸序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 中引入一个或多个核苷酸取代、添加或缺失,来产生编码与附录 A 所示核酸,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 所编码蛋白质序列同源的 LMP 的分离的核酸分子,从而可向编码蛋白质中引入一个或多个氨基酸取代、添加或缺失。可以通过标准技术如定点诱变和 PCR 介导的诱变,向附录 A 所示序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO:

33 或 SEQ ID NO: 35 之一引入突变。优选在一个或多个预测的非必需氨基酸残基上进行保守氨基酸取代。“保守氨基酸取代”是指一个氨基酸残基由另一个具有相似侧链的氨基酸残基代替。本领域中已经确定了具有相似侧链的氨基酸残基家族。这些家族包括具有碱性侧链（如赖氨酸、精氨酸、组氨酸）、酸性侧链（如天冬氨酸、谷氨酸）、不带电的极性侧链（如甘氨酸、天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、酪氨酸、半胱氨酸）、非极性侧链（如丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸）、 β -支链侧链（如苏氨酸、缬氨酸、异亮氨酸）以及芳香侧链（如酪氨酸、苯丙氨酸、色氨酸、组氨酸）的氨基酸残基。因此，优选用来自相同侧链家族的另一氨基酸残基代替 LMP 中预测的非必需氨基酸残基。或者，在另一个实施方案中，可以通过诸如饱和诱变的方法向所有或部分 LMP 编码序列中随机引入突变，并筛查所得突变体中本文所述的 LMP 活性，以鉴定保留 LMP 活性的突变体。诱变附录 A 所示序列，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一后，可以重组表达编码的蛋白质，并通过如本文描述的测定法（见实施例 11-13）测定蛋白质的活性。

优选通过重组 DNA 技术产生 LMP。例如，将编码蛋白质的核酸分子克隆到表达载体（如本文所述）中，将表达载体引入宿主细胞（如本文所述），并在宿主细胞中表达 LMP。随后可利用标准蛋白质纯化技术、通过合适的纯化策略将 LMP 从细胞中分离出来。作为重组表达的备选方案，还可使用标准肽合成技术化学合成 LMP 或其肽。此外，可以例如使用抗 LMP 抗体从细胞中分离天然 LMP，所述抗体可通过标准技术，利用本发明的 LMP 或其片段制备。

本发明也提供了 LMP 嵌合或融合蛋白质。本文所用的 LMP “嵌合蛋白质”或“融合蛋白质”包括与非 LMP 多肽有效连接的 LMP 多肽。“LMP

多肽”或“LMP 蛋白质”是指具有 LMP 相应氨基酸序列的多肽，而“非 LMP 多肽”是指具有与 LMP 非基本同源的蛋白质对应的氨基酸序列的多肽，例如与 LMP 不同且来源于相同或不同生物的蛋白质。就融合蛋白质而言的术语“有效连接”旨在表明 LMP 多肽与非 LMP 多肽相互融合，从而使二者都能行使属于所用序列的预期功能。可以将非 LMP 多肽融合于 LMP 多肽的 N-末端或 C-末端。例如，在一个实施方案中，融合蛋白质为 GST-LMP（谷胱甘肽 S-转移酶）融合蛋白质，其中 LMP 序列融合于 GST 序列的 C-末端。此类融合蛋白可促进重组 LMP 的纯化。在另一实施方案中，融合蛋白为在其 N-末端含有异源信号序列的 LMP。在某些宿主细胞（如哺乳动物宿主细胞）中，使用异源信号序列能够提高 LMP 的表达和/或分泌。

优选通过标准重组 DNA 技术生产本发明的 LMP 嵌合或融合蛋白质。例如，根据常规技术将编码不同多肽序列的 DNA 片断于读框内连接在一起，如使用平末端或粘性末端进行连接、限制性酶消化以提供合适的末端、必要时填平粘性末端、碱性磷酸酶处理以避免不期望的连接，以及酶促连接。在另一个实施方案中，可以通过常规技术包括自动化 DNA 合成仪合成融合基因。或者可以使用锚定引物进行基因片断的 PCR 扩增，所述锚定引物能够在两个连续基因片断之间产生互补突出端，使其可随后退火并再扩增产生嵌合基因序列（例如参见 *Current Protocols in Molecular Biology*, Ausubel 等编, John Wiley & Sons: 1992）。此外，许多已经编码融合部分的表达载体（如 GST 多肽）为商业可得。可以将 LMP 编码核酸克隆到这样的表达载体中，从而使融合部分与 LMP 同读框相连。

除了上述编码 LMP 的核酸分子之外，本发明另一方面涉及与其反义的分离的核酸分子。“反义”核酸含有与编码蛋白质的“正义”核酸互补的核苷酸序列，例如与双链 cDNA 分子的编码链互补或与 mRNA 序列互补。因此，反义核酸可以与正义核酸形成氢键。反义核酸可以与整个 LMP 编码链互补，也可仅与其中一部分互补。在一个实施方案中，反义核酸分子与 LMP 编码核苷酸序列的编码链“编码区”反义。术语“编码区”是指

含有能够翻译为氨基酸残基的密码子的核苷酸序列区域(例如 *TaFAD-01* 的完整编码区包含核苷酸 165-1325)。在另一实施方案中,反义核酸分子与 LMP 编码核苷酸序列的编码链“非编码区”反义。术语“非编码区”是指编码区两侧不被翻译成氨基酸的 5'和 3'序列(即也称为 5'和 3'非翻译区)。

给定本文公开的编码 LMP 的编码链序列(例如附录 A 所示序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35),可以按照 Watson 和 Crick 碱基配对原则设计本发明的反义核酸。反义核酸分子可以与 LMP mRNA 的整个编码区互补,但更优选为仅与 LMP mRNA 编码区或非编码区的一部分反义的寡核苷酸。例如,反义寡核苷酸可以与 LMP mRNA 翻译起始位点周围区域互补。反义寡核苷酸长度可为如 5、10、15、20、25、30、35、40、45 或 50 个核苷酸。可以使用化学合成以及酶促连接反应,通过本领域已知的方法构建本发明的反义或正义核酸。例如,可以使用天然存在的核苷酸或多种修饰核苷酸化学合成反义核酸(如反义寡核苷酸),所述修饰核苷酸是设计为了提高分子的生物学稳定性或提高反义和正义核酸间所形成的双链体的物理稳定性,如可使用硫代磷酸酯衍生物和吡啶取代的核苷酸。可用于产生反义核酸的修饰核苷酸实例包括 5-氟尿嘧啶、5-溴尿嘧啶、5-氯尿嘧啶、5-碘尿嘧啶、次黄嘌呤、黄嘌呤、4-乙酰胞嘧啶核苷、5-(羧基羟甲基)尿嘧啶、5-羧基甲基氨基-甲基-2-硫代尿苷、5-羧基甲基氨基甲基尿嘧啶、双氢尿嘧啶、 β -D-半乳糖 Q 核苷、肌苷、N-6-异戊烯基腺嘌呤、1-甲基鸟嘌呤、1-甲基肌苷、2,2-二甲基鸟嘌呤、2-甲基腺嘌呤、2-甲基鸟嘌呤、3-甲基胞嘧啶、5-甲基-胞嘧啶、N-6-腺嘌呤、7-甲基鸟嘌呤、5-甲基氨基甲基尿嘧啶、5-甲氧基氨基-甲基-2-硫尿嘧啶、 β -D-甘露糖 Q 核苷、5'-甲氧基羧基甲基尿嘧啶、5-甲氧基尿嘧啶、2-甲硫基-N-6-异戊烯基腺嘌呤、尿嘧啶-5-氧基乙酸(v)、

wybutoxosine、假尿嘧啶、Q 核苷、2-硫胞嘧啶、5-甲基-2-硫尿嘧啶、2-硫尿嘧啶、4-硫尿嘧啶、5-甲基尿嘧啶、尿嘧啶-5-氧基乙酸甲酯、尿嘧啶-5-氧基乙酸(v)、5-甲基-2-硫尿嘧啶、3-(3-氨基-3-N-2-羧丙基)尿嘧啶、(acp3)w 和 2,6-二氨基嘌呤。或者,可以使用表达载体生物产生反义核酸,所述表达载体中含有按反义方向亚克隆的核酸(即由插入核酸转录得到的 RNA 为目的靶核酸的反义方向,在以下部分将进一步说明)。

在反义技术的另一种变化形式中,可以使用双链干扰 RNA 构建体引起转基因植物中 LMP mRNA 水平以及 LMP 活性的下调。这需要用嵌合构建体转化植物,所述构建体中含有正义方向的 LMP 序列部分,其与相同的 LMP 序列部分的反义序列融合。可以利用长度不等的 DNA 接头区来分隔该构建体中 LMP 序列的正义及反义片断。

本发明的反义核酸分子一般施用于细胞或原位产生,以使其能够与编码 LMP 的细胞 mRNA 和/或基因组 DNA 杂交或与之结合,从而抑制所述蛋白质的表达,如通过抑制转录和/或翻译。可以通过常规的核苷酸互补而形成稳定的双链体进行杂交,或者,如在反义核酸分子结合于 DNA 双链体的情况下,通过在双螺旋大沟内的特定相互作用进行杂交。通过如将反义核酸分子与可结合于细胞表面受体或抗原的肽或抗体相连,可以对反义分子进行修饰而使其能够特异性地与所选择细胞表面的受体或抗原结合。也可使用本文描述的载体将反义核酸分子递送到细胞中。为达到胞内反义分子的足够浓度,优选将反义核酸分子置于原核、病毒或真核(包括植物)的强启动子调控下的载体构建体。

又在另一个实施方案中,本发明的反义核酸分子为端基异构核酸分子。端基异构核酸分子与互补 RNA 形成特殊双链杂合体,其链不同于通常的单元而是相互平行的(Gaultier 等人,1987, Nucleic Acids Res. 15:6625-6641)。反义核酸分子也可含有 2'-O-甲基核糖核苷(Inoue 等人,1987, Nucleic Acids Res. 15:6131-6148)或嵌合型 RNA-DNA 类似物(Inoue 等人,1987, FEBS Lett. 215: 327-330)。

又在另一个实施方案中,本发明的反义核酸为核酶。核酶为具有核糖

核酸酶活性的催化性 RNA 分子,能够裂解与之具有互补区域的单链核酸如 mRNA。因此,可以使用核酶(如锤头状核酶(描述于 Haselhoff & Gerlach, 1988, Nature 334:585-591))催化裂解 LMP mRNA 转录物,由此抑制 LMP mRNA 的翻译。可以根据本文公开的 LMP cDNA 核苷酸序列(即附录 A 的 Bn01,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35)或根据本发明教导方法分离的异源序列,设计对 LMP 编码核酸具有特异性的核酶。例如,可以构建四膜虫(*Tetrahymena*) L-19 IVS RNA 的衍生物,其中活性位点的核苷酸序列与 LMP 编码 mRNA 中待切割的核苷酸序列互补(见如 Cech 等人的美国专利号 4,987,071 和 Cech 等人的美国专利号 5,116,742)。或者,可以使用 LMP mRNA 从 RNA 分子库中选出具有特异性核糖核酸酶活性的催化性 RNA (见如 Bartel, D. & Szostak J.W. 1993, Science 261: 1411-1418)。

或者,可以通过寻靶与 LMP 核苷酸序列调节区(例如 LMP 启动子和/或增强子)互补的核苷酸序列,形成可阻断靶细胞中 LMP 基因转录的三螺旋结构,从而抑制 LMP 的基因表达(一般参见 Helene C. 1991, Anticancer Drug Des. 6:569-84; Helene C.等人, 1992, Ann. N.Y. Acad. Sci. 660:27-36; 以及 Maher, L.J. 1992, Bioassays 14:807-15)。

又在另一实施方案中,可使用微 RNA 技术(Bartel D., Cell, 116: 281-297, 2004)。可设计微 RNA 前体使其寻靶并下调目的基因的表达。所述前体可显著表达于种子或其他组织中。miRNA (~21-25nt)衍生自较大前体,具有转录自非蛋白质编码基因的茎环结构。miRNA 寻靶特定的 mRNA 以在转录后水平(即降解 mRNA)或翻译水平(即抑制蛋白质合成)上抑制基因表达。

本发明的另一方面涉及含有编码 LMP (或其部分)的核酸的载体,优选表达载体。本文所用的术语“载体”是指能够转运与之相连的其他核酸

的核酸分子。一类载体为“质粒”，是指可以连入额外 DNA 片断的环形双链 DNA 环。另一类载体为病毒载体，其中额外 DNA 片断可以连入病毒基因组。某些载体能够在引入的宿主细胞中自主复制（例如，具有细菌复制起点的细菌载体以及游离型哺乳动物载体）。其他载体（如非游离型哺乳动物载体）则在进入宿主细胞后整合入宿主细胞的基因组，从而可随着宿主基因组一同复制。此外，某些载体能够指导与之有效连接的基因的表达。这类载体在本文被称为“表达载体”。一般而言，用于重组 DNA 技术的表达载体通常以质粒的形式存在。在本说明书中，“质粒”和“载体”可以互换使用，因为质粒是最常用的载体形式。然而，本发明旨在包括此类具有等同功能的其他形式的表达载体，例如病毒载体（如复制缺陷型逆转录病毒、腺病毒和腺相关病毒）。

本发明的重组表达载体包含以适于在宿主细胞中表达的形式存在的本发明的核酸，这意味着重组表达载体中包含一个或多个与待表达的核酸序列有效连接的调节序列，所述调节序列的选择根据用于表达的宿主细胞而定。就重组表达载体而言的“有效连接”，是指目的核苷酸序列与调节序列以允许所述核苷酸序列表达的形式连接，从而两个序列互相融合而各自行使其预期功能（如在体外转录/翻译系统或在引入载体的宿主细胞中）。术语“调节序列”包括启动子、增强子以及其他表达控制元件（例如多聚腺苷化信号）。这些调节序列描述于如 Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990)或见于 Gruber 和 Crosby 的 *Methods in Plant Molecular Biology and Biotechnology*, CRC Press, Boca Raton, Florida 编辑: Glick & Thompson, 第 7 章, 89-108 页，包括其中的参考文献。调节序列包括那些能够在许多类型宿主细胞中指导核苷酸序列组成型表达的序列，以及那些只在某些宿主细胞中或某些条件下指导核苷酸序列表达的序列。本领域技术人员将会意识到，表达载体的设计取决于所要转化的宿主细胞的选择、期望达到的蛋白质表达水平等因素。可以将本发明的表达载体引入宿主细胞中，从而产生本文所述核酸所编码的蛋白质或肽，包括融合蛋白质或肽（例如 LMP、LMP 的突变

形式、融合蛋白质等)。

可设计本发明的重组表达载体用于在原核或真核细胞中表达 LMP。例如, 可将 LMP 基因表达于细菌细胞、昆虫细胞(使用杆状病毒表达载体)、酵母及其他真菌细胞(见 Romanos 等人, 1992, *Foreign gene expression in yeast: a review*, *Yeast* 8:423-488; 在 *More Gene Manipulations in Fungi*, Bennet & Lasure 编, 396-428 页: Academic Press: San Diego 中的 van den Hondel, C.A.M.J.J. 等人, 1991, “Heterologous gene expression in filamentous fungi”; 以及 *Applied Molecular Genetics of Fungi*, Peberdy 等人编, 1-28 页, Cambridge University Press: Cambridge 中的 van den Hondel 和 Punt, 1991, “Gene transfer systems and vector development for filamentous fungi”); 藻类 (Falciatore 等人, 1999, *Marine Biotechnology* 1:239-251); 全毛亚纲 (Holotrichia)、缘毛亚纲 (Peritrichia)、旋唇亚纲 (Spirotrichia)、吸管虫类纲 (Suctoria)、四膜虫 (Tetrahymena)、草履虫 (Paramecium)、豆形虫 (Colpidium)、瞬目虫 (Glaucocystis)、匙口虫 (Platyophrya)、Potomacus、假康纤虫 (Pseudocohnilembus)、游仆虫 (Euplotes)、Engelmanniella 以及尾棘虫 (Stylonychia) 属的纤毛虫, 尤其是遵循 WO 98/01572 所述转化方法的带有载体的浮萍棘尾虫 (*Stylonychia lemnae*) 属的纤毛虫; 以及多细胞植物细胞(见 Schmidt & Willmitzer, 1988, *Plant Cell Rep.*: 583-586; *Plant Molecular Biology and Biotechnology*, C Press, Boca Raton, Florida, 6/7 章, S.71-119 (1993); 发表于 *Transgenic Plants*, 卷 1, Engineering and Utilization, Kung 和 Wu 编, Academic Press 1993, 128-43 的 White, Jenes 等人, “Techniques for Gene Transfer”; Potrykus, 1991, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42:205-225(及其中引用的参考文献)) 或哺乳动物细胞。在 Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA 1990 中有对合适宿主细胞的进一步讨论。或者, 可以在体外转录并翻译重组表达载体, 例如使用 T7 启动子调节序列和 T7 聚合酶。

通常使用含有能够调控融合或非融合蛋白质表达的组成型或诱导型启

动子的载体在原核细胞中表达蛋白质。融合载体向其所编码蛋白质中添加若干氨基酸，通常是添加在重组蛋白质的氨基端，但也可以添加到 C-末端或融合入蛋白质的适宜区域。这样的融合载体一般可以实现以下一种或多种目的：1)提高重组蛋白的表达；2)提高重组蛋白的溶解性；以及 3)作为亲和纯化中的配体协助重组蛋白的纯化。在融合表达载体中，通常在融合部分与重组蛋白的连接处引入蛋白水解裂解位点，以使重组蛋白能够在融合蛋白纯化后与融合部分分离。此类酶及其同源识别序列包括 Xa 因子、凝血酶和肠激酶。

典型的融合表达载体包括分别将谷胱甘肽 S-转移酶(GST)、麦芽糖 E 结合蛋白或蛋白质 A 与目的重组蛋白融合的 pGEX (Pharmacia Biotech Inc; Smith 和 Johnson, 1988, Gene 67:31-40)、pMAL (New England Biolabs, Beverly, MA)以及 pRIT5 (Pharmacia, Piscataway, NJ)。在一个实施方案中，将 LMP 编码序列克隆到 pGEX 表达载体中以产生编码融合蛋白的载体，所述融合蛋白从 N-末端到 C-末端包括：GST-凝血酶裂解位点-X 蛋白质。可以使用谷胱甘肽-琼脂糖树脂，通过亲和层析法纯化融合蛋白。可以利用凝血酶裂解融合蛋白来回收与 GST 去融合的重组 LMP。

合适的非融合大肠杆菌(*E.coli*)诱导型表达载体的实例包括 pTrc (Amann 等人, 1988, Gene 69:301-315)和 pET 11d (Studier 等人, 1990, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, California 60-89)。pTrc 载体中的靶基因表达有赖于杂合 trp-lac 融合启动子启动的宿主 RNA 聚合酶转录。而 pET 11d 载体中的靶基因表达则有赖于共表达病毒 RNA 聚合酶(T7 gn1)介导的 T7 gn10-lac 融合启动子启动的转录。此病毒聚合酶由宿主株 BL21 (DE3)或 HMS174 (DE3)从具有处于 lacUV 5 启动子转录控制下的 T7 gn1 基因的定居噬菌体提供。

最大化重组蛋白表达的策略之一是在对所述重组蛋白质的蛋白水解裂解能力缺陷的宿主细菌中表达所述蛋白质(Gottesman S., 1990, Gene Expression Technology: Methods in Enzymology 185: 119-128, Academic Press, San Diego, California)。另一策略则为改变欲插入表达载体的核酸的

核酸序列,使每一氨基酸的个体密码子为选用于表达的细菌中偏好使用的密码子(Wada 等人,1992, *Nucleic Acids Res.* 20:2111-2118)。可以通过标准 DNA 合成技术完成对本发明核酸序列的此类改动。

在另一实施方案中,LMP 表达载体为酵母表达载体。可用于在酿酒酵母(*S. cerevisiae*)中表达的载体实例包括 pYepSec1 (Baldari 等人,1987, *Embo J.* 6:229-234)、pMFa (Kurjan 和 Herskowitz,1982, *Cell* 30:933-943)、pJRY88 (Schultz 等人,1987, *Gene* 54:113-123)以及 pYES2 (Invitrogen Corporation, San Diego, CA)。载体以及构建适用于其他真菌(如丝状真菌)的载体的方法,包括详述于 *Applied Molecular Genetics of Fungi*, Peberdy 等人编,1-28 页, Cambridge University Press: Cambridge 中 van den Hondel 和 Punt,1991,“Gene transfer systems and vector development for filamentous fungi”中的载体和方法。

或者,可使用杆状病毒表达载体于昆虫细胞中表达本发明的 LMP。可在培养的昆虫细胞(如 Sf 9 细胞)中表达的杆状病毒载体包括 pAc 系列(Smith 等人,1983, *Mol. Cell Biol.* 3:2156-2165)和 pVL 系列(Lucklow 和 Summers,1989, *Virology* 170:31-39)。

而在另一实施方案中,使用哺乳动物表达载体在哺乳动物细胞中表达本发明的核酸。哺乳动物表达载体的实例包括 pCDM8 (Seed,1987, *Nature* 329: 840)和 pMT2PC (Kaufman 等人,1987, *EMBO J.* 6: 187-195)。当用于哺乳动物细胞时,通常使用病毒调节元件来提供表达载体的控制功能。例如,常用启动子来源于多瘤病毒、腺病毒 2、巨细胞病毒以及猿病毒 40。其他可同时用于原核及真核细胞的合适表达系统,见 Sambrook, Fritsh 和 Maniatis, *Molecular Cloning: A Laboratory Manual*, 第二版, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989 第 16 及 17 章。

在另一实施方案中,可在单细胞植物细胞(如藻类,见 Falciatore 等人,1999, *Marine Biotechnology* 1:239-251 及其中的参考文献)和高等植物的植物细胞(例如种子植物,如作物植物)中表达本发明的 LMP。植物表

达载体的实例包括 Becker, Kemper, Schell 和 Masterson (1992, "New plant binary vectors with selectable markers located proximal to the left border", *Plant Mol. Biol.* 20:1195-1197) 以及 Bevan (1984, "Binary *Agrobacterium* vectors for plant transformation, *Nucleic Acids Res.* 12:8711-8721; *Transgenic Plants*, 卷 1, Engineering and Utilization, Kung 和 R. Wu 编辑, Academic Press, 1993, S. 15-38 中的 Vectors for Gene Transfer in Higher Plants) 中详述的那些。

植物表达盒中优选含有有效连接的能够驱动植物细胞中基因表达的调节序列, 从而使各序列可行使其功能(如终止转录), 包括多聚腺苷化信号。优选的多聚腺苷化序列来源于根癌农杆菌(*Agrobacterium tumefaciens*)的 t-DNA, 如 Ti-质粒 pTiACH5 章鱼碱合成酶的基因 3 (Gielen 等人, 1984, *EMBO J.* 3: 835) 或其功能等效物, 而且在植物中具有功能活性的所有其他终止子也适用。

由于植物基因表达常常并不仅限于转录水平, 植物表达盒优选含有其他有效连接的序列如翻译增强子, 如含有烟草花叶病毒 5' 非翻译前导序列的超驱动序列, 从而提高蛋白质/RNA 比率 (Gallie 等人, 1987, *Nucleic Acids Res.* 15:8693-8711)。

植物基因表达必须与合适的启动子有效连接, 所述启动子使基因以适时的、细胞或组织特异性方式表达。优选可驱动组成型表达的启动子 (Benfey 等人, 1989, *EMBO J.* 8:2195-2202), 如那些来源于植物病毒的启动子, 如 35S CAMV (Franck 等人, 1980, *Cell* 21:285-294)、19S CaMV (也见 US 5,352,605 及 WO 84/02913), 或如来源于 1,5-二磷酸核酮糖羧化酶/加氧酶 (Rubisco) 小亚基的植物启动子 (US 4,962,028 所述)。更优选的为能在种子发育的所有或选择的阶段驱动 LMP 蛋白质表达的种子特异性启动子。种子特异性植物启动子为本领域普通技术人员公知, 并可用种子特异性 mRNA 文库及表达分型技术对其进行鉴定和表征。种子特异性启动子包括油菜籽的 napin 基因启动子 (US 5,608,152)、蚕豆 (*Vicia faba*) USP 启动子 (Baeumlein 等人, 1991, *Mol. Gen. Genetics* 225:459-67)、拟南芥油

体蛋白启动子(WO 98/45461)、菜豆(*Phaseolus vulgaris*)的菜豆蛋白启动子(US 5,504,200)、芸苔属 Bce4-启动子(WO9113980)或豆球蛋白 B4 启动子(LeB4; Baeumlein 等, 1992, Plant J. 2:233-239), 以及可赋予单子叶植物如玉米、大麦、小麦、裸麦、稻等种子特异性表达的启动子。需要指出的合适启动子为大麦的 lpt2 或 lpt1 基因启动子(WO 95/15389 和 WO 95/23230), 或 WO 99/16890 中所述的启动子(大麦的大麦醇溶蛋白基因启动子、稻谷蛋白基因启动子、稻水稻素基因启动子、小麦醇溶蛋白基因启动子、小麦谷蛋白基因启动子、玉米醇溶蛋白基因启动子、燕麦谷蛋白基因启动子、高粱 kasirin 基因启动子以及裸麦黑麦碱(secalin)基因启动子)。

也可以通过诱导型启动子促进植物基因表达(综述可见 Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol. 48:89-108)。化学诱导型启动子特别适用于需要以时间特异性方式进行基因表达的情形。此类启动子的实例为水杨酸诱导的启动子(WO 95/19443)、四环素诱导的启动子(Gatz 等人, 1992, Plant J. 2:397-404), 以及乙醇诱导的启动子(WO93/21334)。

应答生物或非生物胁迫条件的启动子也是适用的启动子, 例如病原诱导的 PRP-1 基因启动子(Ward 等人, 1993, Plant. Mol. Biol. 22:361-366)、热诱导的番茄 hsp80 启动子(US 5,187,267)、冷诱导的马铃薯 α -淀粉酶启动子(WO 96/12814), 以及创伤诱导的 pinII-启动子(EP 375091)。

其他可用于植物基因表达盒的优选序列为将基因产物定向于其适当的细胞区室所必需的靶向序列(综述见 Kermode 1996, Crit. Rev. Plant Sci. 15: 285-423 及其引用的参考文献), 所述细胞区室如液泡、细胞核、所有类型的质体如造粉体、叶绿体、有色体、胞外空间、线粒体、内质网、油体、过氧化物酶体以及其他植物细胞区室。特别合适的启动子为那些赋予启动质体特异性基因表达的启动子, 这是因为质体是脂质生物合成中前体及某些终产物合成的场所。合适的启动子如病毒 RNA 聚合酶启动子, 描述于 WO 95/16783 和 WO 97/06250, 而拟南芥的 clpP 启动子描述于 WO 99/46394。

本发明还提供了重组表达载体, 其中含有以反义方向克隆到表达载体

中的本发明 DNA 分子。即 DNA 分子以允许与 LMP mRNA 反义的 RNA 分子表达（通过 DNA 分子转录）的方式与调节序列有效连接。可以选择可在多种细胞类型中指导反义 RNA 分子持续表达的序列，作为与反义方向克隆的核酸有效连接的调节序列，如病毒启动子和/或增强子；也可以选择指导反义 RNA 的组成型、组织特异性或细胞类型特异性表达的调控序列。反义表达载体的形式可为重组质粒、噬菌粒或减毒病毒，其中反义核酸的产生受到高效调节区域的调控，其活性是由引入载体的细胞类型决定的。利用反义基因调节基因表达的讨论见 Weintraub 等人(1986, *Antisense RNA as a molecular tool for genetic analysis, Reviews-Trends in Genetics*, 卷 1)以及 Mol 等人(1990, *FEBS Lett.* 268:427-430)。

本发明的另一方面涉及引入了本发明重组表达载体的宿主细胞。本文术语“宿主细胞”和“重组宿主细胞”可互换使用。应当理解这些术语不仅仅指具体的受试细胞，也指该细胞的后代或潜在的后代。在后续世代中由于突变或环境影响会产生某些修饰，这样的后代实际上可能与其亲代细胞不相同，但仍包括在本文所用的术语范围内。宿主细胞可以为任何原核或真核细胞。例如，可以在细菌细胞、昆虫细胞、真菌细胞、哺乳动物细胞（如中国仓鼠卵巢细胞(CHO)或 COS 细胞）、藻类、纤毛虫或植物细胞中表达 LMP。其他合适的宿主细胞为本领域技术人员所公知。

可以通过常规转化或转染技术将载体 DNA 引入原核或真核细胞中。本文所用的术语“转化”、“转染”、“接合”与“转导”是指将外源核酸（如 DNA）引入宿主细胞的多种本领域公认的技术，包括磷酸钙或氯化钙共沉淀、DEAE-葡聚糖介导的转染、脂转染法、自然感受态、化学介导的转移或电穿孔法。用于转化或转染宿主细胞包括植物细胞的合适方法可见于 Sambrook 等人(1989, *Molecular Cloning: A Laboratory Manual*, 第二版, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY)以及其他实验手册如 *Methods in Molecular Biology*, 1995, 卷 44, “*Agrobacterium protocols*”, Gartland 和 Davey 编, Humana Press, Totowa, New Jersey。

对于稳定转染哺乳动物及植物细胞而言, 已知取决于所用表达载体和转染技术的不同, 仅有一小部分细胞可能将外源 DNA 整合进自身基因组中。为了鉴定并选出这些整合体, 一般将编码可选择标记(如抗生素抗性)的基因与目的基因一起引入宿主细胞。优选的可选择标记包括那些能赋予药物如 G418、潮霉素、卡那霉素和氨甲蝶呤抗性的标记, 或能在植物中赋予除草剂如草甘膦或草胺磷抗性的标记。可以在编码 LMP 的同一载体上向宿主细胞中引入编码可选择标记的核酸, 或也可使用单独的载体引入。可以通过如药物选择(如整合选择标记基因的细胞能够存活, 而其他细胞死亡)的方法鉴定被引入的核酸稳定转染的细胞。

为产生同源重组微生物, 制备含有至少一部分 LMP 基因的载体, 所述 LMP 基因中引入了缺失、添加或取代, 以改变(如功能性破坏) LMP 基因。优选的 LMP 基因为拟南芥、大豆、稻、玉米、亚麻、大麦或小麦 LMP 基因, 但也可为相关植物或甚至哺乳动物、酵母或昆虫来源的同源物。在优选实施方案中, 设计载体使内源性 LMP 基因在同源重组时被功能性破坏(即不再编码有功能蛋白质; 也称为敲除载体)。或者, 可以设计载体, 使内源性 LMP 基因在同源重组时发生突变或其他改变但仍然编码有功能蛋白质(如可以改变上游调节区域而改变内源 LMP 的表达)。为通过同源重组产生点突变, 可在嵌合修复技术中使用 DNA-RNA 杂合体(Cole-Strauss 等人, 1999, *Nucleic Acids Res.* 27:1323-1330 以及 Kmiec 1999, *American Scientist* 87:240-247)。拟南芥或其他作物中的同源重组方法也为本领域所公知, 是本文考虑使用的。

在同源重组载体中, LMP 基因改变了的部分在其 5'和 3'末端侧接额外的 LMP 基因核酸, 从而使同源重组可以发生在载体所携带的外源性 LMP 基因与微生物或植物的内源性 LMP 基因之间。额外的侧翼 LMP 核酸的长度足够与内源性基因进行成功的同源重组。一般而言, 载体中含有数百碱基对至数千碱基对的侧翼 DNA(同时在 5'和 3'末端)(见如 Thomas 和 Capecchi, 1987, *Cell* 51:503 中有关同源重组载体的描述)。将载体引入微生物或植物细胞中(如通过聚乙二醇介导的 DNA)。使用本领域公知的

技术选出引入的 LMP 基因与内源性 LMP 基因发生了同源重组的细胞。

在另一实施方案中，可以产生含有能够调节性表达引入基因的选择系统的重组微生物。例如，将载体中的 LMP 基因置于 lac 操纵子的调控下，可以使 LMP 基因仅在 IPTG 存在时表达。此类调节系统为本领域所熟知。

可以使用本发明的宿主细胞（如培养中的原核或真核宿主细胞）生产（即表达）LMP。因此，本发明还提供利用本发明的宿主细胞生产 LMP 的方法。在一个实施方案中，该方法包括在适当的培养基中培养本发明的宿主细胞（其中已经引入编码 LMP 的重组表达载体，或该细胞基因组中含有野生型或改变的 LMP 基因），直至产生 LMP。在另一实施方案中，该方法还包括从培养基或宿主细胞中分离 LMP。

本发明的另一方面涉及分离的 LMP 及其生物学活性部分。“分离的”或“纯化的”蛋白质或其生物学活性部分基本不含有使用重组 DNA 技术生产时的细胞材料，或采用化学合成时的化学前体或其他化学物。用语“基本不含有细胞材料”包括这样的 LMP 制品，其中该蛋白质与天然或重组产生该蛋白质的细胞的细胞成分分开。在一个实施方案中，用语“基本不含有细胞材料”包括这样的 LMP 制品，其中非 LMP（本文也称为“杂质蛋白质”）含量少于约 30%（干重），更优选非 LMP 含量少于约 20%，更优选非 LMP 含量少于约 10%，而最优选非 LMP 含量少于约 5%。当重组生产 LMP 或其生物活性部分时，也优选其基本不含有培养基，即培养基少于蛋白质制品总量的约 20%，更优选少于约 10%，而最优选少于约 5%。用语“基本不含有化学前体或其他化学物”包括这样的 LMP 制品，其中该蛋白质与该蛋白质合成中涉及的化学前体或其他化学物分开。在一个实施方案中，用语“基本不含有化学前体或其他化学物”包括这样的 LMP 制品，其中含有少于约 30%（干重）化学前体或非 LMP 化学物，更优选化学前体或非 LMP 化学物少于约 20%，更优选化学前体或非 LMP 化学物少于约 10%，而最优选化学前体或非 LMP 化学物少于约 5%。在优选实施方案中，分离的蛋白质或其生物学活性部分中不含有 LMP 来源相同生物的杂质蛋白质。一般而言，通过如在拟南芥、大豆、稻、玉米、亚麻、

大麦或小麦以外的植物或微生物、藻类或真菌中重组表达拟南芥、大豆、稻、玉米、亚麻、大麦或小麦 LMP，来制备此类蛋白质。

本发明的分离 LMP 或其生物活性部分能够参与产生拟南芥、大豆、稻、玉米、亚麻、大麦或小麦中种子贮存化合物或细胞膜所必需的化合物代谢，或具有表 3 列出的一种或多种活性。在优选实施方案中，蛋白质或其部分包含与附录 A 所示的核酸，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 编码的氨基酸序列基本同源的氨基酸序列，从而使该蛋白质或其部分保留了参与拟南芥、大豆、稻、玉米、亚麻、大麦或小麦细胞膜构建或跨膜分子转运所必需的化合物代谢。蛋白质部分优选如本文所述的生物活性部分。在另一优选的实施方案中，本发明的 LMP 具有由附录 A 所示核酸，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 编码的氨基酸序列。在另一优选实施方案中，LMP 具有由核苷酸序列编码的氨基酸序列，所述核苷酸序列可与附录 A 所示核苷酸序列，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 杂交，例如在严紧条件下杂交。在另一优选实施方案中，LMP 具有由核苷酸序列编码的氨基酸序列，所述氨基酸序列与附录 A 所示核酸，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、

SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 编码的氨基酸序列之间的同源性至少为约 50-60%，优选至少约 60-70%，更优选至少约 70-80%、80-90%或 90-95%，也优选至少约 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%，甚至最优选与附录 A 所示核酸，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 编码的氨基酸序列之间的同源性至少约 96%、97%、98%、99%或更高。优选利用 Vector NTI Suite 9.0 的 Align X (2003 年 8 月 22 日)确定多肽序列的同源性，设定参数如下：空位开放罚分：10；空位延伸罚分：0.05；空位分离罚分范围：8；比对延迟的%一致性：40。

本发明的优选 LMP 也优选具有本文所述的至少一种 LMP 活性。例如，本发明的优选 LMP 包括由可与附录 A 所示核苷酸序列，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 杂交，例如在严格条件下杂交的核苷酸序列编码的氨基酸序列，且能够参与拟南芥、大豆、稻、玉米、亚麻、大麦或小麦细胞膜构建或跨膜分子转运所必需的化合物代谢，或具有表 3 列出的一种或多种活性。

在其他实施方案中，LMP 与附录 A 所示核酸，如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 编码的氨基酸序列基本上同源，正如前文所详述，虽然在氨基酸序列上由于天然变异或诱变有

所不同,但还是保留了由附录 A 核酸序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一编码蛋白质的功能活性。因此,在另一实施方案中,LMP 为含有氨基酸序列的蛋白质,所述氨基酸序列与附录 A 所示核酸序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 之一的同源性至少为约 50-60%,优选至少约 60-70%,更优选至少约 70-80%、80-90%或 90-95%,也优选至少约 85%、86%、87%、88%、89%、90%、91%、92%、93%、94%或 95%,甚至更为优选与完整氨基酸序列之编码核酸所编码的序列之间的同源性至少约 96%、97%、98%、99%或更高,并具有本文所述的至少一种 LMP 活性。优选利用 Vector NTI Suite 9.0 的 Align X (2003 年 8 月 22 日)确定多肽序列的同源性,设定参数如下:空位开放罚分: 10;空位延伸罚分: 0.05;空位分离罚分范围: 8;比对延迟的%一致性: 40。在另一实施方案中,本发明涉及全长拟南芥、大豆、稻、玉米、亚麻、大麦或小麦蛋白质,其与附录 A 所示的核酸序列,如优选实施方案中公开的 SEQ ID NO: 5、SEQ ID NO: 7、SEQ ID NO: 9、SEQ ID NO: 11、SEQ ID NO: 13、SEQ ID NO: 15、SEQ ID NO: 17、SEQ ID NO: 19、SEQ ID NO: 21、SEQ ID NO: 23、SEQ ID NO: 25、SEQ ID NO: 27、SEQ ID NO: 29、SEQ ID NO: 31、SEQ ID NO: 33 或 SEQ ID NO: 35 编码的完整氨基酸序列基本上同源。

可以使用显性失活突变或反式显性抑制来降低 LMP 在转基因种子中的活性,以改变种子贮存化合物的水平。为此,产生能够消除 LMP 活性的突变,并在转基因植物中过表达无活性的非功能性 LMP 基因。无活性的反式显性 LMP 蛋白质与活性内源性 LMP 蛋白质竞争底物或其他蛋白

质发生反应，而稀释活性 LMP 的活性。通过这一途径可以降低 LMP 的生物学活性而不必实际改变内源性 LMP 的基因表达。Pontier 等人使用这一策略以调节植物转录因子的活性 (Pontier D, Miao ZH, Lam E, Plant J. 2001 Sep. 27(6):529-38, Trans-dominant suppression of plant TGA factors reveals their negative and positive roles in plant defense responses)。

LMP 同源物可以通过诱变产生，例如不连续点突变或截短 LMP。本文所用的术语“同源物”是指用作 LMP 活性激动剂或拮抗剂的 LMP 变体形式。LMP 激动剂能够保留与 LMP 基本相同的生物学活性或其活性的子集。LMP 拮抗剂能够通过如竞争性结合包括 LMP 在内的细胞膜成分代谢级联反应中的上游或下游成员、或通过与介导化合物跨膜转运的 LMP 结合来抑制 LMP 天然形式的一种或多种活性，从而阻止易位的发生。

在可选的实施方案中，可以通过在 LMP 突变体（如截短突变体）组合文库中筛查 LMP 激动剂或拮抗剂活性，确定 LMP 的同源物。在一个实施方案中，通过在核酸水平组合诱变构建 LMP 变体杂合文库，该 LMP 变体杂合文库由一个杂合的基因文库编码而成。例如，可以通过将合成寡核苷酸混合物以酶连接于基因序列的方法产生 LMP 变体杂合文库，从而使一系列简并的潜在 LMP 序列以单个多肽的形式表达，或者以一系列更大的含所述系列 LMP 序列的融合蛋白（如对于噬菌体展示文库而言）的形式表达。有很多种方法可以用于从简并的寡核苷酸序列中产生潜在的 LMP 同源物文库。可以在自动 DNA 合成仪中完成简并基因序列的化学合成，然后将合成基因连入合适的表达载体中。使用基因的简并集合可以在一个混合物中提供编码所需潜在 LMP 序列集合的所有编码序列。合成简并寡核苷酸的方法为本领域所公知（见如 Narang, 1983, Tetrahedron 39:3; Itakura 等人, 1984, Annu. Rev. Biochem. 53:323; Itakura 等人, 1984, Science 198:1056; Ike 等人, 1983, Nucleic Acids Res. 11:477）。

此外，可以利用 LMP 编码序列的片断文库建立 LMP 片断的杂合群，用于筛选及随后选择 LMP 的同源物。在一个实施方案中，可以用核酸酶在每分子仅发生约一次切口形成的条件下处理 LMP 编码序列的双链 PCR

片断, 将该双链 DNA 变性, 复性 DNA 以形成含有来自不同切口产物正义/反义对的双链 DNA, 用 S1 核酸酶处理去除再形成双链体中的单链部分, 并将所产生的片断文库连入表达载体, 从而建立编码序列的片断文库。通过这种方法, 可以产生编码 LMP N-末端、C-末端及大小不等的内部片断的表达文库。

已知本领域有若干技术可用于筛选通过点突变或截短产生的组合文库的基因产物, 以及在 cDNA 文库中筛选具有所选性质的基因产物。这些技术适用于快速筛选由组合诱变 LMP 同源物产生的基因文库。最广泛用于高通量分析筛查大型基因库的技术一般包括将基因文库克隆到可复制的表达载体中, 用所得的载体文库转化适当的细胞, 并在一定条件下表达组合基因, 所述条件下, 目的活性的检测有助于分离编码其产物被检测的基因的载体。递归整体诱变(recursive ensemble mutagenesis, REM), 一项能够增加文库中功能性突变体频率的新技术, 也可与筛查测定法组合使用鉴定 LMP 同源物 (Arkin & Yourvan, 1992, Proc. Natl. Acad. Sci. USA 89:7811-7815; Delgrave 等人, 1993, Protein Engineering 6: 327-331)。

在另一实施方案中, 可以使用本领域熟知的方法, 利用基于细胞的测定法分析杂合 LMP 文库。

本文所述的核酸分子、蛋白质、蛋白质同源物、融合蛋白质、引物、载体以及宿主细胞可以用于以下一种或多种方法: 鉴定拟南芥、大豆、稻、玉米、亚麻、大麦或小麦及相关生物; 绘制拟南芥、大豆、稻、玉米、亚麻、大麦或小麦相关生物的基因组图谱; 鉴定并定位拟南芥、大豆、稻、玉米、亚麻、大麦或小麦目的序列; 进化研究; 确定 LMP 功能所需的区域; 调控 LMP 活性; 调控一种或多种细胞功能的代谢; 调控一种或多种化合物的跨膜转运; 以及调控种子贮存化合物的积累。

植物拟南芥代表了高等(或种子)植物的一页。它与如欧洲油菜、大豆、稻、玉米、亚麻、大麦或小麦等其他植物相关, 都需要光来驱动光合作用及生长。植物如拟南芥、欧洲油菜、大豆、稻、玉米、亚麻、大麦或小麦在 DNA 序列和多肽水平上具有高度同源性, 从而能够利用其他植物

或生物来源的探针对 DNA 分子进行异源筛选，进而能够导出共有序列，其适于在第三物种中异源筛选或解释功能并预测基因功能。因此，鉴定这类功能的能力具有重要实用性，如预测酶的底物特异性。此外，这些核酸分子可以用作绘制拟南芥基因组或相关生物基因组图谱的参照点。

本发明的 LMP 核酸分子具有多种用途。首先，本发明的核酸及蛋白质分子可作为基因组特定区域的标志。这不仅可用于绘制基因组图谱，也可用于拟南芥、欧洲油菜、大豆、稻、玉米、亚麻、大麦或小麦蛋白质的功能研究。例如，为了鉴定某一特定的拟南芥、欧洲油菜、大豆、稻、玉米、亚麻、大麦或小麦 DNA 结合蛋白的基因组结合区域，可以消化拟南芥、欧洲油菜、大豆、稻、玉米、亚麻、大麦或小麦的基因组，并将其片断与 DNA 结合蛋白孵育。可另外用本发明的核酸分子（优选具有易检测标记的核酸分子）作为探针与那些与蛋白质结合的片断杂交；此类核酸分子与基因组片断的结合能够在拟南芥、欧洲油菜、大豆、稻、玉米、亚麻、大麦或小麦的基因组图谱上定位该片断，而且当用不同的酶进行多次重复时，还有助于快速确定蛋白质结合的核酸序列。此外，本发明的核酸分子可能与相关物种的序列有充分的同源性，使得这些核酸分子可以用作构建相关植物基因组图谱的标记物。

本发明的 LMP 核酸分子也可用于进化和蛋白质结构研究。多种原核及真核细胞使用本发明分子所参与的代谢和转运过程；通过比较本发明的核酸分子序列与其他生物中那些编码相似酶的核酸分子序列，能够评价生物之间的进化相关性。与之类似，这样的比较能够评价序列中哪些为保守区域而哪些不是，可能有助于确定那些对于酶功能至关重要的蛋白质区域。这类确定对于蛋白质遗传工程研究具有一定价值，可能会提示该蛋白质在不丧失功能的前提下能够耐受怎样的诱变。

对于本发明 LMP 核酸分子的操作可能会导致产生与野生型 LMP 功能有所差异的 LMP。这些蛋白质可能在效用或活性上有所改善、可能在细胞中以高于正常的量存在，也可能效用或活性有所降低。

本发明 LMP 的改变可以通过几种机制直接影响种子贮存化合物的累

积和/或成分。在表达 LMP 的植物中, 转运的增加会导致化合物累积的改变和/或植物组织及器官中的溶质分配(partitioning), 最终在种子发育过程中影响一种或多种种子贮存化合物的累积。Mitsukawa 等人(1997, Proc. Natl. Acad. Sci. USA 94:7098-7102)提供了一个实例, 其中在烟草培养细胞中过表达拟南芥高亲和力磷酸盐转运蛋白基因, 增强了磷酸盐限制条件下的细胞生长。磷酸盐的可利用性也显著影响糖和代谢中间物的产生(Hurry 等人, 2000, Plant J. 24:383-396)以及叶和根中的脂质组成(Härtel 等人, 2000, Proc. Natl. Acad. Sci. USA 97:10649-10654)。与之类似, 已经证实植物乙酰辅酶 A 羧化酶 (ACCase) 的活性受到磷酸化调节 (Savage 和 Ohlrogge, 1999, Plant J. 18:521-527), 且作用于 ACCase 的激酶和磷酸酶 (LMP) 的活性改变, 会导致种子脂质累积水平的提高或降低。此外, 叶绿体包膜上存在脂质激酶活性提示信号转导路径和/或膜蛋白质的调节发生在包膜上 (见如 Müller 等人, 2000, J. Biol. Chem. 275:19475-19481 及其中所引文献)。ABI1 和 ABI2 基因编码的两个蛋白质丝氨酸/苏氨酸磷酸酶 2C 为脱落酸代谢途径的调节物, 因此见于种子发育的早期和晚期阶段 (如 Merlot 等人, 2001, Plant J. 25:295-303)。更多实例也可见前文“发明背景”章节。

本发明也提供了可与本文公开或描述的核酸编码的 LMP 多肽或其部分特异性结合的抗体。

可以通过许多众所周知的方法制备抗体 (见如 Harlow 和 Lane, Antibodies; A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1988)。简言之, 可将纯化抗原以足以激发免疫反应的量和间隔注射到动物体内。可直接纯化抗体, 也可从动物中获取脾细胞。然后将这些细胞与永生化细胞系融合并筛查其抗体分泌。这些抗体可以用于筛选核酸克隆文库以寻找分泌该抗原的细胞。随后可对阳性克隆进行测序 (见如 Kelly 等人, 1992, Bio/Technology 10:163-167; Bebbington 等人, 1992, Bio/Technology 10:169-175)。

短语与多肽“选择性结合”是指能对蛋白质在异源蛋白质和其他生物

制品群中是否存在具有确定性的结合反应。因此，在指定的免疫测定条件下，与特定蛋白质结合的具体抗体并不会与该样品中的其他蛋白质发生显著量的结合。在此类条件下与抗体的选择性结合可能需要选择一种对特定蛋白质有特异性的抗体。有很多免疫测定方法可用于选择能够与特定蛋白质选择性结合的抗体。例如，常规使用固相 ELISA 免疫测定法选择可与蛋白质发生选择性免疫反应的抗体。见 Harlow 和 Lane, “Antibodies, A Laboratory Manual” Cold Spring Harbor Publications, New York (1988) 中关于可用于测定选择性结合的免疫测定方法和条件的描述。

某些情况下需要从多种宿主中制备单克隆抗体。制备此类单克隆抗体的技术描述可见于 Stites 等人编 “Basic and Clinical Immunology,” (Lange Medical Publications, Los Altos, Calif., 第四版) 及其所引参考文献，以及 Harlow 和 Lane (“Antibodies, A Laboratory Manual” Cold Spring Harbor Publications, New York, 1988)。

本申请参考了多种出版物。所有这些出版物及其中所引用的参考文献的公开内容在此申请中全文引用作为参考，以便更详细地说明本发明所属领域的状况。

对本领域技术人员而言，显然可以对本发明进行多种修饰和改变而不背离本发明的范围或精神。参考本文公开的说明和本发明的实施，本发明的其他实施方案对于本领域技术人员而言是显而易见的。应当理解说明书和实施例均为例示性的，本发明的确切范围和精神参见本文的权利要求书。

实施例

实施例 1：一般步骤

a) 一般克隆步骤：

根据 Sambrook 等人 (1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6) 或 Kaiser、Michaelis 和 Mitchell (1994, Methods in Yeast Genetics, Cold Spring Harbor Laboratory Press: ISBN 0-87969-451-3) 中所述进行克隆步骤，例如限制性切割、琼脂糖凝胶电泳、

DNA 片断纯化、转运核酸至硝化纤维素和尼龙膜上、DNA 片断连接、大肠杆菌和酵母细胞的转化、细菌生长以及重组 DNA 的序列分析。

b) 化学物质:

若非文中另有说明, 所用化学物为来自 Fluka (Neu-Ulm)、Merck (Darmstadt)、Roth (Karlsruhe)、Serva (Heidelberg) 以及 Sigma (Deisenhofen) 公司的 p.a. 级产品。使用来自 Milli-Q 水系统水纯化工厂 (Millipore, Eschborn) 的纯化的无致热原的水 (下文称为 H_2O) 配制溶液。限制性内切酶、DNA 修饰酶以及分子生物学试剂盒获自 AGS (Heidelberg)、Amersham (Braunschweig)、Biometra (Göttingen)、Boehringer (Mannheim)、Genomed (Bad Oeynhausen)、New England Biolabs (Schwalbach/Taunus)、Novagen (Madison, Wisconsin, USA)、Perkin-Elmer (Weiterstadt)、Pharmacia (Freiburg)、Qiagen (Hilden) 及 Stratagene (Amsterdam, Netherlands) 公司。除非另有说明, 均根据生产商的说明使用它们。

c) 植物材料和生长:

拟南芥植物

对于该研究, 使用拟南芥野生型和 *fad2* 突变体 (如 Miquel & Browse, 1992, J Biol Chem 267: 1502-1509 中所述) 植物的根材料、叶、长角果和种子。将野生型和 *fad2* 突变体拟南芥种子在 $4^{\circ}C$ 黑暗中预温育三天后置于 $60-80 \mu mol m^{-2} s^{-1}$ 光子通量密度、16 小时 ($22^{\circ}C$) 光照期和 8 小时黑暗期 ($18^{\circ}C$) 的培养箱 (AR-75, Percival Scientific, Boone, IA) 中。所有的植物在 pH 6.2、2% 蔗糖和 1.2% 琼脂的半浓度 MS 培养基 (Murashige & Skoog, 1962, *Physiol. Plant.* 15:473-497) 上发芽。种子在 20% 漂白剂 0.5% triton $\times$ 100 中灭菌 20 分钟并用过量的无菌水冲洗 6 次。植物如上所述或在如 Klaus 等人 (2002, *Plant Physiol.* 128:885-895) 所述标准条件下在土壤中生长。

大豆

本研究使用大豆变种 Resnick 创建 cDNA 文库。从某些情况下经暗、盐、热和旱处理的植物中收集种子、种荚、花、叶、茎及根组织。在某些情况下，植物也经线虫感染。然而，本研究集中研究种子和种荚组织在 cDNA 文库中的用途。对植物打上标签，在开花期后固定日期收获种子：5-15 天、15-25 天、25-35 天及 33-50 天。

稻

本研究使用粳稻日本晴变种 (*Oryza sativa* ssp. Japonica cv. Nippon-barre) 创建 cDNA 文库。从某些情况下经暗、盐、热和旱处理的植物中收集种子、种荚、花、叶、茎及根组织。本研究集中研究种子胚组织在 cDNA 文库中的用途。分别收集胚和胚乳以防胚乳组织可能干扰 RNA 提取。植物生长于 85°F 温室 Wisconsin 土壤中(高有机物含量)，光周期为 14 h。从发育的种子中切下稻胚。

玉米

使用玉米杂交株 B73 × Mo17 及 B73 近交系(杂交株 B73 × Mo17 的雌性近交亲本)创建 cDNA 文库。收集植物的果实或种子(受粉后 1 至 9 天的受精卵/幼粒；受粉后 23 天的乳熟期谷粒[R3, 早期淀粉生成]；大田生长植物受粉后 30 天的蜡熟初期谷粒(R4)，具有淀粉颗粒及发育良好的胚；发泡期(blister)的极幼粒[R2, 水胚乳]；受粉后 36 天的顶凹初期(early dent stage)谷粒(R5)，胚乳变得坚实；B73 近交，受粉后 9 天及 19 天的谷粒)、花(穗发育：6 cm (V10) 及以上的穗，包括开花，受粉后 44 至 70 天；耳穗发育：耳穗抽穗 2 cm (V13) 及以上，包括抽穗丝(未受粉)，受粉后 51 至 70 天)、叶/芽/花结(所有种子填充前的混合期；包括大田中抽穗前的 a) 3 叶植物(V3)、b) 6 叶植物(V6)及 c) 较老叶源(地面第 3)的叶)、茎(2 至 5 叶植物的地下部分；受粉后 13 至 29 天的大田生长植物，除去根和大部分叶组织；大田生长植物受粉后 56 至 84 天，抽穗期和种子填充期(乳熟期)穗

附近的茎组织)和根组织(从幼年至中年植物: 来自受粉后 12 至 35 天的幼苗、6 叶植物和 9 叶植物)。

亚麻

本研究使用亚麻变种 00-44427 及变种 00-44338 (来自 Svalöf Weibull 收藏) 创建 cDNA 文库。植物在 19℃ 生长于冷却温室腔内 2 升罐的罐装土壤内, 其中含有 5 ml 奥绿肥 Osmocote/升土壤。在 15 daa (胚处于急速增长期, 约占据种子的 2/3; 鱼雷晚期胚为绿色)、25 daa (胚完全伸长且宽度增加, 整个种子变得充实; 胚仍为复绿色) 及 33 daa (种子开始成熟; 胚的颜色变为淡绿且尖部为黄色) 时从发育的种子中收集材料。

大麦

本研究使用大麦变种 Morex 创建 cDNA 文库。植物在 15 h 光周期下生长于温室的 metromix 中, 日间温度 23℃ 而夜间 18℃。谷粒处于水熟期至乳熟晚期。谷穗中至上部基本被收获。在种植 75 天后收集种子材料。

小麦

本研究使用小麦变种 Galeon 创建 cDNA 文库。从某些情况下经暗、盐、热和旱处理的植物中收集种子、花、果实、叶、茎及根组织。植物在 12 h 光周期下生长于温室的 metromix 中, 日间温度 72°F 而夜间 65°F。

实施例 2: 从植物中分离总 DNA

分离总 DNA 的细节涉及处理 1g 鲜重的植物材料。

CTAB 缓冲液: 2% (w/v) N-鲸蜡基-N,N,N-三甲基溴化铵(CTAB); 100 mM Tris HCl pH 8.0; 1.4 M NaCl; 20 mM EDTA。

N-十二烷基肌氨酸缓冲液: 10% (w/v) N-十二烷基肌氨酸; 100 mM Tris HCl pH 8.0; 20 mM EDTA。

将植物材料于液氮中用研钵研磨为细粉, 并将其转移至 2 ml

Eppendorf 管中。随后用一层 1 ml 的分解缓冲液 (1ml CTAB 缓冲液, 100 μ l N-十二烷基肌氨酸缓冲液, 20 μ l β -巯基乙醇以及 10 mg/ml 的蛋白酶 K 溶液 10 μ l) 覆盖冰冻植物材料, 并在 60℃ 连续振荡孵育 1 小时。将所得匀浆分装到两个 Eppendorf 管(2 ml)中, 并用相同体积的氯仿/异戊醇(24:1)振荡抽提两次。每次均在 8000 g、室温离心 15 分钟, 使两相分开。然后使用冰冷的异丙醇在 -70℃ 沉淀 DNA 30 分钟。将沉淀的 DNA 在 4℃、10,000 g 离心 30 分钟沉淀, 并重悬于 180 μ l TE 缓冲液中 (Sambrook 等人, 1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6)。为进一步纯化, 用 NaCl (终浓度 1.2 M) 处理 DNA, 并用两倍体积的无水乙醇在 -70℃ 再次沉淀 30 分钟。用 70% 乙醇洗涤后, 干燥 DNA 并随即溶于 50 μ l H₂O + RNase (终浓度 50 mg/ml)。DNA 在 4℃ 溶解过夜, 并于其后在 37℃ 进行 RNase 消化 1 小时。4℃ 保存 DNA。

实施例 3: 从植物中分离总 RNA 及 poly-(A)+RNA

拟南芥

分离总 RNA 及 poly-(A)+RNA 以研究转录物。根据以下方法从拟南芥植物的长角果中分离 RNA:

从拟南芥种子中制备 RNA——“热”提取:

缓冲液、酶及溶液:

2M KCl

蛋白酶 K

苯酚 (RNA 用)

氯仿:异戊醇

(苯酚:氯仿 1:1; 调节 pH 适于 RNA)

4 M LiCl, DEPC 处理的

DEPC 处理水

3 M NaOAc, pH 5, DEPC 处理的

异丙醇

70%乙醇（DEPC处理水配制）

**重悬浮缓冲液：0.5% SDS, 10 mM Tris pH 7.5, DEPC处理水配制的
1 mM EDTA（因为该溶液无法用 DEPC 处理）**

提取缓冲液：

0.2M 硼酸钠

30 mM EDTA

30 mM EGTA

1% SDS（每 2.5ml 缓冲液 250 μ l 10% SDS 溶液）

1%脱氧胆酸盐（2.5 ml 缓冲液 25 mg）

2% PVPP（不可溶——每 2.5 ml 缓冲液 50 mg）

2% PVP 40K（每 2.5 ml 缓冲液 50 mg）

10 mM DTT

**100 mM β -巯基乙醇（现配，烟熏罩下处理——每 5ml 缓冲液使用
35 μ l 14.3 M 的溶液）**

提取

加热提取缓冲液至 80℃。在液氮冷却的研钵中研磨组织，并将组织粉末转移到 1.5 ml 管中。冷冻保存组织直至加入缓冲液；转移样品应当使用预冷的匙突；而管应当时刻保留在液氮中。然后向管中加入 350 μ l 预热的提取缓冲液（对于较大样品，每 100 mg 组织加入缓冲液体积可大至 500 μ l），振荡，将管加热至 80℃约 1 分钟后置于冰上。振荡样品并用电研钵再次研磨。

消化

加入蛋白酶 K（0.15 mg/100 mg 组织），振荡并在 37℃放置 1 小时。

一次纯化

加入 27 μ l 2M KCl，在冰上冰冻 10 分钟，随后以 12,000 rpm 于室温

离心 10 分钟。然后将上清液转移到不含 RNAase 的新管中，并进行一次苯酚抽提，然后用氯仿:异戊醇抽提。向上清液中加入 1 体积异丙醇，在冰上冰冻混合物 10 分钟。通过离心（室温 7000 rpm，10 分钟）沉淀 RNA。振荡混合物 10 至 15 分钟，将沉淀溶解于 1 ml 4M LiCl 溶液中。离心 5 分钟沉淀 RNA。

二次纯化

将沉淀重悬于 500 μ l 重悬缓冲液中。加入 500 μ l 苯酚并振荡之。然后加入 250 μ l 氯仿:异戊醇并振荡；离心 5 分钟。将上清液转移至新管中。重复氯仿:异戊醇抽提直至界面清晰。将上清液转移到新管中并加入 1/10 体积 pH 为 5 的 3M NaOAc 以及 600 μ l 异丙醇。在 -20℃ 保存混合物 20 分钟或更久。离心 10 分钟沉淀 RNA，并用 70% 乙醇将沉淀洗涤一次。除去所有剩余乙醇后将沉淀溶解在 15 至 20 μ l DEPC 处理水中。在 260 nm 和 280 nm 处测量 1:200 稀释液的吸光度，对 RNA 定量定性(40 μ g RNA/ml = 1 OD₂₆₀)。

如(Hosein, 2001, *Plant Mol. Biol. Rep.* 19, 65a-65e; Ruuska, S.A., Girke, T., Benning, C., & Ohlrogge, J.B., 2002, *Plant Cell* 14, 1191-1206)所述，分离来自野生型拟南芥的 RNA。

利用 Amersham Pharmacia Biotech mRNA 纯化试剂盒（运用 oligo(dT)-纤维素柱）从总 RNA 中制备 mRNA。

按照生产商说明使用 Dyna Beads[®] (Dyna, Oslo, Norway) 分离 poly-(A)+RNA。确定 RNA 或 poly-(A)+RNA 的浓度后，加入 1/10 体积 pH 4.6 的 3M 醋酸钠及 2 倍体积乙醇沉淀 RNA，并于 -70℃ 保存。

大豆、稻、玉米、亚麻、大麦或小麦

从种荚中分离出大豆和亚麻的种子以产生种子和种荚 cDNA 文库所需的匀浆材料。在液氮下用研钵和研杵将组织研磨为细粉，并转移到 50 ml 管中。将组织样品贮存于 -80℃ 直至进行提取。

对于稻而言，通过切割分离 5K-10K 胚和胚乳。切割时将组织置于冰上的小管或 petri 盘中。容器置于干冰上，然后保存于-80℃。

对于玉米而言，在液氮下用研钵和研杵将组织研磨为细粉，并转移到 50 ml 管中。将组织样品贮存于 -80℃ 直至进行提取。

对于大麦而言，割下麦穗（约 2 英寸部分）以得到指定阶段的小花。修整去掉所有的芒。所选的阶段为种子填充早/中期。种子组织 cDNA 文库麦粒或处于水熟期或处于乳熟期。

对于小麦而言，将 Galeon 小麦种子的种子萌发样品种植于 20" × 12" 的 metromix 上深度 2" 处。土壤充分吸收水分，然后每天施水两次。当 3-4 天后胚芽鞘为约 1 cm 时，用水洗涤幼苗并吸干水分（blotted）。为建立花 cDNA 文库，于麦穗从鞘中出现 30%、60% 及 100% 时的两天中每一天均收集等数量的麦穗。此时尚未出现花粉囊。为了建立种子组织 cDNA 文库，麦粒为水熟期或乳熟期，根据麦粒在麦穗上的位置而定；对于晚期种子发育阶段仅收集麦穗。对于根文库仅收集根。植物具有一主干和三根坚实的分蘖。植物生长于罐中，洗去培养基并保留根作为样品。对植物不作处理。

根据生产商说明利用 RNeasy Maxi 试剂盒(Qiagen)从组织中提取总 RNA，同时也根据生产商说明利用 Oligotex mRNA 纯化系统试剂盒(Qiagen)从总 RNA 中加工出 mRNA。mRNA 送往 Hyseq Pharmaceuticals Incorporated (Sunnyville, CA)，以进一步将每一组织类型来源的 mRNA 加工为 cDNA 文库，并用于其特有的方法，其中基于杂交类型对于类似的质粒插入物进行聚类化。

实施例 4：构建 cDNA 文库

为构建 cDNA 文库，使用鼠白血病病毒逆转录酶(Roche, Mannheim, Germany)以及 oligo(dT)引物合成第一链，于 12℃（2 小时）、16℃（1 小时）和 22℃（1 小时）与 DNA 聚合酶 I、Klenow 酶和 RNA 酶 H 消化产物孵育，制备第二链。在 65℃ 孵育（10 分钟）终止反应后转移至冰上。利用 T4 DNA 聚合酶(Roche, Mannheim)在 37℃（30 分钟）钝化双链 DNA

分子。通过酚/氯仿抽提以及 Sephadex G50 旋转柱除去核苷酸。用 T4 DNA 连接酶(Roche, 12℃ 过夜)将 EcoRI 接纳体(Pharmacia, Freiburg, Germany) 连接到 cDNA 末端上, 并与多核苷酸激酶孵育(Roche, 37℃, 30 分钟)进行磷酸化。用低熔点琼脂糖凝胶分离该混合物。从凝胶中洗脱大于 300 碱基对的 DNA 分子、苯酚抽提、在 Elutip-D 柱(Schleicher 和 Schuell, Dassel, Germany)上浓缩、与载体臂连接, 并根据生产商的说明, 利用 Gigapack Gold 试剂盒(Stratagene, Amsterdam, Netherlands)及材料将之装入 λ ZAPII 噬菌体或 λ ZAP 表达噬菌体。

大豆、稻、玉米、亚麻、大麦和小麦 cDNA 文库产生于 Hyseq Pharmaceuticals Incorporated (Sunnyville, CA)。文库建立中未使用扩增步骤以保留表达信息。Hyseq 的基因组方法包括将这些基因分组聚类, 并测序每一聚类的代表成员。从 oligo dT 柱纯化的 mRNA 中产生 cDNA 文库。随机挑取 cDNA 文库转化大肠杆菌的菌落, 通过 PCR 扩增其 cDNA 插入物, 并点样于尼龙膜上。用一套 $^{33}\text{-P}$ 放射性标记的寡核苷酸与克隆杂交, 并确定产生杂交的类型中特定的克隆属于哪一聚类。获取 cDNA 克隆及其 DNA 序列, 用于在转基因植物中过表达及本文所述的其他分子生物学方法。

实施例 5: 鉴定 *FAD2* 样目的 LMP 基因

拟南芥

使用拟南芥野生型和拟南芥 *fad2* 突变体鉴定 LMP 编码基因。*FAD2* 基因已克隆并描述(J Okuley, J Lightner, K Feldmann, N Yadav, E Lark, & J Browse, Plant Cell 6:147-158, 1994)。*FAD2* 编码微粒体脂肪酸 $\Delta 6$ -去饱和酶, 所述酶向油酸($\text{C}18:1^{\Delta 9}$)结合卵磷脂的 $\Delta 12$ 位插入双键以产生亚油酸($\text{C}18:2^{\Delta 9, 12}$), 因此也称为 12-去饱和酶。*FAD2* 是拟南芥基因组中的单个基因。

大豆、稻、玉米、亚麻、大麦和小麦

本实施例说明了如何鉴定并分离编码大豆、稻、玉米、亚麻、大麦和小麦 FAD2 样多肽的 cDNA 克隆。

为在适当数据库中鉴定 FAD2 样基因，利用 BLAST 软件(Basic Local Alignment Search Tool, 2.2.6 版本, Altschul 等人, 1997, Nucleic Acid Res. 25: 3389-3402)进行相似性分析。除 e 值阈值($1e-10$)外均使用默认设置，所有蛋白质搜索均利用 BLOSUM62 矩阵完成。使用拟南芥 FAD2 多肽的氨基酸序列为搜索查询条件，并利用 TBLASTN 算法比对大豆、稻、玉米、亚麻、大麦及小麦中在所有六个读框中翻译的 DNA 数据。此类对 BPS 内部数据库的相似性分析鉴定到众多的 EST 和 cDNA 重叠群。

使用从 Hyseq 聚类方法得到的 RNA 表达谱数据确定器官特异性。选择种子文库相对其他组织文库表达更多的克隆为 LMP 候选基因。根据其表达谱，选择大豆、稻、玉米、亚麻、大麦和小麦克隆在拟南芥以及特定作物植物中过表达。

实施例 6: 克隆鉴定的 LMP 基因的全长 cDNA 及直系同源物

已在内部专有 Hyseq 数据库中鉴定了对应拟南芥、大豆、稻、玉米、亚麻、大麦或小麦全长序列和部分 cDNA 的克隆。利用 ABI 377 板式凝胶测序仪及 BigDye Terminator Ready Reaction 试剂盒(PE Biosystems, Foster City, CA)对大豆、稻、玉米、亚麻、大麦及小麦的 Hyseq 克隆中 DNA 标志处进行测序。进行序列比对以确定 Hyseq 克隆是否为全长或部分克隆。当确定 Hyseq 克隆为部分 cDNA 时，使用以下方法分离全长序列。利用在 5'和 3'端均可进行 cDNA 末端快速扩增(RACE)的 Clontech 的 SMART RACE cDNA 扩增试剂盒，通过 RACE PCR 分离全长 cDNA。根据 Hyseq 克隆序列设计 RACE PCR 引物。根据生产商的条件进行全长 cDNA 的分离和 RACE PCR 方法的使用。利用 QIAquick 凝胶提取试剂盒(Qiagen)从琼脂糖凝胶中提取 RACE 产物片断，并按照生产商的说明连入 TOPO pCR 2.1 载体(Invitrogen)。在标准条件(Sambrook 等人, 1989)下将重组载体转化入 TOP10 细胞(Invitrogen)。转化细胞在 37℃、含 50 μ g/ml

卡那霉素的 LB 琼脂中生长过夜，并涂布 40 μ l 40 mg/ml X-gal 的二甲基甲酰胺贮液，进行蓝白筛选。选择单个白色菌落并用其接种 3 ml 含 50 μ g/ml 卡那霉素的液体 LB，在 37℃ 生长过夜。根据生产商说明，使用 QIAprep Spin Miniprep 试剂盒(Qiagen)提取质粒 DNA。根据标准分子生物学技术进行后续的克隆分析和限制性酶切图谱绘制 (Sambrook 等人, 1989)。

通过使用下列方法分离全长 cDNA 并克隆进二元载体：使用得自 Hyseq 克隆的全长序列或后来的 RACE 扩增产物设计基因特异引物。以 Hyseq 克隆或 cDNA 文库为 DNA 模板使用降落 PCR (touch-down PCR) 扩增全长序列和基因。在一些情况下，设计引物向基因起始密码子紧邻上游添加“AACA” Kozak 样序列，且在一些情况下下游的两个碱基改变为 GC 以促进提高基因表达水平 (Chandrashekhar 等人, 1997, Plant Molecular Biology 35:993-1001)。PCR 反应循环为：94℃, 5 分钟; 94℃, 1 分钟, 65℃, 1 分钟, 72℃, 4 分钟, 9 个循环，其中退火温度每循环降低 1℃; 94℃, 1 分钟, 55℃, 1 分钟, 72℃, 4 分钟, 20 个循环; 72℃, 10 分钟结束 PCR 循环。使用 GenElute-EtBr 旋转柱(Sigma)在 1% 琼脂糖凝胶上凝胶纯化扩增的 PCR 产物，并在标准酶促消化后连接至植物二元载体 pBPS-GBI 上，用于转化拟南芥。二元载体通过在 LB 培养基和合适的抗生素中在大肠杆菌 DH5 中过夜培养扩增，并使用 Qiagen MiniPrep DNA 制备试剂盒制备质粒用于后续步骤。在多个克隆步骤中均通过用限制性消化确定插入物大小来确认插入物，并测序插入物以保证拟南芥转化中使用预期的基因。

基因序列可用于鉴定 cDNA 或基因组文库中的同源或异源基因 (直向同源物，来自其他植物的相同 LMP 基因)。可以通过设计由多重序列比对确定的保守序列的 PCR 引物完成该鉴定。常通过设计目的基因的全长或部分序列的简并引物鉴定直向同源物。

基因序列可用于鉴定 cDNA 或基因组文库中的同源物或直向同源物。利用如 cDNA 文库等通过核酸杂交，可以分离同源基因 (如全长 cDNA 克隆)：根据目的基因的丰度不同，铺板 100,000 至 1,000,000 个重组噬菌体并转移至尼龙膜。碱变性后，通过如紫外线交联将 DNA 固定在膜上。在

高度严谨条件下进行杂交。在 1 M NaCl 离子强度、温度 68℃ 进行水性溶液杂交和洗涤。通过如放射性(³²P)缺刻转录标记制备杂交探针(High Prime, Roche, Mannheim, Germany)。自显影法检测信号。

可以通过类似上述的方法, 在低严谨杂交和洗涤条件下鉴定相关但不相同的部分同源或异源基因。对于水相杂交, 离子强度一般保持在 1M NaCl, 而温度则逐渐从 68℃ 降至 42℃。

可以使用合成的放射性标记寡核苷酸探针, 分离仅在一个截然不同的结构域(如 10-20 个氨基酸)上具有同源性(或序列同一性/相似性)的基因序列。通过用 T4 多核苷酸激酶磷酸化两条互补的寡核苷酸的 5-引物端, 制备放射性标记的寡核苷酸。退火并连接互补的寡核苷酸以形成串联体。然后通过如缺刻转录放射性标记双链串联体。通常使用高寡核苷酸浓度, 在低严谨条件下进行杂交。

寡核苷酸杂交溶液:

6 × SSC

M 磷酸钠

mM EDTA (pH 8)

0.5 % SDS

100 µg/ml 变性鲑精 DNA

0.1%脱脂奶粉

在杂交过程中, 逐步将温度降至估算的寡核苷酸 T_m 以下 5-10℃ 或降至室温, 然后进行洗涤步骤和放射自显影。在低严谨条件下洗涤, 例如使用 4 × SSC 进行 3 步洗涤。更多细节述于 Sambrook 等人(1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press), 或 Ausubel 等人(1994, "Current Protocols in Molecular Biology", John Wiley & Sons)。

实施例 7: 用抗体筛选表达文库来鉴定目的基因

cDNA 克隆可用于在如大肠杆菌(如 Qiagen QIAexpress pQE 系统)

中产生重组蛋白。然后一般使用 Ni-NTA 亲和层析(Qiagen)对重组蛋白进行亲和纯化。通过使用如兔免疫标准技术,可将重组蛋白用于产生特异性抗体。使用以重组抗原饱和的 Ni-NTA 柱(如 Gu 等人, 1994, BioTechniques 17:257-262 所述), 对抗体进行亲和纯化。此后可通过免疫筛选(Sambrook 等人, 1989, "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press; 或 Ausubel 等人, 1994, "Current Protocols in Molecular Biology," John Wiley & Sons), 将抗体用于筛选表达 cDNA 文库以鉴定同源或异源基因。

实施例 8: Northern 杂交

为进行 RNA 杂交, 如 Amasino (1986, Anal. Biochem. 152:304)所述, 使用甲醛在 1.25%的琼脂糖凝胶中电泳分离 20 μg 总 RNA 或 1 μg poly-(A)+RNA, 使用 10 $\times$ SSC, 通过毛细吸引将其转移到带正电的尼龙膜 (Hybond N+, Amersham, Braunschweig)上, 用紫外线固定, 并使用杂交缓冲液 (10%硫酸右旋糖酐 w/v, 1 M NaCl, 1% SDS, 100 $\mu\text{g}/\text{ml}$ 鲑精 DNA) 在 68 $^{\circ}\text{C}$ 预杂交 3 小时。在预杂交中使用 α - ^{32}P dCTP (Amersham, Braunschweig, Germany)、Highprime DNA 标记试剂盒(Roche, Mannheim, Germany)标记 DNA 探针。在相同缓冲液中加入经标记的 DNA 探针, 于 68 $^{\circ}\text{C}$ 杂交过夜。68 $^{\circ}\text{C}$ 使用 2 $\times$ SSC、15 分钟洗涤两次, 1 $\times$ SSC、1% SDS、30 分钟洗涤两次。密封滤膜于 -70 $^{\circ}\text{C}$ 曝光, 为期 1 至 14 天。

实施例 9: DNA 测序和计算机功能分析

cDNA 文库可用于根据标准方法进行 DNA 测序, 特别是通过使用 ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction 试剂盒 (Perkin-Elmer, Weiterstadt, Germany)的链终止法。从 cDNA 文库回收制备质粒后通过体内大量切割进行随机测序、重转化, 随后在琼脂平板上涂布 DH10B (材料和方案细节来自 Stratagene, Amsterdam, Netherlands)。可从用含氨苄青霉素的 Luria-Broth 培养基过夜培养的大肠杆菌培养物中

制备质粒 DNA (参阅 Sambrook 等人(1989, Cold Spring Harbor Laboratory Press: ISBN 0-87969-309-6)关于根据生产商方案的 Qiagen DNA 制备机器人(Qiagen, Hilden)).可使用由 Bio-Max (Munich, Germany) 商业提供的软件包 EST-MAX 处理并注释序列。该程序包含了所有对蛋白质序列功能及结构表征有重要意义的生物信息学方法。参见 <http://pedant.mips.biochem.mpg.de>。

EST-MAX 中并入的最为重要的算法有 FASTA: 高度敏感的蛋白质序列数据库检索, 附统计学显著性评价(Pearson W.R. 1990, Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods Enzymol.* 183:63-98); BLAST: 高度敏感的蛋白质序列数据库检索, 附统计学显著性评价 (Altschul S.F., Gish W., Miller W., Myers E.W.和 Lipman D.J. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410); PREDATOR: 高精度度单个及多个序列的二级结构预测(Frishman 和 Argos 1997, 75% accuracy in protein secondary structure prediction. *Proteins* 27:329-335); CLUSTALW: 多重序列比对(Thompson, J.D., Higgins, D.G.和 Gibson, T.J. 1994, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice, *Nucleic Acids Res.*22:4673-4680); TMAP: 多重比对序列的跨膜区域预测 (Persson B.和 Argos P. 1994, Prediction of transmembrane segments in proteins utilizing multiple sequence alignments, *J. Mol. Biol.* 237:182-192); ALOM2: 单个序列的跨膜区域预测(Klein P., Kanehisa M.和 DeLisi C. 1984, Prediction of protein function from sequence properties: A discriminant analysis of a database. *Biochim. Biophys. Acta* 787:221-226. Dr. K. Nakai 第 2 版); PROSEARCH: PROSITE 蛋白质序列模式检测(Kolakowski L.F. Jr., Leunissen J.A.M.和 Smith J.E. 1992, ProSearch: fast searching of protein sequences with regular expression patterns related to protein structure and function. *Biotechniques* 13:919-921); BLIMPS: 针对无空位构件(ungapped blocks)

数据库的相似性搜索(Wallace 和 Henikoff 1992, PATMAT: A searching and extraction program for sequence, pattern and block queries and databases, CABIOS 8:249-254. Bill Alford 执笔)。

实施例 10: 进行植物转化的质粒

可使用诸如 pBinAR 的二元载体转化植物(Höfgen & Willmitzer 1990, Plant Sci. 66:221-230)。将 cDNA 以正义或反义方向连入 T-DNA, 构建二元载体。cDNA 5'端的植物启动子激活 cDNA 的转录。多聚腺苷化序列定位于 cDNA 的 3'端。可使用组织特异性启动子进行组织特异性表达。例如将 napin 或 LeB4 或 USP 启动子克隆入 cDNA 5'端, 可进行种子特异性表达。也可使用其他任何种子特异性启动子元件。整个植物中的组成型表达也可使用 CaMV35 启动子。也可以使用信号肽将表达蛋白质定位于细胞区室内, 如质体、线粒体或内质网(Kermode, 1996, Crit. Rev. Plant Sci. 15:285-423)。信号肽于读框内克隆入 cDNA 5'端以实现融合蛋白的亚细胞定位。

使用的植物二元载体实例为克隆进 LMP 候选基因的 pBPS-GB1、pSUN2-GW 或 pBPS-GB047 载体。这些二元载体包含 AtAct2-I 启动子控制下的抗生素抗性基因和候选基因前的 USP 或其他种子特异性启动子或组成型启动子, 所述候选基因后带有 NOSpA 终止子或 OCS 终止子。将部分或全长 LMP cDNA 以正义或反义方向克隆入植物二元载体 USP 或种子特异性的、或其他种子特异性的或组成型启动子后的多克隆位点。可用于不同作物物种的其他启动子见实施例 11。

在标准条件下将含有目的基因的重组载体转化到 Top10 细胞(Invitrogen)中。在含有 50 µg/ml 卡那霉素的 LB 琼脂上选择转化细胞, 并使细胞在 37°C 生长过夜。按照生产商说明使用 QIAprep Spin Miniprep Kit (Qiagen)提取质粒 DNA。根据标准分子生物学技术进行后续克隆的分析以及限制性图谱的绘制(Sambrook 等人, 1989, Molecular Cloning, A Laboratory Manual, 第二版, Cold Spring Harbor Laboratory Press. Cold

Spring Harbor, NY).

实施例 11: 农杆菌介导的植物转化

可以使用标准转化和再生技术 (Gelvin, Stanton B. & Schilperoort R.A, *Plant Molecular Biology Manual*, 第二版. Kluwer Academic Publ., Dordrecht 1995, 见于 Sect., Ringbuc Zentrale Signatur:BT11-P; Glick, Bernard R.和 Thompson, John E. *Methods in Plant Molecular Biology and Biotechnology*, S. 360, CRC Press, Boca Raton 1993), 利用本文所述的 LMP 核酸进行农杆菌介导的植物转化。例如, 可以使用 GV3 (pMP90) (Koncz & Schell, 1986, *Mol. Gen. Genet.* 204:383-396) 或 LBA4404 (Clontech)根癌农杆菌菌株进行农杆菌介导的植物转化。

在标准条件下 (Bechtold, 1993, *Acad. Sci. Paris.* 316:1194-1199; Bent 等人, 1994, *Science* 265:1856-1860) 种植并转化拟南芥。此外, 可以通过子叶或胚轴转化 (Moloney 等人, 1989, *Plant Cell Report* 8:238-242; De Block 等人, 1989, *Plant Physiol.* 91:694-701), 利用本发明的 LMP 核酸转化油菜籽。选择农杆菌及植物的抗生素使用取决于转化所用的二元载体和农杆菌菌株。通常使用可选择植物标记进行油菜籽选择。此外, 可以使用如 Mlynarova 等人所述的技术(1994, *Plant Cell Report* 13:282-285), 进行农杆菌介导至亚麻的基因转移。

拟南芥 *FAD2* 或 *FAD2* 样基因可被克隆入二元载体, 转化并通过以下启动子表达: 组成型启动子如超启动子(Stanton B. Gelvin, USP#5,428,147 和 USP#5,217,903)、或种子特异性启动子如蚕豆 USP (未知种子蛋白质) (Baeumlein 等人. 1991, *Mol. Gen. Genetics* 225:459-67)、或豆球蛋白 B4 启动子(LeB4; Baeumlein 等人. 1992, *Plant J.* 2:233-239)以及可导致单子叶植物如玉米、大麦、小麦、裸麦、稻进行种子特异性表达的启动子。

拟南芥 AHAS (AtAHAS)基因可用作这些构建体的可选择标记。

可使用如 EP 0424 047、美国专利号 5,322,783 (Pioneer Hi-Bred International)或 EP 0397 687、美国专利号 5,376,543 或美国专利号

5,169,770 (University Toledo)所述的技术、或利用本领域所公知的多种其他转化方法中的任一种转化大豆。用 70%乙醇在室温对大豆种子连续振荡表面消毒 4 分钟,然后加入补充有 0.05% (v/v) Tween 的 20% (v/v) Clorox 连续振荡 20 分钟。用蒸馏水漂洗种子 4 次,并将其置于 Petri 盘中的润湿无菌滤纸上,室温留置 6 至 39 小时。剥去种皮,将子叶从胚轴上分离下来。检查胚轴确保分生组织区没有损伤。收集切下的胚轴于半开无菌 Petri 盘中,并风干至湿度小于 20% (鲜重),置于密封 Petri 盘中备用。

这一植物转化方法也适用于欧洲油菜、亚麻和其他作物。具体而言,用 70%乙醇在室温对芸苔的种子连续振荡表面消毒 4 分钟,然后加入补充有 0.05% (v/v) Tween 的 20% (v/v) Clorox 连续振荡 20 分钟。用蒸馏水漂洗种子 4 次,并将其置于 Petri 盘中的润湿无菌滤纸上,室温留置 18 小时。剥去种皮,种子于半开无菌 Petri 盘中风干过夜。在此期间,种子丧失了约 85%水分。然后将种子保存于室温密封 Petri 盘中备用。

取含适当抗生素(如 100 mg/l 链霉素、50 mg/l 卡那霉素)的 LB 固态培养基中的单菌落,在液态 LB 培养基中长至 600 nm 处光密度为 0.8,制备根癌农杆菌培养物。然后在室温、7000 rpm、7 分钟离心沉淀细菌培养物,并重悬浮于补加 100 mM 乙酰丁香酮的 MS 培养基(Murashige & Skoog, 1962, *Physiol. Plant.* 15:473-497)中。在此预诱导培养基中室温孵育细菌培养物 2 小时,之后使用。大豆合子种子的胚轴在约 44%湿度、室温用预诱导的农杆菌悬浮培养物吸胀 2 小时(干胚在农杆菌培养物中吸胀对于玉米胚轴也适用)。将胚从吸胀培养物中取出并转移到含有固态 MS 培养基(补加 2%蔗糖)的 Petri 盘中,在室温暗处孵育 2 天。还可以将胚置于 Petri 盘中的润湿(液态 MS 培养基)无菌滤纸上,在前述相同条件下孵育。此后将胚转移到固态或液态含有 500 mg/l 羧苄青霉素或 300 mg/l 头孢噻肟的 MS 培养基中以杀死农杆菌。使用液态培养基润湿无菌滤纸。将胚在 25℃、 $440\ \mu\text{mol m}^{-2}\text{s}^{-1}$ 和 12 小时光周期中孵育 4 周。一旦幼苗生根,将其转移到无菌 metromix 土壤中。转移植物至土壤之前洗去体外植物的培养基。植物在塑料罩下维持 1 周以促进累积过程。随后将植物转移到生长室内,在

25℃、440 $\mu\text{mol m}^{-2}\text{s}^{-1}$ 光强度和 12 小时光周期下孵育大约 80 天。

通过 PCR 分析初级转基因植物 (T_0) 样品以证实 T-DNA 的存在。Southern 杂交证实了这些结果, 其中在 1% 琼脂糖凝胶上进行 DNA 电泳并转移到带正电的尼龙膜(Roche Diagnostics)上。使用 PCR DIG 探针合成试剂盒(Roche Diagnostics)、按照生产商建议通过 PCR 制备地高辛标记的探针。

作为单子叶植物转化的实例, 可使用组成型或种子特异性启动子以及玉米泛素内含子和 *FAD2* 或 *FAD2* 样核酸分子。例如, 用 *PacI* 和 *XmaI* 消化 pUC 中的 *PtxA-FAD2* 直系同源物基因构建体。用 *PacI* 和 *XmaI* 消化 pBPSMM348, 以分离玉米泛素内含子(ZmUbi 内含子), 然后进行电泳和 QIAEX II 凝胶提取试剂盒(cat# 20021)处理。将 ZmUbi 内含子连接于 pUC 中的 *PtxA-FAD2* 或 *FAD2* 样核酸分子, 产生基于 pUC 的 *PtxA-ZmUbi* 内含子-*FAD2* 或 *FAD2* 样核酸分子构建体, 然后以 *AfeI* 和 *PmeI* 进行限制性酶消化。电泳后将 *PtxA-ZmUbi* 内含子 *FAD2* 或 *FAD2* 样基因盒从 Seaplaque 低熔点琼脂糖凝胶(SeaPlaque® GTG® Agarose 目录号 50110)中切下。用 *PmeI* 消化含有可选择标记盒(单子叶型载体)的单子叶植物型载体。将含有种子特异性启动子 - ZmUbi 内含子的 *FAD2* 或 *FAD2* 样核酸分子表达盒连于单子叶型载体。随后, 根据本领域常规方法, 利用电穿孔法将该构建体转化入含有 pSB1 (超级 *vir* 质粒)的重组 LBA4404 菌株中。根据 US 5,591,616 所述的方法使用未成熟胚进行农杆菌介导的玉米转化。应用咪唑酮除草剂选择获得转基因玉米株系。

用于玉米的启动子实例也是玉米醇溶蛋白, 后者为一组见于玉米胚乳的贮存蛋白质。已经分离了玉米醇溶蛋白基因的基因组克隆 (Pedersen 等人, Cell 29:1015-1026, 1982), 并且也可使用来自这些克隆的启动子, 包括 15 kD、16 kD、19 kD、22 kD、27 kD 基因。已知在玉米中起作用的其他启动子为淀粉合酶、分支酶、去分支酶、油体蛋白、谷蛋白和蔗糖合酶启动子。特别优选用于玉米胚乳表达的启动子为稻的谷蛋白基因的启动子, 特别是 *Osgt-1* 启动子(Zheng 等人, Mole. Cell Biol. 13:5829-5842, 1993)。

适用于小麦中表达的启动子实例包括 ADP 葡萄糖焦磷酸化酶 (ADP glucose pyrosynthase) 亚基启动子、淀粉粒结合淀粉合酶及其他淀粉合酶启动子、分支和去分支酶启动子、胚胎发生丰富蛋白启动子、麦醇溶蛋白启动子和谷蛋白启动子。

适用于大麦中表达的启动子实例包括 ADP 葡萄糖焦磷酸化酶亚基启动子、淀粉粒结合淀粉合酶及其他淀粉合酶启动子、分支和去分支酶启动子、蔗糖合酶启动子、大麦醇溶蛋白启动子、胚球蛋白启动子和糊粉特异性蛋白质启动子。

适用于大豆中表达的启动子实例包括本文已经提及的启动子。然而，其他可用的启动子为大豆 7S 启动子和大豆 7S' β 伴大豆球蛋白启动子。

一般而言，可将植物启动子如 USP 控制下的稻 (或其他单子叶植物) *FAD2* 基因或 *FAD2* 样基因可转化入玉米或另一作物植物中，以在其他单子叶植物中产生单子叶 *FAD2* 基因的功能，或在其他双子叶植物中产生双子叶 *FAD2* 基因的功能，或在双子叶植物中产生单子叶基因的功能，或反之亦然。可通过与本文所述构建其他质粒方法类似的方法构建含有这些 *FAD2* 或 *FAD2* 样编码序列、5'端为启动子和 3'端为终止子的质粒。适用于稻中表达的启动子实例包括 ADP 葡萄糖焦磷酸化酶亚基启动子、淀粉粒结合淀粉合酶及其他淀粉合酶启动子、分支酶启动子、去分支酶启动子、蔗糖合酶启动子、谷蛋白启动子。特别优选的启动子为稻谷蛋白启动子，*Osgt-1*。

实施例 12: 体内诱变

可以通过用大肠杆菌或其他微生物 (如芽孢杆菌属(*Bacillus spp.*)或酵母, 如酿酒酵母) 掺入并传代质粒 (或其他载体) DNA, 进行微生物的体内诱变, 其中所述微生物保持其遗传信息完整性的能力受损。典型的突变菌株中 DNA 修复系统的基因有突变 (如 *mutHLS*, *mutD*, *mutT* 等; 参考 Rupp W.D. 1996, DNA repair mechanisms, in: *Escherichia coli* and *Salmonella*, 第 2277-2294 页, ASM: Washington)。此类菌株为本领域技术

人员所熟知，其用途在例如 Greener 和 Callahan, 1994, *Strategies* 7:32-34 中有所阐释。优选在微生物中选择并检测之后再将突变 DNA 分子转移到植物中。根据本文件中的多个实施例产生转基因植物。

实施例 13: 评估转化生物中的 mRNA 表达和重组基因产物的活性

转化宿主生物中重组基因产物的活性可在转录水平或/和翻译水平上测量。确定基因转录水平（可翻译为基因产物的 mRNA 量的指标）的有效方法之一是进行 Northern 印迹（参见如 Ausubel 等人, 1988, *Current Protocols in Molecular Biology*, Wiley: New York），其中用可检测的标签（通常具有放射活性或可化学发光）标记为结合目的基因而设计的引物，从而使得在生物培养物总 RNA 提取、胶上电泳、转移到稳定基质并与该探针孵育时，探针的结合和结合量能够提示该基因 mRNA 的存在和量。这一信息至少部分证实了转化基因的转录程度。可以通过本领域已知的若干方法，例如 Bormann 等人, 1992, *Mol. Microbiol.* 6:317-326 所述的方法，从植物细胞、组织或器官中制备总细胞 RNA。

可以使用诸如 Western 印迹的标准技术（见如 Ausubel 等人, 1988, *Current Protocols in Molecular Biology*, Wiley: New York），评估从 mRNA 翻译的蛋白质的存在或相对量。这一方法中，提取总细胞蛋白质，通过凝胶电泳分离之，转移到诸如硝化纤维的基质上，并与可特异性结合于目的蛋白质的探针（如抗体）一同孵育。通常使用易于检测的化学发光物或比色标记标记此探针。观测到的标记存在与否及其数量提示了细胞中期望突变蛋白的存在与否及其数量。

可以通过若干充分确立的方法，如 DNA 条带移位测定法（也称为凝胶迟滞测定法），测量结合于 DNA 的 LMP 活性。可以使用报告基因测定法（如 Kolmar H. 等人, 1995, *EMBO J.* 14:3895-3904 及其引用参考文献所述）测量此类 LMP 对其他分子表达的影响。报告基因检测系统众所周知，使用如 β -半乳糖苷酶、绿色荧光蛋白以及其他酶，在原核和真核细胞中都可应用该系统。

根据如 Gennis R.B. 所述的技术 (1989 Pores, Channels and Transporters, in Biomembranes, Molecular Structure and Function, Springer: Heidelberg, 85-137、199-234 和 270-322 页), 进行脂质代谢膜转运蛋白活性的测定。

实施例 14: 转基因植物中拟南芥、大豆、稻、玉米、亚麻、大麦或小麦中 *FAD2* 或 *FAD2* 样基因功能的体外分析

确定酶的活性和动力学参数是本领域的成熟技术。确定任何给定的改变蛋白质活性的实验必须根据野生型酶的特异性活性量体而做, 这也完全在本领域技术人员的能力范围内。为确定多种酶的活性, 酶的大致情况以及涉及其结构、动力学、原则、方法、应用和实例的具体细节可见于以下参考文献, 如 Dixon, M. & Webb, E.C. 1979, *Enzymes*. Longmans: London; Fersht, (1985) *Enzyme Structure and Mechanism*. Freeman: New York; Walsh (1979) *Enzymatic Reaction Mechanisms*. Freeman: San Francisco; Price, N.C., Stevens, L. (1982) *Fundamentals of Enzymology*. Oxford Univ. Press: Oxford; Boyer, P.D. 编. (1983) *The Enzymes*, 第三版. Academic Press: New York; Bisswanger, H., 1994, *Enzymkinetik*, 第二版, VCH: Weinheim (ISBN 3527300325); Bergmeyer, H.U., Bergmeyer, J., Graßl, M. 编, (1983-1986) *Methods of Enzymatic Analysis*, 第三版, I-XII 卷, Verlag Chemie: Weinheim; 以及 Ullmann Encyclopedia of Industrial Chemistry (1987) 卷 A9, *Enzymes*. VCH: Weinheim, 352-363 页。

实施例 15: 分析重组蛋白对期望种子贮存化合物产量的影响

作为种子油改变的实例, 以气相色谱 (GC) 分析经转化拟南芥植物种子的脂肪酸谱。图 10 和 11 中的每一柱形代表从野生型植物或 *fad2* 突变体或 *fad2* 突变体背景上独立的转基因事件的 5mg 大量种子中得到的数值。如图 10 所示, 拟南芥的 *fad2* 突变体中油酸 (C18:1) 含量从野生型的 16% 增加至 *fad2* 突变体的约 60%。与之类似, 亚油酸 (C18:2) 的比例大幅下降,

从野生型的约 31% 降至 *fad2* 突变体的 2%。用稻基因 OsFAD-01 (Seq ID No. 17) 转化的拟南芥 *fad2* 突变体种子中出现了由若干 *fad2* 突变体背景上独立转基因株系的脂肪酸模式(见图 10 的事件 FAD1172、FAD1214、FAD1246 和 FAD1294)向野生型组成的重建, 即使在分离的 T2 种子群中也是如此。使用大麦基因 HvFAD-01 (Seq ID No. 23) 得到了相似的反应, 如图 11 中事件 FAD0503、FAD0483、FAD0475 和 FAD0465 所示。这一结果提示, 稻基因(OsFAD-01)和大麦基因(HvFAD-01)均能够互补(complement) *fad2* 拟南芥突变体中种子脂肪酸谱, 提示其作为脂肪酸去饱和酶的功能。本申请中提及的不同作物的所有其他脂肪酸去饱和酶基因在 *fad2* 拟南芥突变体背景中也均显示类似的表现(数据未显示)。

通过在合适的条件下种植经修饰的植物, 并分析其种子或其他任何植物器官中期望产物(即脂质或脂肪酸)产量的增加, 可以评价植物中遗传修饰对期望种子贮存化合物(如糖、脂质或脂肪酸)的影响。此类分析技术为本领域技术人员所熟知, 包括质谱法、薄层层析法、多种染色法、酶学和微生物学方法以及分析色谱法, 如高效液相色谱法(见如 Ullman, 1985, Encyclopedia of Industrial Chemistry, 卷 A2, 第 89-90 及 443-613 页, VCH: Weinheim; Fallon, A. 等人, 1987, Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology, 卷 17; Rehm 等人, 1993 Product recovery and purification, Biotechnology, 卷 3, 第 III 章, 469-714 页, VCH: Weinheim; Belter, P.A. 等人, 1988 Bioseparations: downstream processing for biotechnology, John Wiley & Sons; Kennedy J.F. & Cabral J.M.S., 1992, Recovery processes for biological materials, John Wiley and Sons; Shaeiwitz J.A. & Henry J.D., 1988, "Biochemical separations", 见 Ulmann's Encyclopedia of Industrial Chemistry, "Separation and purification techniques in biotechnology", 卷 B3, 第 11 章 1-27 页, VCH: Weinheim; 以及 Dechow F.J. 1989)。

除上述方法外, 如 Cahoon 等人(1999, Proc. Natl. Acad. Sci. USA 96, 22:12935-12940)和 Browse 等人(1986, Anal. Biochemistry 442:141-145)所

述, 从植物材料中提取植物脂质。脂质或脂肪酸的定性及定量分析描述于 Christie, William W. *Advances in Lipid Methodology*, Ayr/Scotland: Oily Press. - (Oily Press Lipid Library; Christie, William W., *Gas Chromatography and Lipids. A Practical Guide* -Ayr, Scotland: Oily Press, 1989 Repr. 1992. -IX, 307 S. - (Oily Press Lipid Library; 以及 *Progress in Lipid Research*, Oxford: Pergamon Press, 1 (1952) - 16 (1977) *Progress in the Chemistry of Fats and Other Lipids* CODEN)。

可以按照标准分析方法: GC、GC-MS 或 TLC, 通过分析转基因植物, 获取脂肪酸产物存在的明确证据, 分别见于 Christie 的不同论述及其参考文献 (1997, *Advances on Lipid Methodology*, 第 4 版: Christie, Oily Press, Dundee, 119-169 页; 1998)。以叶为例的详细方法见于 Lemieux 等人(1990, *Theor. Appl. Genet.* 80:234-240), 以种子为例的见于 Focks & Benning (1998, *Plant Physiol.* 118:91-101)。

通过脂酶消化对甘油主链的 sn-1、sn-2 或 sn-3 位进行脂肪酸组成的位置分析 (见如 Siebertz & Heinz 1977, *Z. Naturforsch.* 32c:193-205, 以及 Christie, 1987, 《*Lipid Analysis*》第 2 版, Pergamon Press, Exeter, ISBN 0-08-023791-6)。

可通过任何适当的方法测量总种子油水平。通常使用常规方法如近红外分析(NIR)或核磁共振成像(NMR)进行种子油含量的定量。NIR 光谱法业已成为用于筛选种子样品的标准方法, 只要目的样品适用于该技术即可。研究的样品包括芸苔、大豆、玉米、小麦、稻等等。可使用单个种子的 NIR 分析 (参见例如 Velasco 等人 “Estimation of seed weight, oil content and fatty acid composition in intact single seeds of rapeseed (*Brassica napus* L.) by near-infrared reflectance spectroscopy”, *Euphytica*, 卷 106, 1999, 79-85 页)。NMR 也用于分析种子的油含量 (参见例如 Robertson & Morrison, “Analysis of oil content of sunflower seed by wide-line NMR,” *Journal of the American Oil Chemists Society*, 1979, 卷 56, 1979, 961-964 页, 其整体在本文引用作为参考)。

收集蛋白质活性增加或减少对脂质和糖生物合成路径影响的相关信息的典型方法是例如用 ^{14}C -醋酸盐或 ^{14}C -丙酮酸盐的标记, 研究分析叶或种子的碳流通 (见如 Focks & Benning, 1998, *Plant Physiol.* 118:91-101; Eccleston & Ohlrogge, 1998, *Plant Cell* 10: 613-621)。包括标准物如 ^{14}C -蔗糖和 ^{14}C -苹果酸的相应分离 (如在 TLC 板上) 后, 通过液体闪烁计数可以确定 ^{14}C 在脂质和水性可溶成分间的分布 (Eccleston & Ohlrogge, 1998, *Plant Cell* 10:613-621)。

可以通过声波破碎、玻璃磨、液氮和研磨或其他适用的方法碎裂待分析的材料。碎裂后的材料需进行离心。将沉淀重悬于蒸馏水中, 在 100°C 加热 10 分钟, 冰上冷却并再次离心后, 90°C 用含 0.5M 硫酸、2%二甲氧基丙烷的甲醇提取 1 小时, 产生可生成转甲基脂质的水解油和脂质化合物。在 petrolether 中抽提这些脂肪酸甲酯, 并最终使用毛细柱 (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25m, 0.32mm) 在 170°C - 240°C 的温度梯度下进行 20 分钟 GC 分析, 并于 240°C 进行 5 分钟。通过使用商业来源可得的标准 (即 Sigma), 鉴定产生的脂肪酸甲酯。

当无法得到脂肪酸的标准物时, 通过衍生化以及此后的 GC-MS 分析完成分子鉴定。例如, 在 4,4-二甲氧基噁唑啉衍生物衍生反应后, 通过 GS-MS 显示三键脂肪酸的定位 (Christie, *Oily Press*, Dundee, 1998)。

分析糖 (特别是淀粉) 的常用标准方法, 由 Stitt M., Lilley R.Mc.C., Gerhardt R. 和 Heldt M.W. 发表 (1989, "Determination of metabolite levels in specific cells and subcellular compartments of plant leaves," *Methods Enzymol.* 174:518-552; 其他方法也可见 Härtel 等人, 1998, *Plant Physiol. Biochem.* 36:407-417 及 Focks & Benning, 1998, *Plant Physiol.* 118:91-101)。

为抽提可溶性糖和淀粉, 在 1.5 ml 聚丙烯试管中将 50 粒种子于 500 μl 80% (v/v) 乙醇中匀浆, 在 70°C 孵育 90 分钟。16,000 g 离心 5 分钟后, 将上清液转移至新的试管中。用 500 μl 80% 乙醇抽提沉淀两次。真空室温蒸发混合上清液中的溶剂。将残留物溶解于 50 μl 水, 即为可溶性糖类级分。乙醇抽提后留下的沉淀中含有包括淀粉在内的不可溶性糖类, 在 200 μl 0.2

N KOH 中匀浆, 将悬浮物在 95℃ 孵育 1 小时以溶解淀粉。加入 35 μ l 1N 醋酸后, 在 16,000 g 离心 5 分钟, 使用上清液进行淀粉定量。

为量化可溶性糖, 向含有 100 mM 咪唑、pH 6.9、5 mM $MgCl_2$ 、2 mM NADP、1 mM ATP 和 2 单位 2 ml^{-1} 葡萄糖-6-磷酸脱氢酶的 990 μ l 反应缓冲液中加入 10 μ l 糖提取物。依次向其中加入 4.5 单位己糖激酶、1 单位葡萄糖磷酸异构酶以及 2 μ l 饱和果糖苷酶溶液, 以酶促测定葡萄糖、果糖和蔗糖。在波长 340 nm 处光度法监测 NADPH 的产生。与之相似, 使用 Boehringer Mannheim 的试剂盒分析 30 μ l 不可溶糖类级分中的淀粉。

分析叶和种子中蛋白质含量的实例可见于 Bradford M.M. (1976, "A rapid and sensitive method for the quantification of microgram quantities of protein using the principle of protein dye binding" *Anal. Biochem.* 72:248-254)。为定量总种子蛋白质, 在 1.5 ml 聚丙烯试管中将 15-20 粒种子于 250 μ l 丙酮中匀浆。16,000g 离心后弃去上清液, 将真空干燥的沉淀重悬于 250 μ l 含有 50 mM Tris-HCl (pH8.0)、250 mM NaCl、1 mM EDTA 和 1%(w/v)SDS 的提取缓冲液中。25℃ 孵育 2 小时后, 将匀浆在 16,000 g 离心 5 分钟, 使用 200 μ l 上清液进行蛋白质测量。在该测定中用 γ -球蛋白进行校准。使用 Lowry DC 蛋白质测定法(Bio-Rad)或 Bradford 测定法(Bio-Rad)进行蛋白质测量。

根据 Renz 等人(1993, *Planta* 190:156-165), 分光光度法酶法测定己糖激酶和果糖激酶; 根据 Burrell 等人(1994, *Planta* 194:95-101), 酶法测定葡萄糖磷酸异构酶、ATP 依赖性 6-磷酸果糖激酶、焦磷酸盐依赖性 6-磷酸果糖激酶、果糖-1,6-二磷酸醛缩酶、丙糖磷酸异构酶、甘油-3-磷酸脱氢酶、磷酸甘油激酶、磷酸甘油变位酶、烯醇酶和丙酮酸激酶; 根据 Zrenner 等人 (1995, *Plant J.* 7:97-107), 酶法测定 UDP-葡萄糖-焦磷酸化酶。

如 Härtel 等人所述(1998, *Plant Physiol. Biochem.* 36:407-417), 测量糖类代谢的中间产物, 如葡萄糖-1-磷酸、葡萄糖-6-磷酸、果糖-6-磷酸、磷酸烯醇丙酮酸、丙酮酸和 ATP; 根据 Jelitto 等人(1992, *Planta* 188:238-244), 测量代谢产物。

除测量最终种子贮存化合物（即脂质、淀粉或贮存蛋白质）外，也可以分析产生期望种子贮存化合物所用的代谢路径的其他成分（如中间产物和副产物），以确定化合物产生的总效率（Fiehn 等人, 2000, *Nature Biotech.* 18:1447-1161）。

例如，可使用标准程序构建含有本文公开的核酸或其片断的酵母表达载体，并将其转化至酿酒酵母中。可分析产生的转基因细胞在糖、油、脂质或脂肪酸含量上的变化。

与之相似，可使用标准程序构建含有本文公开的核酸或其片断的植物表达载体，并将其转化至合适的植物细胞中，如拟南芥、大豆、油菜籽、稻、玉米、小麦、蒺藜状苜蓿(*Medicago truncatula*)等。可分析产生的转基因细胞和/或其衍生植物在糖、油、脂质或脂肪酸含量上的变化。

此外，本文公开的序列或其片断可用于在多种生物基因组中产生敲除突变，如细菌、哺乳动物细胞、酵母细胞和植物细胞(Girke 等人, 1998, *Plant J.* 15:39-48)。然后可评价产生的敲除细胞中种子贮存化合物的组成和含量，以及突变对表型和/或基因型的影响。其他基因失活的方法包括 US 6,004,804 “非嵌合突变载体”和 Puttaraju 等人 (1999, "Spliceosome-mediated RNA *trans*-splicing as a tool for gene therapy" *Nature Biotech.* 17:246-252)所述的那些。

实施例 16: 从转化生物中纯化期望产物

可以通过本领域已知的多种方法从植物材料中回收 LMP。从植物器官中分离种子贮存化合物前，先将植物器官与其他组织或器官机械分离。组织匀浆后，离心除去细胞碎片，保留含有可溶性蛋白质的上清液级分，以进一步纯化期望化合物。若产物是由培养的细胞分泌的，则通过低速离心将细胞从培养物中去除，保留上清液级分，作进一步纯化。

在适当的树脂上对来自任一种纯化方法的上清液级分进行色谱分析，其中期望分子而非样品中的多种杂质留存于色谱树脂上，或杂质而非样品留存于树脂上。必要时可使用相同或不同的色谱树脂重复此类色谱步骤。

本领域技术人员擅长选择适当的色谱树脂，并将其最有效地运用于纯化具体的待纯化分子。可以通过过滤或超滤法浓缩纯化产物，并于可最大化产物稳定性的温度保存。

本领域公知多种纯化方法，因而前述纯化方法并非限制性的。此类纯化方法的描述见于如 Bailey & Ollis, 1986, *Biochemical Engineering Fundamentals*, McGraw-Hill: New York.

根据本领域标准技术鉴定分离化合物并评估其纯度。这些技术包括高效液相色谱法(HPLC)、分光光度法、染色法、薄层层析法、分析色谱法如高效液相色谱法、NIRS、酶法测定或微生物学测定。此类分析方法可见于 Patek 等人(1994, *Appl. Environ. Microbiol.* 60:133-140)、Malakhova 等人(1996, *Biotekhnologiya* 11: 27-32), 以及 Schmidt 等人(1998, *Bioprocess Engineer* 19: 67-70)、Ulmann's *Encyclopedia of Industrial Chemistry* (1996, 卷 A27, VCH: Weinheim, 第 89-90、521-540、540-547、559-566、575-581 和 581-587 页), 以及 Michal G. (1999, *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, John Wiley and Sons; Fallon, A.等人, 1987, *Applications of HPLC in Biochemistry, Laboratory Techniques in Biochemistry and Molecular Biology*, 卷 17)。

实施例 17: 通过设计微 RNA 前体下调基因表达

微 RNA (miRNA)是植物和动物中进化保守的基于 RNA 的基因表达调节物。miRNA (~ 21 至 25nt)来源于转录自非蛋白质编码基因的具有茎环结构的较大前体。miRNA 寻靶特定的 mRNA，以在转录后水平（即降解 mRNA）或翻译水平（即抑制蛋白质合成）上抑制基因表达(Bartel D 2004, *Cell* 116, 281-297)。

可以设计 miRNA 前体（前 miRNA），使由前 miRNA 编码的内源 miRNA 能被 miRNA 取代，以寻靶目的基因如 dsRed 报告基因。

选用玉米 miR166 前体进行设计。编码 miR166 前体的核苷酸序列如 SEQ ID NO: 47 所示。通过多点通路克隆方法(multi-site Gateway cloning

approach) (Invitrogen, Carlsbad, CA)产生两个二元表达构建体。如 SEQ ID NO: 41 所示的 RLM323 用作对照, 即处于 ScBV (甘蔗杆状病毒) 启动子及 NOS (胭脂碱合酶) 终止子调控下的天然玉米 miR166 表达。如 SEQ ID NO: 42 所示的 RLM325 与 RLM323 相同, 只是前体中如 SEQ ID NO: 37 和 SEQ ID NO: 38 所述的天然 miR166 (5' tcggaccagggttcattcccc 3') 被如 SEQ ID NO: 39 所示的 miRNA 靶向 dsRed (5' ttgtagatgaagcagccgtcc 3') 所取代。MiR dsRed 与 dsRed mRNA 的 3' 区域互补。

通过农杆菌将 RLM323 及 RLM325 转化入纯合子玉米 SDM10828 中, 其中已经携有二元载体 RLM185, 以在 ScBV 启动子和 NOS 终止子的调控下表达 dsRed。收集 3 个携带 RLM323 的独立 T0 事件和 29 个携带 RLM325 的独立 T0 事件的叶样品。然后使用 Typhoon 9400 (General Engineering, 配套的成像系统) 分析该样品, 以检测 dsRed 荧光。RLM325 事件的荧光强度与对照 RLM323 事件的强度相比, 降低了 90% 以上。

在 Northern 印迹分析中证实了 RLM325 事件中 miR dsRed 的产生。利用与 miR dsRed 互补的放射活性标记探针, 能在 RLM325 事件中检测到 ~21nt 的特异性条带, 而 RLM323 事件(对照)中则没有。通过 qRT-PCR 证实了 RLM325 事件中 dsRed mRNA 相对于对照 RLM323 的减少 (近 90%)。总之, 这些数据证实可以设计 miRNA 前体以寻靶目的基因。

在包括种子在内的多种组织中表达编码脂肪酸去饱和酶的玉米基因。使用 19 至 21nt (如 SEQ ID NO: 40 所述的 ACCAGACCCCGAACGCCGC) 的玉米去饱和酶编码区或 mRNA 中 5'UTR 及 3'UTR 互补物, 取代 Zm miR166 前体中如 SEQ ID NO: 37 及 SEQ ID NO: 38 所述的 Zm miR166 (5' tcggaccagggttcattcccc 3')。然后将该转基因转化入玉米。该基因改造的 Zm miR166 基因的表达受玉米种子特异性启动子 (如胚乳特异性 10 KD 玉米醇溶蛋白启动子或 Glob1 胚特异性启动子) 的调控。

当处理基因改造的 Zm miR166 前体时, 在种子中产生了微 RNA (如 SEQ ID NO: 40 所述的 ACCAGACCCCGAACGCCGC)。该 miRNA 特异性结合玉米脂肪酸去饱和酶 mRNA 中与该 miRNA 互补的区域, 从而通过

基因沉默机制导致种子中转录或翻译水平上该靶定玉米去饱和酶的表达减少。从而使转基因玉米具有期望的脂肪酸水平和组成，例如种子中低亚麻酸和/或中等或高油酸水平。

实施例 18：有关种子大小增加的筛选

FAD2 和作物 *FAD2* 样基因的条件性表达可导致种子大小增加。将会产生表达 *FAD2* 或 *FAD2* 样基因的转基因拟南芥或作物植物。利用显微镜确定种子大于野生型的转基因植物。此外，还测量转基因株系的种子重量。例如，*fad2* 突变体种子与野生型相比种子重量减少了 20%。在独立的拟南芥转基因株系 pFAD2RT-7 及 pFAD2RT-5 的分离 T2 种子世代中，100 粒种子的重量分别增加了 30% 及 40%。纯合子 T3 种子中，种子重量与空载体对照相比增加高达 60%（数据未示）。种子重量的增加体现在拟南芥中 *FAD2* 基因过表达种子大小的增加。种子大小的增加使得多种经济上重要的作物植物的产率增加。因此，种子大小的增加是使用本申请所述的 *FAD2* 或 *FAD2* 样核酸分子进行遗传工程和选择的目标之一。

表 1. 植物脂质类别

中性脂质	三酰基甘油(TAG)
	二酰基甘油(DAG)
	单酰基甘油(MAG)
极性脂质	单半乳糖基二酰基甘油(MGDG)
	双半乳糖基二酰基甘油(DGDG)
	磷脂酰甘油(PG)
	磷脂酰胆碱(PC)
	磷脂酰乙醇胺(PE)
	磷脂酰肌醇(PI)
	磷脂酰丝氨酸(PS)
	硫代异鼠李糖基甘油二酯

表 2. 常见植物脂肪酸

16:0	软脂酸
16:1	棕榈油酸
16:3	棕榈亚油酸 (palmitolenic acid)
18:0	硬脂酸
18:1	油酸
18:2	亚油酸
18:3	亚麻酸
γ -18:3	γ -亚麻酸*
20:0	花生酸
20:1	二十烯酸
22:6	二十二碳六烯酸(DHA) *
20:2	二十碳二烯酸
20:4	花生四烯酸(AA) *
20:5	二十碳五烯酸(EPA) *
22:1	芥酸

这些脂肪酸不常规见于植物种子油中，但其在转基因植物种子油中的产生在植物生物技术中至关重要。

表 3. 推定的 *FAD2* 样 LMP 功能表
(全长核酸序列可使用序列代号在附录 A 中找到)

SEQ ID NO:	序列名称	物种	功能	ORF 位置
1	<i>AtFAD-01</i>	拟南芥	ω -6 脂肪酸去饱和酶, 内质网 (FAD2) / Δ -12 去饱和酶	157-1305
5	<i>GmFAD-01</i>	大豆	ω -6 脂肪酸去饱和酶, 内质网 (FAD2) / Δ -12 去饱和酶	115-1275
9	<i>GmFAD-02</i>	大豆	ω -6 脂肪酸去饱和酶, 内质网 (FAD2) / Δ -12 去饱和酶	96-1244
13	<i>GmFAD-03</i>	大豆	ω -6 脂肪酸去饱和酶, 内质网 (FAD2) / Δ -12 去饱和酶	96-749
17	<i>ZmFAD-01</i>	玉米	玉米 Δ -12 去饱和酶 <i>fad2-2</i>	176-1351
21	<i>OsFAD-01</i>	稻	推断的 Δ -12 油酸去饱和酶	150-1313
25	<i>LuFAD-01</i>	亚麻	Δ -12 脂肪酸去饱和酶	48-1070
29	<i>HvFAD-01</i>	大麦	推断的 Δ -12 油酸去饱和酶	25-1185
33	<i>TaFAD-01</i>	小麦	推断的 Δ -12 油酸去饱和酶	165-1325

本领域技术人员通过常规试验, 能识别或能确定本文所述本发明具体实施方案的多种等效物。此类等效物旨在涵盖于本文公开和要求的本发明权利要求书中。

附录 A

图 1A: SEQ ID NO: 1 -*AtFAD-01* 的核酸序列

图 1B: SEQ ID NO: 3 - *AtFAD-01* 开放读框的核酸序列

图 1C: SEQ ID NO: 4 - *AtFAD-01* 开放读框的氨基酸序列

图 2A: SEQ ID NO: 5 - *GmFAD-01* 的核酸序列

图 2B: SEQ ID NO: 7 - *GmFAD-01* 开放读框的核酸序列

- 图 2C: SEQ ID NO: 8 - GmFAD-01 开放读框的氨基酸序列
- 图 3A: SEQ ID NO: 9 - GmFAD-02 的核酸序列
- 图 3B: SEQ ID NO: 11 - GmFAD-02 开放读框的核酸序列
- 图 3C: SEQ ID NO: 12 - GmFAD-02 开放读框的氨基酸序列
- 图 4A: SEQ ID NO: 12 - GmFAD-03 的核酸序列
- 图 4B: SEQ ID NO: 15 - GmFAD-03 开放读框的核酸序列
- 图 4C: SEQ ID NO: 16 - GmFAD-03 开放读框的氨基酸序列
- 图 5A: SEQ ID NO: 17 - ZmFAD-01 的核酸序列
- 图 5B: SEQ ID NO: 19 - ZmFAD-01 开放读框的核酸序列
- 图 5C: SEQ ID NO: 20 - ZmFAD-01 开放读框的氨基酸序列
- 图 6A: SEQ ID NO: 21 - OsFAD-01 的核酸序列
- 图 6B: SEQ ID NO: 23 - OsFAD-01 开放读框的核酸序列
- 图 6C: SEQ ID NO: 24 - OsFAD-01 开放读框的氨基酸序列
- 图 7A: SEQ ID NO: 25 - LuFAD-01 的核酸序列
- 图 7B: SEQ ID NO: 27 - LuFAD-01 开放读框的核酸序列
- 图 7C: SEQ ID NO: 28 - LuFAD-01 开放读框的氨基酸序列
- 图 8A: SEQ ID NO: 29 - HvFAD-01 的核酸序列
- 图 8B: SEQ ID NO: 31 - HvFAD-01 开放读框的核酸序列
- 图 8C: SEQ ID NO: 32 - HvFAD-01 开放读框的氨基酸序列
- 图 9A: SEQ ID NO: 33 - TaFAD-01 的核酸序列
- 图 9B: SEQ ID NO: 35 - TaFAD-01 开放读框的核酸序列
- 图 9C: SEQ ID NO: 36 - TaFAD-01 开放读框的氨基酸序列。

序列表

<110> 巴斯福植物科学有限公司

<120> 编码植物脂肪酸去饱和酶基因的核酸分子及其使用方法

<130> PF2063015115

<150> US 60/637531

<151> 2004-12-20

<160> 47

<170> PatentIn version 3.3

<210> 1

<211> 1615

<212> DNA

<213> 拟南芥 (*Arabidopsis thaliana*)

<220>

<221> misc_feature

<222> (1)..(156)

<220>

<221> CDS

<222> (157)..(1305)

<220>

<221> misc_feature

<222> (1306)..(1615)

<400> 1

gaccaccaga agaagagcca cacactcaca aattaaaaag agagagagag agagagagac 60

agagagagag agagattctg cggaggagct tcttcttcgt aggggtgttca tcgttattaa 120

cgttatcgcc cctacgtcag ctccatctcc agaaac atg ggt gca ggt gga aga 174

Met Gly Ala Gly Gly Arg

1

5

atg ccg gtt cct act tct tcc aag aaa tcg gaa acc gac acc aca aag 222

Met Pro Val Pro Thr Ser Ser Lys Lys Ser Glu Thr Asp Thr Thr Lys

10

15

20

cgt gtg ccg tgc gag aaa ccg cct ttc tcg gtg gga gat ctg aag aaa Arg Val Pro Cys Glu Lys Pro Pro Phe Ser Val Gly Asp Leu Lys Lys 25 30 35	270
gca atc ccg ccg cat tgt ttc aaa cgc tca atc cct cgc tct ttc tcc Ala Ile Pro Pro His Cys Phe Lys Arg Ser Ile Pro Arg Ser Phe Ser 40 45 50	318
tac ctt atc agt gac atc att ata gcc tca tgc ttc tac tac gtc gcc Tyr Leu Ile Ser Asp Ile Ile Ile Ala Ser Cys Phe Tyr Tyr Val Ala 55 60 65 70	366
acc aat tac ttc tct ctc ctc cct cag cct ctc tct tac ttg gct tgg Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro Leu Ser Tyr Leu Ala Trp 75 80 85	414
cca ctc tat tgg gcc tgt caa ggc tgt gtc cta act ggt atc tgg gtc Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val Leu Thr Gly Ile Trp Val 90 95 100	462
ata gcc cac gaa tgc ggt cac cac gca ttc agc gac tac caa tgg ctg Ile Ala His Glu Cys Gly His His Ala Phe Ser Asp Tyr Gln Trp Leu 105 110 115	510
gat gac aca gtt ggt ctt atc ttc cat tcc ttc ctc ctc gtc cct tac Asp Asp Thr Val Gly Leu Ile Phe His Ser Phe Leu Leu Val Pro Tyr 120 125 130	558
ttc tcc tgg aag tat agt cat cgc cgt cac cat tcc aac act gga tcc Phe Ser Trp Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser 135 140 145 150	606
ctc gaa aga gat gaa gta ttt gtc cca aag cag aaa tca gca atc aag Leu Glu Arg Asp Glu Val Phe Val Pro Lys Gln Lys Ser Ala Ile Lys 155 160 165	654
tgg tac ggg aaa tac ctc aac aac cct ctt gga cgc atc atg atg tta Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu Gly Arg Ile Met Met Leu 170 175 180	702
acc gtc cag ttt gtc ctc ggg tgg ccc ttg tac tta gcc ttt aac gtc Thr Val Gln Phe Val Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val 185 190 195	750

tct ggc aga ccg tat gac ggg ttc gct tgc cat ttc ttc ccc aac gct Ser Gly Arg Pro Tyr Asp Gly Phe Ala Cys His Phe Phe Pro Asn Ala 200 205 210	798
ccc atc tac aat gac cga gaa cgc ctc cag ata tac ctc tct gat gcg Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln Ile Tyr Leu Ser Asp Ala 215 220 225 230	846
ggt att cta gcc gtc tgt ttt ggt ctt tac cgt tac gct gct gca caa Gly Ile Leu Ala Val Cys Phe Gly Leu Tyr Arg Tyr Ala Ala Ala Gln 235 240 245	894
ggg atg gcc tcg atg atc tgc ctc tac gga gta ccg ctt ctg ata gtg Gly Met Ala Ser Met Ile Cys Leu Tyr Gly Val Pro Leu Leu Ile Val 250 255 260	942
aat gcg ttc ctc gtc ttg atc act tac ttg cag cac act cat ccc tcg Asn Ala Phe Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ser 265 270 275	990
ttg cct cac tac gat tca tca gag tgg gac tgg ctc agg gga gct ttg Leu Pro His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu 280 285 290	1038
gct acc gta gac aga gac tac gga atc ttg aac aag gtg ttc cac aac Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His Asn 295 300 305 310	1086
att aca gac aca cac gtg gct cat cac ctg ttc tcg aca atg ccg cat Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His 315 320 325	1134
tat aac gca atg gaa gct aca aag gcg ata aag cca att ctg gga gac Tyr Asn Ala Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly Asp 330 335 340	1182
tat tac cag ttc gat gga aca ccg tgg tat gta gcg atg tat agg gag Tyr Tyr Gln Phe Asp Gly Thr Pro Trp Tyr Val Ala Met Tyr Arg Glu 345 350 355	1230
gca aag gag tgt atc tat gta gaa ccg gac agg gaa ggt gac aag aaa Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp Arg Glu Gly Asp Lys Lys 360 365 370	1278

```

ggt gtg tac tgg tac aac aat aag tta tgaggatgat ggtgaagaaa      1325
Gly Val Tyr Trp Tyr Asn Asn Lys Leu
375                          380

ttgtcgacct ttctcttgtc tgtttgtctt ttgttaaaga agctatgctt cgttttaata 1385

atcttattgt ccattttgtt gtgttatgac attttggtg ctcattatgt tatgtgggaa 1445

gtagtggttc aaatgttttg tgtcggtatt gttcttctca tcgctgtttt gttgggatcg 1505

tagaaatgtg accttcggac agtaaaactc ttgtactaaa actatctccc tattggcatt 1565

tcttaaactt ttaatagtta cgtgctcgta gtgaatcttg acttgagtca          1615

```

<210> 2
 <211> 383
 <212> PRT
 <213> 拟南芥

<400> 2

```

Met Gly Ala Gly Gly Arg Met Pro Val Pro Thr Ser Ser Lys Lys Ser
1          5          10          15

Glu Thr Asp Thr Thr Lys Arg Val Pro Cys Glu Lys Pro Pro Phe Ser
          20          25          30

Val Gly Asp Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser
          35          40          45

Ile Pro Arg Ser Phe Ser Tyr Leu Ile Ser Asp Ile Ile Ile Ala Ser
          50          55          60

Cys Phe Tyr Tyr Val Ala Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro
65          70          75          80

Leu Ser Tyr Leu Ala Trp Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val
          85          90          95

Leu Thr Gly Ile Trp Val Ile Ala His Glu Cys Gly His His Ala Phe
          100         105         110

Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser
          115         120         125

```

Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His
130 135 140

His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys
145 150 155 160

Gln Lys Ser Ala Ile Lys Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu
165 170 175

Gly Arg Ile Met Met Leu Thr Val Gln Phe Val Leu Gly Trp Pro Leu
180 185 190

Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Phe Ala Cys
195 200 205

His Phe Phe Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln
210 215 220

Ile Tyr Leu Ser Asp Ala Gly Ile Leu Ala Val Cys Phe Gly Leu Tyr
225 230 235 240

Arg Tyr Ala Ala Ala Gln Gly Met Ala Ser Met Ile Cys Leu Tyr Gly
245 250 255

Val Pro Leu Leu Ile Val Asn Ala Phe Leu Val Leu Ile Thr Tyr Leu
260 265 270

Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp
275 280 285

Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu
290 295 300

Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu
305 310 315 320

Phe Ser Thr Met Pro His Tyr Asn Ala Met Glu Ala Thr Lys Ala Ile
325 330 335

Lys Pro Ile Leu Gly Asp Tyr Tyr Gln Phe Asp Gly Thr Pro Trp Tyr
340 345 350

Val Ala Met Tyr Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp
355 360 365

Arg Glu Gly Asp Lys Lys Gly Val Tyr Trp Tyr Asn Asn Lys Leu
 370 375 380

<210> 3
 <211> 1149
 <212> DNA
 <213> 拟南芥

<220>
 <221> CDS
 <222> (1)..(1149)

<400> 3
 atg ggt gca ggt gga aga atg ccg gtt cct act tct tcc aag aaa tcg 48
 Met Gly Ala Gly Gly Arg Met Pro Val Pro Thr Ser Ser Lys Lys Ser
 1 5 10 15
 gaa acc gac acc aca aag cgt gtg ccg tgc gag aaa ccg cct ttc tcg 96
 Glu Thr Asp Thr Thr Lys Arg Val Pro Cys Glu Lys Pro Pro Phe Ser
 20 25 30
 gtg gga gat ctg aag aaa gca atc ccg ccg cat tgt ttc aaa cgc tca 144
 Val Gly Asp Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser
 35 40 45
 atc cct cgc tct ttc tcc tac ctt atc agt gac atc att ata gcc tca 192
 Ile Pro Arg Ser Phe Ser Tyr Leu Ile Ser Asp Ile Ile Ile Ala Ser
 50 55 60
 tgc ttc tac tac gtc gcc acc aat tac ttc tct ctc ctc cct cag cct 240
 Cys Phe Tyr Tyr Val Ala Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro
 65 70 75 80
 ctc tct tac ttg gct tgg cca ctc tat tgg gcc tgt caa ggc tgt gtc 288
 Leu Ser Tyr Leu Ala Trp Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val
 85 90 95
 cta act ggt atc tgg gtc ata gcc cac gaa tgc ggt cac cac gca ttc 336
 Leu Thr Gly Ile Trp Val Ile Ala His Glu Cys Gly His His Ala Phe
 100 105 110

agc gac tac caa tgg ctg gat gac aca gtt ggt ctt atc ttc cat tcc Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser 115 120 125	384
ttc ctc ctc gtc cct tac ttc tcc tgg aag tat agt cat cgc cgt cac Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His 130 135 140	432
cat tcc aac act gga tcc ctc gaa aga gat gaa gta ttt gtc cca aag His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys 145 150 155 160	480
cag aaa tca gca atc aag tgg tac ggg aaa tac ctc aac aac cct ctt Gln Lys Ser Ala Ile Lys Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu 165 170 175	528
gga cgc atc atg atg tta acc gtc cag ttt gtc ctc ggg tgg ccc ttg Gly Arg Ile Met Met Leu Thr Val Gln Phe Val Leu Gly Trp Pro Leu 180 185 190	576
tac tta gcc ttt aac gtc tct ggc aga ccg tat gac ggg ttc gct tgc Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Phe Ala Cys 195 200 205	624
cat ttc ttc ccc aac gct ccc atc tac aat gac cga gaa cgc ctc cag His Phe Phe Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln 210 215 220	672
ata tac ctc tct gat gcg ggt att cta gcc gtc tgt ttt ggt ctt tac Ile Tyr Leu Ser Asp Ala Gly Ile Leu Ala Val Cys Phe Gly Leu Tyr 225 230 235 240	720
cgt tac gct gct gca caa ggg atg gcc tcg atg atc tgc ctc tac gga Arg Tyr Ala Ala Ala Gln Gly Met Ala Ser Met Ile Cys Leu Tyr Gly 245 250 255	768
gta ccg ctt ctg ata gtg aat gcg ttc ctc gtc ttg atc act tac ttg Val Pro Leu Leu Ile Val Asn Ala Phe Leu Val Leu Ile Thr Tyr Leu 260 265 270	816
cag cac act cat ccc tcg ttg cct cac tac gat tca tca gag tgg gac Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp 275 280 285	864

tgg ctc agg gga gct ttg gct acc gta gac aga gac tac gga atc ttg 912
 Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu
 290 295 300

aac aag gtg ttc cac aac att aca gac aca cac gtg gct cat cac ctg 960
 Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu
 305 310 315 320

ttc tcg aca atg ccg cat tat aac gca atg gaa gct aca aag gcg ata 1008
 Phe Ser Thr Met Pro His Tyr Asn Ala Met Glu Ala Thr Lys Ala Ile
 325 330 335

aag cca att ctg gga gac tat tac cag ttc gat gga aca ccg tgg tat 1056
 Lys Pro Ile Leu Gly Asp Tyr Tyr Gln Phe Asp Gly Thr Pro Trp Tyr
 340 345 350

gta gcg atg tat agg gag gca aag gag tgt atc tat gta gaa ccg gac 1104
 Val Ala Met Tyr Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp
 355 360 365

agg gaa ggt gac aag aaa ggt gtg tac tgg tac aac aat aag tta 1149
 Arg Glu Gly Asp Lys Lys Gly Val Tyr Trp Tyr Asn Asn Lys Leu
 370 375 380

<210> 4

<211> 383

<212> PRT

<213> 拟南芥

<400> 4

Met Gly Ala Gly Gly Arg Met Pro Val Pro Thr Ser Ser Lys Lys Ser
 1 5 10 15

Glu Thr Asp Thr Thr Lys Arg Val Pro Cys Glu Lys Pro Pro Phe Ser
 20 25 30

Val Gly Asp Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys Arg Ser
 35 40 45

Ile Pro Arg Ser Phe Ser Tyr Leu Ile Ser Asp Ile Ile Ile Ala Ser
 50 55 60

Cys Phe Tyr Tyr Val Ala Thr Asn Tyr Phe Ser Leu Leu Pro Gln Pro
65 70 75 80

Leu Ser Tyr Leu Ala Trp Pro Leu Tyr Trp Ala Cys Gln Gly Cys Val
85 90 95

Leu Thr Gly Ile Trp Val Ile Ala His Glu Cys Gly His His Ala Phe
100 105 110

Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Phe His Ser
115 120 125

Phe Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His
130 135 140

His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys
145 150 155 160

Gln Lys Ser Ala Ile Lys Trp Tyr Gly Lys Tyr Leu Asn Asn Pro Leu
165 170 175

Gly Arg Ile Met Met Leu Thr Val Gln Phe Val Leu Gly Trp Pro Leu
180 185 190

Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly Phe Ala Cys
195 200 205

His Phe Phe Pro Asn Ala Pro Ile Tyr Asn Asp Arg Glu Arg Leu Gln
210 215 220

Ile Tyr Leu Ser Asp Ala Gly Ile Leu Ala Val Cys Phe Gly Leu Tyr
225 230 235 240

Arg Tyr Ala Ala Ala Gln Gly Met Ala Ser Met Ile Cys Leu Tyr Gly
245 250 255

Val Pro Leu Leu Ile Val Asn Ala Phe Leu Val Leu Ile Thr Tyr Leu
260 265 270

Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp
275 280 285

Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu
290 295 300

Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu
305 310 315 320

Phe Ser Thr Met Pro His Tyr Asn Ala Met Glu Ala Thr Lys Ala Ile
325 330 335

Lys Pro Ile Leu Gly Asp Tyr Tyr Gln Phe Asp Gly Thr Pro Trp Tyr
340 345 350

Val Ala Met Tyr Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp
355 360 365

Arg Glu Gly Asp Lys Lys Gly Val Tyr Trp Tyr Asn Asn Lys Leu
370 375 380

<210> 5
<211> 1491
<212> DNA
<213> 大豆 (Glycine max)

<220>
<221> misc_feature
<222> (1)..(114)

<220>
<221> CDS
<222> (115)..(1275)

<220>
<221> misc_feature
<222> (1276)..(1491)

<400> 5
taggcaccta gctagtagct acaatatcag cacttctctc tattgataaa caattggctg 60

taatgccgca gtagaggacg atcacaacat ttcgtgctgg atactttttg tttt atg 117
Met
1

ggt cta gca aag gaa aca ata atg gga ggt gga ggc cgt gtg gcc aaa 165
Gly Leu Ala Lys Glu Thr Ile Met Gly Gly Gly Gly Arg Val Ala Lys
5 10 15

gtt gaa att cag cag aag aag cct ctc tca agg gtt cca aac aca aag Val Glu Ile Gln Gln Lys Lys Pro Leu Ser Arg Val Pro Asn Thr Lys 20 25 30	213
cca cca ttc act gtt ggc caa ctc aag aaa gcc att cca ccg cac tgc Pro Pro Phe Thr Val Gly Gln Leu Lys Lys Ala Ile Pro Pro His Cys 35 40 45	261
ttt cag cgt tcc ctc ctc act tca ttg tcc tat gtt gtt tat gac ctt Phe Gln Arg Ser Leu Leu Thr Ser Leu Ser Tyr Val Val Tyr Asp Leu 50 55 60 65	309
tca ttg gct ttc att ttc tac att gcc acc acc tac ttc cac ctc ctc Ser Leu Ala Phe Ile Phe Tyr Ile Ala Thr Thr Tyr Phe His Leu Leu 70 75 80	357
cct cac ccc ttt tcc ctc att gca tgg cca atc tat tgg gtt ctc caa Pro His Pro Phe Ser Leu Ile Ala Trp Pro Ile Tyr Trp Val Leu Gln 85 90 95	405
ggc tgc att ctt act ggc gtg tgg gtg att gct cac gag tgt ggt cac Gly Cys Ile Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His 100 105 110	453
cat gcc ttc agc aag tac cca tgg gtt gat gat gtt atg ggt ttg acc His Ala Phe Ser Lys Tyr Pro Trp Val Asp Asp Val Met Gly Leu Thr 115 120 125	501
gtt cac tca gca ctt tta gtc cct tat ttc tca tgg aaa ata agc cat Val His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Ile Ser His 130 135 140 145	549
cgc cgc cac cac tcc aac acg ggt tcc ctt gac cgt gat gaa gtg ttt Arg Arg His His Ser Asn Thr Gly Ser Leu Asp Arg Asp Glu Val Phe 150 155 160	597
gtc cca aaa cca aaa tcc aaa gtt gca tgg tac acc aag tac ctg aac Val Pro Lys Pro Lys Ser Lys Val Ala Trp Tyr Thr Lys Tyr Leu Asn 165 170 175	645
aac cct cta gga agg gct gct tct ctt ctc atc aca ctc aca ata ggg Asn Pro Leu Gly Arg Ala Ala Ser Leu Leu Ile Thr Leu Thr Ile Gly 180 185 190	693

tgg cct atg tat tta gcc ttc aat gtc tct ggc aga ccc tat gat ggt	741
Trp Pro Met Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Gly	
195 200 205	
ttt gca agc cac tac cac cct tat gct ccc ata tat tct aac cgt gag	789
Phe Ala Ser His Tyr His Pro Tyr Ala Pro Ile Tyr Ser Asn Arg Glu	
210 215 220 225	
agg ctt ctg atc tat gtc tct gat gtt gct ttg ttt tct gtg act tac	837
Arg Leu Leu Ile Tyr Val Ser Asp Val Ala Leu Phe Ser Val Thr Tyr	
230 235 240	
tct ctc tac cgt gtt gca act atg aaa ggg ttg gtt tgg ctg cta tgt	885
Ser Leu Tyr Arg Val Ala Thr Met Lys Gly Leu Val Trp Leu Leu Cys	
245 250 255	
gtt tat ggg gtg cca ttg ctc att gtg aac ggt ttt ctt gtg act atc	933
Val Tyr Gly Val Pro Leu Leu Ile Val Asn Gly Phe Leu Val Thr Ile	
260 265 270	
aca tat ttg cag cac aca cac ttt gcc ttg cct cat tac gat tca tca	981
Thr Tyr Leu Gln His Thr His Phe Ala Leu Pro His Tyr Asp Ser Ser	
275 280 285	
gaa tgg gac tgg ctg aag gga gct ttg gca act atg gac aga gat tat	1029
Glu Trp Asp Trp Leu Lys Gly Ala Leu Ala Thr Met Asp Arg Asp Tyr	
290 295 300 305	
ggg att ctg aac aag gtg ttt cat cac ata act gat act cat gtg gct	1077
Gly Ile Leu Asn Lys Val Phe His His Ile Thr Asp Thr His Val Ala	
310 315 320	
cac cat ctc ttc tct aca atg cca cat tac cat gca atg gag gca acc	1125
His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr	
325 330 335	
aat gca atc aag cca ata ttg ggt gag tac tac caa ttt gat gac aca	1173
Asn Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Asp Thr	
340 345 350	
cca ttt tac aag gca ctg tgg aga gaa gcg aga gag tgc ctc tat gtg	1221
Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Arg Glu Cys Leu Tyr Val	
355 360 365	

gag cca gat gaa gga aca tcc gag aag ggc gtg tat tgg tac agg aac 1269
 Glu Pro Asp Glu Gly Thr Ser Glu Lys Gly Val Tyr Trp Tyr Arg Asn
 370 375 380 385

aag tat tgatggacca agcaatgggc catagtggga gttatggaag tttgtcact 1325
 Lys Tyr

tatcacttaa ttagtagaat gttataaata agtggatttg ccgcgtaatg acttgtgtgc 1385

attgtgaaac agcttgtagc gatccatggg tataatgtaa aaatatgtgg aaaggggtct 1445

ggttaaaaaa aaaaaaaaaa aagcggccgt tttaaaggaa acaagg 1491

<210> 6

<211> 387

<212> PRT

<213> 大豆

<400> 6

Met Gly Leu Ala Lys Glu Thr Ile Met Gly Gly Gly Gly Arg Val Ala
 1 5 10 15

Lys Val Glu Ile Gln Gln Lys Lys Pro Leu Ser Arg Val Pro Asn Thr
 20 25 30

Lys Pro Pro Phe Thr Val Gly Gln Leu Lys Lys Ala Ile Pro Pro His
 35 40 45

Cys Phe Gln Arg Ser Leu Leu Thr Ser Leu Ser Tyr Val Val Tyr Asp
 50 55 60

Leu Ser Leu Ala Phe Ile Phe Tyr Ile Ala Thr Thr Tyr Phe His Leu
 65 70 75 80

Leu Pro His Pro Phe Ser Leu Ile Ala Trp Pro Ile Tyr Trp Val Leu
 85 90 95

Gln Gly Cys Ile Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly
 100 105 110

His His Ala Phe Ser Lys Tyr Pro Trp Val Asp Asp Val Met Gly Leu
 115 120 125

Thr Val His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Ile Ser
 130 135 140

His Arg Arg His His Ser Asn Thr Gly Ser Leu Asp Arg Asp Glu Val
 145 150 155 160

Phe Val Pro Lys Pro Lys Ser Lys Val Ala Trp Tyr Thr Lys Tyr Leu
 165 170 175

Asn Asn Pro Leu Gly Arg Ala Ala Ser Leu Leu Ile Thr Leu Thr Ile
 180 185 190

Gly Trp Pro Met Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp
 195 200 205

Gly Phe Ala Ser His Tyr His Pro Tyr Ala Pro Ile Tyr Ser Asn Arg
 210 215 220

Glu Arg Leu Leu Ile Tyr Val Ser Asp Val Ala Leu Phe Ser Val Thr
 225 230 235 240

Tyr Ser Leu Tyr Arg Val Ala Thr Met Lys Gly Leu Val Trp Leu Leu
 245 250 255

Cys Val Tyr Gly Val Pro Leu Leu Ile Val Asn Gly Phe Leu Val Thr
 260 265 270

Ile Thr Tyr Leu Gln His Thr His Phe Ala Leu Pro His Tyr Asp Ser
 275 280 285

Ser Glu Trp Asp Trp Leu Lys Gly Ala Leu Ala Thr Met Asp Arg Asp
 290 295 300

Tyr Gly Ile Leu Asn Lys Val Phe His His Ile Thr Asp Thr His Val
 305 310 315 320

Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala
 325 330 335

Thr Asn Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Asp
 340 345 350

Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Arg Glu Cys Leu Tyr
 355 360 365

Val Glu Pro Asp Glu Gly Thr Ser Glu Lys Gly Val Tyr Trp Tyr Arg
 370 375 380

Asn Lys Tyr
 385

<210> 7

<211> 1161

<212> DNA

<213> 大豆

<220>

<221> CDS

<222> (1)..(1161)

<400> 7

atg ggt cta gca aag gaa aca ata atg gga ggt gga ggc cgt gtg gcc 48
 Met Gly Leu Ala Lys Glu Thr Ile Met Gly Gly Gly Gly Arg Val Ala
 1 5 10 15

aaa gtt gaa att cag cag aag aag cct ctc tca agg gtt cca aac aca 96
 Lys Val Glu Ile Gln Gln Lys Lys Pro Leu Ser Arg Val Pro Asn Thr
 20 25 30

aag cca cca ttc act gtt ggc caa ctc aag aaa gcc att cca ccg cac 144
 Lys Pro Pro Phe Thr Val Gly Gln Leu Lys Lys Ala Ile Pro Pro His
 35 40 45

tgc ttt cag cgt tcc ctc ctc act tca ttg tcc tat gtt gtt tat gac 192
 Cys Phe Gln Arg Ser Leu Leu Thr Ser Leu Ser Tyr Val Val Tyr Asp
 50 55 60

ctt tca ttg gct ttc att ttc tac att gcc acc acc tac ttc cac ctc 240
 Leu Ser Leu Ala Phe Ile Phe Tyr Ile Ala Thr Thr Tyr Phe His Leu
 65 70 75 80

ctc cct cac ccc ttt tcc ctc att gca tgg cca atc tat tgg gtt ctc 288
 Leu Pro His Pro Phe Ser Leu Ile Ala Trp Pro Ile Tyr Trp Val Leu
 85 90 95

caa ggt tgc att ctt act ggc gtg tgg gtg att gct cac gag tgt ggt 336
 Gln Gly Cys Ile Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly
 100 105 110

cac cat gcc ttc agc aag tac cca tgg gtt gat gat gtt atg ggt ttg His His Ala Phe Ser Lys Tyr Pro Trp Val Asp Asp Val Met Gly Leu 115 120 125	384
acc gtt cac tca gca ctt tta gtc cct tat ttc tca tgg aaa ata agc Thr Val His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Ile Ser 130 135 140	432
cat cgc cgc cac cac tcc aac acg ggt tcc ctt gac cgt gat gaa gtg His Arg Arg His His Ser Asn Thr Gly Ser Leu Asp Arg Asp Glu Val 145 150 155 160	480
ttt gtc cca aaa cca aaa tcc aaa gtt gca tgg tac acc aag tac ctg Phe Val Pro Lys Pro Lys Ser Lys Val Ala Trp Tyr Thr Lys Tyr Leu 165 170 175	528
aac aac cct cta gga agg gct gct tct ctt ctc atc aca ctc aca ata Asn Asn Pro Leu Gly Arg Ala Ala Ser Leu Leu Ile Thr Leu Thr Ile 180 185 190	576
ggg tgg cct atg tat tta gcc ttc aat gtc tct ggc aga ccc tat gat Gly Trp Pro Met Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp 195 200 205	624
ggt ttt gca agc cac tac cac cct tat gct ccc ata tat tct aac cgt Gly Phe Ala Ser His Tyr His Pro Tyr Ala Pro Ile Tyr Ser Asn Arg 210 215 220	672
gag agg ctt ctg atc tat gtc tct gat gtt gct ttg ttt tct gtg act Glu Arg Leu Leu Ile Tyr Val Ser Asp Val Ala Leu Phe Ser Val Thr 225 230 235 240	720
tac tct ctc tac cgt gtt gca act atg aaa ggg ttg gtt tgg ctg cta Tyr Ser Leu Tyr Arg Val Ala Thr Met Lys Gly Leu Val Trp Leu Leu 245 250 255	768
tgt gtt tat ggg gtg cca ttg ctc att gtg aac ggt ttt ctt gtg act Cys Val Tyr Gly Val Pro Leu Leu Ile Val Asn Gly Phe Leu Val Thr 260 265 270	816
atc aca tat ttg cag cac aca cac ttt gcc ttg cct cat tac gat tca Ile Thr Tyr Leu Gln His Thr His Phe Ala Leu Pro His Tyr Asp Ser 275 280 285	864

tca gaa tgg gac tgg ctg aag gga gct ttg gca act atg gac aga gat Ser Glu Trp Asp Trp Leu Lys Gly Ala Leu Ala Thr Met Asp Arg Asp 290 295 300	912
tat ggg att ctg aac aag gtg ttt cat cac ata act gat act cat gtg Tyr Gly Ile Leu Asn Lys Val Phe His His Ile Thr Asp Thr His Val 305 310 315 320	960
gct cac cat ctc ttc tct aca atg cca cat tac cat gca atg gag gca Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala 325 330 335	1008
acc aat gca atc aag cca ata ttg ggt gag tac tac caa ttt gat gac Thr Asn Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Asp 340 345 350	1056
aca cca ttt tac aag gca ctg tgg aga gaa gcg aga gag tgc ctc tat Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Arg Glu Cys Leu Tyr 355 360 365	1104
gtg gag cca gat gaa gga aca tcc gag aag ggc gtg tat tgg tac agg Val Glu Pro Asp Glu Gly Thr Ser Glu Lys Gly Val Tyr Trp Tyr Arg 370 375 380	1152
aac aag tat Asn Lys Tyr 385	1161
<210> 8 <211> 387 <212> PRT <213> 大豆	
<400> 8	
Met Gly Leu Ala Lys Glu Thr Ile Met Gly Gly Gly Gly Arg Val Ala 1 5 10 15	
Lys Val Glu Ile Gln Gln Lys Lys Pro Leu Ser Arg Val Pro Asn Thr 20 25 30	
Lys Pro Pro Phe Thr Val Gly Gln Leu Lys Lys Ala Ile Pro Pro His 35 40 45	

Cys Phe Gln Arg Ser Leu Leu Thr Ser Leu Ser Tyr Val Val Tyr Asp
 50 55 60

Leu Ser Leu Ala Phe Ile Phe Tyr Ile Ala Thr Thr Tyr Phe His Leu
 65 70 75 80

Leu Pro His Pro Phe Ser Leu Ile Ala Trp Pro Ile Tyr Trp Val Leu
 85 90 95

Gln Gly Cys Ile Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly
 100 105 110

His His Ala Phe Ser Lys Tyr Pro Trp Val Asp Asp Val Met Gly Leu
 115 120 125

Thr Val His Ser Ala Leu Leu Val Pro Tyr Phe Ser Trp Lys Ile Ser
 130 135 140

His Arg Arg His His Ser Asn Thr Gly Ser Leu Asp Arg Asp Glu Val
 145 150 155 160

Phe Val Pro Lys Pro Lys Ser Lys Val Ala Trp Tyr Thr Lys Tyr Leu
 165 170 175

Asn Asn Pro Leu Gly Arg Ala Ala Ser Leu Leu Ile Thr Leu Thr Ile
 180 185 190

Gly Trp Pro Met Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp
 195 200 205

Gly Phe Ala Ser His Tyr His Pro Tyr Ala Pro Ile Tyr Ser Asn Arg
 210 215 220

Glu Arg Leu Leu Ile Tyr Val Ser Asp Val Ala Leu Phe Ser Val Thr
 225 230 235 240

Tyr Ser Leu Tyr Arg Val Ala Thr Met Lys Gly Leu Val Trp Leu Leu
 245 250 255

Cys Val Tyr Gly Val Pro Leu Leu Ile Val Asn Gly Phe Leu Val Thr
 260 265 270

Ile Thr Tyr Leu Gln His Thr His Phe Ala Leu Pro His Tyr Asp Ser
 275 280 285

Ser Glu Trp Asp Trp Leu Lys Gly Ala Leu Ala Thr Met Asp Arg Asp			
290	295	300	
Tyr Gly Ile Leu Asn Lys Val Phe His His Ile Thr Asp Thr His Val			
305	310	315	320
Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala			
	325	330	335
Thr Asn Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Asp			
	340	345	350
Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Arg Glu Cys Leu Tyr			
	355	360	365
Val Glu Pro Asp Glu Gly Thr Ser Glu Lys Gly Val Tyr Trp Tyr Arg			
	370	375	380
Asn Lys Tyr			
385			
<210>	9		
<211>	1616		
<212>	DNA		
<213>	大豆		
<220>			
<221>	misc_feature		
<222>	(1).. (95)		
<220>			
<221>	CDS		
<222>	(96).. (1244)		
<220>			
<221>	misc_feature		
<222>	(1245).. (1616)		
<400>	9		
gcgtgtcggt ctctctctct ctctcaccct cctcttcaca cattttctgt gcgctctaac			60
aaacattctc gttcacactt tcagattgtg tgaag atg ggg gcg ggt ggc cga			113
	Met Gly Ala Gly Gly Arg		
	1	5	

act gat gtt cct cct gcc aac agg aag tca gag gtt gac cct ttg aag	161
Thr Asp Val Pro Pro Ala Asn Arg Lys Ser Glu Val Asp Pro Leu Lys	
10 15 20	
cgg gtg cca ttt gaa aaa cct cca ttt agt ctc agc caa atc aag aag	209
Arg Val Pro Phe Glu Lys Pro Pro Phe Ser Leu Ser Gln Ile Lys Lys	
25 30 35	
gtc att cca cct cac tgt ttc cag cgt tct gtt ttc cgc tca ttc tcc	257
Val Ile Pro Pro His Cys Phe Gln Arg Ser Val Phe Arg Ser Phe Ser	
40 45 50	
tat gtt gtt tac gac ctc acc ata gcc ttc tgc ctc tat tat gtt gcc	305
Tyr Val Val Tyr Asp Leu Thr Ile Ala Phe Cys Leu Tyr Tyr Val Ala	
55 60 65 70	
acc cat tac ttc cac ctc ctt ccc agc cct ctc tct ttc ttg gca tgg	353
Thr His Tyr Phe His Leu Leu Pro Ser Pro Leu Ser Phe Leu Ala Trp	
75 80 85	
cca atc tac tgg gct gtc caa ggt tgc atc ctt act gga gtt tgg gtc	401
Pro Ile Tyr Trp Ala Val Gln Gly Cys Ile Leu Thr Gly Val Trp Val	
90 95 100	
att gcc cat gag tgt ggc cac cat gca ttc agt gac tac cag ttg ctt	449
Ile Ala His Glu Cys Gly His His Ala Phe Ser Asp Tyr Gln Leu Leu	
105 110 115	
gat gat att gtt ggc ctt gtc ctc cac tcc ggt ctc cta gtc cca tac	497
Asp Asp Ile Val Gly Leu Val Leu His Ser Gly Leu Leu Val Pro Tyr	
120 125 130	
ttt tca tgg aaa tac agc cat cgc cgt cac cac tcc aac act ggt tct	545
Phe Ser Trp Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser	
135 140 145 150	
ctt gag cgg gat gaa gta ttt gtg cca aag cag aag tcc tgt atc aag	593
Leu Glu Arg Asp Glu Val Phe Val Pro Lys Gln Lys Ser Cys Ile Lys	
155 160 165	
tgg tac tct aaa tac ctt aac aat cct cca ggc aga gtc ctc act ctt	641
Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Val Leu Thr Leu	
170 175 180	

gct gtc acc ctc aca ctt ggt tgg ccc ttg tac ttg gct tta aat gtt Ala Val Thr Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Leu Asn Val 185 190 195	689
tct gga agg cct tat gat aga ttt gct tgc cac tat gac cca tat ggt Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro Tyr Gly 200 205 210	737
ccc att tac tct gat cgt gaa cga ctt caa ata tat ata tca gat gca Pro Ile Tyr Ser Asp Arg Glu Arg Leu Gln Ile Tyr Ile Ser Asp Ala 215 220 225 230	785
gga gta ctt gca gta tgc tat ggc ctt ttc cgt ctt gcc atg gca aaa Gly Val Leu Ala Val Cys Tyr Gly Leu Phe Arg Leu Ala Met Ala Lys 235 240 245	833
gga ctt gcc tgg gtg gtg tgt gtt tat gga gtt cca ttg cta gtg gtc Gly Leu Ala Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu Val Val 250 255 260	881
aat gga ttt ttg gtg ttg att aca ttc ttg cag cat act cac cct gca Asn Gly Phe Leu Val Leu Ile Thr Phe Leu Gln His Thr His Pro Ala 265 270 275	929
ttg cca cat tac act tcc tct gag tgg gac tgg ttg aga gga gct tta Leu Pro His Tyr Thr Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu 280 285 290	977
gca aca gtg gat aga gat tat gga atc ctg aac aag gtc ttc cat aat Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His Asn 295 300 305 310	1025
att aca gac act cat gta gca cat cac ttg ttc tcc aca atg cca cat Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His 315 320 325	1073
tat cat gca atg gag gct aca aag gca ata aaa ccc att ttg gga gag Tyr His Ala Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly Glu 330 335 340	1121
tat tat cgg ttt gat gag act cca ttt gtc aag gca atg tgg aga gag Tyr Tyr Arg Phe Asp Glu Thr Pro Phe Val Lys Ala Met Trp Arg Glu 345 350 355	1169

gca aga gag tgt att tat gtg gag cca gat caa agt acc gag agc aaa 1217
 Ala Arg Glu Cys Ile Tyr Val Glu Pro Asp Gln Ser Thr Glu Ser Lys
 360 365 370

ggt gta ttt tgg tac aac aat aag ttg tgatgattaa ttagccgag 1264
 Gly Val Phe Trp Tyr Asn Asn Lys Leu
 375 380

gcttctttga actttccctt gtgactgttt agtatcatgg ttgcttattg ggaataattt 1324

tggtgaaccc tgatgttggt agtaagtatc tagacagttg catagcggtt ttgtttacag 1384

aataagatat agcctctctg aacagtttga ttattgcacc atggtttgca atcggtgcat 1444

gtcgaccaag tttctcaaga ctgtggagaa gcttattctt gttccagttc ttgaatccaa 1504

gttgttaccg tattctgtta ttattgactt agaatcctta accttttctg ctgttttctc 1564

atgatcgtca ctgcgaaatg aatcacattt caaaccaaaa aaaaaaaaaa aa 1616

<210> 10

<211> 383

<212> PRT

<213> 大豆

<400> 10

Met Gly Ala Gly Gly Arg Thr Asp Val Pro Pro Ala Asn Arg Lys Ser
 1 5 10 15

Glu Val Asp Pro Leu Lys Arg Val Pro Phe Glu Lys Pro Pro Phe Ser
 20 25 30

Leu Ser Gln Ile Lys Lys Val Ile Pro Pro His Cys Phe Gln Arg Ser
 35 40 45

Val Phe Arg Ser Phe Ser Tyr Val Val Tyr Asp Leu Thr Ile Ala Phe
 50 55 60

Cys Leu Tyr Tyr Val Ala Thr His Tyr Phe His Leu Leu Pro Ser Pro
 65 70 75 80

Leu Ser Phe Leu Ala Trp Pro Ile Tyr Trp Ala Val Gln Gly Cys Ile
 85 90 95

Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His Ala Phe
 100 105 110

Ser Asp Tyr Gln Leu Leu Asp Asp Ile Val Gly Leu Val Leu His Ser
 115 120 125

Gly Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His
 130 135 140

His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys
 145 150 155 160

Gln Lys Ser Cys Ile Lys Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro
 165 170 175

Gly Arg Val Leu Thr Leu Ala Val Thr Leu Thr Leu Gly Trp Pro Leu
 180 185 190

Tyr Leu Ala Leu Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys
 195 200 205

His Tyr Asp Pro Tyr Gly Pro Ile Tyr Ser Asp Arg Glu Arg Leu Gln
 210 215 220

Ile Tyr Ile Ser Asp Ala Gly Val Leu Ala Val Cys Tyr Gly Leu Phe
 225 230 235 240

Arg Leu Ala Met Ala Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly
 245 250 255

Val Pro Leu Leu Val Val Asn Gly Phe Leu Val Leu Ile Thr Phe Leu
 260 265 270

Gln His Thr His Pro Ala Leu Pro His Tyr Thr Ser Ser Glu Trp Asp
 275 280 285

Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu
 290 295 300

Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu
 305 310 315 320

Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr Lys Ala Ile
 325 330 335

Lys Pro Ile Leu Gly Glu Tyr Tyr Arg Phe Asp Glu Thr Pro Phe Val
 340 345 350

Lys Ala Met Trp Arg Glu Ala Arg Glu Cys Ile Tyr Val Glu Pro Asp
 355 360 365

Gln Ser Thr Glu Ser Lys Gly Val Phe Trp Tyr Asn Asn Lys Leu
 370 375 380

<210> 11

<211> 1149

<212> DNA

<213> 大豆

<220>

<221> CDS

<222> (1)..(1149)

<400> 11

atg ggg gcg ggt ggc cga act gat gtt cct cct gcc aac agg aag tca 48
 Met Gly Ala Gly Gly Arg Thr Asp Val Pro Pro Ala Asn Arg Lys Ser
 1 5 10 15

gag gtt gac cct ttg aag cgg gtg cca ttt gaa aaa cct cca ttt agt 96
 Glu Val Asp Pro Leu Lys Arg Val Pro Phe Glu Lys Pro Pro Phe Ser
 20 25 30

ctc agc caa atc aag aag gtc att cca cct cac tgt ttc cag cgt tct 144
 Leu Ser Gln Ile Lys Lys Val Ile Pro Pro His Cys Phe Gln Arg Ser
 35 40 45

gtt ttc cgc tca ttc tcc tat gtt gtt tac gac ctc acc ata gcc ttc 192
 Val Phe Arg Ser Phe Ser Tyr Val Val Tyr Asp Leu Thr Ile Ala Phe
 50 55 60

tgc ctc tat tat gtt gcc acc cat tac ttc cac ctc ctt ccc agc cct 240
 Cys Leu Tyr Tyr Val Ala Thr His Tyr Phe His Leu Leu Pro Ser Pro
 65 70 75 80

ctc tct ttc ttg gca tgg cca atc tac tgg gct gtc caa ggt tgc atc 288
 Leu Ser Phe Leu Ala Trp Pro Ile Tyr Trp Ala Val Gln Gly Cys Ile
 85 90 95

ctt act gga gtt tgg gtc att gcc cat gag tgt ggc cac cat gca ttc Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His Ala Phe 100 105 110	336
agt gac tac cag ttg ctt gat gat att gtt ggc ctt gtc ctc cac tcc Ser Asp Tyr Gln Leu Leu Asp Asp Ile Val Gly Leu Val Leu His Ser 115 120 125	384
ggc ctc cta gtc cca tac ttt tca tgg aaa tac agc cat cgc cgt cac Gly Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His 130 135 140	432
cac tcc aac act ggt tct ctt gag cgg gat gaa gta ttt gtg cca aag His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys 145 150 155 160	480
cag aag tcc tgt atc aag tgg tac tct aaa tac ctt aac aat cct cca Gln Lys Ser Cys Ile Lys Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro 165 170 175	528
ggc aga gtc ctc act ctt gct gtc acc ctc aca ctt ggt tgg ccc ttg Gly Arg Val Leu Thr Leu Ala Val Thr Leu Thr Leu Gly Trp Pro Leu 180 185 190	576
tac ttg gct tta aat gtt tct gga agg cct tat gat aga ttt gct tgc Tyr Leu Ala Leu Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys 195 200 205	624
cac tat gac cca tat ggt ccc att tac tct gat cgt gaa cga ctt caa His Tyr Asp Pro Tyr Gly Pro Ile Tyr Ser Asp Arg Glu Arg Leu Gln 210 215 220	672
ata tat ata tca gat gca gga gta ctt gca gta tgc tat ggc ctt ttc Ile Tyr Ile Ser Asp Ala Gly Val Leu Ala Val Cys Tyr Gly Leu Phe 225 230 235 240	720
cgt ctt gcc atg gca aaa gga ctt gcc tgg gtg gtg tgt gtt tat gga Arg Leu Ala Met Ala Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly 245 250 255	768
gtt cca ttg cta gtg gtc aat gga ttt ttg gtg ttg att aca ttc ttg Val Pro Leu Leu Val Val Asn Gly Phe Leu Val Leu Ile Thr Phe Leu 260 265 270	816

cag cat act cac cct gca ttg cca cat tac act tcc tct gag tgg gac	864
Gln His Thr His Pro Ala Leu Pro His Tyr Thr Ser Ser Glu Trp Asp	
275 280 285	
tggttg agaggatgacacagtgat agagat tat gga atc ctg	912
Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu	
290 295 300	
aac aag gtc ttc cat aat att aca gac act cat gta gca cat cac ttg	960
Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu	
305 310 315 320	
ttc tcc aca atg cca cat tat cat gca atg gag gct aca aag gca ata	1008
Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr Lys Ala Ile	
325 330 335	
aaa ccc att ttg gga gag tat tat cgg ttt gat gag act cca ttt gtc	1056
Lys Pro Ile Leu Gly Glu Tyr Tyr Arg Phe Asp Glu Thr Pro Phe Val	
340 345 350	
aag gca atg ttg aga gag gca aga gag tgt att tat gtg gag cca gat	1104
Lys Ala Met Trp Arg Glu Ala Arg Glu Cys Ile Tyr Val Glu Pro Asp	
355 360 365	
caa agt acc gag agc aaa ggt gta ttt ttggtac aac aat aag ttg	1149
Gln Ser Thr Glu Ser Lys Gly Val Phe Trp Tyr Asn Asn Lys Leu	
370 375 380	

<210> 12

<211> 383

<212> PRT

<213> 大豆

<400> 12

Met Gly Ala Gly Gly Arg Thr Asp Val Pro Pro Ala Asn Arg Lys Ser
1 5 10 15

Glu Val Asp Pro Leu Lys Arg Val Pro Phe Glu Lys Pro Pro Phe Ser
20 25 30

Leu Ser Gln Ile Lys Lys Val Ile Pro Pro His Cys Phe Gln Arg Ser
35 40 45

Val Phe Arg Ser Phe Ser Tyr Val Val Tyr Asp Leu Thr Ile Ala Phe
50 55 60

Cys Leu Tyr Tyr Val Ala Thr His Tyr Phe His Leu Leu Pro Ser Pro
65 70 75 80

Leu Ser Phe Leu Ala Trp Pro Ile Tyr Trp Ala Val Gln Gly Cys Ile
85 90 95

Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His Ala Phe
100 105 110

Ser Asp Tyr Gln Leu Leu Asp Asp Ile Val Gly Leu Val Leu His Ser
115 120 125

Gly Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His
130 135 140

His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys
145 150 155 160

Gln Lys Ser Cys Ile Lys Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro
165 170 175

Gly Arg Val Leu Thr Leu Ala Val Thr Leu Thr Leu Gly Trp Pro Leu
180 185 190

Tyr Leu Ala Leu Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys
195 200 205

His Tyr Asp Pro Tyr Gly Pro Ile Tyr Ser Asp Arg Glu Arg Leu Gln
210 215 220

Ile Tyr Ile Ser Asp Ala Gly Val Leu Ala Val Cys Tyr Gly Leu Phe
225 230 235 240

Arg Leu Ala Met Ala Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly
245 250 255

Val Pro Leu Leu Val Val Asn Gly Phe Leu Val Leu Ile Thr Phe Leu
260 265 270

Gln His Thr His Pro Ala Leu Pro His Tyr Thr Ser Ser Glu Trp Asp
275 280 285

Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu
290 295 300

Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu
305 310 315 320

Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr Lys Ala Ile
325 330 335

Lys Pro Ile Leu Gly Glu Tyr Tyr Arg Phe Asp Glu Thr Pro Phe Val
340 345 350

Lys Ala Met Trp Arg Glu Ala Arg Glu Cys Ile Tyr Val Glu Pro Asp
355 360 365

Gln Ser Thr Glu Ser Lys Gly Val Phe Trp Tyr Asn Asn Lys Leu
370 375 380

<210> 13
<211> 1053
<212> DNA
<213> 大豆

<220>
<221> misc_feature
<222> (1)..(95)

<220>
<221> CDS
<222> (96)..(749)

<220>
<221> misc_feature
<222> (750)..(1053)

<400> 13

acttttcatg gaaatacagc catgccgctc accactccaa cacaggttct cttgagcgag 60

atgaagtatt tgtgccaaag cagaagtcca gtatc atg tgg tac tct aaa tac 113
Met Trp Tyr Ser Lys Tyr
1 5

ctt aac aat cca cca ggc aga gtc ctc act ctt gcc gtc acc ctc acg Leu Asn Asn Pro Pro Gly Arg Val Leu Thr Leu Ala Val Thr Leu Thr 10 15 20	161
ctt ggt tgg ccc ttg tac ttg gct ttt aat gtt tct gga agg cct tat Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr 25 30 35	209
gat aga ttt gct tgc cac tat gac cct tat ggt ccc att tac tct gac Asp Arg Phe Ala Cys His Tyr Asp Pro Tyr Gly Pro Ile Tyr Ser Asp 40 45 50	257
cga gaa cga ctt caa ata tat ata tca gat gca gga gta ctt gca gta Arg Glu Arg Leu Gln Ile Tyr Ile Ser Asp Ala Gly Val Leu Ala Val 55 60 65 70	305
tgc tat ggc ctt ttc tgt ctt gcc atg gca aaa ggg ctt gcc tgg gtg Cys Tyr Gly Leu Phe Cys Leu Ala Met Ala Lys Gly Leu Ala Trp Val 75 80 85	353
gtg tgt gtt tat gga gtt cca ttg ctt gtg gtc aat gga ttt ttg gtg Val Cys Val Tyr Gly Val Pro Leu Leu Val Val Asn Gly Phe Leu Val 90 95 100	401
ttg att aca ttt ttg cag cac act cac cct gca ttg cca cac tac act Leu Ile Thr Phe Leu Gln His Thr His Pro Ala Leu Pro His Tyr Thr 105 110 115	449
tcc tct gag tgg gac tgg ttg aga gga gct tta gca aca gtg gat aga Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg 120 125 130	497
gat tat gga atc ctg aac aag gtc ttc cat aat att aca gac act cat Asp Tyr Gly Ile Leu Asn Lys Val Phe His Asn Ile Thr Asp Thr His 135 140 145 150	545
gta gct cat cac ttg ttc tcc aca atg cca cat tat cat gca atg gag Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu 155 160 165	593
gcg aca aag gca ata aag ccc atc ttg gga gag tat tat cgg ttt gat Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Arg Phe Asp 170 175 180	641

```

ggg act cca ttt gtc aag gca atg tgg aga gag gca aga gag tgt att      689
Gly Thr Pro Phe Val Lys Ala Met Trp Arg Glu Ala Arg Glu Cys Ile
      185              190              195

tat gtg gag cca gat caa agt act cag agc aaa ggt gta ttt tgg tac      737
Tyr Val Glu Pro Asp Gln Ser Thr Gln Ser Lys Gly Val Phe Trp Tyr
      200              205              210

aac aat aag ttg tgatgattaa ttagccggg ggcttcttgt ccgtcaggtg      789
Asn Asn Lys Leu
215

ggatggtttg aactttcctt tgtgactgtt tagtatcatg cttgcttatt gggaataatt      849

ttgttgaacc ctgatgttgg tagtagtata tagaaagtag catagcgttt ttgtttgcag      909

aataagatat agcatctctg aacagtttga ttattgcacc atgttttgca atcagtgcac      969

gtcgaccggg tttctcaaga ttgtggggat gcttattctt gttccagttc ttgaatccaa    1029

gttgttatca tattctgtta ttga                                           1053

```

<210> 14
 <211> 218
 <212> PRT
 <213> 大豆

<400> 14

```

Met Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Val Leu Thr
1              5              10              15

Leu Ala Val Thr Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn
      20              25              30

Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro Tyr
      35              40              45

Gly Pro Ile Tyr Ser Asp Arg Glu Arg Leu Gln Ile Tyr Ile Ser Asp
      50              55              60

Ala Gly Val Leu Ala Val Cys Tyr Gly Leu Phe Cys Leu Ala Met Ala
65              70              75              80

```

Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu Val
85 90 95

Val Asn Gly Phe Leu Val Leu Ile Thr Phe Leu Gln His Thr His Pro
100 105 110

Ala Leu Pro His Tyr Thr Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala
115 120 125

Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His
130 135 140

Asn Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro
145 150 155 160

His Tyr His Ala Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly
165 170 175

Glu Tyr Tyr Arg Phe Asp Gly Thr Pro Phe Val Lys Ala Met Trp Arg
180 185 190

Glu Ala Arg Glu Cys Ile Tyr Val Glu Pro Asp Gln Ser Thr Gln Ser
195 200 205

Lys Gly Val Phe Trp Tyr Asn Asn Lys Leu
210 215

<210> 15
<211> 654
<212> DNA
<213> 大豆

<220>
<221> CDS
<222> (1).. (654)

<400> 15
atg tgg tac tct aaa tac ctt aac aat cca cca ggc aga gtc ctc act 48
Met Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Val Leu Thr
1 5 10 15

ctt gcc gtc acc ctc acg ctt ggt tgg ccc ttg tac ttg gct ttt aat 96
Leu Ala Val Thr Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn
20 25 30

gtt tct gga agg cct tat gat aga ttt gct tgc cac tat gac cct tat	144
Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro Tyr	
35 40 45	
ggt ccc att tac tct gac cga gaa cga ctt caa ata tat ata tca gat	192
Gly Pro Ile Tyr Ser Asp Arg Glu Arg Leu Gln Ile Tyr Ile Ser Asp	
50 55 60	
gca gga gta ctt gca gta tgc tat ggc ctt ttc tgt ctt gcc atg gca	240
Ala Gly Val Leu Ala Val Cys Tyr Gly Leu Phe Cys Leu Ala Met Ala	
65 70 75 80	
aaa ggg ctt gcc tgg gtg gtg tgt gtt tat gga gtt cca ttg ctt gtg	288
Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu Val	
85 90 95	
gtc aat gga ttt ttg gtg ttg att aca ttt ttg cag cac act cac cct	336
Val Asn Gly Phe Leu Val Leu Ile Thr Phe Leu Gln His Thr His Pro	
100 105 110	
gca ttg cca cac tac act tcc tct gag tgg gac tgg ttg aga gga gct	384
Ala Leu Pro His Tyr Thr Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala	
115 120 125	
tta gca aca gtg gat aga gat tat gga atc ctg aac aag gtc ttc cat	432
Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His	
130 135 140	
aat att aca gac act cat gta gct cat cac ttg ttc tcc aca atg cca	480
Asn Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro	
145 150 155 160	
cat tat cat gca atg gag gcg aca aag gca ata aag ccc atc ttg gga	528
His Tyr His Ala Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly	
165 170 175	
gag tat tat cgg ttt gat ggg act cca ttt gtc aag gca atg tgg aga	576
Glu Tyr Tyr Arg Phe Asp Gly Thr Pro Phe Val Lys Ala Met Trp Arg	
180 185 190	
gag gca aga gag tgt att tat gtg gag cca gat caa agt act cag agc	624
Glu Ala Arg Glu Cys Ile Tyr Val Glu Pro Asp Gln Ser Thr Gln Ser	
195 200 205	

aaa ggt gta ttt tgg tac aac aat aag ttg
 Lys Gly Val Phe Trp Tyr Asn Asn Lys Leu
 210 215

654

<210> 16
 <211> 218
 <212> PRT
 <213> 大豆

<400> 16

Met Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Val Leu Thr
 1 5 10 15

Leu Ala Val Thr Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn
 20 25 30

Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Tyr Asp Pro Tyr
 35 40 45

Gly Pro Ile Tyr Ser Asp Arg Glu Arg Leu Gln Ile Tyr Ile Ser Asp
 50 55 60

Ala Gly Val Leu Ala Val Cys Tyr Gly Leu Phe Cys Leu Ala Met Ala
 65 70 75 80

Lys Gly Leu Ala Trp Val Val Cys Val Tyr Gly Val Pro Leu Leu Val
 85 90 95

Val Asn Gly Phe Leu Val Leu Ile Thr Phe Leu Gln His Thr His Pro
 100 105 110

Ala Leu Pro His Tyr Thr Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala
 115 120 125

Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His
 130 135 140

Asn Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro
 145 150 155 160

His Tyr His Ala Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly
 165 170 175

Glu Tyr Tyr Arg Phe Asp Gly Thr Pro Phe Val Lys Ala Met Trp Arg
180 185 190

Glu Ala Arg Glu Cys Ile Tyr Val Glu Pro Asp Gln Ser Thr Gln Ser
195 200 205

Lys Gly Val Phe Trp Tyr Asn Asn Lys Leu
210 215

<210> 17
<211> 1750
<212> DNA
<213> 玉米 (*Zea mays*)

<220>
<221> misc_feature
<222> (1)..(175)

<220>
<221> CDS
<222> (176)..(1351)

<220>
<221> misc_feature
<222> (1352)..(1750)

<400> 17

ctccctcctc cttcctcctc cctgcaaadc ctgcaggcac caccgctcgt tttcctgtcc 60

gggggacagg agagaagggg agagaccgag agagggtgag gcgcggcgtc cgccgatctg 120

ctccgcccc cgaagcagcc tgtcacgtcg tcctcactct cagcaaccag cgaaa atg 178
Met
1

ggt gcc gga ggc agg atg acc gag aag gag cgg gag aag cat gag cag 226
Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Lys His Glu Gln
5 10 15

gag cag gtc gcc cgt gct acc ggc ggt ggc gcg gca gtg cag cgg tcg 274
Glu Gln Val Ala Arg Ala Thr Gly Gly Gly Ala Ala Val Gln Arg Ser
20 25 30

ccg gtg gag aag ccg ccg ttc acg ttg ggg cag atc aag aag gcg atc Pro Val Glu Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile 35 40 45	322
ccg ccg cac tgc ttc gag cgc tcc gtg ctg agg tcc ttc tcg tac gtg Pro Pro His Cys Phe Glu Arg Ser Val Leu Arg Ser Phe Ser Tyr Val 50 55 60 65	370
gcc cac gac ctg tcg ctc gcc gcg gcg ctc ctc tac ctc gcg gtg gcc Ala His Asp Leu Ser Leu Ala Ala Ala Leu Leu Tyr Leu Ala Val Ala 70 75 80	418
gtg ata ccg gcg cta ccc tgc ccg ctc cgc tac gcg gcc tgg ccg ctg Val Ile Pro Ala Leu Pro Cys Pro Leu Arg Tyr Ala Ala Trp Pro Leu 85 90 95	466
tac tgg gtg gcc cag ggg tgc gtg tgc acg ggc gtg tgg gtg atc gcg Tyr Trp Val Ala Gln Gly Cys Val Cys Thr Gly Val Trp Val Ile Ala 100 105 110	514
cac gag tgc ggc cac cac gcc ttc tcc gac cac gcg ctc ctg gac gac His Glu Cys Gly His His Ala Phe Ser Asp His Ala Leu Leu Asp Asp 115 120 125	562
gcc gtc ggc ctg gcg ctg cac tcg gcg ctg ctg gtg ccc tac ttc tcg Ala Val Gly Leu Ala Leu His Ser Ala Leu Leu Val Pro Tyr Phe Ser 130 135 140 145	610
tgg aag tac agc cac cgg cgc cac cac tcc aac acg ggg tcc ctg gag Trp Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu 150 155 160	658
cgc gac gag gtg ttc gtg ccg agg acc aag gag gcg ctg ccg tgg tac Arg Asp Glu Val Phe Val Pro Arg Thr Lys Glu Ala Leu Pro Trp Tyr 165 170 175	706
gcc ccg tac gtg cac ggc agc ccc gcg ggc cgg ctg gcg cac gtc gcc Ala Pro Tyr Val His Gly Ser Pro Ala Gly Arg Leu Ala His Val Ala 180 185 190	754
gtg cag ctc acc ctc ggc tgg ccg ctg tac ctg gcc acc aac gcg tcg Val Gln Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Thr Asn Ala Ser 195 200 205	802

ggg cgg ccg tac ccg cgc ttc gcc tgc cac ttc gac ccc tac ggc ccc	850
Gly Arg Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro	
210 215 220 225	
atc tac aac gac cgg gag cgc gcc cag atc ttc gtc tcg gac gcc ggc	898
Ile Tyr Asn Asp Arg Glu Arg Ala Gln Ile Phe Val Ser Asp Ala Gly	
230 235 240	
gtc gtg gcc gtg gcg ttc ggg ctg tac aag ctg gcg gcg gcg ttc ggg	946
Val Val Ala Val Ala Phe Gly Leu Tyr Lys Leu Ala Ala Ala Phe Gly	
245 250 255	
gtc tgg tgg gtg gtg cgc gtg tac gcc gtg ccg ctg ctg atc gtc aac	994
Val Trp Trp Val Val Arg Val Tyr Ala Val Pro Leu Leu Ile Val Asn	
260 265 270	
gcg tgg ctg gtg ctc atc acg tac ctg cag cac acc cac ccg gcg ctg	1042
Ala Trp Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala Leu	
275 280 285	
ccc cac tac gac tcg ggc gag tgg gac tgg ctg cgc ggc gcg ctc gcc	1090
Pro His Tyr Asp Ser Gly Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala	
290 295 300 305	
acc gtc gac cga gac tac ggc gtc ctc aac cgc gtg ttc cac cac atc	1138
Thr Val Asp Arg Asp Tyr Gly Val Leu Asn Arg Val Phe His His Ile	
310 315 320	
acg gac acg cac gtc gcg cac cac ctc ttc tcc acc atg ccg cac tac	1186
Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr	
325 330 335	
cac gcc gtg gag gcc acc agg gcg atc agg ccc gtc ctc ggc gag tac	1234
His Ala Val Glu Ala Thr Arg Ala Ile Arg Pro Val Leu Gly Glu Tyr	
340 345 350	
tac cag ttc gac ccg acc cct gtc gcc aag gcc acc tgg cgc gag gcc	1282
Tyr Gln Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala	
355 360 365	
agg gag tgc atc tac gtc gag cct gag aac cgc aac cgc aag ggc gtc	1330
Arg Glu Cys Ile Tyr Val Glu Pro Glu Asn Arg Asn Arg Lys Gly Val	
370 375 380 385	

```

ttc tgg tac aac agc aag ttc tagccgccgc ttgctttttc cctaggaatg      1381
Phe Trp Tyr Asn Ser Lys Phe
                               390

ggaggagaaa tcaggatgag aagatggtaa tgtctgcac tacctgtcta atggtagtc    1441

accagtcttt agacaggaag agagcatttg ggcttcagaa aaggaggctt actgcactac    1501

tgcagtgcc a tcgctagatt taaggcaa at tcagtgtgtc tgtgcccatg gctgtgagct    1561

ttgggtactc tcaagtagtc aagttctctt gtttttgttt ttagtcgtcg ctgtttagg    1621

cttgccggcg gcggtcgttg cgtggccgcg ccttgtcgtg tgcgtcttgc catctcttcg    1681

tgctcccttc gtgttgttgt aaaacactag tctggtgtct ttggcggaat aacagatcgt    1741

cgaacgaca                                                                1750

```

```

<210> 18
<211> 392
<212> PRT
<213> 玉米

```

```

<400> 18
Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Lys His Glu
1           5           10          15

Gln Glu Gln Val Ala Arg Ala Thr Gly Gly Gly Ala Ala Val Gln Arg
          20          25          30

Ser Pro Val Glu Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala
          35          40          45

Ile Pro Pro His Cys Phe Glu Arg Ser Val Leu Arg Ser Phe Ser Tyr
          50          55          60

Val Ala His Asp Leu Ser Leu Ala Ala Ala Leu Leu Tyr Leu Ala Val
65          70          75          80

Ala Val Ile Pro Ala Leu Pro Cys Pro Leu Arg Tyr Ala Ala Trp Pro
          85          90          95

Leu Tyr Trp Val Ala Gln Gly Cys Val Cys Thr Gly Val Trp Val Ile
100         105         110

```

Ala His Glu Cys Gly His His Ala Phe Ser Asp His Ala Leu Leu Asp
115 120 125

Asp Ala Val Gly Leu Ala Leu His Ser Ala Leu Leu Val Pro Tyr Phe
130 135 140

Ser Trp Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu
145 150 155 160

Glu Arg Asp Glu Val Phe Val Pro Arg Thr Lys Glu Ala Leu Pro Trp
165 170 175

Tyr Ala Pro Tyr Val His Gly Ser Pro Ala Gly Arg Leu Ala His Val
180 185 190

Ala Val Gln Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Thr Asn Ala
195 200 205

Ser Gly Arg Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly
210 215 220

Pro Ile Tyr Asn Asp Arg Glu Arg Ala Gln Ile Phe Val Ser Asp Ala
225 230 235 240

Gly Val Val Ala Val Ala Phe Gly Leu Tyr Lys Leu Ala Ala Ala Phe
245 250 255

Gly Val Trp Trp Val Val Arg Val Tyr Ala Val Pro Leu Leu Ile Val
260 265 270

Asn Ala Trp Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala
275 280 285

Leu Pro His Tyr Asp Ser Gly Glu Trp Asp Trp Leu Arg Gly Ala Leu
290 295 300

Ala Thr Val Asp Arg Asp Tyr Gly Val Leu Asn Arg Val Phe His His
305 310 315 320

Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His
325 330 335

Tyr His Ala Val Glu Ala Thr Arg Ala Ile Arg Pro Val Leu Gly Glu
340 345 350

Tyr Tyr Gln Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu
355 360 365

Ala Arg Glu Cys Ile Tyr Val Glu Pro Glu Asn Arg Asn Arg Lys Gly
370 375 380

Val Phe Trp Tyr Asn Ser Lys Phe
385 390

<210> 19

<211> 1176

<212> DNA

<213> 玉米

<220>

<221> CDS

<222> (1)..(1176)

<400> 19

atg ggt gcc gga ggc agg atg acc gag aag gag cgg gag aag cat gag 48
Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Lys His Glu
1 5 10 15

cag gag cag gtc gcc cgt gct acc ggc ggt ggc gcg gca gtg cag cgg 96
Gln Glu Gln Val Ala Arg Ala Thr Gly Gly Gly Ala Ala Val Gln Arg
20 25 30

tcg ccg gtg gag aag ccg ccg ttc acg ttg ggg cag atc aag aag gcg 144
Ser Pro Val Glu Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala
35 40 45

atc ccg ccg cac tgc ttc gag cgc tcc gtg ctg agg tcc ttc tcg tac 192
Ile Pro Pro His Cys Phe Glu Arg Ser Val Leu Arg Ser Phe Ser Tyr
50 55 60

gtg gcc cac gac ctg tcg ctc gcc gcg gcg ctc ctc tac ctc gcg gtg 240
Val Ala His Asp Leu Ser Leu Ala Ala Ala Leu Leu Tyr Leu Ala Val
65 70 75 80

gcc gtg ata ccg gcg cta ccc tgc ccg ctc cgc tac gcg gcc tgg ccg 288
Ala Val Ile Pro Ala Leu Pro Cys Pro Leu Arg Tyr Ala Ala Trp Pro
85 90 95

ctg tac tgg gtg gcc cag ggg tgc gtg tgc acg ggc gtg tgg gtg atc Leu Tyr Trp Val Ala Gln Gly Cys Val Cys Thr Gly Val Trp Val Ile 100 105 110	336
gcg cac gag tgc ggc cac cac gcc ttc tcc gac cac gcg ctc ctg gac Ala His Glu Cys Gly His His Ala Phe Ser Asp His Ala Leu Leu Asp 115 120 125	384
gac gcc gtc ggc ctg gcg ctg cac tcg gcg ctg ctg gtg ccc tac ttc Asp Ala Val Gly Leu Ala Leu His Ser Ala Leu Leu Val Pro Tyr Phe 130 135 140	432
tcg tgg aag tac agc cac cgg cgc cac cac tcc aac acg ggg tcc ctg Ser Trp Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu 145 150 155 160	480
gag cgc gac gag gtg ttc gtg ccg agg acc aag gag gcg ctg ccg tgg Glu Arg Asp Glu Val Phe Val Pro Arg Thr Lys Glu Ala Leu Pro Trp 165 170 175	528
tac gcc ccg tac gtg cac ggc agc ccc gcg ggc cgg ctg gcg cac gtc Tyr Ala Pro Tyr Val His Gly Ser Pro Ala Gly Arg Leu Ala His Val 180 185 190	576
gcc gtg cag ctc acc ctc ggc tgg ccg ctg tac ctg gcc acc aac gcg Ala Val Gln Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Thr Asn Ala 195 200 205	624
tcg ggg cgg ccg tac ccg cgc ttc gcc tgc cac ttc gac ccc tac ggc Ser Gly Arg Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly 210 215 220	672
ccc atc tac aac gac cgg gag cgc gcc cag atc ttc gtc tcg gac gcc Pro Ile Tyr Asn Asp Arg Glu Arg Ala Gln Ile Phe Val Ser Asp Ala 225 230 235 240	720
ggc gtc gtg gcc gtg gcg ttc ggg ctg tac aag ctg gcg gcg gcg ttc Gly Val Val Ala Val Ala Phe Gly Leu Tyr Lys Leu Ala Ala Ala Phe 245 250 255	768
ggg gtc tgg tgg gtg gtg cgc gtg tac gcc gtg ccg ctg ctg atc gtc Gly Val Trp Trp Val Val Arg Val Tyr Ala Val Pro Leu Leu Ile Val 260 265 270	816

aac gcg tgg ctg gtg ctc atc acg tac ctg cag cac acc cac ccg gcg Asn Ala Trp Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala 275 280 285	864
ctg ccc cac tac gac tcg ggc gag tgg gac tgg ctg cgc ggc gcg ctc Leu Pro His Tyr Asp Ser Gly Glu Trp Asp Trp Leu Arg Gly Ala Leu 290 295 300	912
gcc acc gtc gac cga gac tac ggc gtc ctc aac cgc gtg ttc cac cac Ala Thr Val Asp Arg Asp Tyr Gly Val Leu Asn Arg Val Phe His His 305 310 315 320	960
atc acg gac acg cac gtc gcg cac cac ctc ttc tcc acc atg ccg cac Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His 325 330 335	1008
tac cac gcc gtg gag gcc acc agg gcg atc agg ccc gtc ctc ggc gag Tyr His Ala Val Glu Ala Thr Arg Ala Ile Arg Pro Val Leu Gly Glu 340 345 350	1056
tac tac cag ttc gac ccg acc cct gtc gcc aag gcc acc tgg cgc gag Tyr Tyr Gln Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu 355 360 365	1104
gcc agg gag tgc atc tac gtc gag cct gag aac cgc aac cgc aag ggc Ala Arg Glu Cys Ile Tyr Val Glu Pro Glu Asn Arg Asn Arg Lys Gly 370 375 380	1152
gtc ttc tgg tac aac agc aag ttc Val Phe Trp Tyr Asn Ser Lys Phe 385 390	1176

<210> 20

<211> 392

<212> PRT

<213> 玉米

<400> 20

Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Lys His Glu
1 5 10 15

Gln Glu Gln Val Ala Arg Ala Thr Gly Gly Gly Ala Ala Val Gln Arg
20 25 30

Ser Pro Val Glu Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala
 35 40 45

Ile Pro Pro His Cys Phe Glu Arg Ser Val Leu Arg Ser Phe Ser Tyr
 50 55 60

Val Ala His Asp Leu Ser Leu Ala Ala Ala Leu Leu Tyr Leu Ala Val
 65 70 75 80

Ala Val Ile Pro Ala Leu Pro Cys Pro Leu Arg Tyr Ala Ala Trp Pro
 85 90 95

Leu Tyr Trp Val Ala Gln Gly Cys Val Cys Thr Gly Val Trp Val Ile
 100 105 110

Ala His Glu Cys Gly His His Ala Phe Ser Asp His Ala Leu Leu Asp
 115 120 125

Asp Ala Val Gly Leu Ala Leu His Ser Ala Leu Leu Val Pro Tyr Phe
 130 135 140

Ser Trp Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu
 145 150 155 160

Glu Arg Asp Glu Val Phe Val Pro Arg Thr Lys Glu Ala Leu Pro Trp
 165 170 175

Tyr Ala Pro Tyr Val His Gly Ser Pro Ala Gly Arg Leu Ala His Val
 180 185 190

Ala Val Gln Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Thr Asn Ala
 195 200 205

Ser Gly Arg Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly
 210 215 220

Pro Ile Tyr Asn Asp Arg Glu Arg Ala Gln Ile Phe Val Ser Asp Ala
 225 230 235 240

Gly Val Val Ala Val Ala Phe Gly Leu Tyr Lys Leu Ala Ala Ala Phe
 245 250 255

Gly Val Trp Trp Val Val Arg Val Tyr Ala Val Pro Leu Leu Ile Val
 260 265 270

Asn Ala Trp Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala
275 280 285

Leu Pro His Tyr Asp Ser Gly Glu Trp Asp Trp Leu Arg Gly Ala Leu
290 295 300

Ala Thr Val Asp Arg Asp Tyr Gly Val Leu Asn Arg Val Phe His His
305 310 315 320

Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His
325 330 335

Tyr His Ala Val Glu Ala Thr Arg Ala Ile Arg Pro Val Leu Gly Glu
340 345 350

Tyr Tyr Gln Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu
355 360 365

Ala Arg Glu Cys Ile Tyr Val Glu Pro Glu Asn Arg Asn Arg Lys Gly
370 375 380

Val Phe Trp Tyr Asn Ser Lys Phe
385 390

<210> 21
<211> 1863
<212> DNA
<213> 稻 (*Oryza sativa*)

<220>
<221> misc_feature
<222> (1)..(149)

<220>
<221> CDS
<222> (150)..(1313)

<220>
<221> misc_feature
<222> (1314)..(1863)

<400> 21

ctcctctcct cctccccctgc acagaccact cgtttctctcc acaaagaggg agggaacaag	60
ggaaggggtgt cgcccgcccc ccaccccgat ctgcctccgc cgctccgctc ctccgcgcct	120
gcgaaatcta ccaacgctaa ctcagcaag atg ggt gcc ggc ggc agg atg acg	173
Met Gly Ala Gly Gly Arg Met Thr	
1 5	
gag aag gag cgg gag gag cag cag aag ctg ctc ggc cgc gcc ggc aat	221
Glu Lys Glu Arg Glu Glu Gln Gln Lys Leu Leu Gly Arg Ala Gly Asn	
10 15 20	
ggc gcg gcc gtg cag cgg tcg ccg acg gac aag ccg ccg ttc acg ctg	269
Gly Ala Ala Val Gln Arg Ser Pro Thr Asp Lys Pro Pro Phe Thr Leu	
25 30 35 40	
ggg cag atc aag aag gcc atc ccg cct cac tgc ttc cag cgc tcg gtg	317
Gly Gln Ile Lys Lys Ala Ile Pro Pro His Cys Phe Gln Arg Ser Val	
45 50 55	
atc aag tcc ttc tcc tac gtg gtc cat gac ctc gtg atc gtc gcc gcg	365
Ile Lys Ser Phe Ser Tyr Val Val His Asp Leu Val Ile Val Ala Ala	
60 65 70	
ctg ctc tac ttc gcg ctg gtc atg atc ccc gtg ctg ccg agc ggg atg	413
Leu Leu Tyr Phe Ala Leu Val Met Ile Pro Val Leu Pro Ser Gly Met	
75 80 85	
gag ttc gcg gca tgg ccg ctc tac tgg atc gcg cag ggc tgc gtg ctc	461
Glu Phe Ala Ala Trp Pro Leu Tyr Trp Ile Ala Gln Gly Cys Val Leu	
90 95 100	
acc ggc gtg tgg gtc atc gcg cac gag tgc ggc cac cat gcc ttc tcc	509
Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His Ala Phe Ser	
105 110 115 120	
gac tac tcg gtg ctc gac gac atc gtc ggc ctc gtg ctg cac tcg tcg	557
Asp Tyr Ser Val Leu Asp Asp Ile Val Gly Leu Val Leu His Ser Ser	
125 130 135	
ctg ctc gtc ccc tac ttc tcg tgg aag tac agc cac cgg cgc cac cac	605
Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His His	
140 145 150	

tcc aac acc ggg tcg ctg gag cgc gac gag gtg ttc gtc ccg aag cag Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys Gln 155 160 165	653
aag tcg gcg atg gcg tgg tac acc ccg tac gtg tac cac aac ccg atc Lys Ser Ala Met Ala Trp Tyr Thr Pro Tyr Val Tyr His Asn Pro Ile 170 175 180	701
ggc cgg ctg gtg cac atc ttc gtg cag ctc acc ctc ggg tgg ccg ctg Gly Arg Leu Val His Ile Phe Val Gln Leu Thr Leu Gly Trp Pro Leu 185 190 195 200	749
tac ctg gcg ttc aac gtg tcc ggc cgc ccg tac ccg cgc ttc gcg tgc Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Pro Arg Phe Ala Cys 205 210 215	797
cac ttc gac ccc tac ggc ccg atc tac aac gac cgg gag cgc gtc cag His Phe Asp Pro Tyr Gly Pro Ile Tyr Asn Asp Arg Glu Arg Val Gln 220 225 230	845
atc ttc atc tcc gac gtc ggc gtc gtg tcc gcg ggg ctc gcc ctg ttc Ile Phe Ile Ser Asp Val Gly Val Val Ser Ala Gly Leu Ala Leu Phe 235 240 245	893
aag ctg tcg tcg gcg ttc ggg ttc tgg tgg gtg gtg cgc gtc tac ggc Lys Leu Ser Ser Ala Phe Gly Phe Trp Trp Val Val Arg Val Tyr Gly 250 255 260	941
gtg ccg ctg ctg atc gtg aac gcg tgg ctg gtg ctc atc acc tac ctg Val Pro Leu Leu Ile Val Asn Ala Trp Leu Val Leu Ile Thr Tyr Leu 265 270 275 280	989
cag cac acc cac ccg gcg ctg ccg cac tac gac tcg agc gag tgg gac Gln His Thr His Pro Ala Leu Pro His Tyr Asp Ser Ser Glu Trp Asp 285 290 295	1037
tgg ctc cgc ggc gcg ctg gcc acc gtg gac cgc gac tac ggc atc ctc Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu 300 305 310	1085
aac aag gtg ttc cac aac atc acg gac acg cac gtc gcg cac cac ctc Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu 315 320 325	1133

ttc tcc acc atg ccg cac tac cac gcc atg gag gcc act aag gcg atc 1181
 Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr Lys Ala Ile
 330 335 340

cgc ccc atc ctc ggc gag tac tac cag ttc gac ccg acg ccc gtc gcc 1229
 Arg Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Pro Thr Pro Val Ala
 345 350 355 360

aag gcg aca tgg cgc gag gcc aag gag tgc atc tac gtc gag cct gag 1277
 Lys Ala Thr Trp Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Glu
 365 370 375

gac aac aag ggc gtc ttc tgg tac aac aac aag ttc taactgctgc 1323
 Asp Asn Lys Gly Val Phe Trp Tyr Asn Asn Lys Phe
 380 385

tgctgtgaaa tcagcatcac acatccatag ccaagcagca aacaaatttg aagaagaaga 1383

ttacaaggga agagaagata gtgtcttcgg aaatcgtcgt agcaagtatc catccatcca 1443

tccaacccat gaacaatcgt ctatctatcc atgcatctat ctatggttag tctctttaga 1503

taggagaggg cacttgggca cagaggaagg ctattgcagt gccattgcta gagttgccat 1563

caagtgcaaa gtaggcggat caggcgtgtg tcatgcctgt ggatttgtag tctatgtatg 1623

tgtcagctgc tgagctccgg tgtcgcagcc ttggtccctg tcgtgttatt tccatcgttt 1683

ttttccctcc gccattgttc ggtttaggtg ttgtcatggt cggcgtccgt gtggacgacg 1743

tgtcttgtgc ggtttgtctg tcattgagtt ggctccgtcc gttgcttget gttgtaaaac 1803

gcttgtggtg ttcatggcgg aataactaaa cgtcgaatgg aatgacaact tttttgcgta 1863

<210> 22

<211> 388

<212> PRT

<213> 稻

<400> 22

Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Glu Gln Gln
 1 5 10 15

Lys Leu Leu Gly Arg Ala Gly Asn Gly Ala Ala Val Gln Arg Ser Pro
 20 25 30

Thr Asp Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro
 35 40 45

Pro His Cys Phe Gln Arg Ser Val Ile Lys Ser Phe Ser Tyr Val Val
 50 55 60

His Asp Leu Val Ile Val Ala Ala Leu Leu Tyr Phe Ala Leu Val Met
 65 70 75 80

Ile Pro Val Leu Pro Ser Gly Met Glu Phe Ala Ala Trp Pro Leu Tyr
 85 90 95

Trp Ile Ala Gln Gly Cys Val Leu Thr Gly Val Trp Val Ile Ala His
 100 105 110

Glu Cys Gly His His Ala Phe Ser Asp Tyr Ser Val Leu Asp Asp Ile
 115 120 125

Val Gly Leu Val Leu His Ser Ser Leu Leu Val Pro Tyr Phe Ser Trp
 130 135 140

Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg
 145 150 155 160

Asp Glu Val Phe Val Pro Lys Gln Lys Ser Ala Met Ala Trp Tyr Thr
 165 170 175

Pro Tyr Val Tyr His Asn Pro Ile Gly Arg Leu Val His Ile Phe Val
 180 185 190

Gln Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly
 195 200 205

Arg Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile
 210 215 220

Tyr Asn Asp Arg Glu Arg Val Gln Ile Phe Ile Ser Asp Val Gly Val
 225 230 235 240

Val Ser Ala Gly Leu Ala Leu Phe Lys Leu Ser Ser Ala Phe Gly Phe
 245 250 255

Trp Trp Val Val Arg Val Tyr Gly Val Pro Leu Leu Ile Val Asn Ala
260 265 270

Trp Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro
275 280 285

His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr
290 295 300

Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His Asn Ile Thr
305 310 315 320

Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His
325 330 335

Ala Met Glu Ala Thr Lys Ala Ile Arg Pro Ile Leu Gly Glu Tyr Tyr
340 345 350

Gln Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Lys
355 360 365

Glu Cys Ile Tyr Val Glu Pro Glu Asp Asn Lys Gly Val Phe Trp Tyr
370 375 380

Asn Asn Lys Phe
385

<210> 23

<211> 1164

<212> DNA

<213> 稻

<220>

<221> CDS

<222> (1)..(1164)

<400> 23

atg ggt gcc ggc ggc agg atg acg gag aag gag cgg gag gag cag cag 48
Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Glu Gln Gln
1 5 10 15

aag ctg ctc ggc cgc gcc ggc aat ggc gcg gcc gtg cag cgg tcg ccg 96
Lys Leu Leu Gly Arg Ala Gly Asn Gly Ala Ala Val Gln Arg Ser Pro
20 25 30

acg gac aag ccg ccg ttc acg ctg ggg cag atc aag aag gcc atc ccg Thr Asp Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro 35 40 45	144
cct cac tgc ttc cag cgc tcg gtg atc aag tcc ttc tcc tac gtg gtc Pro His Cys Phe Gln Arg Ser Val Ile Lys Ser Phe Ser Tyr Val Val 50 55 60	192
cat gac ctc gtg atc gtc gcc gcg ctg ctc tac ttc gcg ctg gtc atg His Asp Leu Val Ile Val Ala Ala Leu Leu Tyr Phe Ala Leu Val Met 65 70 75 80	240
atc ccc gtg ctg ccg agc ggg atg gag ttc gcg gca tgg ccg ctc tac Ile Pro Val Leu Pro Ser Gly Met Glu Phe Ala Ala Trp Pro Leu Tyr 85 90 95	288
tgg atc gcg cag ggc tgc gtg ctc acc ggc gtg tgg gtc atc gcg cac Trp Ile Ala Gln Gly Cys Val Leu Thr Gly Val Trp Val Ile Ala His 100 105 110	336
gag tgc ggc cac cat gcc ttc tcc gac tac tcg gtg ctc gac gac atc Glu Cys Gly His His Ala Phe Ser Asp Tyr Ser Val Leu Asp Asp Ile 115 120 125	384
gtc ggc ctc gtg ctg cac tcg tcg ctg ctc gtc ccc tac ttc tcg tgg Val Gly Leu Val Leu His Ser Ser Leu Leu Val Pro Tyr Phe Ser Trp 130 135 140	432
aag tac agc cac cgg cgc cac cac tcc aac acc ggg tcg ctg gag cgc Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg 145 150 155 160	480
gac gag gtg ttc gtc ccg aag cag aag tcg gcg atg gcg tgg tac acc Asp Glu Val Phe Val Pro Lys Gln Lys Ser Ala Met Ala Trp Tyr Thr 165 170 175	528
ccg tac gtg tac cac aac ccg atc ggc cgg ctg gtg cac atc ttc gtg Pro Tyr Val Tyr His Asn Pro Ile Gly Arg Leu Val His Ile Phe Val 180 185 190	576
cag ctc acc ctc ggg tgg ccg ctg tac ctg gcg ttc aac gtg tcc ggc Gln Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly 195 200 205	624

cgc ccg tac ccg cgc ttc gcg tgc cac ttc gac ccc tac ggc ccg atc Arg Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile 210 215 220	672
tac aac gac cgg gag cgc gtc cag atc ttc atc tcc gac gtc ggc gtc Tyr Asn Asp Arg Glu Arg Val Gln Ile Phe Ile Ser Asp Val Gly Val 225 230 235 240	720
gtg tcc gcg ggg ctc gcc ctg ttc aag ctg tcg tcg gcg ttc ggg ttc Val Ser Ala Gly Leu Ala Leu Phe Lys Leu Ser Ser Ala Phe Gly Phe 245 250 255	768
tgg tgg gtg gtg cgc gtc tac ggc gtg ccg ctg ctg atc gtg aac gcg Trp Trp Val Val Arg Val Tyr Gly Val Pro Leu Leu Ile Val Asn Ala 260 265 270	816
tgg ctg gtg ctc atc acc tac ctg cag cac acc cac ccg gcg ctg ccg Trp Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro 275 280 285	864
cac tac gac tcg agc gag tgg gac tgg ctc cgc ggc gcg ctg gcc acc His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr 290 295 300	912
gtg gac cgc gac tac ggc atc ctc aac aag gtg ttc cac aac atc acg Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His Asn Ile Thr 305 310 315 320	960
gac acg cac gtc gcg cac cac ctc ttc tcc acc atg ccg cac tac cac Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His 325 330 335	1008
gcc atg gag gcc act aag gcg atc cgc ccc atc ctc ggc gag tac tac Ala Met Glu Ala Thr Lys Ala Ile Arg Pro Ile Leu Gly Glu Tyr Tyr 340 345 350	1056
cag ttc gac ccg acg ccc gtc gcc aag gcg aca tgg cgc gag gcc aag Gln Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Lys 355 360 365	1104
gag tgc atc tac gtc gag cct gag gac aac aag ggc gtc ttc tgg tac Glu Cys Ile Tyr Val Glu Pro Glu Asp Asn Lys Gly Val Phe Trp Tyr 370 375 380	1152

aac aac aag ttc
Asn Asn Lys Phe
385

1164

<210> 24
<211> 388
<212> PRT
<213> 稻

<400> 24

Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Glu Gln Gln
1 5 10 15

Lys Leu Leu Gly Arg Ala Gly Asn Gly Ala Ala Val Gln Arg Ser Pro
20 25 30

Thr Asp Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro
35 40 45

Pro His Cys Phe Gln Arg Ser Val Ile Lys Ser Phe Ser Tyr Val Val
50 55 60

His Asp Leu Val Ile Val Ala Ala Leu Leu Tyr Phe Ala Leu Val Met
65 70 75 80

Ile Pro Val Leu Pro Ser Gly Met Glu Phe Ala Ala Trp Pro Leu Tyr
85 90 95

Trp Ile Ala Gln Gly Cys Val Leu Thr Gly Val Trp Val Ile Ala His
100 105 110

Glu Cys Gly His His Ala Phe Ser Asp Tyr Ser Val Leu Asp Asp Ile
115 120 125

Val Gly Leu Val Leu His Ser Ser Leu Leu Val Pro Tyr Phe Ser Trp
130 135 140

Lys Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg
145 150 155 160

Asp Glu Val Phe Val Pro Lys Gln Lys Ser Ala Met Ala Trp Tyr Thr
165 170 175

Pro Tyr Val Tyr His Asn Pro Ile Gly Arg Leu Val His Ile Phe Val
 180 185 190

Gln Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly
 195 200 205

Arg Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile
 210 215 220

Tyr Asn Asp Arg Glu Arg Val Gln Ile Phe Ile Ser Asp Val Gly Val
 225 230 235 240

Val Ser Ala Gly Leu Ala Leu Phe Lys Leu Ser Ser Ala Phe Gly Phe
 245 250 255

Trp Trp Val Val Arg Val Tyr Gly Val Pro Leu Leu Ile Val Asn Ala
 260 265 270

Trp Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro
 275 280 285

His Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr
 290 295 300

Val Asp Arg Asp Tyr Gly Ile Leu Asn Lys Val Phe His Asn Ile Thr
 305 310 315 320

Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His
 325 330 335

Ala Met Glu Ala Thr Lys Ala Ile Arg Pro Ile Leu Gly Glu Tyr Tyr
 340 345 350

Gln Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Lys
 355 360 365

Glu Cys Ile Tyr Val Glu Pro Glu Asp Asn Lys Gly Val Phe Trp Tyr
 370 375 380

Asn Asn Lys Phe
 385

<210> 25

<211> 1519

<212>	DNA	
<213>	亚麻 (Linum usitatissimum)	
<220>		
<221>	misc_feature	
<222>	(1)..(47)	
<220>		
<221>	CDS	
<222>	(48)..(1070)	
<220>		
<221>	misc_feature	
<222>	(1071)..(1519)	
<400>	25	
gctgtaacaa tatacacagg aagaagaaaa atgggtgccg gcgcaga atg tca gtg		56
Met Ser Val		
1		
cct cca tca tcc aaa cct atg aag agg tct cct tac tca aag cca cca		104
Pro Pro Ser Ser Lys Pro Met Lys Arg Ser Pro Tyr Ser Lys Pro Pro		
5 10 15		
ttc acg ctc ggt gag ctc aag aag gcc att cct cca cac tgt ttc aaa		152
Phe Thr Leu Gly Glu Leu Lys Lys Ala Ile Pro Pro His Cys Phe Lys		
20 25 30 35		
cgc tca atc ccc cga tcg ttc gcc tac gtg gcg tac gac ctc acc att		200
Arg Ser Ile Pro Arg Ser Phe Ala Tyr Val Ala Tyr Asp Leu Thr Ile		
40 45 50		
gca gca atc ttc tac tac atc gcc acc act tac ttc cac ctc ctc cct		248
Ala Ala Ile Phe Tyr Tyr Ile Ala Thr Thr Tyr Phe His Leu Leu Pro		
55 60 65		
agc cct ctc aac tac ctc gcc tgg ccg gtc tac tgg gcc tgc cag ggc		296
Ser Pro Leu Asn Tyr Leu Ala Trp Pro Val Tyr Trp Ala Cys Gln Gly		
70 75 80		
tgc atc ctc act gga gta tgg gtg ttg gct cac gaa tgc ggt cac cat		344
Cys Ile Leu Thr Gly Val Trp Val Leu Ala His Glu Cys Gly His His		
85 90 95		

gcc ttc agc gac tac cag tgg ctc gac gac atg gtt ggc ttc gtc ctc Ala Phe Ser Asp Tyr Gln Trp Leu Asp Asp Met Val Gly Phe Val Leu 100 105 110 115	392
cat tcg tcc ctc ctt gtt cct tac ttc tcc tgg aag cac agc cac cgc His Ser Ser Leu Leu Val Pro Tyr Phe Ser Trp Lys His Ser His Arg 120 125 130	440
cgc cac cat tcc aac acg gga tcg ctt gat cgt gat gag gtg ttt gtc Arg His His Ser Asn Thr Gly Ser Leu Asp Arg Asp Glu Val Phe Val 135 140 145	488
ccc aag cag aag gcc gaa atc ggg tgg tac tcc aag tac ctt aac aac Pro Lys Gln Lys Ala Glu Ile Gly Trp Tyr Ser Lys Tyr Leu Asn Asn 150 155 160	536
cca cct ggc cgt gtg atc aca ttg gcc gtc aca tta acg ctc ggt tgg Pro Pro Gly Arg Val Ile Thr Leu Ala Val Thr Leu Thr Leu Gly Trp 165 170 175	584
cct ctg tac ttg gca ttc aac gtc tcc ggg aga cca tat gac cgg ttc Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr Asp Arg Phe 180 185 190 195	632
gca tgc cat ttt gac cct cac ggt ccg att tac aat gat cgc gag cgt Ala Cys His Phe Asp Pro His Gly Pro Ile Tyr Asn Asp Arg Glu Arg 200 205 210	680
atg gag ata tac cta tcc gac gca ggg ata ttc acc gtg tgc tac atc Met Glu Ile Tyr Leu Ser Asp Ala Gly Ile Phe Thr Val Cys Tyr Ile 215 220 225	728
cta tac aga ctc gtc ctc acg aaa gga ctc gtt tgg gtc gtg tcc ata Leu Tyr Arg Leu Val Leu Thr Lys Gly Leu Val Trp Val Val Ser Ile 230 235 240	776
tac gga gtc cca cta ttg ata gtg aat gga ttc cta gtc ctc atc act Tyr Gly Val Pro Leu Leu Ile Val Asn Gly Phe Leu Val Leu Ile Thr 245 250 255	824
ttc ttg cag cac acg cat cct tct ctt ccg cac tac aaa gtc ctc cga Phe Leu Gln His Thr His Pro Ser Leu Pro His Tyr Lys Val Leu Arg 260 265 270 275	872

atg ggg act gga tgc gag gcg ccc tct cga ccg tgg atc gag act acg Met Gly Thr Gly Cys Glu Ala Pro Ser Arg Pro Trp Ile Glu Thr Thr 280 285 290	920
ggt tac tca aca ccg tgt tcc aca aca tca ccg ata cac atg tcg cgc Gly Tyr Ser Thr Pro Cys Ser Thr Thr Ser Pro Ile His Met Ser Arg 295 300 305	968
acc atc tct tct cca cga tgc ctc att acc acg cga tgg agg cta cca Thr Ile Ser Ser Pro Arg Cys Leu Ile Thr Thr Arg Trp Arg Leu Pro 310 315 320	1016
agg cga tca agc cgg ttc tcg ggg agt att acc agt tcg atg gga ctc Arg Arg Ser Ser Arg Phe Ser Gly Ser Ile Thr Ser Ser Met Gly Leu 325 330 335	1064
cct ttg tgaaggccat gtggaggag gcaaaggagt gcatctatgt cgagccggat Pro Leu 340	1120
gaaggcgacc ccagccaagg cgtgttctgg tacaacaaca agctgtgagg gtcttcgaaa	1180
tttgcagagg tttgtagtgt ttgttcttaa tgggtgttacc agaaaaatgt ttgaagaaag	1240
aagctgcaat agctagtgca gaactgggtgt atgtttctgt aatgtttgtt aagttatgtc	1300
cctagtggtc gttaatgtta ctgtacttct ctgttcttct ccatcgagcc aacatacctt	1360
cactcctctg ttaatgtact gagttggtcg agtttaaact taacggacca ccaggctcaa	1420
attcgagtca ccgggttggc cgagtttaga ctgcattgac cacaatgatg caatcgcaaa	1480
actgaagtga ctacaatcgc aaaacttaat tcccagtca	1519
<210> 26	
<211> 341	
<212> PRT	
<213> 亚麻	
<400> 26	

Met Ser Val Pro Pro Ser Ser Lys Pro Met Lys Arg Ser Pro Tyr Ser
1 5 10 15

Lys Pro Pro Phe Thr Leu Gly Glu Leu Lys Lys Ala Ile Pro Pro His
20 25 30

Cys Phe Lys Arg Ser Ile Pro Arg Ser Phe Ala Tyr Val Ala Tyr Asp
35 40 45

Leu Thr Ile Ala Ala Ile Phe Tyr Tyr Ile Ala Thr Thr Tyr Phe His
50 55 60

Leu Leu Pro Ser Pro Leu Asn Tyr Leu Ala Trp Pro Val Tyr Trp Ala
65 70 75 80

Cys Gln Gly Cys Ile Leu Thr Gly Val Trp Val Leu Ala His Glu Cys
85 90 95

Gly His His Ala Phe Ser Asp Tyr Gln Trp Leu Asp Asp Met Val Gly
100 105 110

Phe Val Leu His Ser Ser Leu Leu Val Pro Tyr Phe Ser Trp Lys His
115 120 125

Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Asp Arg Asp Glu
130 135 140

Val Phe Val Pro Lys Gln Lys Ala Glu Ile Gly Trp Tyr Ser Lys Tyr
145 150 155 160

Leu Asn Asn Pro Pro Gly Arg Val Ile Thr Leu Ala Val Thr Leu Thr
165 170 175

Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr
180 185 190

Asp Arg Phe Ala Cys His Phe Asp Pro His Gly Pro Ile Tyr Asn Asp
195 200 205

Arg Glu Arg Met Glu Ile Tyr Leu Ser Asp Ala Gly Ile Phe Thr Val
210 215 220

Cys Tyr Ile Leu Tyr Arg Leu Val Leu Thr Lys Gly Leu Val Trp Val
225 230 235 240

Val Ser Ile Tyr Gly Val Pro Leu Leu Ile Val Asn Gly Phe Leu Val
245 250 255

Leu Ile Thr Phe Leu Gln His Thr His Pro Ser Leu Pro His Tyr Lys
260 265 270

Val Leu Arg Met Gly Thr Gly Cys Glu Ala Pro Ser Arg Pro Trp Ile
275 280 285

Glu Thr Thr Gly Tyr Ser Thr Pro Cys Ser Thr Thr Ser Pro Ile His
290 295 300

Met Ser Arg Thr Ile Ser Ser Pro Arg Cys Leu Ile Thr Thr Arg Trp
305 310 315 320

Arg Leu Pro Arg Arg Ser Ser Arg Phe Ser Gly Ser Ile Thr Ser Ser
325 330 335

Met Gly Leu Pro Leu
340

<210> 27

<211> 1023

<212> DNA

<213> 亚麻

<220>

<221> CDS

<222> (1)..(1023)

<400> 27

atg tca gtg cct cca tca tcc aaa cct atg aag agg tct cct tac tca 48
Met Ser Val Pro Pro Ser Ser Lys Pro Met Lys Arg Ser Pro Tyr Ser
1 5 10 15

aag cca cca ttc acg ctc ggt gag ctc aag aag gcc att cct cca cac 96
Lys Pro Pro Phe Thr Leu Gly Glu Leu Lys Lys Ala Ile Pro Pro His
20 25 30

tgt ttc aaa cgc tca atc ccc cga tcg ttc gcc tac gtg gcg tac gac 144
Cys Phe Lys Arg Ser Ile Pro Arg Ser Phe Ala Tyr Val Ala Tyr Asp
35 40 45

ctc acc att gca gca atc ttc tac tac atc gcc acc act tac ttc cac 192
Leu Thr Ile Ala Ala Ile Phe Tyr Tyr Ile Ala Thr Thr Tyr Phe His
50 55 60

ctc ctc cct agc cct ctc aac tac ctc gcc tgg ccg gtc tac tgg gcc Leu Leu Pro Ser Pro Leu Asn Tyr Leu Ala Trp Pro Val Tyr Trp Ala 65 70 75 80	240
tgc cag ggc tgc atc ctc act gga gta tgg gtg ttg gct cac gaa tgc Cys Gln Gly Cys Ile Leu Thr Gly Val Trp Val Leu Ala His Glu Cys 85 90 95	288
ggc cac cat gcc ttc agc gac tac cag tgg ctc gac gac atg gtt ggc Gly His His Ala Phe Ser Asp Tyr Gln Trp Leu Asp Asp Met Val Gly 100 105 110	336
ttc gtc ctc cat tcg tcc ctc ctt gtt cct tac ttc tcc tgg aag cac Phe Val Leu His Ser Ser Leu Leu Val Pro Tyr Phe Ser Trp Lys His 115 120 125	384
agc cac cgc cgc cac cat tcc aac acg gga tcg ctt gat cgt gat gag Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Asp Arg Asp Glu 130 135 140	432
gtg ttt gtc ccc aag cag aag gcc gaa atc ggg tgg tac tcc aag tac Val Phe Val Pro Lys Gln Lys Ala Glu Ile Gly Trp Tyr Ser Lys Tyr 145 150 155 160	480
ctt aac aac cca cct ggc cgt gtg atc aca ttg gcc gtc aca tta acg Leu Asn Asn Pro Pro Gly Arg Val Ile Thr Leu Ala Val Thr Leu Thr 165 170 175	528
ctc ggt tgg cct ctg tac ttg gca ttc aac gtc tcc ggg aga cca tat Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr 180 185 190	576
gac cgg ttc gca tgc cat ttt gac cct cac ggt ccg att tac aat gat Asp Arg Phe Ala Cys His Phe Asp Pro His Gly Pro Ile Tyr Asn Asp 195 200 205	624
cgc gag cgt atg gag ata tac cta tcc gac gca ggg ata ttc acc gtg Arg Glu Arg Met Glu Ile Tyr Leu Ser Asp Ala Gly Ile Phe Thr Val 210 215 220	672
tgc tac atc cta tac aga ctc gtc ctc acg aaa gga ctc gtt tgg gtc Cys Tyr Ile Leu Tyr Arg Leu Val Leu Thr Lys Gly Leu Val Trp Val 225 230 235 240	720

gtg tcc ata tac gga gtc cca cta ttg ata gtg aat gga ttc cta gtc 768
Val Ser Ile Tyr Gly Val Pro Leu Leu Ile Val Asn Gly Phe Leu Val
245 250 255

ctc atc act ttc ttg cag cac acg cat cct tct ctt ccg cac tac aaa 816
Leu Ile Thr Phe Leu Gln His Thr His Pro Ser Leu Pro His Tyr Lys
260 265 270

gtc ctc cga atg ggg act gga tgc gag gcg ccc tct cga ccg tgg atc 864
Val Leu Arg Met Gly Thr Gly Cys Glu Ala Pro Ser Arg Pro Trp Ile
275 280 285

gag act acg ggt tac tca aca ccg tgt tcc aca aca tca ccg ata cac 912
Glu Thr Thr Gly Tyr Ser Thr Pro Cys Ser Thr Thr Ser Pro Ile His
290 295 300

atg tcg cgc acc atc tct tct cca cga tgc ctc att acc acg cga tgg 960
Met Ser Arg Thr Ile Ser Ser Pro Arg Cys Leu Ile Thr Thr Arg Trp
305 310 315 320

agg cta cca agg cga tca agc cgg ttc tcg ggg agt att acc agt tcg 1008
Arg Leu Pro Arg Arg Ser Ser Arg Phe Ser Gly Ser Ile Thr Ser Ser
325 330 335

atg gga ctc cct ttg 1023
Met Gly Leu Pro Leu
340

<210> 28

<211> 341

<212> PRT

<213> 亚麻

<400> 28

Met Ser Val Pro Pro Ser Ser Lys Pro Met Lys Arg Ser Pro Tyr Ser
1 5 10 15

Lys Pro Pro Phe Thr Leu Gly Glu Leu Lys Lys Ala Ile Pro Pro His
20 25 30

Cys Phe Lys Arg Ser Ile Pro Arg Ser Phe Ala Tyr Val Ala Tyr Asp
35 40 45

Leu Thr Ile Ala Ala Ile Phe Tyr Tyr Ile Ala Thr Thr Tyr Phe His
50 55 60

Leu Leu Pro Ser Pro Leu Asn Tyr Leu Ala Trp Pro Val Tyr Trp Ala
65 70 75 80

Cys Gln Gly Cys Ile Leu Thr Gly Val Trp Val Leu Ala His Glu Cys
85 90 95

Gly His His Ala Phe Ser Asp Tyr Gln Trp Leu Asp Asp Met Val Gly
100 105 110

Phe Val Leu His Ser Ser Leu Leu Val Pro Tyr Phe Ser Trp Lys His
115 120 125

Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Asp Arg Asp Glu
130 135 140

Val Phe Val Pro Lys Gln Lys Ala Glu Ile Gly Trp Tyr Ser Lys Tyr
145 150 155 160

Leu Asn Asn Pro Pro Gly Arg Val Ile Thr Leu Ala Val Thr Leu Thr
165 170 175

Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn Val Ser Gly Arg Pro Tyr
180 185 190

Asp Arg Phe Ala Cys His Phe Asp Pro His Gly Pro Ile Tyr Asn Asp
195 200 205

Arg Glu Arg Met Glu Ile Tyr Leu Ser Asp Ala Gly Ile Phe Thr Val
210 215 220

Cys Tyr Ile Leu Tyr Arg Leu Val Leu Thr Lys Gly Leu Val Trp Val
225 230 235 240

Val Ser Ile Tyr Gly Val Pro Leu Leu Ile Val Asn Gly Phe Leu Val
245 250 255

Leu Ile Thr Phe Leu Gln His Thr His Pro Ser Leu Pro His Tyr Lys
260 265 270

Val Leu Arg Met Gly Thr Gly Cys Glu Ala Pro Ser Arg Pro Trp Ile
275 280 285

Glu Thr Thr Gly Tyr Ser Thr Pro Cys Ser Thr Thr Ser Pro Ile His
290 295 300

Met Ser Arg Thr Ile Ser Ser Pro Arg Cys Leu Ile Thr Thr Arg Trp
305 310 315 320

Arg Leu Pro Arg Arg Ser Ser Arg Phe Ser Gly Ser Ile Thr Ser Ser
325 330 335

Met Gly Leu Pro Leu
340

<210> 29
<211> 1545
<212> DNA
<213> 大麦 (Hordeum vulgare)

<220>
<221> misc_feature
<222> (1)..(24)

<220>
<221> CDS
<222> (25)..(1185)

<220>
<221> misc_feature
<222> (1186)..(1545)

<400> 29

accaccacca cccctaccag catc atg ggt gcc ggc ggc ggc atg acc gag	51
Met Gly Ala Gly Gly Gly Met Thr Glu	
1 5	
aag gag cgg gag aag cag gag cag ctc ggc cgc gcc ggc ggc ggc gca	99
Lys Glu Arg Glu Lys Gln Glu Gln Leu Gly Arg Ala Gly Gly Gly Ala	
10 15 20 25	
gcc ttc cag cgc tcg ccg acg gac aag ccg ccg ttc acg ctc ggt cag	147
Ala Phe Gln Arg Ser Pro Thr Asp Lys Pro Pro Phe Thr Leu Gly Gln	
30 35 40	

atc aag aag gcg atc ccg cct cac tgc ttc cag cgc tcc atc atc aag Ile Lys Lys Ala Ile Pro Pro His Cys Phe Gln Arg Ser Ile Ile Lys 45 50 55	195
tcc ttc tcc tac gtg gtt cat gac ctg gtc atc atc gcc gcc ctg ctg Ser Phe Ser Tyr Val Val His Asp Leu Val Ile Ile Ala Ala Leu Leu 60 65 70	243
tac gcc gct ctg gtc tgg atc ccc acc ctc cct acc gtg ttg cag ctg Tyr Ala Ala Leu Val Trp Ile Pro Thr Leu Pro Thr Val Leu Gln Leu 75 80 85	291
ggc gcg tgg ccg ctc tac tgg atc gtt cag ggc tgc gtc atg acc ggc Gly Ala Trp Pro Leu Tyr Trp Ile Val Gln Gly Cys Val Met Thr Gly 90 95 100 105	339
gtc tgg gtc atc gcg cac gag tgc ggc cac cat gcc ttc tct gac tac Val Trp Val Ile Ala His Glu Cys Gly His His Ala Phe Ser Asp Tyr 110 115 120	387
tcg ctg ctc gac gac acc gtc ggc ctg gtg ctc cac tcg tgg ctg ctc Ser Leu Leu Asp Asp Thr Val Gly Leu Val Leu His Ser Trp Leu Leu 125 130 135	435
gtc cca tac ttc tcg tgg aag tac agc cac cgt cgc cac cac tcc aac Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His His Ser Asn 140 145 150	483
acc ggg tcg ctg gag cgc gac gag gtg ttt gtc ccc aag cag aag gag Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys Gln Lys Glu 155 160 165	531
gcg ctg gca tgg tac act ccc tac atc tac aac aac ccc atc ggc cgt Ala Leu Ala Trp Tyr Thr Pro Tyr Ile Tyr Asn Asn Pro Ile Gly Arg 170 175 180 185	579
ctg gtg cac atc gtg gtg cag ctc acc ctc ggg tgg ccg ctg tac ctg Leu Val His Ile Val Val Gln Leu Thr Leu Gly Trp Pro Leu Tyr Leu 190 195 200	627
gcg ctc aac gcc tca ggc cgt ccg tac ccg cgc ttc gcc tgc cac ttc Ala Leu Asn Ala Ser Gly Arg Pro Tyr Pro Arg Phe Ala Cys His Phe 205 210 215	675

gac ccc tac ggc ccg atc tac aac gac cgg gag cga gcc cag att ttc Asp Pro Tyr Gly Pro Ile Tyr Asn Asp Arg Glu Arg Ala Gln Ile Phe 220 225 230	723
atc tcg gat gtc ggc gtg ttg gcc gtc tcc ttg gcc ctg ctc aag ctt Ile Ser Asp Val Gly Val Leu Ala Val Ser Leu Ala Leu Leu Lys Leu 235 240 245	771
gtg tcg tcg ttt ggg ttc tgg tgg gtg gtg cgg gtc tac ggc gtg ccg Val Ser Ser Phe Gly Phe Trp Trp Val Val Arg Val Tyr Gly Val Pro 250 255 260 265	819
ctg ctg atc gtg aac gcg tgg ctg gtc ctg atc acc tac ctg cag cac Leu Leu Ile Val Asn Ala Trp Leu Val Leu Ile Thr Tyr Leu Gln His 270 275 280	867
acc cac cca gcg ctg ccg cac tac gac tcg acg gag tgg gac tgg ctg Thr His Pro Ala Leu Pro His Tyr Asp Ser Thr Glu Trp Asp Trp Leu 285 290 295	915
cgg ggg gcg ctc gcc acc atg gac cgg gac tac ggc att ctc aac cgc Arg Gly Ala Leu Ala Thr Met Asp Arg Asp Tyr Gly Ile Leu Asn Arg 300 305 310	963
gtg ttc cac aac atc acg gac acg cac gtg gcg cac cac ctc ttc tcc Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu Phe Ser 315 320 325	1011
aac atg ccg cac tac cac gcc atg gag gcc acc aag gcg atc aag ccc Asn Met Pro His Tyr His Ala Met Glu Ala Thr Lys Ala Ile Lys Pro 330 335 340 345	1059
atc ctc ggc gag tac tac cag ttt gac ggc acc ccg gtc gcc aag gcc Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Gly Thr Pro Val Ala Lys Ala 350 355 360	1107
aca tgg cgc gag gcc aag gag tgc atc tac gtt gag ccc gag gac cgc Thr Trp Arg Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Glu Asp Arg 365 370 375	1155
aag ggg gtc ttc tgg tac agc aac aag ttc tagccgcaag gatcgtcatc Lys Gly Val Phe Trp Tyr Ser Asn Lys Phe 380 385	1205
agccgtgttc caggaagaac tcagagaaga ggctcctaca agtaatccat ccattctacct	1265

acatatggtt agtttttaga tagcagaggg catttgggca caaacaagac tactattacc 1325
 gtgccaatgc tagaaagagt tgagtgggtgc aaggaggagt agcgtgtccg tgacttttgt 1385
 cagttccttc ttacttttcc tcctgcgtct tagtcgccgg cggtcgttgt tgggtgccgt 1445
 ggccattgac atggccgtgt gtgttgtgtg tgcgtctgtc attgcattgg cgtcatctcc 1505
 ccccggtccgt gtcattgtgt ttagaccat ttcgtgtttt 1545

<210> 30

<211> 387

<212> PRT

<213> 大麦

<400> 30

Met Gly Ala Gly Gly Gly Met Thr Glu Lys Glu Arg Glu Lys Gln Glu
 1 5 10 15

Gln Leu Gly Arg Ala Gly Gly Gly Ala Ala Phe Gln Arg Ser Pro Thr
 20 25 30

Asp Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro
 35 40 45

His Cys Phe Gln Arg Ser Ile Ile Lys Ser Phe Ser Tyr Val Val His
 50 55 60

Asp Leu Val Ile Ile Ala Ala Leu Leu Tyr Ala Ala Leu Val Trp Ile
 65 70 75 80

Pro Thr Leu Pro Thr Val Leu Gln Leu Gly Ala Trp Pro Leu Tyr Trp
 85 90 95

Ile Val Gln Gly Cys Val Met Thr Gly Val Trp Val Ile Ala His Glu
 100 105 110

Cys Gly His His Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp Thr Val
 115 120 125

Gly Leu Val Leu His Ser Trp Leu Leu Val Pro Tyr Phe Ser Trp Lys
 130 135 140

Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp			
145	150	155	160
Glu Val Phe Val Pro Lys Gln Lys Glu Ala Leu Ala Trp Tyr Thr Pro			
	165	170	175
Tyr Ile Tyr Asn Asn Pro Ile Gly Arg Leu Val His Ile Val Val Gln			
	180	185	190
Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Leu Asn Ala Ser Gly Arg			
	195	200	205
Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr			
	210	215	220
Asn Asp Arg Glu Arg Ala Gln Ile Phe Ile Ser Asp Val Gly Val Leu			
225	230	235	240
Ala Val Ser Leu Ala Leu Leu Lys Leu Val Ser Ser Phe Gly Phe Trp			
	245	250	255
Trp Val Val Arg Val Tyr Gly Val Pro Leu Leu Ile Val Asn Ala Trp			
	260	265	270
Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro His			
	275	280	285
Tyr Asp Ser Thr Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met			
	290	295	300
Asp Arg Asp Tyr Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp			
305	310	315	320
Thr His Val Ala His His Leu Phe Ser Asn Met Pro His Tyr His Ala			
	325	330	335
Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln			
	340	345	350
Phe Asp Gly Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Lys Glu			
	355	360	365
Cys Ile Tyr Val Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Ser			
	370	375	380

Asn Lys Phe
385

<210> 31

<211> 1161

<212> DNA

<213> 大麦

<220>

<221> CDS

<222> (1)..(1161)

<400> 31

atg ggt gcc ggc ggc ggg atg acc gag aag gag cgg gag aag cag gag	48
Met Gly Ala Gly Gly Gly Met Thr Glu Lys Glu Arg Glu Lys Gln Glu	
1 5 10 15	

cag ctc ggc cgc gcc ggc ggc ggc gca gcc ttc cag cgc tcg ccg acg	96
Gln Leu Gly Arg Ala Gly Gly Gly Ala Ala Phe Gln Arg Ser Pro Thr	
20 25 30	

gac aag ccg ccg ttc acg ctc ggt cag atc aag aag gcg atc ccg cct	144
Asp Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro	
35 40 45	

cac tgc ttc cag cgc tcc atc atc aag tcc ttc tcc tac gtg gtt cat	192
His Cys Phe Gln Arg Ser Ile Ile Lys Ser Phe Ser Tyr Val Val His	
50 55 60	

gac ctg gtc atc atc gcc gcc ctg ctg tac gcc gct ctg gtc tgg atc	240
Asp Leu Val Ile Ile Ala Ala Leu Leu Tyr Ala Ala Leu Val Trp Ile	
65 70 75 80	

ccc acc ctc cct acc gtg ttg cag ctg ggc gcg tgg ccg ctc tac tgg	288
Pro Thr Leu Pro Thr Val Leu Gln Leu Gly Ala Trp Pro Leu Tyr Trp	
85 90 95	

atc gtt cag ggc tgc gtc atg acc ggc gtc tgg gtc atc gcg cac gag	336
Ile Val Gln Gly Cys Val Met Thr Gly Val Trp Val Ile Ala His Glu	
100 105 110	

tgc ggc cac cat gcc ttc tct gac tac tcg ctg ctc gac gac acc gtc	384
Cys Gly His His Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp Thr Val	
115 120 125	

ggc ctg gtg ctc cac tcg tgg ctg ctc gtc cca tac ttc tcg tgg aag Gly Leu Val Leu His Ser Trp Leu Leu Val Pro Tyr Phe Ser Trp Lys 130 135 140	432
tac agc cac cgt cgc cac cac tcc aac acc ggg tcg ctg gag cgc gac Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp 145 150 155 160	480
gag gtg ttt gtc ccc aag cag aag gag gcg ctg gca tgg tac act ccc Glu Val Phe Val Pro Lys Gln Lys Glu Ala Leu Ala Trp Tyr Thr Pro 165 170 175	528
tac atc tac aac aac ccc atc ggc cgt ctg gtg cac atc gtg gtg cag Tyr Ile Tyr Asn Asn Pro Ile Gly Arg Leu Val His Ile Val Val Gln 180 185 190	576
ctc acc ctc ggg tgg ccg ctg tac ctg gcg ctc aac gcc tca ggc cgt Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Leu Asn Ala Ser Gly Arg 195 200 205	624
ccg tac ccg cgc ttc gcc tgc cac ttc gac ccc tac ggc ccg atc tac Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr 210 215 220	672
aac gac cgg gag cga gcc cag att ttc atc tcg gat gtc ggc gtg ttg Asn Asp Arg Glu Arg Ala Gln Ile Phe Ile Ser Asp Val Gly Val Leu 225 230 235 240	720
gcc gtc tcc ttg gcc ctg ctc aag ctt gtg tcg tcg ttt ggg ttc tgg Ala Val Ser Leu Ala Leu Leu Lys Leu Val Ser Ser Phe Gly Phe Trp 245 250 255	768
tgg gtg gtg cgg gtc tac ggc gtg ccg ctg ctg atc gtg aac gcg tgg Trp Val Val Arg Val Tyr Gly Val Pro Leu Leu Ile Val Asn Ala Trp 260 265 270	816
ctg gtc ctg atc acc tac ctg cag cac acc cac cca gcg ctg ccg cac Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro His 275 280 285	864
tac gac tcg acg gag tgg gac tgg ctg cgg ggg gcg ctc gcc acc atg Tyr Asp Ser Thr Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met 290 295 300	912

gac cgg gac tac ggc att ctc aac cgc gtg ttc cac aac atc acg gac 960
 Asp Arg Asp Tyr Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp
 305 310 315 320

acg cac gtg gcg cac cac ctc ttc tcc aac atg ccg cac tac cac gcc 1008
 Thr His Val Ala His His Leu Phe Ser Asn Met Pro His Tyr His Ala
 325 330 335

atg gag gcc acc aag gcg atc aag ccc atc ctc ggc gag tac tac cag 1056
 Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln
 340 345 350

ttt gac ggc acc ccg gtc gcc aag gcc aca tgg cgc gag gcc aag gag 1104
 Phe Asp Gly Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Lys Glu
 355 360 365

tgc atc tac gtt gag ccc gag gac cgc aag ggg gtc ttc tgg tac agc 1152
 Cys Ile Tyr Val Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Ser
 370 375 380

aac aag ttc 1161
 Asn Lys Phe
 385

<210> 32
 <211> 387
 <212> PRT
 <213> 大麦

<400> 32
 Met Gly Ala Gly Gly Gly Met Thr Glu Lys Glu Arg Glu Lys Gln Glu
 1 5 10 15
 Gln Leu Gly Arg Ala Gly Gly Gly Ala Ala Phe Gln Arg Ser Pro Thr
 20 25 30
 Asp Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro
 35 40 45
 His Cys Phe Gln Arg Ser Ile Ile Lys Ser Phe Ser Tyr Val Val His
 50 55 60
 Asp Leu Val Ile Ile Ala Ala Leu Leu Tyr Ala Ala Leu Val Trp Ile
 65 70 75 80

Pro Thr Leu Pro Thr Val Leu Gln Leu Gly Ala Trp Pro Leu Tyr Trp
 85 90 95

Ile Val Gln Gly Cys Val Met Thr Gly Val Trp Val Ile Ala His Glu
 100 105 110

Cys Gly His His Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp Thr Val
 115 120 125

Gly Leu Val Leu His Ser Trp Leu Leu Val Pro Tyr Phe Ser Trp Lys
 130 135 140

Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp
 145 150 155 160

Glu Val Phe Val Pro Lys Gln Lys Glu Ala Leu Ala Trp Tyr Thr Pro
 165 170 175

Tyr Ile Tyr Asn Asn Pro Ile Gly Arg Leu Val His Ile Val Val Gln
 180 185 190

Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Leu Asn Ala Ser Gly Arg
 195 200 205

Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr
 210 215 220

Asn Asp Arg Glu Arg Ala Gln Ile Phe Ile Ser Asp Val Gly Val Leu
 225 230 235 240

Ala Val Ser Leu Ala Leu Leu Lys Leu Val Ser Ser Phe Gly Phe Trp
 245 250 255

Trp Val Val Arg Val Tyr Gly Val Pro Leu Leu Ile Val Asn Ala Trp
 260 265 270

Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro His
 275 280 285

Tyr Asp Ser Thr Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met
 290 295 300

Asp Arg Asp Tyr Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp
 305 310 315 320

Thr His Val Ala His His Leu Phe Ser Asn Met Pro His Tyr His Ala
325 330 335

Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln
340 345 350

Phe Asp Gly Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Lys Glu
355 360 365

Cys Ile Tyr Val Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Ser
370 375 380

Asn Lys Phe
385

<210> 33
<211> 1749
<212> DNA
<213> 小麦 (Triticum aestivum)

<220>
<221> misc_feature
<222> (1)..(164)

<220>
<221> CDS
<222> (165)..(1325)

<220>
<221> misc_feature
<222> (1326)..(1749)

<400> 33

cgcggtccgcg gctcctcccc cccgcacaaa ccactcggtc gtcccgtaa caagaggagc 60

agaggcgccg gagagggaag aggggtgcgcg cgctcgcgtg tgtggtgtcc gccggcccga 120

tctgccctgc tcccgcgcgc tcgaccacca cccctatcag catc atg ggt gcc ggc 176

Met Gly Ala Gly

1

ggc agg atg acg gag aag gag cgg gag aag cag gag cag ctc ggc cgc Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu Gln Leu Gly Arg 5 10 15 20	224
gcc aac ggc ggc gca gcc tac cag cgc tcg ccg acg gac aag ccg ccg Ala Asn Gly Gly Ala Ala Tyr Gln Arg Ser Pro Thr Asp Lys Pro Pro 25 30 35	272
ttc acg ctg ggt cag atc aag aag gca atc ccg cct cac tgc ttc cag Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro His Cys Phe Gln 40 45 50	320
cgc tcg atc atc aag tcc ttc tcc tac gtg gtc cat gac ctg gtc atc Arg Ser Ile Ile Lys Ser Phe Ser Tyr Val Val His Asp Leu Val Ile 55 60 65	368
gtc gcg gcc ctg ctg tac gcg gcg ctg gtt tgg atc cct acc ctc ccg Val Ala Ala Leu Leu Tyr Ala Ala Leu Val Trp Ile Pro Thr Leu Pro 70 75 80	416
acc gtg ctg cag ctg ggc gcc tgg ccg ctc tac tgg atc gtg cag ggc Thr Val Leu Gln Leu Gly Ala Trp Pro Leu Tyr Trp Ile Val Gln Gly 85 90 95 100	464
tgc gtc atg acc ggc gtc tgg gtc atc gcc cac gag tgc ggc cac cac Cys Val Met Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His 105 110 115	512
gcc ttc tcc gac tac tcg ctg ctc gac gac acc gtc ggc ctg gtg ctc Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp Thr Val Gly Leu Val Leu 120 125 130	560
cac tcg tgg ctg ctc gtc ccc tac ttc tcg tgg aag tac agc cac cgt His Ser Trp Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg 135 140 145	608
cgc cac cac tcc aac acc ggg tcg ctg gag cgt gat gag gtg ttc gtc Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val 150 155 160	656
ccc aag cag aag gag gcg ctg gcg tgg tac acc cct tac atc tac aac Pro Lys Gln Lys Glu Ala Leu Ala Trp Tyr Thr Pro Tyr Ile Tyr Asn 165 170 175 180	704

aac ccc gtc ggc cgt ctg gtg cac atc gtc gtg cag ctc acc ctc ggg Asn Pro Val Gly Arg Leu Val His Ile Val Val Gln Leu Thr Leu Gly 185 190 195	752
tgg ccg ctg tac ctg gcg ctc aac gcc tca ggc cgc ccg tac ccg cgg Trp Pro Leu Tyr Leu Ala Leu Asn Ala Ser Gly Arg Pro Tyr Pro Arg 200 205 210	800
ttc gcc tgc cac ttc gac ccc tac ggc ccg atc tac aac gac cgg gag Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr Asn Asp Arg Glu 215 220 225	848
cga gcc cag att ttc atc tca gac gtc gga gtg ctg gcc gtg tca ttg Arg Ala Gln Ile Phe Ile Ser Asp Val Gly Val Leu Ala Val Ser Leu 230 235 240	896
gct ctg ctg aag ctc gtg tcg tcg ttc ggg ttc tgg tgg gtg gtg cgg Ala Leu Leu Lys Leu Val Ser Ser Phe Gly Phe Trp Trp Val Val Arg 245 250 255 260	944
gtc tac ggc gtg ccg ctg ctg atc gtg aac gct tgg ctg gtc ctg atc Val Tyr Gly Val Pro Leu Leu Ile Val Asn Ala Trp Leu Val Leu Ile 265 270 275	992
acc tac ctg cag cac acc cac ccg gcg ctg ccg cac tac gac tcg acg Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro His Tyr Asp Ser Thr 280 285 290	1040
gag tgg gac tgg ctg cgc ggg gcg ctc gcc acc atg gac cgc gac tac Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met Asp Arg Asp Tyr 295 300 305	1088
ggc atc ctc aac cgc gtg ttc cac aac atc acg gac acg cac gtg gcg Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp Thr His Val Ala 310 315 320	1136
cac cac ctc ttc tcc acc atg ccg cac tac cac gcc atg gag gcc acc His His Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr 325 330 335 340	1184
aag gcg atc aag ccc atc ctc ggc gag tac tac cag ttc gac ccc acc Lys Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Pro Thr 345 350 355	1232

```

ccc gtc gcc aag gcc aca tgg cgc gag gcc aag gag tgc atc tac gtc      1280
Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Lys Glu Cys Ile Tyr Val
          360              365              370

gag ccc gag gac cgc aag ggg gtc ttc tgg tac agc aac aag ttc      1325
Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Ser Asn Lys Phe
          375              380              385

tagccgccaa gatccatcaa ctgtgctgga gaaagaactc agagaagaga tcctaccaag      1385

taattccatc catctaccta cagtccatat ggtagtctt tagatagcag agggcatttg      1445

ggcacaaaag aagactacta ttaccgtgcc aatgctagaa gagctgagtg gtgcaaggaa      1505

gagtagcgtg tccgtgactt tggtcagttc cgtctttact ttttctctgc gttctagtcg      1565

tcggcttagg tttggccggc ggtcatcggtt ggtgtccgtg gccgtggaca tggccgcgtg      1625

tgttgtgtgt gcgtctgtca ttgcatggc gtcattctcc cccgtccgtg tcatgttgtt      1685

gtagaccatt tcgtgtttat ggcggaataa ctgatcgctg aaggaagggc aacttttttg      1745

agta                                                                    1749

```

<210> 34
 <211> 387
 <212> PRT
 <213> 小麦

```

<400> 34
Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu
1              5              10              15

Gln Leu Gly Arg Ala Asn Gly Gly Ala Ala Tyr Gln Arg Ser Pro Thr
          20              25              30

Asp Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro
          35              40              45

His Cys Phe Gln Arg Ser Ile Ile Lys Ser Phe Ser Tyr Val Val His
          50              55              60

Asp Leu Val Ile Val Ala Ala Leu Leu Tyr Ala Ala Leu Val Trp Ile
65              70              75              80

```

Pro Thr Leu Pro Thr Val Leu Gln Leu Gly Ala Trp Pro Leu Tyr Trp
 85 90 95

Ile Val Gln Gly Cys Val Met Thr Gly Val Trp Val Ile Ala His Glu
 100 105 110

Cys Gly His His Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp Thr Val
 115 120 125

Gly Leu Val Leu His Ser Trp Leu Leu Val Pro Tyr Phe Ser Trp Lys
 130 135 140

Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp
 145 150 155 160

Glu Val Phe Val Pro Lys Gln Lys Glu Ala Leu Ala Trp Tyr Thr Pro
 165 170 175

Tyr Ile Tyr Asn Asn Pro Val Gly Arg Leu Val His Ile Val Val Gln
 180 185 190

Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Leu Asn Ala Ser Gly Arg
 195 200 205

Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr
 210 215 220

Asn Asp Arg Glu Arg Ala Gln Ile Phe Ile Ser Asp Val Gly Val Leu
 225 230 235 240

Ala Val Ser Leu Ala Leu Leu Lys Leu Val Ser Ser Phe Gly Phe Trp
 245 250 255

Trp Val Val Arg Val Tyr Gly Val Pro Leu Leu Ile Val Asn Ala Trp
 260 265 270

Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro His
 275 280 285

Tyr Asp Ser Thr Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met
 290 295 300

Asp Arg Asp Tyr Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp
 305 310 315 320

Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala
 325 330 335

Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln
 340 345 350

Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Lys Glu
 355 360 365

Cys Ile Tyr Val Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Ser
 370 375 380

Asn Lys Phe
 385

<210> 35
 <211> 1161
 <212> DNA
 <213> 小麦

<220>
 <221> CDS
 <222> (1)..(1161)

<400> 35

atg ggt gcc ggc ggc agg atg acg gag aag gag cgg gag aag cag gag 48
 Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu
 1 5 10 15

cag ctc ggc cgc gcc aac ggc ggc gca gcc tac cag cgc tcg ccg acg 96
 Gln Leu Gly Arg Ala Asn Gly Gly Ala Ala Tyr Gln Arg Ser Pro Thr
 20 25 30

gac aag ccg ccg ttc acg ctg ggt cag atc aag aag gca atc ccg cct 144
 Asp Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro
 35 40 45

cac tgc ttc cag cgc tcg atc atc aag tcc ttc tcc tac gtg gtc cat 192
 His Cys Phe Gln Arg Ser Ile Ile Lys Ser Phe Ser Tyr Val Val His
 50 55 60

gac ctg gtc atc gtc gcg gcc ctg ctg tac gcg gcg ctg gtt tgg atc Asp Leu Val Ile Val Ala Ala Leu Leu Tyr Ala Ala Leu Val Trp Ile 65 70 75 80	240
cct acc ctc ccg acc gtg ctg cag ctg ggc gcc tgg ccg ctc tac tgg Pro Thr Leu Pro Thr Val Leu Gln Leu Gly Ala Trp Pro Leu Tyr Trp 85 90 95	288
atc gtg cag ggc tgc gtc atg acc ggc gtc tgg gtc atc gcc cac gag Ile Val Gln Gly Cys Val Met Thr Gly Val Trp Val Ile Ala His Glu 100 105 110	336
tgc ggc cac cac gcc ttc tcc gac tac tcg ctg ctc gac gac acc gtc Cys Gly His His Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp Thr Val 115 120 125	384
ggc ctg gtg ctc cac tcg tgg ctg ctc gtc ccc tac ttc tcg tgg aag Gly Leu Val Leu His Ser Trp Leu Leu Val Pro Tyr Phe Ser Trp Lys 130 135 140	432
tac agc cac cgt cgc cac cac tcc aac acc ggg tcg ctg gag cgt gat Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp 145 150 155 160	480
gag gtg ttc gtc ccc aag cag aag gag gcg ctg gcg tgg tac acc cct Glu Val Phe Val Pro Lys Gln Lys Glu Ala Leu Ala Trp Tyr Thr Pro 165 170 175	528
tac atc tac aac aac ccc gtc ggc cgt ctg gtg cac atc gtc gtg cag Tyr Ile Tyr Asn Asn Pro Val Gly Arg Leu Val His Ile Val Val Gln 180 185 190	576
ctc acc ctc ggg tgg ccg ctg tac ctg gcg ctc aac gcc tca ggc cgc Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Leu Asn Ala Ser Gly Arg 195 200 205	624
ccg tac ccg cgg ttc gcc tgc cac ttc gac ccc tac ggc ccg atc tac Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr 210 215 220	672
aac gac cgg gag cga gcc cag att ttc atc tca gac gtc gga gtg ctg Asn Asp Arg Glu Arg Ala Gln Ile Phe Ile Ser Asp Val Gly Val Leu 225 230 235 240	720

gcc gtg tca ttg gct ctg ctg aag ctc gtg tcg tcg ttc ggg ttc tgg Ala Val Ser Leu Ala Leu Leu Lys Leu Val Ser Ser Phe Gly Phe Trp 245 250 255	768
tgg gtg gtg cgg gtc tac ggc gtg ccg ctg ctg atc gtg aac gct tgg Trp Val Val Arg Val Tyr Gly Val Pro Leu Leu Ile Val Asn Ala Trp 260 265 270	816
ctg gtc ctg atc acc tac ctg cag cac acc cac ccg gcg ctg ccg cac Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro His 275 280 285	864
tac gac tcg acg gag tgg gac tgg ctg cgc ggg gcg ctc gcc acc atg Tyr Asp Ser Thr Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met 290 295 300	912
gac cgc gac tac ggc atc ctc aac cgc gtg ttc cac aac atc acg gac Asp Arg Asp Tyr Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp 305 310 315 320	960
acg cac gtg gcg cac cac ctc ttc tcc acc atg ccg cac tac cac gcc Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala 325 330 335	1008
atg gag gcc acc aag gcg atc aag ccc atc ctc ggc gag tac tac cag Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln 340 345 350	1056
ttc gac ccc acc ccc gtc gcc aag gcc aca tgg cgc gag gcc aag gag Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Lys Glu 355 360 365	1104
tgc atc tac gtc gag ccc gag gac cgc aag ggg gtc ttc tgg tac agc Cys Ile Tyr Val Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Ser 370 375 380	1152
aac aag ttc Asn Lys Phe 385	1161

<210> 36
 <211> 387
 <212> PRT
 <213> 小麦

<400> 36

Met Gly Ala Gly Gly Arg Met Thr Glu Lys Glu Arg Glu Lys Gln Glu
1 5 10 15

Gln Leu Gly Arg Ala Asn Gly Gly Ala Ala Tyr Gln Arg Ser Pro Thr
20 25 30

Asp Lys Pro Pro Phe Thr Leu Gly Gln Ile Lys Lys Ala Ile Pro Pro
35 40 45

His Cys Phe Gln Arg Ser Ile Ile Lys Ser Phe Ser Tyr Val Val His
50 55 60

Asp Leu Val Ile Val Ala Ala Leu Leu Tyr Ala Ala Leu Val Trp Ile
65 70 75 80

Pro Thr Leu Pro Thr Val Leu Gln Leu Gly Ala Trp Pro Leu Tyr Trp
85 90 95

Ile Val Gln Gly Cys Val Met Thr Gly Val Trp Val Ile Ala His Glu
100 105 110

Cys Gly His His Ala Phe Ser Asp Tyr Ser Leu Leu Asp Asp Thr Val
115 120 125

Gly Leu Val Leu His Ser Trp Leu Leu Val Pro Tyr Phe Ser Trp Lys
130 135 140

Tyr Ser His Arg Arg His His Ser Asn Thr Gly Ser Leu Glu Arg Asp
145 150 155 160

Glu Val Phe Val Pro Lys Gln Lys Glu Ala Leu Ala Trp Tyr Thr Pro
165 170 175

Tyr Ile Tyr Asn Asn Pro Val Gly Arg Leu Val His Ile Val Val Gln
180 185 190

Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Leu Asn Ala Ser Gly Arg
195 200 205

Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr
210 215 220

Asn Asp Arg Glu Arg Ala Gln Ile Phe Ile Ser Asp Val Gly Val Leu
225 230 235 240

Ala Val Ser Leu Ala Leu Leu Lys Leu Val Ser Ser Phe Gly Phe Trp
245 250 255

Trp Val Val Arg Val Tyr Gly Val Pro Leu Leu Ile Val Asn Ala Trp
260 265 270

Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ala Leu Pro His
275 280 285

Tyr Asp Ser Thr Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met
290 295 300

Asp Arg Asp Tyr Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp
305 310 315 320

Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala
325 330 335

Met Glu Ala Thr Lys Ala Ile Lys Pro Ile Leu Gly Glu Tyr Tyr Gln
340 345 350

Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Lys Glu
355 360 365

Cys Ile Tyr Val Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Ser
370 375 380

Asn Lys Phe
385

- <210> 37
- <211> 21
- <212> DNA
- <213> 玉米

<400> 37
tcggaccagg cttcattccc c

<210> 38

<211> 21
<212> DNA
<213> 合成

<220>
<221> misc_feature
<222> (1)..(21)
<223> 微 RNA

<220>
<221> misc_feature
<222> (1)..(21)
<223> 编码微 RNA 的 DNA 序列

<400> 38
tcggaccagg cttcattccc c 21

<210> 39
<211> 21
<212> DNA
<213> 人工序列

<220>
<223> 编码微 RNA 的 DNA 序列

<220>
<221> misc_feature
<222> (1)..(21)
<223> 微 RNA

<220>
<221> misc_feature
<222> (1)..(21)
<223> 编码微 RNA 的 DNA 序列

<220>
<221> misc_feature
<222> (1)..(21)
<223> 编码微 RNA 的 DNA 序列

<400> 39
ttgtagatga agcagccgtc c 21

<210> 40
<211> 19
<212> RNA
<213> 人工

<220>
<223> 微 RNA

<220>
<221> misc_feature
<222> (1)..(19)
<223> 微 RNA

<400> 40
accagacccc gaacgccgc

19

<210> 41
<211> 11722
<212> DNA
<213> 人工

<220>
<223> 带有玉米 miR166 前体的二元载体

<220>
<221> misc_feature
<222> (171)..(195)
<223> T-DNA 左边界重复区域

<220>
<221> 启动子
<222> (286)..(2273)
<223> 玉米泛素启动子

<220>
<221> CDS
<222> (2316)..(3644)
<223> 大肠杆菌 D-丝氨酸脱水酶[dsdA]基因

<220>
<221> 终止子
<222> (3784)..(4394)
<223> 章鱼碱(Octapine)合酶 3 终止子

<220>
 <221> 启动子
 <222> (4486)..(5906)
 <223> 甘蔗杆状病毒[ScBV]启动子

 <220>
 <221> 前体 RNA
 <222> (5977)..(6669)
 <223> 玉米 miR166 前体

 <220>
 <221> 终止子
 <222> (6732)..(6984)
 <223> 用作为终止子的根癌农杆菌 T-DNA 胭脂碱合酶的 3' UTR

<220>
 <221> misc_feature
 <222> (7079)..(7102)
 <223> T-DNA 右边界重复区域

<220>
 <221> 启动子
 <222> (7695)..(7793)
 <223> [bla] β -内酰胺酶启动子

<220>
 <221> 复制起点
 <222> (9460)..(10141)
 <223> 用于增殖的 ColE1 大肠杆菌复制起点

<220>
 <221> CDS
 <222> (10600)..(11415)
 <223> 卡那霉素抗性选择基因

<400> 41
 ttcaaaccgc gcagcttagt tgccgttctt ccgaatagca tcggtaacat gagcaaagtc 60

 tgccgcctta caacggctct cccgctgacg ccgtcccga ctgatgggct gcctgtatcg 120

 agtgggtgatt ttgtccgag ctgccggtcg gggagctgtt ggctggctgg tggcaggata 180

 tatttggtg taaacaaatt gacgcttaga caacttaata acacattgcg gacgttttta 240

 atgtactgaa ttgatccgc ccgggcggta ccaagcttc gcggctgcag tgcagcgtga 300

cccggtcgtg cccctctcta gagataatga gcattgcatg tctaagttat aaaaaattac	360
cacatatttt tttgtgcaca cttgtttgaa gtgcagttta tctatcttta tacatatatt	420
taaactttac tctacgaata atataatcta tagtactaca ataatatcag tgtttttagag	480
aatcatataa atgaacagtt agacatggtc taaaggacaa ttgagtattt tgacaacagg	540
actctacagt tttatctttt tagtgtgcat gtgttctcct ttttttttgc aaatagcttc	600
acctatataa tacttcatcc attttattag tacatccatt tagggtttag ggtaaatggt	660
ttttatagac taattttttt agtacatcta tttatttcta ttttagcctc taaattaaga	720
aaactaaaac tctatttttag tttttttatt taatagttta gatataaaat agaataaaat	780
aaagtgacta aaaattaaac aaataccctt taagaaatta aaaaaactaa ggaaacattt	840
ttcttgtttc gagtagataa tgccagcctg ttaaagccg tcgacgagtc taacggacac	900
caaccagcga accagcagcg tcgcgtcggg ccaagcgaag cagacggcac ggcatctctg	960
tcgctgcctc tggaccctc tcgagagttc cgctccaccg ttggacttgc tccgctgtcg	1020
gcatccagaa attgcgtggc ggagcggcag acgtgagccg gcacggcagg cggcctcctc	1080
ctcctctcac ggcaccggca gctacggggg attcctttcc caccgctcct tcgctttccc	1140
ttcctcgecc gccgtaataa atagacaccc cctccacacc ctctttcccc aacctcgtgt	1200
tgttcggagc gcacacacac acaaccagat ctccccaaa tccaccgtc ggcacctccg	1260
cttcaaggta cgccgctcgt cctccccccc cccccccctc tctaccttct ctagatcggc	1320
gttccggtcc atggttaggg cccggtagtt ctacttctgt tcatgtttgt gttagatccg	1380
tgtttgtgtt agatccgtgc tgctagcgtt cgtacacgga tgcgacctgt acgtcagaca	1440
cgttctgatt gctaacttgc cagtgtttct ctttggggaa tcctgggatg gctctagccg	1500
ttccgcagac gggatcgatt tcatgatttt tttgtttcg ttgcataggg tttggtttgc	1560
ccttttcctt tatttcaata tatgccgtgc acttgtttgt cgggtcatct tttcatgctt	1620

ttttttgtct tggttgtgat gatgtggtct ggttgggcgg tcgttctaga tcggagtaga	1680
attctgtttc aaactacctg gtggatttat taattttgga tctgtatgtg tgtgccatac	1740
atattcatag ttacgaattg aagatgatgg atggaaatat cgatctagga taggtataca	1800
tgttgatgcg ggttttactg atgcatatac agagatgctt tttgttcgct tggttgtgat	1860
gatgtggtgt ggttgggcgg tcgttcattc gttctagatc ggagtagaat actgtttcaa	1920
actacctggt gtatttatta attttgaac tgtatgtgtg tgtcatacat ctccatagtt	1980
acgagittaa gatggatgga aatatcgatc taggataggt atacatgttg atgtgggttt	2040
tactgatgca tatacatgat ggcatacgca gcatctattc atatgctcta accttgagta	2100
cctatctatt ataataaaca agtatgtttt ataattatct cgatcttgat atacttgat	2160
gatggcatat gcagcagcta tatgtggatt tttttagccc tgccttcata cgctatttat	2220
ttgcttggtta ctgtttcttt tgctgatgct caccctgttg tttggtgtta cttctgcagg	2280
gtacggatcc tcactaagc gcaaagagac gtact atg gaa aac gct aaa atg	2333
Met Glu Asn Ala Lys Met	
1 5	
aac tcg ctc atc gcc cag tat ccg ttg gta aag gat ctg gtt gct ctt	2381
Asn Ser Leu Ile Ala Gln Tyr Pro Leu Val Lys Asp Leu Val Ala Leu	
10 15 20	
aaa gaa acc acc tgg ttt aat cct ggc acg acc tca ttg gct gaa ggt	2429
Lys Glu Thr Thr Trp Phe Asn Pro Gly Thr Thr Ser Leu Ala Glu Gly	
25 30 35	
tta cct tat gtt ggc ctg acc gaa cag gat gtt cag gac gcc cat gcg	2477
Leu Pro Tyr Val Gly Leu Thr Glu Gln Asp Val Gln Asp Ala His Ala	
40 45 50	
cgc tta tcc cgt ttt gca ccc tat ctg gca aaa gca ttt cct gaa act	2525
Arg Leu Ser Arg Phe Ala Pro Tyr Leu Ala Lys Ala Phe Pro Glu Thr	
55 60 65 70	

gct gcc act ggg ggg att att gaa tca gaa ctg gtt gcc att cca gct Ala Ala Thr Gly Gly Ile Ile Glu Ser Glu Leu Val Ala Ile Pro Ala 75 80 85	2573
atg caa aaa cgg ctg gaa aaa gaa tat cag caa ccg atc agc ggg caa Met Gln Lys Arg Leu Glu Lys Glu Tyr Gln Gln Pro Ile Ser Gly Gln 90 95 100	2621
ctg tta ctg aaa aaa gat agc cat ttg ccc att tcc ggc tcc ata aaa Leu Leu Leu Lys Lys Asp Ser His Leu Pro Ile Ser Gly Ser Ile Lys 105 110 115	2669
gca cgc ggc ggg att tat gaa gtc ctg gca cac gca gaa aaa ctg gct Ala Arg Gly Gly Ile Tyr Glu Val Leu Ala His Ala Glu Lys Leu Ala 120 125 130	2717
ctg gaa gcg ggg ttg ctg acg ctt gat gat gac tac agc aaa ctg ctt Leu Glu Ala Gly Leu Leu Thr Leu Asp Asp Asp Tyr Ser Lys Leu Leu 135 140 145 150	2765
tct ccg gag ttt aaa cag ttc ttt agc caa tac agc att gct gtg ggc Ser Pro Glu Phe Lys Gln Phe Phe Ser Gln Tyr Ser Ile Ala Val Gly 155 160 165	2813
tca acc gga aat ctg ggg tta tca atc ggc att atg agc gcc cgc att Ser Thr Gly Asn Leu Gly Leu Ser Ile Gly Ile Met Ser Ala Arg Ile 170 175 180	2861
ggc ttt aag gtg aca gtt cat atg tct gct gat gcc cgg gca tgg aaa Gly Phe Lys Val Thr Val His Met Ser Ala Asp Ala Arg Ala Trp Lys 185 190 195	2909
aaa gcg aaa ctg cgc agc cat ggc gtt acg gtc gtg gaa tat gag caa Lys Ala Lys Leu Arg Ser His Gly Val Thr Val Val Glu Tyr Glu Gln 200 205 210	2957
gat tat ggt gtt gcc gtc gag gaa gga cgt aaa gca gcg cag tct gac Asp Tyr Gly Val Ala Val Glu Glu Gly Arg Lys Ala Ala Gln Ser Asp 215 220 225 230	3005
ccg aac tgt ttc ttt att gat gac gaa aat tcc cgc acg ttg ttc ctt Pro Asn Cys Phe Phe Ile Asp Asp Glu Asn Ser Arg Thr Leu Phe Leu 235 240 245	3053

ggg tat tcc gtc gct ggc cag cgt ctt aaa gcg caa ttt gcc cag caa Gly Tyr Ser Val Ala Gly Gln Arg Leu Lys Ala Gln Phe Ala Gln Gln 250 255 260	3101
ggc cgt atc gtc gat gct gat aac cct ctg ttt gtc tat ctg ccg tgt Gly Arg Ile Val Asp Ala Asp Asn Pro Leu Phe Val Tyr Leu Pro Cys 265 270 275	3149
ggt gtt ggc ggt ggt cct ggt ggc gtc gca ttc ggg ctt aaa ctg gcg Gly Val Gly Gly Gly Pro Gly Gly Val Ala Phe Gly Leu Lys Leu Ala 280 285 290	3197
ttt ggc gat cat gtt cac tgc ttt ttt gcc gaa cca acg cac tcc cct Phe Gly Asp His Val His Cys Phe Phe Ala Glu Pro Thr His Ser Pro 295 300 305 310	3245
tgt atg ttg tta ggc gtc cat aca gga tta cac gat cag att tct gtt Cys Met Leu Leu Gly Val His Thr Gly Leu His Asp Gln Ile Ser Val 315 320 325	3293
cag gat att ggt atc gac aac ctt acc gca gcg gat ggc ctt gca gtt Gln Asp Ile Gly Ile Asp Asn Leu Thr Ala Ala Asp Gly Leu Ala Val 330 335 340	3341
ggt cgc gca tca ggc ttt gtc ggg cgg gca atg gag cgt ctg ctg gat Gly Arg Ala Ser Gly Phe Val Gly Arg Ala Met Glu Arg Leu Leu Asp 345 350 355	3389
ggc ttc tat acc ctt agc gat caa acc atg tat gac atg ctt ggc tgg Gly Phe Tyr Thr Leu Ser Asp Gln Thr Met Tyr Asp Met Leu Gly Trp 360 365 370	3437
ctg gcg cag gaa gaa ggt att cgt ctt gaa cct tcg gca ctg gcg ggt Leu Ala Gln Glu Glu Gly Ile Arg Leu Glu Pro Ser Ala Leu Ala Gly 375 380 385 390	3485
atg gcc gga cct cag cgc gtg tgt gca tca gta agt tac caa cag atg Met Ala Gly Pro Gln Arg Val Cys Ala Ser Val Ser Tyr Gln Gln Met 395 400 405	3533
cac ggt ttc agc gca gaa caa ctg cgt aat acc act cat ctg gtg tgg His Gly Phe Ser Ala Glu Gln Leu Arg Asn Thr Thr His Leu Val Trp 410 415 420	3581

gcg acg gga ggt gga atg gtg ccg gaa gaa gag atg aat caa tat ctg Ala Thr Gly Gly Gly Met Val Pro Glu Glu Glu Met Asn Gln Tyr Leu 425 430 435	3629
gca aaa ggc cgt taa taacgtttca acgcagcatg gatcgtaccg agctcaatcg Ala Lys Gly Arg 440	3684
atcctgcttt aatgagatat gcgagacgcc tatgatcgca tgatatttgc tttcaattct	3744
gttgtgcacg ttgtaaaaaa cctgagcatg tgtagctcag atccttaccg ccggtttcgg	3804
ttcattctaa tgaatatatc acccgttact atcgtatitt tatgaataat attctccgtt	3864
caatttactg attgtaccct actacttata tgtacaatat taaaatgaaa acaatatatt	3924
gtgctgaata gggtttatagc gacatctatg atagagcgcc acaataacaa acaattgcgt	3984
tttattatta caaatccaat tttaaaaaaa gcggcagaac cgggtcaaacc taaaagactg	4044
attacataaa tcttattcaa atttcaaaag tgccccaggg gctagtattc acgacacacc	4104
gagcggcgaa ctaataacgc tcaactgaagg gaactccggg tccccgccgg cgcgcatggg	4164
tgagattcct tgaagttgag tattggccgt ccgctctacc gaaagttacg ggcaccattc	4224
aaccgggtcc agcacggcgg ccgggtaacc gacttgctgc cccgagaatt atgcagcatt	4284
tttttggtgt atgtgggccc caaatgaagt gcaggtcaaa cttgacagt gacgacaaat	4344
cgttgggagg gtccaggcgg aattttgcga caacatgtcg aggctcagca ggatgggccc	4404
aggtagagaa ttcgcggccg tacaacgcgt accggttaat taaattacgc caagctatca	4464
actttgtata gaaaagttga ggaagttgaa gacaaagaag gtcttaaata ctggctagca	4524
acactgaact atgccagaaa ccacatcaaa gatatgggca agcttcttgg ccattatat	4584
ccaaagacct cagagaaagg tgagcgaagg ctcaattcag aagattggaa gctgatcaat	4644
aggatcaaga caatggtgag aacgcttcca aatctcacta ttccaccaga agatgcatac	4704
attatcattg aaacagatgc atgtgcaact ggatggggag cagtatgcaa gtggaagaaa	4764
aacaaggcag acccaagaaa tacagagcaa atctgtaggt atgccagtgg aaaatttgat	4824

aagccaaaag gaacctgtga tgcagaaatc tatggggtta tgaatggctt agaaaagatg	4884
agattgttct acttggacaa aagagagatc acagtcagaa ctgacagtag tgcaatcgaa	4944
aggttctaca acaagagtgc tgaacacaag ctttctgaga tcagatggat caggttcatg	5004
gactacatca ctggtgcagg accagagata gtcattgaac acataaaagg gaagagcaat	5064
ggttttagctg acatcttgct caggctcaaa gccaaattag ctcagaatga accaacggaa	5124
gagatgatcc tgcttacaca agccataagg gaagtaattc cttatccaga tcatccatac	5184
actgagcaac tcagagaatg gggaaacaaa attctggatc cattccccac attcaagaag	5244
gacatgttcg aaagaacaga gcaagctttt atgctaacag aggaaccagt tctactctgt	5304
gcatgcagga agcctgcaat tcagttagtg tccagaacat ctgccaaccc aggaaggaaa	5364
ttcttcaagt gcgcaatgaa caaatgccat tgctgggtact gggcagatct cattgaagaa	5424
cacattcaag acagaattga tgaatttctc aagaatcttg aagttctgaa gaccggtggc	5484
gtgcaaaciaa tggaggagga acttatgaag gaagtcacca agctgaagat agaagagcag	5544
gagttcgagg aataccaggc cacaccaagg gctatgtcgc cagtagccgc agaagatgtg	5604
ctagatctcc aagacgtaag caatgacgat tgaggaggca ttgacgtcag ggatgaccgc	5664
agcggagagt actgggcccc ttcagtggat gctccactga gttgtattat tgtgtgcttt	5724
tcggacaagt gtgctgtcca ctttcttttg gcacctgtgc cactttattc cttgtctgcc	5784
acgatgcctt tgcttagctt gtaagcaagg atcgcagtgc gtgtgtgaca ccacccccct	5844
tccgacgctc tgcctatata aggcaccgtc tgtaagctct tacgatcatc ggtagttcac	5904
cacaagtttg tacaaaaaag caggctccga attcgccctt ggaattctta attaaactga	5964
gatatcggtc cgcgacgctc ggtggttgct tcatggctgt cgagggaat gacgtccggt	6024
ccgaacaagc cacggctgct gctgcgctac cgccgcggtc tcggaccagg cttcattccc	6084
cacgactcac catggaaacg gcagcgaggt gcgagcttcc agtgcgcgca gtgtttgatc	6144

agcaaacacc	tgcgtggccg	aggctgagcc	ttgcaagggc	aggcgtccat	gttctcgtca	6204
gcaagcgtct	tcttcacatg	tctgaggtct	gagcaattgc	gggtggacat	cgatggatga	6264
gcattgccgc	tgcaagggaa	gacgtccgg	ctcacacttg	agttgggata	ccatgcacta	6324
ctcttgttgt	agttagatta	gataccatgc	gcttgaaccg	tgcaatagca	tagcgaggaa	6384
gaggaaccca	gagaggactc	agaaagagat	gttgtctatt	tcattccagt	ctcatctagt	6444
catctagtct	gcctgcgggt	gtctgcaatg	aacgtgacg	ctgctgccgg	agacctactc	6504
tgctactggt	gtggagtgga	caaacagggc	gggaaggcgt	ctgctagaac	ctagaagacg	6564
gacacagggc	ggaccaacag	gactaaagac	cgctgtacgt	gttgaagcgg	aaggactttc	6624
ttattacgtg	tggcgcttcg	aattaaagcc	agcggttaga	acggcccggg	ccggccaagc	6684
ttgggaaggg	cgaattcgac	ccagctttct	tgtacaaagt	ggagctcgat	cgttcaaaca	6744
tttggcaata	aagtttctta	agattgaatc	ctgttgccgg	tcttgcatg	attatcatat	6804
aatttctgtt	gaattacgtt	aagcatgtaa	taattaacat	gtaatgcatg	acgttattta	6864
tgagatgggt	ttttatgatt	agagtccgc	aattatacat	ttaatacgcg	atagaaaaca	6924
aaatatagcg	cgcaaactag	gataaattat	cgcgcgcggt	gtcatctatg	ttactagatc	6984
ggccggccaa	ctttattata	catagttgat	aattcactgg	gccggccaga	tcttgattgt	7044
cgtttccgc	cttcagtita	aactatcagt	gtttgacagg	atatattggc	gggtaaacct	7104
aagagaaaag	agcgtttatt	agaataatcg	gatattttaa	agggcgtgaa	aaggtttatc	7164
cgttcgtcca	tttgatgtg	catgccaacc	acagggttcc	cctcgggagt	gcttggcatt	7224
ccgtgcgata	atgacttctg	ttcaaccacc	caaacgtcgg	aaagcctgac	gacggagcag	7284
cattccaaaa	agatcccttg	gtcgtcttg	gtcggctaga	aggtcgagt	ggctgctgtg	7344
gcttgatccc	tcaacgcggt	cgcggacgta	gcgcagcgcc	gaaaaatcct	cgatcgcaaa	7404
tccgacgctg	tcgaaaagcg	tgatctgctt	gtcgtctttt	cgcccgacgt	cctggccagt	7464
catcacgcgc	caaagtccg	tcacaggatg	atctggcgcg	agttgctgga	tctcgccttc	7524

aatccgggtc tgtggcgga actccacgaa aatatccgaa cgcagcaaga tcgtcgacca	7584
attcttgaag acgaaagggc ctctgtatac gcctattttt ataggttaat gtcataataa	7644
taatggtttc ttagacgtca ggtggcactt ttccgggaaa tgtgcgcgga acccctattt	7704
gtttattttt ctaaatacat tcaaatatgt atccgctcat gagacaataa ccctgataaa	7764
tgcttcaata atattgaaaa aggaagagta tgagtattca acatttcctg gtcgccctta	7824
ttcccttttt tgccgcattt tgccctcctg tttttgtca ccagaaacg ctggtgaaag	7884
taaaagatgc tgaagatcag ttgggtgcac gagggtgta catcgaactg gatctcaaca	7944
gcgtaagat ccttgagagt ttccgccccg aagaacgttt tccaatgatg agcactttta	8004
aagttctgt atgtggcgcg gtattatccc gtgttgacgc cgggcaagag caactcggtc	8064
gccgcataca ctattctcag aatgacttgg ttgagtactc accagtcaca gaaaagcatc	8124
ttacggatgg catgacagta agagaattat gcagtgtgc cataaccatg agtgataaca	8184
ctgcggccaa cttacttctg acaacgatcg gaggaccgaa ggagctaacc gcttttttgc	8244
acaacatggg ggatcatgta actcgccttg atcgttggga accggagctg aatgaagcca	8304
taccaaacga cgagcgtgac accacgatgc cggggggggg ggggggggac atgaggttgc	8364
cccgtattca gtgtcgctga ttgtattgt ctgaagttgt ttttacgtta agttgatgca	8424
gatcaattaa tacgatacct gcgtcataat tgattatttg acgtggtttg atggcctcca	8484
cgcacgttgt gatatgtaga tgataatcat tatcacttta cgggtccttt ccggtgatcc	8544
gacaggttac ggggcggcga cctcgcgggt ttctgctatt tatgaaaatt ttccggttta	8604
aggcgtttcc gttcttcttc gtcataactt aatgttttta tttaaaatac cctctgaaaa	8664
gaaaggaaac gacaggtgct gaaagcgagc tttttggcct ctgtcgtttc ctttctctgt	8724
ttttgtccgt ggaatgaaca atggaacccc ccccccccc ccctgcagca atggcaacaa	8784
cgttgcgcaa actattaact ggccaactac ttactctagc ttcccggcaa caattaatag	8844

actggatgga ggcggataaa gttgcaggac cacttctgcg ctcgccctt ccggctggct	8904
ggtttattgc tgataaatct ggagccgggtg agcgtgggtc tcgcggtatc attgcagcac	8964
tggggccaga tggtaagccc tcccgtatcg tagttatcta cacgacggg agtcaggcaa	9024
ctatggatga acgaaataga cagatcgctg agatagggtc ctactgatt aagcattggt	9084
aactgtcaga ccaagtttac tcatatatac tttagattga tttaaaactt cattttta	9144
ttaaaaggat ctaggtgaag atcctttttg ataatctcat gaccaaatac ccttaacgtg	9204
agttttcggt ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc	9264
ctttttttct gcgcgtaatc tgctgcttgc aaacaaaaa accaccgcta ccagcgggtg	9324
tttgtttggc ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag	9384
cgcagatacc aaatactgtc cttctagtgt agccgtagt aggccaccac ttcaagaact	9444
ctgtagcacc gcctacatac ctgctctgc taatcctgtt accagtggct gctgccagt	9504
gcgataagtc gtgtcttacc gggttggact caagacgata gttaccgat aaggcgcagc	9564
ggtcgggctg aacgggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg	9624
aactgagata cctacagcgt gagctatgag aaagcgccac gcttccgaa gggagaaagg	9684
cggacaggta tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag	9744
ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc	9804
gatttttgtg atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct	9864
ttttacggtt cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc	9924
ctgattctgt ggataaccgt attaccgct ttgagtgagc tgataccgct cgccgcagcc	9984
gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga agagcgctg atgcggtatt	10044
ttctccttac gcattctgtc ggtatttcac accgcatatg gtgcactctc agtacaatct	10104
gctctgatgc cgcatagtta agccagtata cactccgcta tcgctacgtg actgggtcat	10164
ggctgcgccc cgacaccgcg caacaccgcg tgacgcgccc tgacgggctt gtctgctccc	10224

```

ggcatccgct tacagacaag ctgtgaccgt ctccgggagc tgcattgtgc agaggttttc 10284

accgtcatca ccgaaacgcg cgaggcagca gatccccga tcaagtagat acactacata 10344

tatctacaat agacatcgag ccggaaggtg atgtttactt tcctgaaatc cccagcaatt 10404

ttaggccagt ttttacccaa gacttcgcct ctaacataaa ttatagttac caaatctggc 10464

aaaagggttg accggggggg gggggaaagc cacgttgtgt ctcaaatct ctgatgttac 10524

attgcacaag ataaaaatat atcatcatga acaataaaac tgtctgctta cataaacagt 10584

aatacaaggg gtgtt atg agc cat att caa cgg gaa acg tct tgc tcg agg 10635
      Met Ser His Ile Gln Arg Glu Thr Ser Cys Ser Arg
              445              450

ccg cga tta aat tcc aac atg gat gct gat tta tat ggg tat aaa tgg 10683
Pro Arg Leu Asn Ser Asn Met Asp Ala Asp Leu Tyr Gly Tyr Lys Trp
455              460              465              470

gct cgc gat aat gtc ggg caa tca ggt gcg aca atc tat cga ttg tat 10731
Ala Arg Asp Asn Val Gly Gln Ser Gly Ala Thr Ile Tyr Arg Leu Tyr
              475              480              485

ggg aag ccc gat gcg cca gag ttg ttt ctg aaa cat ggc aaa ggt agc 10779
Gly Lys Pro Asp Ala Pro Glu Leu Phe Leu Lys His Gly Lys Gly Ser
              490              495              500

gtt gcc aat gat gtt aca gat gag atg gtc aga cta aac tgg ctg acg 10827
Val Ala Asn Asp Val Thr Asp Glu Met Val Arg Leu Asn Trp Leu Thr
              505              510              515

gaa ttt atg cct ctt ccg acc atc aag cat ttt atc cgt act cct gat 10875
Glu Phe Met Pro Leu Pro Thr Ile Lys His Phe Ile Arg Thr Pro Asp
              520              525              530

gat gca tgg tta ctc acc act gcg atc ccc ggg aaa aca gca ttc cag 10923
Asp Ala Trp Leu Leu Thr Thr Ala Ile Pro Gly Lys Thr Ala Phe Gln
535              540              545              550

```

gta tta gaa gaa tat cct gat tca ggt gaa aat att gtt gat gcg ctg Val Leu Glu Glu Tyr Pro Asp Ser Gly Glu Asn Ile Val Asp Ala Leu 555 560 565	10971
gca gtg ttc ctg cgc cgg ttg cat tcg att cct gtt tgt aat tgt cct Ala Val Phe Leu Arg Arg Leu His Ser Ile Pro Val Cys Asn Cys Pro 570 575 580	11019
ttt aac agc gac cgc gta ttt cgt ctc gct cag gcg caa tca cga atg Phe Asn Ser Asp Arg Val Phe Arg Leu Ala Gln Ala Gln Ser Arg Met 585 590 595	11067
aat aac ggt ttg gtt gat gcg agt gat ttt gat gac gag cgt aat ggc Asn Asn Gly Leu Val Asp Ala Ser Asp Phe Asp Asp Glu Arg Asn Gly 600 605 610	11115
tgg cct gtt gaa caa gtc tgg aaa gaa atg cat aag ctt ttg cca ttc Trp Pro Val Glu Gln Val Trp Lys Glu Met His Lys Leu Leu Pro Phe 615 620 625 630	11163
tca ccg gat tca gtc gtc act cat ggt gat ttc tca ctt gat aac ctt Ser Pro Asp Ser Val Val Thr His Gly Asp Phe Ser Leu Asp Asn Leu 635 640 645	11211
att ttt gac gag ggg aaa tta ata ggt tgt att gat gtt gga cga gtc Ile Phe Asp Glu Gly Lys Leu Ile Gly Cys Ile Asp Val Gly Arg Val 650 655 660	11259
gga atc gca gac cga tac cag gat ctt gcc atc cta tgg aac tgc ctc Gly Ile Ala Asp Arg Tyr Gln Asp Leu Ala Ile Leu Trp Asn Cys Leu 665 670 675	11307
ggt gag ttt tct cct tca tta cag aaa cgg ctt ttt caa aaa tat ggt Gly Glu Phe Ser Pro Ser Leu Gln Lys Arg Leu Phe Gln Lys Tyr Gly 680 685 690	11355
att gat aat cct gat atg aat aaa ttg cag ttt cat ttg atg ctc gat Ile Asp Asn Pro Asp Met Asn Lys Leu Gln Phe His Leu Met Leu Asp 695 700 705 710	11403
gag ttt ttc taa tcagaattgg ttaattggtt gtaacactgg cagagcatta Glu Phe Phe	11455
cgctgacttg acgggacggc ggctttgttg aataaatcga acttttgctg agttgaagga	11515

tcagatcacg catcttcccg acaacgcaga ccgttccgtg gcaaagcaaa agttcaaaat 11575
 caccaactgg tccacctaca acaaagctct catcaaccgt ggctccctca ctttctggct 11635
 ggatgatggg gcgattcagg gatcacaggc agcaacgctc tgtcatcggt acaatcaaca 11695
 tgctaccctc cgcgagatca tccgtgt 11722

<210> 42
 <211> 442
 <212> PRT
 <213> 人工

<220>
 <223> 合成构建体

<400> 42

Met Glu Asn Ala Lys Met Asn Ser Leu Ile Ala Gln Tyr Pro Leu Val
 1 5 10 15

Lys Asp Leu Val Ala Leu Lys Glu Thr Thr Trp Phe Asn Pro Gly Thr
 20 25 30

Thr Ser Leu Ala Glu Gly Leu Pro Tyr Val Gly Leu Thr Glu Gln Asp
 35 40 45

Val Gln Asp Ala His Ala Arg Leu Ser Arg Phe Ala Pro Tyr Leu Ala
 50 55 60

Lys Ala Phe Pro Glu Thr Ala Ala Thr Gly Gly Ile Ile Glu Ser Glu
 65 70 75 80

Leu Val Ala Ile Pro Ala Met Gln Lys Arg Leu Glu Lys Glu Tyr Gln
 85 90 95

Gln Pro Ile Ser Gly Gln Leu Leu Lys Lys Asp Ser His Leu Pro
 100 105 110

Ile Ser Gly Ser Ile Lys Ala Arg Gly Gly Ile Tyr Glu Val Leu Ala
 115 120 125

His Ala Glu Lys Leu Ala Leu Glu Ala Gly Leu Leu Thr Leu Asp Asp
 130 135 140

Asp Tyr Ser Lys Leu Leu Ser Pro Glu Phe Lys Gln Phe Phe Ser Gln
 145 150 155 160

Tyr Ser Ile Ala Val Gly Ser Thr Gly Asn Leu Gly Leu Ser Ile Gly
 165 170 175

Ile Met Ser Ala Arg Ile Gly Phe Lys Val Thr Val His Met Ser Ala
 180 185 190

Asp Ala Arg Ala Trp Lys Lys Ala Lys Leu Arg Ser His Gly Val Thr
 195 200 205

Val Val Glu Tyr Glu Gln Asp Tyr Gly Val Ala Val Glu Glu Gly Arg
 210 215 220

Lys Ala Ala Gln Ser Asp Pro Asn Cys Phe Phe Ile Asp Asp Glu Asn
 225 230 235 240

Ser Arg Thr Leu Phe Leu Gly Tyr Ser Val Ala Gly Gln Arg Leu Lys
 245 250 255

Ala Gln Phe Ala Gln Gln Gly Arg Ile Val Asp Ala Asp Asn Pro Leu
 260 265 270

Phe Val Tyr Leu Pro Cys Gly Val Gly Gly Gly Pro Gly Gly Val Ala
 275 280 285

Phe Gly Leu Lys Leu Ala Phe Gly Asp His Val His Cys Phe Phe Ala
 290 295 300

Glu Pro Thr His Ser Pro Cys Met Leu Leu Gly Val His Thr Gly Leu
 305 310 315 320

His Asp Gln Ile Ser Val Gln Asp Ile Gly Ile Asp Asn Leu Thr Ala
 325 330 335

Ala Asp Gly Leu Ala Val Gly Arg Ala Ser Gly Phe Val Gly Arg Ala
 340 345 350

Met Glu Arg Leu Leu Asp Gly Phe Tyr Thr Leu Ser Asp Gln Thr Met
 355 360 365

Tyr Asp Met Leu Gly Trp Leu Ala Gln Glu Glu Gly Ile Arg Leu Glu
 370 375 380

Pro Ser Ala Leu Ala Gly Met Ala Gly Pro Gln Arg Val Cys Ala Ser
385 390 395 400

Val Ser Tyr Gln Gln Met His Gly Phe Ser Ala Glu Gln Leu Arg Asn
405 410 415

Thr Thr His Leu Val Trp Ala Thr Gly Gly Gly Met Val Pro Glu Glu
420 425 430

Glu Met Asn Gln Tyr Leu Ala Lys Gly Arg
435 440

<210> 43

<211> 271

<212> PRT

<213> 人工

<220>

<223> 合成构建体

<400> 43

Met Ser His Ile Gln Arg Glu Thr Ser Cys Ser Arg Pro Arg Leu Asn
1 5 10 15

Ser Asn Met Asp Ala Asp Leu Tyr Gly Tyr Lys Trp Ala Arg Asp Asn
20 25 30

Val Gly Gln Ser Gly Ala Thr Ile Tyr Arg Leu Tyr Gly Lys Pro Asp
35 40 45

Ala Pro Glu Leu Phe Leu Lys His Gly Lys Gly Ser Val Ala Asn Asp
50 55 60

Val Thr Asp Glu Met Val Arg Leu Asn Trp Leu Thr Glu Phe Met Pro
65 70 75 80

Leu Pro Thr Ile Lys His Phe Ile Arg Thr Pro Asp Asp Ala Trp Leu
85 90 95

Leu Thr Thr Ala Ile Pro Gly Lys Thr Ala Phe Gln Val Leu Glu Glu
100 105 110

Tyr Pro Asp Ser Gly Glu Asn Ile Val Asp Ala Leu Ala Val Phe Leu
 115 120 125

Arg Arg Leu His Ser Ile Pro Val Cys Asn Cys Pro Phe Asn Ser Asp
 130 135 140

Arg Val Phe Arg Leu Ala Gln Ala Gln Ser Arg Met Asn Asn Gly Leu
 145 150 155 160

Val Asp Ala Ser Asp Phe Asp Asp Glu Arg Asn Gly Trp Pro Val Glu
 165 170 175

Gln Val Trp Lys Glu Met His Lys Leu Leu Pro Phe Ser Pro Asp Ser
 180 185 190

Val Val Thr His Gly Asp Phe Ser Leu Asp Asn Leu Ile Phe Asp Glu
 195 200 205

Gly Lys Leu Ile Gly Cys Ile Asp Val Gly Arg Val Gly Ile Ala Asp
 210 215 220

Arg Tyr Gln Asp Leu Ala Ile Leu Trp Asn Cys Leu Gly Glu Phe Ser
 225 230 235 240

Pro Ser Leu Gln Lys Arg Leu Phe Gln Lys Tyr Gly Ile Asp Asn Pro
 245 250 255

Asp Met Asn Lys Leu Gln Phe His Leu Met Leu Asp Glu Phe Phe
 260 265 270

<210> 44

<211> 11722

<212> DNA

<213> 人工

<220>

<223> 带有玉米 miR166 前体的二元载体

<220>

<221> promoter

<222> (325)..(2312)

<223> 玉米泛素启动子

<220>

-
- <221> CDS
<222> (2355)..(3683)
<223> 大肠杆菌 D-丝氨酸脱水酶[dsdA]基因
- <220>
<221> 终止子
<222> (3721)..(4433)
<223> 章鱼碱合酶 3 终止子
- <220>
<221> 启动子
<222> (3823)..(4433)
<223> 甘蔗杆状病毒[ScBV]启动子
- <220>
<221> 前体 RNA
<222> (6016)..(6708)
<223> 设计的玉米 miR166 前体, 其中 21-nt 的 miR166 序列被 21-nt 的 dsRed 靶向 dsRed mRNA 所取代
- <220>
<221> 终止子
<222> (6771)..(7023)
<223> 用作为终止子的根癌农杆菌 T-DNA 胭脂碱合酶的 3' UTR
- <220>
<221> misc_feature
<222> (6771)..(7023)
<223> T-DNA 右边界重复区域
- <220>
<221> 启动子
<222> (7118)..(7141)
<223> [bla] β -内酰胺酶启动子
- <220>
<221> 复制起点
<222> (9500)..(9700)
<223> 用于增殖的 ColE1 大肠杆菌复制起点
- <220>
<221> CDS
<222> (10639)..(11454)
<223> 卡那霉素抗性选择基因

<220>

<221> misc_feature

<222> (10639)..(11454)

<223> T-DNA 左边界重复区域

<400> 44

ttacaatcaa catgctaccc tccgcgagat catccgtggt tcaaaccgg cagcttagtt	60
gccgttcttc cgaatagcat cggtaacatg agcaaagtct gccgccttac aacggctctc	120
ccgctgacgc cgtcccggac tgatgggctg cctgtatcga gtggtgattt tgtgccgagc	180
tgccggtcgg ggagctgttg gctggctggt ggcaggatat attgtggtgt aaacaaattg	240
acgcttagac aacttaataa cacattgcgg acgttttttaa tgtactgaat tggatccgcc	300
cgggcggtac caagcttccg cggctgcagt gcacgctgac ccggtcgtgc ccctctctag	360
agataatgag cattgcatgt ctaagttata aaaaattacc acatatTTTT tttgtcacac	420
ttgtttgaag tgcagtttat ctatctttat acatatattt aaactttact ctacgaataa	480
tataatctat agtactacaa taatatcagt gttttagaga atcatataaa tgaacagtta	540
gacatggtct aaaggacaat tgagtatttt gacaacagga ctctacagtt ttatcttttt	600
agtgatcatg tgttctcctt tttttttgca aatagcttca cctatataat acttcatcca	660
ttttattagt acatccattt agggtttagg gttaatggtt tttatagact aattttttta	720
gtacatctat tttattctat tttagcctct aaattaagaa aactaaaact ctattttagt	780
ttttttattt aatagtttag atataaaata gaataaaata aagtgactaa aaattaaaca	840
aatacccttt aagaaattaa aaaaactaag gaaacatttt tcttgtttcg agtagataat	900
gccagcctgt taaacgccgt cgacgagtct aacggacacc aaccagcgaa ccagcagcgt	960
cgcgtcgggc caagcgaagc agacggcacg gcattctctgt cgctgcctct ggacccctct	1020
cgagagtcc gctccaccgt tggacttgct ccgctgtcgg catccagaaa ttgcgtggcg	1080
gagcggcaga cgtgagccgg cacggcaggc ggcctcctcc tcctctcacg gcaccggcag	1140

ctacggggga ttcctttccc accgttcctt cgctttccct tcctcgcccg ccgtaataaa	1200
tagacacccc ctccacaccc tctttcccca acctcgtgtt gttcggagcg cacacacaca	1260
caaccagatc tccccaaat ccaccgctcg gcacctccgc ttcaaggtag gccgctcgtc	1320
ctcccccccc cccccctct ctacctctc tagatcggcg ttccggtcca tggtagggc	1380
cggtagttc tactttctgt catgtttgtg ttagatccgt gtttgtgtta gatccgtgct	1440
gctagcgttc gtacacggat gcgacctgta cgtcagacac gttctgattg ctaacttgcc	1500
agtgtttctc ttgggggaat cctgggatgg ctctagccgt tccgcagacg ggatcgattt	1560
catgattttt ttgtttcgt tgcatagggt ttggtttgcc cttttccttt atttcaatat	1620
atgccgtgca cttgtttgtc gggcatcctt ttcatgcttt tttttgtctt ggttgtgatg	1680
atgtggtctg gttgggcggg cgttctagat cggagtagaa ttctgtttca aactacctgg	1740
tggatttatt aattttggat ctgtatgtgt gtgccataca tattcatagt tacgaattga	1800
agatgatgga tggaaatac gatctaggat aggtatacat gttgatgcgg gttttactga	1860
tgcataatac gagatgcttt ttgttcgctt ggttgtgatg atgtggtgtg gttgggcggg	1920
cgttcattcg ttctagatcg gagtagaata ctgtttcaaa ctacctggtg tatttattaa	1980
ttttggaact gtatgtgtgt gtcataacac ttcatagtta cgagtttaag atggatggaa	2040
atatcgatct aggataggta tacatgttga tgtgggtttt actgatgcat atacatgatg	2100
gcatatgcag catctattca tatgtcttaa ccttgagtac ctatctatta taataaaca	2160
gtatgtttta taattatttc gatcttgata tacttgatg atggcatatg cagcagctat	2220
atgtggattt ttttagccct gccttcatac gctatttatt tgcttggtac tgtttctttt	2280
gtcgatgctc accctgttgt ttggtgttac ttctgcaggg tacggatcct catctaagcg	2340
caaagagacg tact atg gaa aac gct aaa atg aac tcg ctc atc gcc cag	2390
Met Glu Asn Ala Lys Met Asn Ser Leu Ile Ala Gln	
1 5 10	

tat ccg ttg gta aag gat ctg gtt gct ctt aaa gaa acc acc tgg ttt	2438
Tyr Pro Leu Val Lys Asp Leu Val Ala Leu Lys Glu Thr Thr Trp Phe	
15 20 25	
aat cct ggc acg acc tca ttg gct gaa ggt tta cct tat gtt ggc ctg	2486
Asn Pro Gly Thr Thr Ser Leu Ala Glu Gly Leu Pro Tyr Val Gly Leu	
30 35 40	
acc gaa cag gat gtt cag gac gcc cat gcg cgc tta tcc cgt ttt gca	2534
Thr Glu Gln Asp Val Gln Asp Ala His Ala Arg Leu Ser Arg Phe Ala	
45 50 55 60	
ccc tat ctg gca aaa gca ttt cct gaa act gct gcc act ggg ggg att	2582
Pro Tyr Leu Ala Lys Ala Phe Pro Glu Thr Ala Ala Thr Gly Gly Ile	
65 70 75	
att gaa tca gaa ctg gtt gcc att cca gct atg caa aaa cgg ctg gaa	2630
Ile Glu Ser Glu Leu Val Ala Ile Pro Ala Met Gln Lys Arg Leu Glu	
80 85 90	
aaa gaa tat cag caa ccg atc agc ggg caa ctg tta ctg aaa aaa gat	2678
Lys Glu Tyr Gln Gln Pro Ile Ser Gly Gln Leu Leu Leu Lys Lys Asp	
95 100 105	
agc cat ttg ccc att tcc ggc tcc ata aaa gca cgc ggc ggg att tat	2726
Ser His Leu Pro Ile Ser Gly Ser Ile Lys Ala Arg Gly Gly Ile Tyr	
110 115 120	
gaa gtc ctg gca cac gca gaa aaa ctg gct ctg gaa gcg ggg ttg ctg	2774
Glu Val Leu Ala His Ala Glu Lys Leu Ala Leu Glu Ala Gly Leu Leu	
125 130 135 140	
acg ctt gat gat gac tac agc aaa ctg ctt tct ccg gag ttt aaa cag	2822
Thr Leu Asp Asp Asp Tyr Ser Lys Leu Leu Ser Pro Glu Phe Lys Gln	
145 150 155	
ttc ttt agc caa tac agc att gct gtg ggc tca acc gga aat ctg ggg	2870
Phe Phe Ser Gln Tyr Ser Ile Ala Val Gly Ser Thr Gly Asn Leu Gly	
160 165 170	
tta tca atc ggc att atg agc gcc cgc att ggc ttt aag gtg aca gtt	2918
Leu Ser Ile Gly Ile Met Ser Ala Arg Ile Gly Phe Lys Val Thr Val	
175 180 185	

cat atg tct gct gat gcc cgg gca tgg aaa aaa gcg aaa ctg cgc agc His Met Ser Ala Asp Ala Arg Ala Trp Lys Lys Ala Lys Leu Arg Ser 190 195 200	2966
cat ggc gtt acg gtc gtg gaa tat gag caa gat tat ggt gtt gcc gtc His Gly Val Thr Val Val Glu Tyr Glu Gln Asp Tyr Gly Val Ala Val 205 210 215 220	3014
gag gaa gga cgt aaa gca gcg cag tct gac ccg aac tgt ttc ttt att Glu Glu Gly Arg Lys Ala Ala Gln Ser Asp Pro Asn Cys Phe Phe Ile 225 230 235	3062
gat gac gaa aat tcc cgc acg ttg ttc ctt ggg tat tcc gtc gct ggc Asp Asp Glu Asn Ser Arg Thr Leu Phe Leu Gly Tyr Ser Val Ala Gly 240 245 250	3110
cag cgt ctt aaa gcg caa ttt gcc cag caa ggc cgt atc gtc gat gct Gln Arg Leu Lys Ala Gln Phe Ala Gln Gln Gly Arg Ile Val Asp Ala 255 260 265	3158
gat aac cct ctg ttt gtc tat ctg ccg tgt ggt gtt ggc ggt ggt cct Asp Asn Pro Leu Phe Val Tyr Leu Pro Cys Gly Val Gly Gly Gly Pro 270 275 280	3206
ggt ggc gtc gca ttc ggg ctt aaa ctg gcg ttt ggc gat cat gtt cac Gly Gly Val Ala Phe Gly Leu Lys Leu Ala Phe Gly Asp His Val His 285 290 295 300	3254
tgc ttt ttt gcc gaa cca acg cac tcc cct tgt atg ttg tta ggc gtc Cys Phe Phe Ala Glu Pro Thr His Ser Pro Cys Met Leu Leu Gly Val 305 310 315	3302
cat aca gga tta cac gat cag att tct gtt cag gat att ggt atc gac His Thr Gly Leu His Asp Gln Ile Ser Val Gln Asp Ile Gly Ile Asp 320 325 330	3350
aac ctt acc gca gcg gat ggc ctt gca gtt ggt cgc gca tca ggc ttt Asn Leu Thr Ala Ala Asp Gly Leu Ala Val Gly Arg Ala Ser Gly Phe 335 340 345	3398
gtc ggg cgg gca atg gag cgt ctg ctg gat ggc ttc tat acc ctt agc Val Gly Arg Ala Met Glu Arg Leu Leu Asp Gly Phe Tyr Thr Leu Ser 350 355 360	3446

gat caa acc atg tat gac atg ctt ggc tgg ctg gcg cag gaa gaa ggt Asp Gln Thr Met Tyr Asp Met Leu Gly Trp Leu Ala Gln Glu Glu Gly 365 370 375 380	3494
att cgt ctt gaa cct tcg gca ctg gcg ggt atg gcc gga cct cag cgc Ile Arg Leu Glu Pro Ser Ala Leu Ala Gly Met Ala Gly Pro Gln Arg 385 390 395	3542
gtg tgt gca tca gta agt tac caa cag atg cac ggt ttc agc gca gaa Val Cys Ala Ser Val Ser Tyr Gln Gln Met His Gly Phe Ser Ala Glu 400 405 410	3590
caa ctg cgt aat acc act cat ctg gtg tgg gcg acg gga ggt gga atg Gln Leu Arg Asn Thr Thr His Leu Val Trp Ala Thr Gly Gly Gly Met 415 420 425	3638
gtg ccg gaa gaa gag atg aat caa tat ctg gca aaa ggc cgt taa Val Pro Glu Glu Glu Met Asn Gln Tyr Leu Ala Lys Gly Arg 430 435 440	3683
taacgtttca acgcagcatg gatcgtagcg agctcaatcg atcctgcttt aatgagatat	3743
gcgagacgcc tatgatcgca tgatatattgc ttccaattct gttgtgcacg ttgtaaaaaa	3803
cctgagcatg tgtagctcag atccttaccg ccggtttcgg ttcattctaa tgaatatatc	3863
accggttact atcgatatatt tatgaataat attctccgtt caatttactg attgtaccct	3923
actacttata tgtacaatat taaaatgaaa acaatatatt gtgctgaata ggtttatagc	3983
gacatctatg atagagcgcc acaataacaa acaattgcgt ttattattta caaatccaat	4043
tttaaaaaaa gcggcagaac cgttcaaacc taaaagactg attacataaa tcttattcaa	4103
atttcaaaaag tgccccaggg gctagtatct acgacacacc gagcggcgaa ctaataacgc	4163
tcactgaagg gaactccggt tccccgccgg cgcgcatggg tgagattcct tgaagttgag	4223
tattggccgt ccgctctacc gaaagttacg ggcaccattc aaccgggtcc agcacggcgg	4283
ccgggtaacc gacttgctgc cccgagaatt atgcagcatt tttttggtgt atgtggggcc	4343
caaatgaagt gcaggtcaaa ccttgacagt gacgacaaat cggtgggcgg gtccagggcg	4403
aattttgcga caacatgtcg aggctcagca ggatggggcc aggtacagaa ttcgcgcccg	4463

tacaacgcgt accggttaat taaattacgc caagctatca actttgtata gaaaagttga	4523
ggaagttgaa gacaaagaag gtcttaaadc ctggctagca aactgaact atgccagaaa	4583
ccacatcaaa gatatgggca agcttcttgg ccattatat ccaaagacct cagagaaagg	4643
tgagcgaagg ctcaattcag aagattggaa gctgatcaat aggatcaaga caatggtgag	4703
aacgcttcca aatctcacta ttccaccaga agatgcatac attatcattg aaacagatgc	4763
atgtgcaact ggatggggag cagtatgcaa gtggaagaaa aacaaggcag acccaagaaa	4823
tacagagcaa atctgtaggt atgccagtgg aaaatttgat aagccaaaag gaacctgtga	4883
tgcagaaatc tatggggtta tgaatggctt agaaaagatg agattgttct acttgacaa	4943
aagagagatc acagtcagaa ctgacagtag tgcaatcgaa aggttctaca acaagagtgc	5003
tgaacacaag ccttctgaga tcagatggat caggttcattg gactacatca ctggtgcagg	5063
accagagata gtcattgaac acataaaagg gaagagcaat ggtttagctg acatcttgc	5123
caggctcaaa gccaaattag ctccagaatga accaacggaa gagatgatcc tgcttacaca	5183
agccataagg gaagtaattc cttatccaga tcatccatac actgagcaac tcagagaatg	5243
gggaaacaaa attctggatc cattccccac attcaagaag gacatgttcg aaagaacaga	5303
gcaagctttt atgctaacag aggaaccagt tctactctgt gcatgcagga agcctgcaat	5363
tcagttagtg tccagaacat ctgccaaccc aggaaggaaa ttcttcaagt gcgcaatgaa	5423
caaatgcat tgctggtact gggcagatct cattgaagaa cacattcaag acagaattga	5483
tgaatttctc aagaatcttg aagtcttgaa gaccggtggc gtgcaaaca tggaggagga	5543
acttatgaag gaagtcacca agctgaagat agaagagcag gagttcgagg aataccaggc	5603
cacaccaagg gctatgtcgc cagtagccgc agaagatgtg ctagatctcc aagacgtaag	5663
caatgacgat tgaggaggca ttgacgtcag ggatgaccgc agcggagagt actgggcca	5723
ttcagtggat gctccactga gttgtattat tgtgtgctt tcggacaagt gtgctgtcca	5783

ctttcttttg gcacctgtgc cactttattc cttgtctgcc acgatgcctt tgcttagctt	5843
gtaagcaagg atcgcagtgc gtgtgtgaca ccacccccct tccgacgctc tgcctatata	5903
aggcacccgtc tgtaagctct tacgatcatc ggtagttcac cacaagtttg tacaaaaaag	5963
caggctccga attcgccctt ggaattctta attaaactgca gatatcggtc ccgcagcgtc	6023
ggtggttgct tcatggctgt cgaggacggc ggcttcatct acaacaagc cacggctgct	6083
gtcgcgtac cgcccggtt ttgtagatga agcagccgtc cacgactcac catggaaacg	6143
gcagcgaggt gcgagcttcc agtgcgcga gtgtttgatc agcaaacc tgcgtggccg	6203
aggctgagcc ttgcaagggc aggcgtccat gtctctgtca gcaagcgtct tcttcacatg	6263
tctgaggtct gagcaattgc ggggtggacat cgatggatga gcattgccgc tgcaaggga	6323
gatcgtccgg ctacacttg agttgggata ccatgcacta ctcttgttgt agttagatta	6383
gataccatgc gcttgaaccg tgcaatagca tagcgaggaa gaggaacca gagaggactc	6443
agaaagagat gttgtctatt tcatccagt ctcactagt catctagtct gcctgcgggt	6503
gtctgcaatg aacgctgacg ctgctgccgg agacctactc tgctactggt gtggagtgga	6563
caaacaggga gggaaggcgt ctgctagaac ctagaagacg gacacaggga ggaccaacag	6623
gactaaagac cgtcgtacgt gttgaagcgg aaggactttc ttattacgtg tggcgcttcg	6683
aattaaagcc agcggttaga acggcccggg ccggccaagc ttgggaaggc cgaattcgac	6743
ccagctttct tgtacaaagt ggagctcgat cgttcaaaca ttggcaata aagtttctta	6803
agattgaatc ctggtgccgg tcttgcgatg attatcatat aatttctgtt gaattacgtt	6863
aagcatgtaa taattaacat gtaatgatg acgttattta tgagatgggt ttttatgatt	6923
agagtccgc aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaactag	6983
gataaattat cgcgcgcggt gtcactatg ttactagatc ggccggccaa ctttattata	7043
catagttgat aattcactgg gccggccaga tcttgattgt cgtttccgc cttcagttta	7103

aactatcagt gtttgacagg atatattggc gggtaaacct aagagaaaag agcgtttatt	7163
agaataatcg gatatttaaa agggcgtgaa aaggtttatc cgttcgtcca tttgtatgtg	7223
catgccaacc acagggttcc cctcgggagt gcttggcatt ccgtgcgata atgacttctg	7283
ttcaaccacc caaacgtcgg aaagcctgac gacggagcag cattccaaaa agatcccttg	7343
gctcgtcttg gtcggctaga aggtcgagt ggctgctgtg gcttgatccc tcaacgcggt	7403
cgcggacgta gcgcagcgcc gaaaaatcct cgatcgcaaa tccgacgtg tcgaaaagcg	7463
tgatctgctt gtcgctcttt cggccgacgt cctggccagt catcacgcgc caaagttccg	7523
tcacaggatg atctggcgcg agttgctgga tctcgccttc aatccgggtc tgtggcggga	7583
actccacgaa aatatccgaa cgcagcaaga tcgtcgacca attcttgaag acgaaagggc	7643
ctcgtgatac gcctatTTTT ataggTTaat gtcattgataa taatggTTtc ttagacgtca	7703
ggTggcactt ttcggggaaa tgtgcgcgga acccctatTT gtttattTTT ctaaatacat	7763
tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata atattgaaaa	7823
aggaagagta tgagtattca acatttccgt gtcgccctta ttccctTTTT tgcggcattt	7883
tgccttcctg tttttgtca cccagaaacg ctggtgaaag taaaagatgc tgaagatcag	7943
ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaaag ccttgagagt	8003
tttcgccccg aagaacgttt tccaatgatg agcactTTta aagttctgct atgtggcgcg	8063
gtattatccc gtgttgacgc cgggcaagag caactcggtc gccgcataca ctattctcag	8123
aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg catgacagta	8183
agagaattat gcagtgtgc cataaccatg agtgataaca ctgcggccaa cttacttctg	8243
acaacgatcg gaggaccgaa ggagctaacc gctTTTTTgc acaacatggg ggatcatgta	8303
actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga cgagcgtgac	8363
accacgatgc cggggggggg ggggggggac atgaggttgc cccgtattca gtgtcgtga	8423
tttgtattgt ctgaagttgt ttttacgtta agttgatgca gatcaattaa tacgatacct	8483

gcgtcataat tgattatttg acgtggtttg atggcctcca cgcacgttgt gatatgtaga	8543
tgataatcat taccacttta cgggtccttt ccggtgatcc gacaggttac ggggcggcga	8603
cctcgcggtt tttcgctatt tatgaaaatt ttccggttta aggcgtttcc gttcttcttc	8663
gtcataactt aatgttttta tttaaaatac cctctgaaaa gaaaggaaac gacaggtgct	8723
gaaagcgagc tttttggcct ctgtcgtttc cttctctgt ttttgtccgt ggaatgaaca	8783
atggaacccc ccccccccc ccctgcagca atggcaaca cgttgcgcaa actattaact	8843
ggcgaactac ttactctagc ttcccgcaa caattaatag actggatgga ggcggataaa	8903
gttgaggac cacttctgag ctccggcctt ccggtggct ggtttattgc tgataaatct	8963
ggagccggtg agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggttaagccc	9023
tccggtatcg tagttatcta cagcagggg agtcaggcaa ctatggatga acgaaataga	9083
cagatcgctg agatagggtc ctactgatt aagcattggt aactgtcaga ccaagtttac	9143
tcatatatac tttagattga tttaaaactt catTTTTaat ttaaaggat ctaggtgaag	9203
atcctttttg ataatctcat gacaaaaac ccttaacgtg agttttcgtt cactgagcg	9263
tcagaccccg tagaaaagat caaaggatct tcttgagatc cttttttct gcgcgtaac	9323
tgctgcttgc aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc ggatcaagag	9383
ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtc	9443
cttctagtgt agccgtagt aggccaccac ttcaagaact ctgtagcacc gcctacatac	9503
ctcgctctgc taatcctgtt accagtggct gctgccagt gcgataagtc gtgtcttacc	9563
gggttggact caagacgata gttaccgat aaggcgcagc ggtcgggctg aacggggggt	9623
tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt	9683
gagctatgag aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggttaagc	9743
ggcagggtcg gaacaggaga gcgcacgagg gagcttcag ggggaaacgc ctggtatctt	9803

tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca	9863
ggggggcgga gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt	9923
tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt	9983
attaccgcct ttgagtgagc tgataccgct cgccgcagcc gaacgaccga gcgcagcgag	10043
tcagtgagcg aggaagcgga agagcgctg atgcggtatt ttctccttac gcatctgtgc	10103
ggtatttcac accgcatatg gtgcactctc agtacaatct gctctgatgc cgcatagtta	10163
agccagtata cactccgcta tcgctacgtg actgggtcat ggctgcgccc cgacaccgc	10223
caacaccgcg tgacgcgccc tgacgggctt gtctgctccc ggcatccgct tacagacaag	10283
ctgtgaccgt ctccgggagc tgcattgtgc agaggttttc accgtcatca ccgaaacgcg	10343
cgaggcagca gatccccga tcaagtagat aactacata tatctacaat agacatcgag	10403
ccggaaggtg atgtttactt tcctgaaatc cccagcaatt ttaggccagt ttttaccaa	10463
gacttcgcct ctaacataaa ttatagttac caaatctggc aaaagggttg accggggggg	10523
gggggaaagc cacgttgtgt ctcaaatct ctgatgttac attgcacaag ataaaaatat	10583
atcatcatga acaataaaac tgtctgctta cataaacagt aatacaaggg gtgtt atg	10641
	Met
agc cat att caa cgg gaa acg tct tgc tcg agg ccg cga tta aat tcc	10689
Ser His Ile Gln Arg Glu Thr Ser Cys Ser Arg Pro Arg Leu Asn Ser	
445 450 455	
aac atg gat gct gat tta tat ggg tat aaa tgg gct cgc gat aat gtc	10737
Asn Met Asp Ala Asp Leu Tyr Gly Tyr Lys Trp Ala Arg Asp Asn Val	
460 465 470 475	
ggg caa tca ggt gcg aca atc tat cga ttg tat ggg aag ccc gat gcg	10785
Gly Gln Ser Gly Ala Thr Ile Tyr Arg Leu Tyr Gly Lys Pro Asp Ala	
480 485 490	
cca gag ttg ttt ctg aaa cat ggc aaa ggt agc gtt gcc aat gat gtt	10833
Pro Glu Leu Phe Leu Lys His Gly Lys Gly Ser Val Ala Asn Asp Val	
495 500 505	

aca gat gag atg gtc aga cta aac tgg ctg acg gaa ttt atg cct ctt	10881
Thr Asp Glu Met Val Arg Leu Asn Trp Leu Thr Glu Phe Met Pro Leu	
510 515 520	
ccg acc atc aag cat ttt atc cgt act cct gat gat gca tgg tta ctc	10929
Pro Thr Ile Lys His Phe Ile Arg Thr Pro Asp Asp Ala Trp Leu Leu	
525 530 535	
acc act gcg atc ccc ggg aaa aca gca ttc cag gta tta gaa gaa tat	10977
Thr Thr Ala Ile Pro Gly Lys Thr Ala Phe Gln Val Leu Glu Glu Tyr	
540 545 550 555	
cct gat tca ggt gaa aat att gtt gat gcg ctg gca gtg ttc ctg cgc	11025
Pro Asp Ser Gly Glu Asn Ile Val Asp Ala Leu Ala Val Phe Leu Arg	
560 565 570	
cgg ttg cat tcg att cct gtt tgt aat tgt cct ttt aac agc gac cgc	11073
Arg Leu His Ser Ile Pro Val Cys Asn Cys Pro Phe Asn Ser Asp Arg	
575 580 585	
gta ttt cgt ctc gct cag gcg caa tca cga atg aat aac ggt ttg gtt	11121
Val Phe Arg Leu Ala Gln Ala Gln Ser Arg Met Asn Asn Gly Leu Val	
590 595 600	
gat gcg agt gat ttt gat gac gag cgt aat ggc tgg cct gtt gaa caa	11169
Asp Ala Ser Asp Phe Asp Asp Glu Arg Asn Gly Trp Pro Val Glu Gln	
605 610 615	
gtc tgg aaa gaa atg cat aag ctt ttg cca ttc tca ccg gat tca gtc	11217
Val Trp Lys Glu Met His Lys Leu Leu Pro Phe Ser Pro Asp Ser Val	
620 625 630 635	
gtc act cat ggt gat ttc tca ctt gat aac ctt att ttt gac gag ggg	11265
Val Thr His Gly Asp Phe Ser Leu Asp Asn Leu Ile Phe Asp Glu Gly	
640 645 650	
aaa tta ata ggt tgt att gat gtt gga cga gtc gga atc gca gac cga	11313
Lys Leu Ile Gly Cys Ile Asp Val Gly Arg Val Gly Ile Ala Asp Arg	
655 660 665	
tac cag gat ctt gcc atc cta tgg aac tgc ctc ggt gag ttt tct cct	11361
Tyr Gln Asp Leu Ala Ile Leu Trp Asn Cys Leu Gly Glu Phe Ser Pro	
670 675 680	

tca tta cag aaa cgg ctt ttt caa aaa tat ggt att gat aat cct gat 11409
 Ser Leu Gln Lys Arg Leu Phe Gln Lys Tyr Gly Ile Asp Asn Pro Asp
 685 690 695

atg aat aaa ttg cag ttt cat ttg atg ctc gat gag ttt ttc taa 11454
 Met Asn Lys Leu Gln Phe His Leu Met Leu Asp Glu Phe Phe
 700 705 710

tcagaattgg ttaattgggtt gtaacactgg cagagcatta cgctgacttg acgggacggc 11514

ggctttgttg aataaatcga acttttgctg agttgaagga tcagatcacg catcttcccg 11574

acaacgcaga ccgttccgtg gcaaagcaaa agttcaaaat caccaactgg tccacctaca 11634

acaaagctct catcaaccgt ggctccctca ctttctggct ggatgatggg gcgattcagg 11694

gatcacaggc agcaacgctc tgtcatcg 11722

<210> 45
 <211> 442
 <212> PRT
 <213> 人工

<220>
 <223> 合成构建体

<400> 45
 Met Glu Asn Ala Lys Met Asn Ser Leu Ile Ala Gln Tyr Pro Leu Val
 1 5 10 15

Lys Asp Leu Val Ala Leu Lys Glu Thr Thr Trp Phe Asn Pro Gly Thr
 20 25 30

Thr Ser Leu Ala Glu Gly Leu Pro Tyr Val Gly Leu Thr Glu Gln Asp
 35 40 45

Val Gln Asp Ala His Ala Arg Leu Ser Arg Phe Ala Pro Tyr Leu Ala
 50 55 60

Lys Ala Phe Pro Glu Thr Ala Ala Thr Gly Gly Ile Ile Glu Ser Glu
 65 70 75 80

Leu Val Ala Ile Pro Ala Met Gln Lys Arg Leu Glu Lys Glu Tyr Gln
 85 90 95

Gln Pro Ile Ser Gly Gln Leu Leu Leu Lys Lys Asp Ser His Leu Pro
 100 105 110

Ile Ser Gly Ser Ile Lys Ala Arg Gly Gly Ile Tyr Glu Val Leu Ala
 115 120 125

His Ala Glu Lys Leu Ala Leu Glu Ala Gly Leu Leu Thr Leu Asp Asp
 130 135 140

Asp Tyr Ser Lys Leu Leu Ser Pro Glu Phe Lys Gln Phe Phe Ser Gln
 145 150 155 160

Tyr Ser Ile Ala Val Gly Ser Thr Gly Asn Leu Gly Leu Ser Ile Gly
 165 170 175

Ile Met Ser Ala Arg Ile Gly Phe Lys Val Thr Val His Met Ser Ala
 180 185 190

Asp Ala Arg Ala Trp Lys Lys Ala Lys Leu Arg Ser His Gly Val Thr
 195 200 205

Val Val Glu Tyr Glu Gln Asp Tyr Gly Val Ala Val Glu Glu Gly Arg
 210 215 220

Lys Ala Ala Gln Ser Asp Pro Asn Cys Phe Phe Ile Asp Asp Glu Asn
 225 230 235 240

Ser Arg Thr Leu Phe Leu Gly Tyr Ser Val Ala Gly Gln Arg Leu Lys
 245 250 255

Ala Gln Phe Ala Gln Gln Gly Arg Ile Val Asp Ala Asp Asn Pro Leu
 260 265 270

Phe Val Tyr Leu Pro Cys Gly Val Gly Gly Gly Pro Gly Gly Val Ala
 275 280 285

Phe Gly Leu Lys Leu Ala Phe Gly Asp His Val His Cys Phe Phe Ala
 290 295 300

Glu Pro Thr His Ser Pro Cys Met Leu Leu Gly Val His Thr Gly Leu
 305 310 315 320

His Asp Gln Ile Ser Val Gln Asp Ile Gly Ile Asp Asn Leu Thr Ala
 325 330 335

Ala Asp Gly Leu Ala Val Gly Arg Ala Ser Gly Phe Val Gly Arg Ala
340 345 350

Met Glu Arg Leu Leu Asp Gly Phe Tyr Thr Leu Ser Asp Gln Thr Met
355 360 365

Tyr Asp Met Leu Gly Trp Leu Ala Gln Glu Glu Gly Ile Arg Leu Glu
370 375 380

Pro Ser Ala Leu Ala Gly Met Ala Gly Pro Gln Arg Val Cys Ala Ser
385 390 395 400

Val Ser Tyr Gln Gln Met His Gly Phe Ser Ala Glu Gln Leu Arg Asn
405 410 415

Thr Thr His Leu Val Trp Ala Thr Gly Gly Gly Met Val Pro Glu Glu
420 425 430

Glu Met Asn Gln Tyr Leu Ala Lys Gly Arg
435 440

<210> 46
<211> 271
<212> PRT
<213> 人工

<220>
<223> 合成构建体

<400> 46
Met Ser His Ile Gln Arg Glu Thr Ser Cys Ser Arg Pro Arg Leu Asn
1 5 10 15

Ser Asn Met Asp Ala Asp Leu Tyr Gly Tyr Lys Trp Ala Arg Asp Asn
20 25 30

Val Gly Gln Ser Gly Ala Thr Ile Tyr Arg Leu Tyr Gly Lys Pro Asp
35 40 45

Ala Pro Glu Leu Phe Leu Lys His Gly Lys Gly Ser Val Ala Asn Asp
50 55 60

Val Thr Asp Glu Met Val Arg Leu Asn Trp Leu Thr Glu Phe Met Pro
65 70 75 80

Leu Pro Thr Ile Lys His Phe Ile Arg Thr Pro Asp Asp Ala Trp Leu
85 90 95

Leu Thr Thr Ala Ile Pro Gly Lys Thr Ala Phe Gln Val Leu Glu Glu
100 105 110

Tyr Pro Asp Ser Gly Glu Asn Ile Val Asp Ala Leu Ala Val Phe Leu
115 120 125

Arg Arg Leu His Ser Ile Pro Val Cys Asn Cys Pro Phe Asn Ser Asp
130 135 140

Arg Val Phe Arg Leu Ala Gln Ala Gln Ser Arg Met Asn Asn Gly Leu
145 150 155 160

Val Asp Ala Ser Asp Phe Asp Asp Glu Arg Asn Gly Trp Pro Val Glu
165 170 175

Gln Val Trp Lys Glu Met His Lys Leu Leu Pro Phe Ser Pro Asp Ser
180 185 190

Val Val Thr His Gly Asp Phe Ser Leu Asp Asn Leu Ile Phe Asp Glu
195 200 205

Gly Lys Leu Ile Gly Cys Ile Asp Val Gly Arg Val Gly Ile Ala Asp
210 215 220

Arg Tyr Gln Asp Leu Ala Ile Leu Trp Asn Cys Leu Gly Glu Phe Ser
225 230 235 240

Pro Ser Leu Gln Lys Arg Leu Phe Gln Lys Tyr Gly Ile Asp Asn Pro
245 250 255

Asp Met Asn Lys Leu Gln Phe His Leu Met Leu Asp Glu Phe Phe
260 265 270

<210> 47

<211> 693

<212> DNA

<213> 玉米

<220>

<221> 前体 RNA

<222> (1)..(693)

<400> 47

gcagcgtcgg tggttgcttc atggctgtcg aggggaatga cgtccggtcc gaacaagcca	60
cggtgtctgc tgcgtaccg ccgcggcttc ggaccaggct tcattcccca cgactcacca	120
tggaaacggc agcgaggcgc gagcttcag tgcgcgcagt gtttgatcag caaacacctg	180
cgtggccgag gctgagcctt gcaagggcag gcgtccatgt tctcgtcagc aagcgtcttc	240
ttcacatgtc tgaggctcga gcaattgcgg gtggacatcg atggatgagc attgccgctg	300
caaggaaga tcgtccggct cacacttgag ttgggatacc atgcactact cttgtttag	360
ttagattaga taccatgcgc ttgaaccgtg caatagcata gcgaggaaga ggaaccaga	420
gaggactcag aaagagatgt tgtctatttc atccagtgt catctagtca tctagtctgc	480
ctgcgggtgt ctgcaatgaa cgctgacgct gctgccggag acctactctg ctactggtgt	540
ggagtggaca aacagggagg gaaggcgtct gctagaacct agaagacgga cacagggagg	600
accaacagga ctaaagaccg tcgtacgtgt tgaagcggaa ggactttctt attacgtgtg	660
gcgcttcgaa ttaaagccag cggttagaac ggc	693

附录 A

图 1A: SEQ ID NO: 1-AtFAD-01 核酸序列

GACCACCAGAAGAAGAGCCACACACTCACAAATTA AAAAGAGAGAGAGAGAGAGAGACAGAG
AGAGAGAGAGATTCTGCGGAGGAGCTTCTTCTTCGTAGGGTGTTTCATCGTTATTAACGTTATCGC
CCCTACGTCAGCTCCATCTCCAGAAACATGGGTGCAGGTGGAAGAATGCCGGTTCCTACTTCTT
CCAAGAAATCGGAAACCGACACCACAAAGCGTGTGCCGTGCGAGAAACCGCCTTCTCGGTGG
GAGATCTGAAGAAAGCAATCCCGCCGCTTGTTCAAACGCTCAATCCCTCGCTCTTCTCCTAC
CTTATCAGTGACATCATTATAGCCTCATGCTTCTACTACGTCGCCACCAATTACTTCTCTCCTC
CCTCAGCCTCTCTTACTTGGCTTGGCCACTCTATTGGGCCTGTCAAGGCTGTGTCTTAAGTGG
TATCTGGGTCATAGCCCACGAATGCGGTCAACGCTCAGCGACTACCAATGGCTGGATGAC
ACAGTTGGTCTTATCTTCCATTCCCTCCTCGTCCCTTACTTCTCCTGGAAGTATAGTCATCGC
CGTCACCAATCCAACACTGGATCCCTCGAAAGAGATGAAGTATTTGTCCCAAAGCAGAAATCAGC
AATCAAGTGGTACGGGAAATACCTCAACAACCTCTTGGACGCATCATGATGTTAACCGTCCAGT
TTGTCCTCGGGTGGCCCTTGTACTTAGCCTTAAACGTCTCTGGCAGACCGTATGACGGGTTCCG
TTGCCATTTCTTCCCAACGCTCCCATCTACAATGACCGAGAACGCCTCCAGATATACCTCTCTG
ATGCGGGTATTCTAGCCGTCTGTTTTGGTCTTACCGTTACGCTGCTGCACAAGGGATGGCCTC
GATGATCTGCCTCTACGGAGTACCGCTTCTGATAGTGAATGCGTTCTCCTCGTCTGATCACTTACT
TGCAGCACACTCATCCCTCGTTGCCTCACTACGATTCATCAGAGTGGGACTGGCTCAGGGGAGC
TTTGGCTACCGTAGACAGAGACTACGGAATCTTGAACAAGGTGTTCCACAACATTACAGACACAC
ACGTGGCTCATCACCTGTTCTCGACAATGCCGCATTATAACGCAATGGAAGCTACAAAGGCGATA
AAGCCAATTCTGGGAGACTATTACCAGTTCGATGGAACACCGTGGTATGTAGCGATGTATAGGG
AGGCAAAGGAGTGTATCTATGTAGAACCGGACAGGGAAGGTGACAAGAAAGGTGTGTACTGGTA
CAACAATAAGTTATGAGGATGATGGTGAAGAAATTGTCGACCTTCTCTTGTCTGTTTGTCTTTG
TTAAAGAAGCTATGCTTCGTTTAAATAATCTTATTGTCCATTTTGTGTGTATGACATTTTGGCTG
CTCATTATGTTATGTGGGAAGTTAGTGTCAAATGTTTTGTGTGCGGTATTGTTCTTCTCATCGCTG
TTTTGTTGGGATCGTAGAAATGTGACCTTCGGACAGTAAACTCTTGTACTAAACTATCTCCCTA
TTGGCATTTCTTAAACTTTTAATAGTTACGTGCTCGTAGTGAATCTTGACTTGAGTCA

图 1B: SEQ ID NO: 3-AtFAD-01 开放读框的核酸序列

ATGGGTGCAGGTGGAAGAATGCCGGTTCCTACTTCTTCCAAGAAATCGGAAACCGACACCACAA
AGCGTGTGCCGTGCGAGAAACCGCCTTCTCGGTGGGAGATCTGAAGAAAGCAATCCCGCCGC
ATTGTTTCAAACGCTCAATCCCTCGCTCTTCTCCTACCTTATCAGTGACATCATTATAGCCTCAT
GCTTCTACTACGTCGCCACCAATTACTTCTCTCTCCTCCCTCAGCCTCTCTTACTTGGCTTGGC
CACTCTATTGGGCCTGTCAAGGCTGTGTCTTAAGTGGTATCTGGGTCAAGCCACGAATGCGG
TCACCACGCATTACGCGACTACCAATGGCTGGATGACACAGTTGGTCTTATCTTCCATTCCCTCC
TCCTCGTCCCTTACTTCTCCTGGAAGTATAGTCATCGCCGTCAACATTCCAACACTGGATCCCTC
GAAAGAGATGAAGTATTTGTCCCAAAGCAGAAATCAGCAATCAAGTGGTACGGGAAATACCTCAA
CAACCCTCTTGGACGCATCATGATGTTAACCGTCCAGTTTGTCTCGGGTGGCCCTTGTACTTAG
CCTTTAACGTCTCTGGCAGACCGTATGACGGGTTGCTTGCATTTCTTCCCAACGCTCCCATC
TACAATGACCGAGAACGCCTCCAGATATACCTCTCTGATGCGGGTATTCTAGCCGTCTGTTTTGG
TCTTTACCGTTACGCTGCTGCACAAGGGATGGCCTCGATGATCTGCCTCTACGGAGTACCGCTT
CTGATAGTGAATGCGTTCCTCGTCTTGATCACTTACTTGAGCACACTCATCCCTCGTTGCCTCA
CTACGATTCATCAGAGTGGGACTGGCTCAGGGGAGCTTTGGCTACCGTAGACAGAGACTACGGA
ATCTTGAACAAGGTGTTCCACAACATTACAGACACACAGTGGCTCATCACCTGTTCTCGACAAT
GCCGCATTATAACGCAATGGAAGCTACAAAGGCGATAAAGCCAATTCTGGGAGACTATTACCAGT
TCGATGGAACACCGTGGTATGTAGCGATGTATAGGGAGGCAAAGGAGTGTATCTATGTAGAACC
GGACAGGGAAGGTGACAAGAAAGGTGTGTACTGGTACAACAATAAGTTA

图 1C: SEQ ID NO: 4-AtFAD-01 开放读框的氨基酸序列

MGAGGRMPVPTSSKKSETDTTKRVPCEKPPFSVGLKKAIPPHCFKRSIPRSFYSYLSDIIASCFYYVA
TNYFSLLPQPLSYLAWPLYWACQGCVLTGIWVIAHECGHAFSDYQWLDDTVGLIFHSFLLVPYFSW
KYSHRRHHSNTGSLERDEVFVPKQKSAIKWYGKYLNNPLGRIMMLTVQFVLGWPLYLAFNVSGRPY
DGFACHFFPNAPIYNDRERLQIYLS DAGILAVCFGLYRYAAAQGMASMICLYGVPLLVNAFLVLITYLQ
HTHPSLPHYDSSEWDWLRGALATVDRDYGILNKVFHNITDTHVAHHLFSTMPHYNAMEATKAIPILG
DYYQFDGTPWYVAMYREAKECIYVEPDREGDKKGVYWNKL

图 2A: SEQ ID NO: 5-GmFAD-01 核酸序列

TAGGCACCTAGCTAGTAGCTACAATATCAGCACTTCTCTCTATTGATAACAATTGGCTGTAATGC
CGCAGTAGAGGACGATCACAACATTTCTGCTGGATACTTTTTGTTTTATGGGTCTAGCAAAGGA
AACAATAATGGGAGGTGGAGGCCGTGTGGCCAAAGTTGAAATTCAGCAGAAGAAGCCTCTCTCA
AGGGTTCCAAACACAAAGCCACCATTCACTGTTGGCCAACTCAAGAAAGCCATTCCACCGCACT
GCTTTCAGCGTTCCTCCTCACTTCATTGTCCTATGTTGTTTATGACCTTTCATTGGCTTTCATTTT
CTACATTGCCACCACCTACTTCCACCTCCTCCCTCACCCCTTTTCCCTCATTGCATGGCCAATCTA
TTGGGTTCTCCAAGGTTGCATTCTTACTGGCGTGTGGGTGATTGCTCACGAGTGTGGTCACCAT
GCCTTCAGCAAGTACCCATGGGTTGATGATGTTATGGGTTTGACCGTTCACTCAGCACTTTTAGT
CCCTTATTTCTCATGGAAAATAAGCCATCGCCGCCACCACTCCAACACGGGTTCCCTTGACCGTG
ATGAAGTGTGTTGCCAAAACCAAATCCAAAGTTGCATGGTACACCAAGTACCTGAACAACCT
CTAGGAAGGGCTGCTTCTCTCTCATCACTCACAATAGGGTGGCCTATGTATTTAGCCTTCAA
TGCTCTGGCAGACCCTATGATGGTTTTGCAAGCCACTACCACCCTTATGCTCCCATATATTCTAA
CCGTGAGAGGCTTCTGATCTATGTCTCTGATGTTGCTTTGTTTTCTGTGACTTACTCTCTACCG
TGTTGCAACTATGAAAGGGTTGGTTTGGCTGCTATGTGTTTATGGGGTGCCATTGCTCATTGTGA
ACGGTTTTCTGTGACTATCACATATTTGCAGCACACACACTTTGCCTTGCTCATTACGATTCAT
CAGAATGGGACTGGCTGAAGGGAGCTTTGGCAACTATGGACAGAGATTATGGGATTCTGAACAA
GGTGTTCATCACATAACTGATACTCATGTGGCTCACCATCTTCTCTACAATGCCACATTACCA
TGCAATGGAGGCAACCAATGCAATCAAGCCAATATTGGGTGAGTACTACCAATTTGATGACACAC
CATTTTACAAGGCACTGTGGAGAGAAGCGAGAGAGTGCCTCTATGTGGAGCCAGATGAAGGAAC
ATCCGAGAAGGGCGTGTATTGGTACAGGAACAAGTATTGATGGACCAAGCAATGGGCCATAGTG
GGAGTTATGGAAGTTTTGTCACTTATCACTTAATTAGTAGAATGTTATAAATAAGTGGATTGCGG
CGTAATGACTTGTGTGCATTGTGAAACAGCTTGTAGCGATCCATGGTTATAATGTAAAAATATGTG
GAAAGGGGTCTGGTTAAAAAAAAAAAAAAAAAAGCGGCCGTTTTAAAGGAAACAAGG

图 2B: SEQ ID NO: 7-GmFAD-01 开放读框的核酸序列

ATGGGTCTAGCAAAGGAAACAATAATGGGAGGTGGAGGCCGTGTGGCCAAAGTTGAAATTCAGC
AGAAGAAGCCTCTCTCAAGGGTTCCAAACACAAAGCCACCATTCACTGTTGGCCAACTCAAGAAA
GCCATTCCACCGCACTGCTTTCAGCGTTCCTCCTCACTTCATTGTCCTATGTTGTTTATGACCTT
TCATTGGCTTTCATTTTCTACATTGCCACCACCTACTTCCACCTCCTCCCTCACCCCTTTTCCCTC
ATTGCATGGCCAATCTATTGGGTTCTCCAAGGTTGCATTCTTACTGGCGTGTGGGTGATTGCTCA
CGAGTGTGGTCACCATGCCTTCAGCAAGTACCCATGGGTTGATGATGTTATGGGTTTGACCGTTC
ACTCAGCACTTTTAGTCCCTTATTTCTCATGGAAAATAAGCCATCGCCGCCACCACTCCAACACG
GGTTCCCTTGACCGTGATGAAGTGTGTTGCCAAAACCAAATCCAAAGTTGCATGGTACACCAA
GTACCTGAACAACCCTCTAGGAAGGGCTGCTTCTCTCTCATCACTCACAATAGGGTGGCCTA
TGATTTAGCCTTCAATGTCTCTGGCAGACCCTATGATGGTTTTGCAAGCCACTACCACCCTTATG
CTCCCATATATTCTAACCGTGAGAGGCTTCTGATCTATGTCTCTGATGTTGCTTTGTTTTCTGTGA
CTTACTCTCTACCGTGTGCAACTATGAAAGGGTTGGTTGGCTGCTATGTGTTTATGGGGTG
CCATTGCTCATTGTGAACGGTTTTCTTGTGACTATCACATATTTGCAGCACACACACTTTGCCTTG
CCTCATTACGATTATCAGAATGGGACTGGCTGAAGGGAGCTTTGGCAACTATGGACAGAGATTA
TGGGATTCTGAACAAGGTGTTTCATCACATAACTGATACTCATGTGGCTCACCATCTTCTCTAC
AATGCCACATTACCATGCAATGGAGGCAACCAATGCAATCAAGCCAATATTGGGTGAGTACTACC
AATTTGATGACACACCATTTTACAAGGCACTGTGGAGAGAAGCGAGAGAGTGCCTCTATGTGA
GCCAGATGAAGGAACATCCGAGAAGGGCGTGTATTGGTACAGGAACAAGTAT

图 2C: SEQ ID NO: 8-GmFAD-01 开放读框的氨基酸序列

MGLAKETIMGGGGRVAKVEIQKKPLSRVPNTKPPFTVQQLKKAIPPHCFQRSLLTSLSYVVYDLSLA
FIFYIATTYFHLLPHFSLIAWPIYWVLQGCILTGWVIAHECGHHAFSKYPWVDDVMGLTVHSALLVP
YFSWKISHRRHHSNTGSLDRDEVFVPKPSKVAWYTKYLNPLGRAASLLITLTIGWPMYLA FNVS
RPYDGFASHYHPYAPIYSNRERLLIYVSDVALFSVTYSLYRVATMKGLVWLLCVYGVPLLVNGLVIT
YLQHTHFALPHYDSSEWDWLKALATMDRDRYGILNKVFHHITDTHVAHHLFSTMPHYHAMEATNAIK
PILGEYYQFDDTPFYKALWREARECLYVEPDEGTSEKGVYWRNKY

图 3A: SEQ ID NO: 9-GmFAD-02 核酸序列

GCGTGTGCGGTCTCTCTCTCTCTCACCCCTCCTCTTCACACATTTTCTGTGCGCTCTAACAAACAT
TCTCGTTTCACACTTTTCAGATTGTGTGAAGATGGGGGCGGGTGGCCGAACCTGATGTTCTCCTGC
CAACAGGAAGTCAGAGGTTGACCCCTTGAAGCGGGTGCCATTTGAAAAACCTCCATTTAGTCTCA
GCCAAATCAAGAAGGTCATTCCACCTCACTGTTTCCAGCGTCTGTTTTCCGCTCATTCTCCTATG
TTGTTTACGACCTCACCATAGCCTTCTGCCTCTATTATGTTGCCACCCATTACTTCCACCTCCTTC
CCAGCCCTCTCTTTCTTGGCATGGCCAATCTACTGGGCTGTCCAAGGTTGCATCCTTACTGGA
GTTTGGGTCATTGCCCATGAGTGTGGCCACCATGCATTCAGTGACTACCAGTTGCTTGATGATAT
TGTTGGCCTTGTCTCCACTCCGGTCTCCTAGTCCCATACTTTTCATGGAAATACAGCCATCGCC
GTCACCACTCCAACACTGGTTCTCTTGAGCGGGATGAAGTATTTGTGCCAAAGCAGAAGTCCTGT
ATCAAGTGGTACTCTAAATACCTTAACAATCCTCCAGGCAGAGTCCTCACTCTTGCTGTCACCCT
CACACTTGGTTGGCCCTTGTACTTGGCTTTAAATGTTTCTGGAAGGCCTTATGATAGATTTGCTTG
CCACTATGACCCATATGGTCCCATTACTCTGATCGTGAACGACTTCAAATATATATATCAGATGC
AGGAGTACTTGCAATGCTATGGCCTTTCCGTCTTGCCATGGCAAAGGACTTGCCTGGGTG
GTGTGTGTTTATGGAGTTCATTGCTAGTGGTCAATGGATTTTGGTGTGATTACATTCTTGCAAG
CATACTCACCCCTGCATTGCCACATTACACTTCCTCTGAGTGGGACTGGTTGAGAGGAGCTTTAGC
AACAGTGGATAGAGATTATGGAATCCTGAACAAGGTCTTCCATAATATTACAGACACTCATGTAG
CACATCACTTGTCTCCACAATGCCACATTATCATGCAATGGAGGCTACAAAGGCAATAAAACCC
ATTTTGGGAGAGTATTATCGGTTTGATGAGACTCCATTTGTCAAGGCAATGTGGAGAGAGGCAAG
AGAGTGTATTTATGTGGAGCCAGATCAAAGTACCGAGAGCAAAGGTGTATTTTGGTACAACAATA
AGTTGTGATGATTAATGTAGCCGAGGCTTCTTTGAACCTTCCCTTGTGACTGTTTAGTATCATGGT
TGCTTATTTGGGAATAATTTTGTGAACCCTGATGTTGGTAGTAAGTATCTAGACAGTTGCATAGCG
GTTTTGTTTACAGAATAAGATATAGCCTCTCTGAACAGTTTGATTATTGCACCATGGTTTGCAATC
GGTGCATGTCGACCAAGTTTCTCAAGACTGTGGAGAAGCTTATTCTTGTTCCAGTCTTGAATCC
AAGTTGTTACCGTATTCTGTTATTATTGACTTAGAATCCTTAACCTTTTCTGCTGTTTTCTCATGAT
CGTCACTCGCAAATGAATCACATTTCAAACCAAAAAAAAAAAAAAAAAA

图 3B: SEQ ID NO: 11-GmFAD-02 开放读框的核酸序列

ATGGGGGCGGGTGGCCGAACCTGATGTTCTCCTGCCAACAGGAAGTCAGAGGTTGACCCCTTG
AAGCGGGTGGCATTGAAAAACCTCCATTTAGTCTCAGCCAAATCAAGAAGGTCATTCCACCTCA
CTGTTTCCAGCGTCTGTTTTCCGCTCATTCTCCTATGTTGTTTACGACCTCACCATAGCCTTCTG
CCTCTATTATGTTGCCACCCATTACTTCCACCTCCTTCCAGCCCTCTCTTTTCTTGGCATGGCC
AATCTACTGGGCTGTCCAAGGTTGCATCCTTACTGGAGTTTGGGTCATTGCCCATGAGTGTGGC
CACCATGCATTCACTGACTACCAAGTTGCTTGATGATATTGTTGGCCTTGTCTCCACTCCGGTCT
CCTAGTCCCATACTTTTTCATGGAAATACAGCCATCGCCGTCAACCACTCCAACACTGGTTCTCTTG
AGCGGGATGAAGTATTTGTGCCAAAGCAGAAGTCCTGTATCAAGTGGTACTCTAAATACCTTAAC
AATCCTCCAGGCAGAGTCCTCACTCTTGCTGTCACCCTCACACTTGGTTGGCCCTTGTACTTGGC
TTTAAATGTTTCTGGAAGGCCTTATGATAGATTTGCTTGCCACTATGACCCATATGGTCCCATTTA
CTCTGATCGTGAACGACTTCAAATATATATATCAGATGCAGGAGTACTTGCAGTATGCTATGGCC
TTTTCCGTCTTGCCATGGCAAAGGACTTGCCTGGGTGGTGTGTGTTTATGGAGTTCATTGCTA
GTGGTCAATGGATTTTGGTGTGATTACATTCTTGCAAGCATACTCACCCCTGCATTGCCACATTAC
ACTTCTCTGAGTGGGACTGGTTGAGAGGAGCTTTAGCAACAGTGGATAGAGATTATGGAATCCT
GAACAAGGTCTTCCATAATATTACAGACACTCATGTAGCACATCACTTGTCTCCACAATGCCACA
TTATCATGCAATGGAGGCTACAAAGGCAATAAAACCCATTTTGGGAGAGTATTATCGGTTTGATG
AGACTCCATTTGTCAAGGCAATGTGGAGAGAGGCAAGAGAGTGTATTTATGTGGAGCCAGATCA
AAGTACCGAGAGCAAAGGTGTATTTTGGTACAACAATAAGTTG

图 3C: SEQ ID NO: 12-GmFAD-02 开放读框的氨基酸序列

MGAGGRDTPPANRKSEVDPLKRVPFKPPFSLSQIKKVIPPHCFQRSVFRSFSYVVYDLTIAFLYY
VATHYFHLLPSPLSFLAWPIYWAVQGCLTGWVIAHECGHHAFSDYQLDDIVGLVLHSGLLVPYFS
WKYSHRRHHSNTGSLERDEVFVPKQKSCIKWYSKYLNPPGRVLTAVTLTLGWPLYLALNVSGRP
YDRFACHYDPYGPIYSRERLQIYISDAGVLAVCYGLFRLAMAKGLAWVVCVYGVPLLVNGLVLITF
LQHTHPALPHYTSSEWDWLRGALATVDRDYGILNKVFHNITDTHVAHHLFSTMPHYHAMEATKAIKPI
LGEYYRFDETPFVKAMWREARECIYVEPDQSTESKGVFWYNNKL

图 4A: SEQ ID NO: 13-GmFAD-03 核酸序列

ACTTTTCATGGAAATACAGCCATCGCCGTCACCACTCCAACACAGGTTCTCTTGAGCGAGATGAA
GTATTTGTGCCAAAGCAGAAGTCCAGTATCATGTGGTACTCTAAATACCTTAACAATCCACCAGG
CAGAGTCCTCACTCTTGCCGTCACCCTCACGCTTGGTTGGCCCTTGTA CTGGCTTTAATGTTT
CTGGAAGGCCTTATGATAGATTTGCTTGCCACTATGACCCTTATGGTCCCATTTACTCTGACCGA
GAACGACTTCAAATATATATATCAGATGCAGGAGTACTTGCAGTATGCTATGGCCTTTCTGTCTT
GCCATGGCAAAAGGGCTTGCCTGGGTGGTGTGTGTTTATGGAGTTCCATTGCTTGTGGTCAATG
GATTTTTGGTGTGATTACATTTTTGCAGCACACTCACCTGCATTGCCACACTACACTTCCTCTG
AGTGGGACTGGTTGAGAGGAGCTTTAGCAACAGTGGATAGAGATTATGGAATCCTGAACAAGGT
CTTCCATAATATTACAGACACTCATGTAGCTCATCACTTGTTCTCCACAATGCCACATTATCATGC
AATGGAGGCGACAAAGGCAATAAAGCCCATCTTGGGAGAGTATTATCGGTTTGATGGGACTCCA
TTTGTCAAGGCAATGTGGAGAGAGGCAAGAGAGTGTATTTATGTGGAGCCAGATCAAAGTACTCA
GAGCAAAGGTGTATTTGGTACAACAATAAGTTGTGATGATTAATGTAGCCGGGGGCTTCTTGTC
CGTCAGGTGGGATGGTTTGAACCTTCTTTGTGACTGTTTAGTATCATGCTTGCTTATTGGGAATA
ATTTGTTGAACCCTGATGTTGGTAGTAGTATCTAGAAAGTAGCATAGCGTTTTTGTGTCAGAAT
AAGATATAGCATCTCTGAACAGTTTGATTATTGCACCATGTTTGCAATCAGTGCATGTCGACCGG
GTTTCTCAAGATTGTGGGGATGCTTATTCTTGTTCCAGTTCTTGAATCCAAGTTGTTATCATATTCT
GTTATTGA

图 4B: SEQ ID NO: 15-GmFAD-03 开放读框的核酸序列

ATGTGGTACTCTAAATACCTTAACAATCCACCAGGCAGAGTCCTCACTCTTGCCGTCACCCTCAC
GCTTGGTTGGCCCTTGTA CTGGCTTTAATGTTTCTGGAAGGCCTTATGATAGATTTGCTTGCCA
CTATGACCCTTATGGTCCCATTTACTCTGACCGAGAACGACTTCAAATATATATATCAGATGCAGG
AGTACTTGCAGTATGCTATGGCCTTTCTGTCTTGCCATGGCAAAAGGGCTTGCCTGGGTGGTGT
GTGTTTATGGAGTTCCATTGCTTGTGGTCAATGGATTTTTGGTGTGATTACATTTTTGCAGCACA
CTCACCTGCATTGCCACACTACACTTCCTCTGAGTGGGACTGGTTGAGAGGAGCTTTAGCAAC
AGTGGATAGAGATTATGGAATCCTGAACAAGGTCTTCCATAATATTACAGACACTCATGTAGCTCA
TCACTTGTTCTCCACAATGCCACATTATCATGCAATGGAGGCGACAAAGGCAATAAAGCCCATCT
TGGGAGAGTATTATCGGTTTGATGGGACTCCATTTGTCAAGGCAATGTGGAGAGAGGCAAGAGA
GTGTATTTATGTGGAGCCAGATCAAAGTACTCAGAGCAAAGGTGTATTTGGTACAACAATAAGTT
G

图 4C: SEQ ID NO: 16-GmFAD-03 开放读框的氨基酸序列

MWYSKYLNNPPGRVLTAVTLTLGWPLYLAFNVSGRPYDRFACHYDPYGPISDRERLQIYISDAGVL
AVCYGLFCLAMAKGLAWVVCVYGVPLLVNGLVLITFLQHTHPALPHYTSSEWDWLRGALATVDRD
YGILNKVFHNITDTHVAHHLFSTMPHYHAMEATKAIKPILGEYYRFDGTPFVKAMWREARECIYVEPD
QSTQSKGVFWYNNKL

图 5A: SEQ ID NO: 17-ZmFAD-01 核酸序列

CTCCCTCCTCCTCCTCCTCCCTGCAAATCCTGCAGGCACACCGCTCGTTTTCTGTCCGGGG
GACAGGAGAGAAGGGGAGAGACCGAGAGAGGGTGAGGCGCGCGTCCGCCGATCTGCTCCGC
CCCCGAAGCAGCCTGTACGTCGTCCTCACTCTCAGCAACCAGCGAAAATGGGTGCCGGAGG
CAGGATGACCGAGAAGGAGCGGGAGAAGCATGAGCAGGAGCAGGTGCGCCGTGCTACCGGCG
GTGGCGCGGCAGTGACGCGGTGCGCGGTGAGAAAGCCGCGTTCACGTTGGGGCAGATCAAG
AAGGCGATCCCGCCGCACTGCTTCGAGCGCTCCGTGCTGAGGTCCTTCTCGTACGTGGCCAC
GACCTGTGCTCGCCGCGGCGCTCCTCTACCTCGCGGTGGCCGTGATACCGGCGCTACCCTGC
CCGCTCCGCTACGCGGCGCTGGCCGTGTAAGTGGGTGGCCAGGGGTGCGTGTGCACGGGCGT
GTGGGTGATCGCGCACGAGTGCGGCCACACGCTTCTCCGACCACGCGCTCCTGGACGACGC
CGTCCGCTGGCGCTGCACTCGGCGCTGCTGGTGCCCTACTTCTCGTGGAAGTACAGCCACCG
GCGCCACCACTCCAACACGGGGTCCCTGGAGCGCGACGAGGTGTTCTGTGCCGAGGACCAAGG
AGGCGCTGCCGTGGTACGCCCGGTACGTGCACGGCAGCCCCGCGGGCCGGCTGGCGCACGTC
GCCGTGCAGCTCACCCTCGGCTGGCCGCTGTACCTGGCCACCAACGCGTCGGGGCGGCCGTA
CCCGCGCTTCGCTGCCACTTCGACCCCTACGGCCCCATCTACAACGACCGGGAGCGCGCCCA
GATCTTCGTCTCGGACGCGGCGTGTGGCCGTGGCGTTCGGGCTGTACAAGCTGGCGGCGG
CGTTCGGGGTCTGGTGGGTGGTGCGCGTGTACGCCGTGCCGCTGCTGATCGTCAACGCGTGGC
TGGTGCTCATCACGTACCTGCAGCACACCCACCGGCGCTGCCCACTACGACTCGGGCGAGT
GGGACTGGCTGCGCGGCGCGCTCGCCACCGTCGACCGAGACTACGGCGTCTCAACCGCGT
TTCCACCACATCACGGACACGCACGTGCGCACCACTCTTCTCCACCATGCCGCACTACCAG
CCGTGGAGGCCACAGGGCGATCAGGCCGCTCGGCGAGTACTACCAGTTCGACCCGACCC
CTGTGCCAAGGCCACCTGGCGCGAGGCGAGGAGTGATCTACGTGAGCCTGAGAACCGCA
ACCGCAAGGGCGTCTTCTGGTACAACAGCAAGTCTAGCCGCGCTTGCTTTTCCCTAGGAAT
GGGAGGAGAAATCAGGATGAGAAGATGTAATGTCTGCATCTACCTGTCTAATGGTTAGTACCA
GTCTTTAGACAGGAAGAGAGCATTGGGCTTCAGAAAAGGAGGCTTACTGCACTACTGCAGTGC
CATCGCTAGATTTAAGGCAAATTCAGTGTGTCTGTGCCCATGGCTGTGAGCTTTGGGTACTCTCA
AGTAGTCAAGTCTCTTGTGTTTTGTGTTTAGTCGTGCTGTTGTAGGCTTGCCGGCGGCGGTGCT
TGCGTGGCCGCGCCTTGTGCTGTGCGTCTTGCCATCTCTTCTGCTCCCTCGTGTGTTGTAAA
ACACTAGTCTGGTGTCTTGGCGGAATAACAGATCGTCGAACGACA

图 5B: SEQ ID NO: 19-ZmFAD-01 开放读框的核酸序列

ATGGGTGCCGGAGGCAGGATGACCGAGAAGGAGCGGGAGAAGCATGAGCAGGAGCAGGTGCG
CCGTGCTACCGGCGGTGGCGCGGCAGTGACGCGGTGCGCGGTGGAGAAGCCGCGTTCACGT
TGGGGCAGATCAAGAAGGCGATCCCGCCGCACTGCTTCGAGCGCTCCGTGCTGAGGTCCTTCT
CGTACGTGGCCACGACCTGTGCTCGCCGCGGCGCTCCTCTACCTCGCGGTGGCCGCTGATC
CGGCGCTACCCTGCCGCTCCGCTACGCGCGCTGCGCGTGTACTGGGTGGCCAGGGGTGC
GTGTGCACGGCGGTGGGTGATCGCGCACGAGTGCGGCCACACGCTTCTCCGACCACGC
GCTCCTGGACGCGCGTGGCGCTGGCGCTGCACTCGGCGCTGCTGGTGCCCTACTTCTCGTG
GAAGTACAGCCACCGGCGCCACCACTCCAACACGGGGTCCCTGGAGCGCGACGAGGTGTTCTG
GCCGAGGACCAAGGAGGCGCTGCCGTGGTACGCCCGTACGTGCACGGCAGCCCCGCGGGCC
GGCTGGCGCACGTGCGCGTGCAGCTCACCCTCGGCTGGCCGCTGTACCTGGCCACCAACGCGT
CGGGGCGGCGTACCCGCGCTTCGCTGCCACTTCGACCCCTACGGCCCCATCTACAACGACC
GGGAGCGCGCCAGATCTTCTCGTCTCGGACGCGGCGTGTGGCCGTGGCGTTCGGGCTGTAC
AAGCTGGCGGCGGCGTTCGGGGTCTGGTGGGTGGTGCGCGTGTACGCCGTGCCGCTGCTGAT
CGTCAACGCGTGGTGGTGTGCTCATCACGTACCTGCAGCACACCCACCGGCGTGGCCACTA
CGACTCGGCGAGTGGGACTGGCTGCGCGGCGGCTGCCACCGTCGACCGAGACTACGGCG
TCCTCAACCGCGTGTTCACCAACATCACGGACACGCACGTGCGCACCACTCTTCTCCACCAT
GCCGCACTACCACGCGTGGAGGCCACAGGGCGATCAGGCCGCTCCTCGGCGAGTACTACCA
GTTGACCCGACCCCTGTGCGCAAGGCCACCTGGCGCGAGGCCAGGGAGTGATCTACGTGCA
GCCTGAGAACCGCAACCGCAAGGGCGTCTTCTGGTACAACAGCAAGTTC

图 5C: SEQ ID NO: 20-ZmFAD-01 开放读框的氨基酸序列

MGAGGRMTEKEREKHEQEVARATGGGAHVQRSPVEKPPFTLGQIKKAIPPHCFERSVLRFSYSVA
HDLSLAALLYLAVAVIPALPCPLRYAAWPLYWVAQGCVCTGWVIAHECGHHAFSDHALLDDAVGL
ALHSALLVPYFSWKYSHRRHHSNTGSLRDEVFVPRTEALPWYAPYVHGSPAGRLAHVAVQLTLG
WPLYLATNASGRPYPRFACHFDPIYPIYNDRERAQIFVSDAGVVAVAFGLYKLAFAFGVWVVRVY
AVPLLVNWLVLITYLQHTHPALPHYDSGEWDWLRGALATVDRDYGVLRVFNHITDTHVAHHLFST
MPHYHAVEATRAIRPVLGEYYQFDPTPAKATWREARECIYVEPENRNRKGVFWYNSKF

图 6A: SEQ ID NO: 21-0sFAD-01 核酸序列

CTCCTCTCCTCCTCCCCTGCACAGACCACTCGTTTCCTCCACAAAGAGGGAGGGAACAAGGGAA
 GGGTGTGCGCCGCCCCCACCCTGATCTGCCTCCGCGCTCCGCTCCTCCGCGCCTGCGAAAT
 CTACCAACGCTAACTCAGCAAGATGGGTGCCGGCGGCAGGATGACGGAGAAGGAGCGGGAGG
 AGCAGCAGAAGCTGCTCGGCCGCGCCGGAATGGCGCGGCCGTGCAGCGGTGCGCCAGCGAC
 AAGCCGCCGTTACGCTGGGGCAGATCAAGAAGGCCATCCCGCTCACTGCTTCCAGCGCTCG
 GTGATCAAGTCCTTCTCCTACGTGGTCCATGACCTCGTGATCGTCGCGCGCTGCTCTACTTCG
 CGCTGGTCATGATCCCCGTGCTGCCGAGCGGGATGGAGTTCGCGGCATGGCCGCTCTACTGGA
 TCGCGCAGGGCTGCGTGCTCACCGGCGTGTTGGTTCATCGCGCACGAGTGCGGCCACCATGCC
 TTCTCCGACTACTCGGTGCTCGACGACATCGTCGGCCTCGTGCTGCACTCGTCGCTGCTCGTCC
 CCTACTTCTCGTGGAAGTACAGCCACCGGCGCCACCACTCCAACACCGGGTCGCTGGAGCGCG
 ACGAGGTGTTTCGTCGCCAAGCAGAAAGTCGGCGATGGCGTGGTACACCCCGTACGTGTACCACA
 ACCCGATCGGCCGGCTGGTGCACATCTTCGTGCAGCTCACCTCGGGTGGCCGCTGTACCTGG
 CGTTCAACGTGTCCGGCCGCCGTACCCGCGCTTCGCGTGCCACTTCGACCCCTACGGCCCGA
 TCTACAACGACCGGGAGCGCGTCCAGATCTTCATCTCCGACGTGGCGCTCGTGTCGGCGGGC
 TCGCCCTGTTCAAGTGTCTGTCGGCGTTCGGGTTCTGGTGGTGGTGGCGCTACGGCGTGC
 CGCTGCTGATCGTGAACGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
 TGCCGCACTACGACTCGAGCGAGTGGGACTGGCTCCGCGGCGCGCTGGCCACCGTGGACCGC
 GACTACGGCATCCTCAACAAGGTGTTCCACAACATCACGGACACGCACGTGCGCACCACTCT
 TCTCCACCATGCCGCACTACCACGCCATGGAGGCCACTAAGGCGATCCGCCCCATCCTCGGCG
 AGTACTACAGTTCGACCCGACGCCGTCGCCAAGGCGACATGGCGCGAGGGCCAAGGAGTGCA
 TCTACGTGAGCCTGAGGACAACAAGGGCGTCTTCTGGTACAACAACAAGTTCTAACTGCTGCT
 GCTGTGAAATCAGCATCACACATCCATAGCCAAGCAGCAAACAATTTGAAGAAGAAGATTACAA
 GGAAGAGAAGATAGTGTCTTCGAAATCGTCGTAGCAAGTATCCATCCATCCATCCAACCCATG
 AACATCGTCTATCTATCCATGCATCTATCTATGTTAGTCTTTAGATAGGAGAGGGCACTTGG
 GCACAGAGGAAGGCTATTGCAGTGCCATTGCTAGAGTTGCCATCAAGTGCAAAGTAGGCGGATC
 AGGCGTGTGTCATGCTGTGGATTTGTAGTCTATGTATGTGTCAGCTGCTGAGCTCCGGTGTG
 CAGCCTTGGTCCCTGTCGTGTTATTTCCATCGTTTTTTCCCTCCGCCATTGTTGGTTTAGGTGT
 TGTATGTCGGCGTCCGTGTGGACGACGTGTCTTGTGCGGTTTGTCTGTATTGAGTTGGCTC
 CGTCCGTTGCTTGTGTTGTAACGCTTGTGGTGTTCATGGCGGAATAACTAAACGTGCAATGG
 AATGACAACCTTTTTGCGTA

图 6B: SEQ ID NO: 23-0sFAD-01 开放读框的核酸序列

ATGGGTGCCGGCGGCAGGATGACGGAGAAGGAGCGGGAGGAGCAGCAGAAGCTGCTCGGCCG
 CGCCGGCAATGGCGCGGCCGTGCAGCGGTGCCGACGGACAAGCCGCCGTTACGCTGGGGC
 AGATCAAGAAGGCCATCCCGCCTCACTGCTTCCAGCGCTCGGTGATCAAGTCCTTCTCCTACGT
 GGTCCATGACCTCGTGATCGTCGCCGCGCTGCTCTACTTCGCGCTGGTTCATGATCCCGTGTG
 CCGAGCGGGATGGAGTTCGCGGCATGGCCGCTCTACTGGATCGCGCAGGGCTGCGTGCTCAC
 CGGCGTGTGGGTTCATCGCGCACGAGTGGCGCCACCATGCCTTCTCCGACTACTCGGTGCTCGA
 CGACATCGTCGGCTCGTGCTGCACTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG
 CACCGGCGCCACCTCCAACACCGGGTCGCTGGAGCGCGACGAGGTGTTTCGTCGCCGAAGCA
 GAAGTCGGCGATGGCGTGGTACACCCCGTACGTGTACCACAACCCGATCGGCCGGCTGGTGCA
 CATCTTCGTGCAGCTCACCTCGGGTGGCGCTGTACCTGGCGTTCAACGTGTCCGGCCGCCC
 GTACCCGCGCTTCGCGTGCCACTTCGACCCCTACGGCCCGATCTACAACGACCGGGAGCGCGT
 CCAGATCTTCATCTCCGACGTGCGCGTGTGTCGCGGGGCTCGCCCTGTTCAAGCTGTGCTC
 GCGGTTCCGGTTCTGGTGGGTGGTGGCGCTTACGGCGTGCCGCTGCTGATCGTGAACGCGTG
 GCTGGTGTCTACCTACCTGCAGCACACCCACCCGCGCTGCCGCACTACGACTCGAGCGA
 GTGGGACTGGCTCCGCGGCGCGCTGGCCACCGTGACCGCGACTACGGCATCCTCAACAAGG
 TGTCCACAACATCACGGACACGCACGTGCGCACCACTCTTCTCCACCATGCCGCACTACCA
 CGCCATGGAGGCCACTAAGGCGATCCGCCCATCCTCGGCGAGTACTACCAGTTCGACCCGAC
 GCCGTCGCCAAGGCGACATGGCGCGAGGCCAAGGAGTGCATCTACGTCGAGCCTGAGGACAA
 CAAGGGCGTCTTCTGGTACAACAACAAGTTC

图 6C: SEQ ID NO: 24-0sFAD-01 开放读框的氨基酸序列

MGAGGRMTEKEREEQQKLLGRAGNGAAVQSRPTDKPFTLGQIKKAIPPHCFQRSVIKSFSYVVHDL
 VIVAALLYFALVMIPVLPSPMEFAAWPLYWIAQGCVLTVWVIAHECGHHAFSDYSVLDDIVGLVLHS
 SLLVPYFSWKYSHRRHHSNTGSLRDEVFVPKQKSAMAWYTPYVYHNPIGRLVHIFVQLTLGWPLYL
 AFNVSGRPYPFRACHFDYPGPIYNDREVRVQIFISDVGVVSAGLALFKLSSAFGFWVVRVYGVPLLV
 NAWLVILITYLQHTHPALPHYDSSEWDWLRLGALATVDRDYGILNKVFHNITDTHVAHHLFSTMPHYHA
 MEATKAIRPILGEYYQFDPTPAKATWREAKECIYVEPEDNKGVFVWYNNKF

图 7A: SEQ ID NO: 25-LuFAD-01 核酸序列

GCTGTAACAATATACACAGGAAGAAGAAAAATGGGTGCCGGCGCAGAATGTCAGTGCCTCCATC
 ATCCAAACCTATGAAGAGGTCTCCTTACTCAAAGCCACCATTACGCTCGGTGAGCTCAAGAAGG
 CCATTCTCCACACTGTTTCAAACGCTCAATCCCCGATCGTTTCGCCTACGTGGCGTACGACCTC
 ACCATTGCAGCAATCTTCTACTACATCGCCACCACTTACTTCCACCTCCTCCCTAGCCCTCTCAA
 CTACCTCGCTGGCCGGTCTACTGGGCCTGCCAGGGCTGCATCCTCACTGGAGTATGGGTGTT
 GGCTCACGAATGCGGTCAACATGCCTTCAGCGACTACCAGTGGCTCGACGACATGGTTGGCTTC
 GTCCTCCATTCTCCCTCCTTGTTCCTTACTTCTCCTGGAAGCACAGCCACCGCCGCCACCATT
 CAACACGGGATCGCTTGATCGTGATGAGGTGTTTGTCCCAAGCAGAAGGCCGAAATCGGGTG
 GTACTCCAAGTACCTTAACAACCCACCTGGCCGTGTGATCACATTGGCCGTACATTAAACGCTCG
 GTTGGCCTCTGTACTTGGCATTCAACGTCTCCGGGAGACCATATGACCGGTTCCGATGCCATTTT
 GACCCTCACGGTCCGATTTACAATGATCGCGAGCGTATGGAGATATACCTATCCGACGCAGGGA
 TATTCACCGTGTGCTACATCCTATACAGACTCGTCCTCACGAAAGGACTCGTTTGGGTCTGTCC
 ATATACGGAGTCCCACTATTGATAGTGAATGGATTCTAGTCCTCATCACTTTCTTGCAGCACAC
 GCATCCTTCTCTTCCGCACTACAAAGTCTCCGAATGGGGACTGGATGCGAGGCGCCCTCTCGA
 CCGTGGATCGAGACTACGGGTACTCAACACCGTGTTCCACAACATCACCGATACACATGTCGC
 GCACCATCTCTTCTCCACGATGCCTCATTACCACGCGATGGAGGCTACCAAGGCGATCAAGCCG
 GTTCTCGGGGAGTATTACCAGTTTCGATGGGACTCCCTTTGTGAAGGCCATGTGGAGGGAGGCAA
 AGGAGTGCATCTATGTCGAGCCGGATGAAGGCGACCCAGCCAAGGCGTGTTCTGGTACAACA
 ACAAGCTGTGAGGGTCTTCGAAATTTGCAGAGGTTTGTAGTGTGTTCTTAATGGTGTACCAG
 AAAAAATGTTTGAAGAAAGAAGCTGCAATAGCTAGTGCAGAACTGGTGTATGTTTCTGTAATGTTT
 TTAAGTTATGTCCCTAGTGGTCGTTAATGTTACTGTACTTCTCTGTTCTTCTCCATCGAGCCAACA
 TACCTTCACTCCTCTGTTAATGTACTGAGTTGGTCGAGTTTAACTTAACGGACCACCAGGCTCA
 AATTCGAGTCACCGGGTGGCCGAGTTAGACTGCATTGACCACAATGATGCAATCGAAAACCTG
 AAGTGACTACAATCGAAAACCTAATTCCAGTCA

图 7B: SEQ ID NO: 27-LuFAD-01 开放读框的核酸序列

ATGTCAGTGCCTCCATCATCCAAACCTATGAAGAGGTCTCCTTACTCAAAGCCACCATTACGCT
 CGGTGAGCTCAAGAAGGCCATTCTCCACACTGTTTCAAACGCTCAATCCCCGATCGTTTCGCC
 TACGTGGCGTACGACCTCACCATTGCAGCAATCTTCTACTACATCGCCACCACTTACTTCCACCT
 CCTCCCTAGCCCTCTCAACTACCTCGCCTGGCCGGTCTACTGGGCCTGCCAGGGCTGCATCCTC
 ACTGGAGTATGGGTGTTGGCTCACGAATGCGGTCAACATGCCTTCAGCGACTACCAGTGGCTCG
 ACGACATGGTTGGCTTCGTCTCCTCATTCTGTCCTCCTTGTTCCTTACTTCTCCTGGAAGCACAGC
 CACCGCCGCCACCATTCCAACACGGGATCGCTTGATCGTGATGAGGTGTTTGTCCCAAGCAGA
 AGGCCGAAATCGGGTGGTACTCCAAGTACCTTAACAACCCACCTGGCCGTGTGATCACATTGGC
 CGTCACATTAAACGCTCGGTGGCCTCTGTACTTGGCATTCAACGTCTCCGGGAGACCATATGAC
 CGGTTTCGATGCCATTTGACCCTCACGGTCCGATTTACAATGATCGCGAGCGTATGGAGATATA
 CCTATCCGACGCAGGGATATTCACCGTGTGCTACATCCTATACAGACTCGTCCTCACGAAAGGA
 CTCGTTTGGGTCTGTCCATATACGGAGTCCCACTATTGATAGTGAATGGATTCTAGTCCTCAT
 CACTTTCTTGCAGCACACGCATCCTTCTTCCGCACTACAAAGTCTCCGAATGGGGACTGGAT
 GCGAGGCGCCCTCTCGACCGTGGATCGAGACTACGGGTACTCAACACCGTGTTCCACAACATC
 ACCGATACACATGTCGCGCACCATCTCTTCTCCACGATGCCTCATTACCACGCGATGGAGGCTA
 CCAAGGCGATCAAGCCGGTCTCGGGGAGTATTACCAGTTCGATGGGACTCCCTTG

图 7C: SEQ ID NO: 28-LuFAD-01 开放读框的氨基酸序列

MSVPPSSKPMKRSPYSKPPFTLGELKKAIPPHCFKRSIPRSFAYVAYDLTIAAIFYIATTFHLLPSPLN
 YLAWPVYWACQGCILTGWVLAHECGHHAFSDYQWLDDMVGFVLHSSLLVPYFSWKHSHRRHHSN
 TGS�DRDEVFVPKQKAEIGWYSKYLNPPGRVITLAVTLTLGWPLYLAFNVSGRPYDRFACHFDPHG
 PIYNDRERMEIYLS DAGIFTVCYILYRLVLTGLVWVVSİYGVPLLVNGFLVLITFLQHTHPSLPHYKVL
 RMGTGCEAPSRPWİETTGYSTPCSTTSPIHMSRTİSSPRCLİTRWRLPRRSSRFSGSİTSSMGLPL

图 8A: SEQ ID NO: 29-HvFAD-01 核酸序列

ACCACCACCACCCCTACCAGCATCATGGGTGCCGGCGCGGGATGACCGAGAAGGAGCGGGA
 GAAGCAGGAGCAGCTCGGCCGCGCCGGCGGGCGGCAGCCTTCCAGCGCTCGCCGACGGACA
 AGCCGCCGTTCCAGCTCGGTGAGATCAAGAAGGCGATCCCGCCTCACTGCTTCCAGCGCTCCAT
 CATCAAGTCTTCTCCTACGTGGTTCATGACCTGGTCATCATCGCCGCCCTGCTGTACGCCGCT
 CTGGTCTGGATCCCCACCCTCCCTACCGTGTTCAGCTGGGCGCGTGGCCGCTCTACTGGATC
 GTTCAGGGCTGCGTCATGACCGGCGTCTGGGTATCGCGCACGAGTGCGGCCACCATGCCTTC
 TCTGACTACTCGCTGCTCGACGACACCGTCGGCCTGGTGCTCCACTCGTGGCTGCTCGTCCCAT
 ACTTCTCGTGGAAGTACAGCCACCGTCGCCACCACTCCAACACCGGGTCTGCTGGAGCGCGACG
 AGGTGTTTGTCCCCAAGCAGAAGGAGGCGCTGGCATGGTACACTCCCTACATCTACAACAACCC
 CATCGGCCGCTCTGGTGCATCGTGTGACGCTCACCTCGGGTGGCCGCTGTACCTGGCGCT
 CAACGCCTCAGGCCGTCCTGACCCGCGCTTCGCCTGCCACTTCGACCCCTACGGCCCGATCTA
 CAACGACCGGGAGCGAGCCAGATTTTCATCTCGGATGTGCGCGTGTGGCCGCTCTCCTTGGC
 CCTGCTCAAGCTTGTGTGCTGTTTGGGTTCTGGTGGGTGGTGGCGGTCTACGGCGTGCCGCT
 GCTGATCGTGAACGCGTGGCTGGTCTGATCACCTACCTGCAGCACACCCACCCAGCGCTGCC
 GCACTACGACTCGACGGAGTGGGACTGGCTGCGGGGGGCGCTCGCCACCATGGACCGGGACT
 ACGGCATTCTCAACCGCGTGTTCACAACATCACGGACACGCACGTGGCGCACCACTCTTCTC
 CAACATGCCGCACTACCACGCCATGGAGGCCACCAAGGCGATCAAGCCCATCTCGGCGAGTA
 CTACCAGTTTGACGGCACCCCGGTGCGCAAGGCCACATGGCGCGAGGCCAAGGAGTGCATCTA
 CGTTGAGCCCGAGGACCGCAAGGGGGTCTTCTGGTACAGCAACAAGTTCTAGCCGCAAGGATC
 GTCATCAGCCGTGTTCCAGGAAGAACTCAGAGAAGAGGTCTTACAAGTAATCCATCCATCTACC
 TACATATGGTTAGTTTTAGATAGCAGAGGGCATTTGGGCACAAACAAGACTACTATTACCGTGC
 CAATGCTAGAAAGAGTTGAGTGGTGAAGGAGGAGTAGCGTGTCCGTGACTTTGTGAGTTTCT
 TCTTTACTTTCTCCTCGCTCTAGTCGCCGGCGGTGTTGTTGGTGTCCGTGGCTTACATG
 GCCGTGTGTGTTGTGCGTCTGTGATTGCATTGGCGTCATCTCCCCCGTCCGTGTCATGTT
 GTTGTAGACCATTTCTGTTTT

图 8B: SEQ ID NO: 31-HvFAD-01 开放读框的核酸序列

ATGGGTGCCGGCGGGCGGGATGACCGAGAAGGAGCGGGAGAAGCAGGAGCAGCTCGGCCGCG
 CCGGCGGGCGGGCGCAGCCTTCCAGCGCTCGCCGACGGACAAGCCGCCGTTCCAGCTCGGTGAG
 ATCAAGAAGGCGATCCCGCCTCACTGCTTCCAGCGCTCCATCATCAAGTCTTCTCCTACGTGGT
 TCATGACCTGGTCATCATCGCCGCCCTGCTGTACGCCGCTCTGGTCTGGATCCCCACCCTCCCT
 ACCGTGTTGAGCTGGGCGCGTGGCCGCTCTACTGGATCGTTCAGGGCTGCGTCATGACCGGC
 GTCTGGGTATCGCGCACGAGTGCGGCCACCATGCCTTCTGACTACTCGTGTCTGACGACA
 CCGTCGGCCTGGTGTCTCACTCGTGGTGTCTGCTCCATACTTCTCGTGGAAGTACAGCCACCG
 TCGCCACCACTCCAACACCGGGTCTGAGCGCGAGGAGGTGTTTGTCCCAAGCAGAAGGA
 GCGCTGGCATGGTACACTCCCTACATCTACAACAACCCCATCGGCCGTCTGGTGCATCGTG
 GTGACGCTCACCTCGGGTGGCCGCTGTACCTGGCGCTCAACGCCTCAGGCCGTCCGTACCCG
 CGCTTCGCCTGCCACTTCGACCCCTACGGCCCGATCTACAACGACCGGGAGCGAGCCAGATT
 TTCATCTCGGATGTGCGCGTGTGGCCGCTCTCCTTGGCCCTGCTCAAGCTTGTGTCGTTTTG
 GGTTCTGGTGGTGGTGGCGGTCTACGGCGTGCCGCTGCTGATCGTGAACGCGTGGCTGGTCC
 TGATCACCTACCTGCAGCACACCCACCCAGCGCTGCCGCACTACGACTCGACGGAGTGGGACT
 GGCTGCGGGGGGCGCTCGCCACCATGGACCGGGACTACGGCATTCTCAACCGCGTGTTCACA
 ACATCACGGACACGCACGTGGCGCACCACTCTTCTCAACATGCCGCACTACCACGCCATGGA
 GGCCACCAAGGCGATCAAGCCCATCTCGGCGAGTACTACCAGTTTGACGGCACCCCGGTGCG
 CAAGGCCACATGGCGCGAGGCCAAGGAGTGCATCTACGTTGAGCCCGAGGACCGCAAGGGGG
 TCTTCTGGTACAGCAACAAGTTC

图 8C: SEQ ID NO: 32-HvFAD-01 开放读框的氨基酸序列

MGAGGGMTEKEREKQELGRAGGGAFFQSPDKPFTLGQIKKAIPPHCFQRSIIKSFSYVVDLVI
 IAALLYAALVWIPTLPTVLQLGAWPLYWIVQGCVMGTGWVIAHECGHHAFFSDYSLDDTVGLVLHSLV
 LVPYFSWKYSHRRHHSNTGSLERDEVFVPKQKEALAWYTPYIYNNPIGRLVHIVVQLTLGWPLYLALN
 ASGRPYPRFACHFDYPYGPYNDREARQIFISDVGLAVSLALLKLVSSFGFWVVRVYGVPLLVNAW
 LVLITYLQHTHPALPHYDSTWDWLRGALATMDRDYGLNRVFNITDTHVAHHLFSNMPHYHAMEA
 TKAIPILGEYYQFDGTPVAKATWREAKECIYVEPEDRKGVFWYSNKF

图 9A: SEQ ID NO: 33-TaFAD-01 核酸序列

CGCGTCCGCGGCTCCTCCCCCGCACAAACCACTCGTTCTGTCCTCCGTCAACAAGAGGAGCAGA
GGCGCCGGAGAGGGAAGAGGGTGCGCGCGCTCGCGTGTGTGGTGTCCGCCGGCCCGATCTGC
CCTGCTCCCGCCGCTCGACCACCACCCCTATCAGCATCATGGGTGCCGGCGGCAGGATGACG
GAGAAGGAGCGGGAGAAGCAGGAGCAGCTCGGCCGCGCCAACGGCGGCGCAGCCTACCAGC
GCTCGCCGACGGACAAGCCGCCGTTACGCTGGGTGAGATCAAGAAGGCAATCCCGCCTCACT
GCTTCAGCGCTCGATCATCAAGTCCTTCTCCTACGTGGTCCATGACCTGGTCATCGTCGCGGC
CCTGCTGTACGCGGCGCTGGTTTGGATCCCTACCCTCCCGACCGTGCTGCAGCTGGGCGCCTG
GCCGCTCTACTGGATCGTGCAGGGCTGCGTCATGACCGGCGTCTGGGTGATCGCCACGAGTG
CGGCCACCACGCCTTCTCCGACTACTCGTGTCTCGACGACACCGTCGGCCTGGTGTCCACTC
GTGGCTGCTCGTCCCCTACTTCTCGTGGAAGTACAGCCACCGTCGCCACCACTCCAACACCGGG
TCGCTGGAGCGTGATGAGGTGTTCTGTCCTCCCAAGCAGAAGGAGGCGCTGGCGTGGTACACCCCT
TACATCTACAACAACCCCGTCGGCCGCTCTGGTGCACATCGTCTGTCAGCTCACCTCGGGTGGC
CGCTGTACCTGGCGCTCAACGCCTCAGGCCGCCGTACCCGCGGTTGCCTGCCACTTCGACC
CCTACGGCCCGATCTACAACGACCGGGAGCGAGCCAGATTTTCATCTCAGACGTGCGAGTGCT
GGCCGTGTCTATTGGCTCTGCTGAAGCTCGTGTCTGTCGTTTCGGGTTCTGGTGGTGGTGGTGGT
CTACGGCGTGCCGCTGCTGATCGTGAACGCTTGGCTGGTCTGATCACCTACCTGCAGCACACC
CACCCGGCGCTGCCGCACTACGACTCGACGGAGTGGGACTGGCTGCGCGGGGCGCTCGCCAC
CATGGACCGCGACTACGGCATCCTCAACCGCGTGTTCACAACATCACGGACACGCACGTGGC
GCACCACCTCTTCTCCACCATGCCGCACTACCACGCCATGGAGGCCACCAAGGCGATCAAGCC
CATCCTCGGCGAGTACTACCAGTTCGACCCACCCCGCTCGCCAAGGCCACATGGCGCGAGGC
CAAGGAGTGCATCTACGTCGAGCCCGAGGACCGCAAGGGGGTCTTCTGGTACAGCAACAAGTT
CTAGCCGCCAAGATCCATCAACTGTGCTGGAGAAAGAACTCAGAGAAGAGATCCTACCAAGTAA
TTCCATCCATCTACCTACAGTCCATATGGTTAGTCTTTAGATAGCAGAGGGCATTTGGGCACAAA
AGAAGACTACTATTACCGTGCCCAATGTACAAGAGCTGAGTGGTGCAAGGAAGAGTAGCGTGTCT
CGTGACTTTGGTCAGTTCCGTCTTTACTTTTTCTCTGCGTTCTAGTCGTCGGCTTAGGTTTGGCC
GGCGGTGATCGTTGGTGTCCGTGGCCGTGGACATGGCCGCGTGTGTTGTGTGTGCGTCTGTCA
TTGCATTGGCGTCATCTCCCCCGTCCGTGTATGTTGTTGTAGACCATTTCGTGTTTATGGCGG
AATAACTGATCGTCGAAGGAAGGGCAACTTTTTTGAAGTA

图 9B: SEQ ID NO: 35-TaFAD-01 开放读框的核酸序列

ATGGGTGCCGGCGGCAGGATGACGGAGAAGGAGCGGGAGAAGCAGGAGCAGCTCGGCCGCG
CCAACGGCGGCGCAGCCTACCAGCGCTCGCCGACGGACAAGCCGCCGTTACGCTGGGTGAG
ATCAAGAAGGCAATCCCGCCTCACTGCTTCCAGCGCTCGATCATCAAGTCCTTCTCCTACGTGGT
CCATGACCTGGTCATCGTCGCGGCCCTGCTGTACGCGGCGCTGGTTTGGATCCCTACCCTCCC
GACCGTGCTGCAGCTGGGCGCCTGGCCGCTCTACTGGATCGTGCAGGGCTGCGTCATGACCGG
CGTCTGGGTGATCGCCACGAGTGCGGCCACCACGCCTTCTCCGACTACTCGTGTCTGACGA
CACCGTCGGCCTGGTGTCTCACTCGTGCTGCTCGTCCCCTACTTCTCGTGGAAGTACAGCCAC
CGTCGCCACCACTCCAACACCGGGTGGTGGAGCGTGATGAGGTGTTCTGTCCTCCCAAGCAGAAG
GAGGCGCTGGCGTGGTACACCCCTTACATCTACAACAACCCCGTCGGCCGCTCTGGTGCACATC
GTCGTGCAGCTCACCTCGGGTGGCCGCTGTACCTGGCGCTCAACGCCTCAGGCCGCCGCTAC
CCGCGGTTGCTGCTGCCACTTCGACCCCTACGGCCCGATCTACAACGACCGGGAGCGAGCCAG
ATTTTCATCTCAGACGTGCGAGTGCTGGCCGTGTCTATTGGTCTGCTGAAGCTCGTGTCTGCTGTT
CGGGTTCTGGTGGGTGGTGGCGGTCTACGGCGTGCCGCTGCTGATCGTGAACGCTTGGCTGGT
CCTGATCACCTACCTGCAGCACACCCACCCGGCGCTGCCGCACTACGACTCGACGGAGTGGGA
CTGGCTGCGCGGGGCGCTCGCCACCATGGACCGCGACTACGGCATCCTCAACCGCGTGTTCAC
CAACATCACGGACACGCACGTGGCGCACCACTCTTCTCCACCATGCCGCACTACCACGCCATG
GAGGCCACCAAGGCGATCAAGCCCATCCTCGGCGAGTACTACCAGTTCGACCCACCCCGCTC
GCCAAGGCCACATGGCGCGAGGCCAAGGAGTGCATCTACGTCGAGCCCGAGGACCGCAAGGG
GGTCTTCTGGTACAGCAACAAGTTCT

图 9C: SEQ ID NO: 36-TaFAD-01 开放读框的氨基酸序列

MGAGGRMTEKEREKQEQLGRANGGAAYQRSPTDKPPFTLGQIKKAIPPHCFQRSIIKSFSYVVDLVI
VAALLYAALVWIPTLPTVLQLGAWPLYWIVQGCVMGTGVWVIAHECGHHAFSDYSLDDTVGLVLHSW
LLVPYFSWKYSHRRHHSNTGSLERDEVFPKQKEALAWYTPYIYNNPVGRLVHIVVQLTLGWPLYLA
LNASGRPYPRFACHFDYPYGPYNDREARQIFISDVGVLAVSLALLKLVSSFGFWVVRVYGVPLLVN
AWLVLITYLQHTHPALPHYDSTEWDWLRGALATMDRDYGILNRFVHNITDTHVAHHLFSTMPHYHAM
EATKAIPKILGEYYQFDPTPVAKATWREAKECIYVEPEDRKGVFWYSNKF

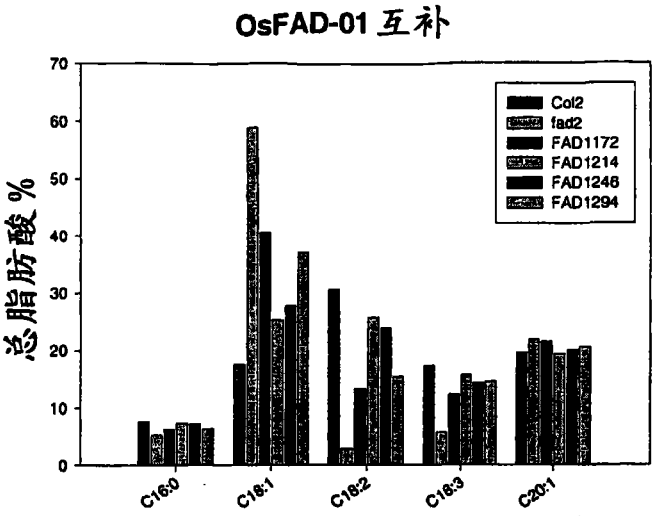


图 10

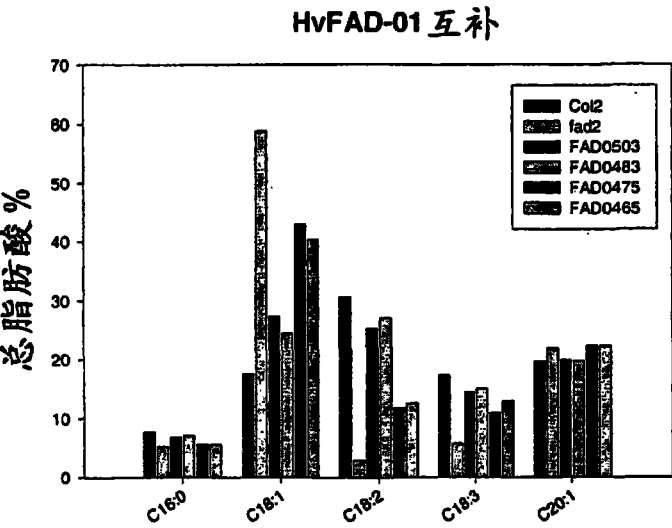


图 11

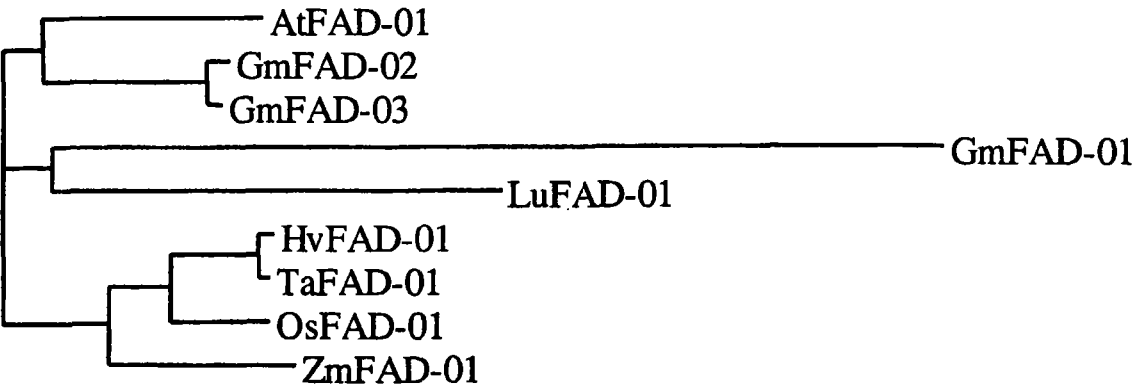


图 12

	AlFAD- 01	GmFAD- 02	GmFAD- 03	GmFAD- 01	LuFAD- 01	HvFAD- 01	TaFAD- 01	OSPAD- 01	ZnFAD- 01
AlFAD- 01	100	78	47	29	56	68	68	68	67
GmFAD- 02		100	58	29	59	72	72	73	70
GmFAD- 03			100	33	34	44	44	45	43
GmFAD- 01				100	24	28	28	28	27
LuFAD- 01					100	50	50	51	49
HvFAD- 01						100	98	89	80
TaFAD- 01							100	89	81
OSPAD- 01								100	80
ZnFAD- 01									100

图 13

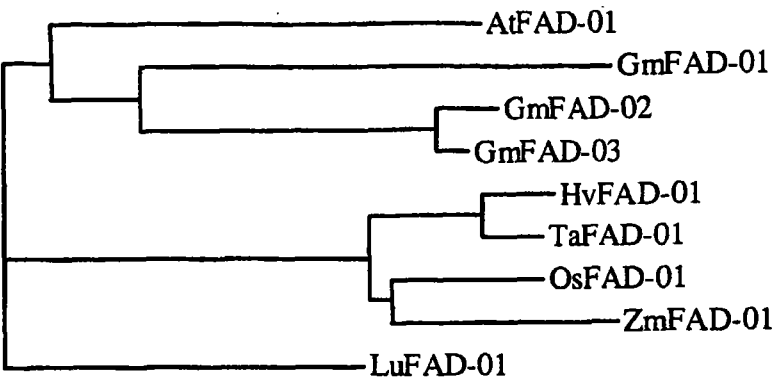


图 14

	AlFAD	GmFA	GmFA	GmFA	HvFAD	LuFAD	OsFAD	TaFAD	ZnFAD
	-01	D-01	D-02	D-03	-01	-01	-01	-01	-01
AlFAD-01	100	65	72	43	62	64	62	63	60
GmFA-D-01		100	71	42	60	59	60	59	58
GmFA-D-02			100	56	61	64	61	61	60
GmFA-D-03				100	36	39	36	36	35
HvFAD-01					100	60	88	95	82
LuFAD-01						100	61	61	59
OsFAD-01							100	89	86
TaFAD-01								100	83
ZnFAD-01									100

图 15

1 50

SEQ ID NO:2 (1) MGAGGRMPVPTSSKKS-----E--TDTTK-RVPCEKPPFSVGD LKKAIP

SEQ ID NO:6 (1) -----

SEQ ID NO:10(1) MGAGGRTDVPPANIKS-----E--VDPLK-RVPFEKPPFSISQKKVIP

SEQ ID NO:14(1) -----

SEQ ID NO:26(1) -----MSVPPSSKPMK-----RSPYSKPPFGLGLKKAIP

SEQ ID NO:18(1) MGAGGRMTEKEREKHEQEQVARATGGGAQVRSVPVEKPPFGLGQKKKAIP

SEQ ID NO:30(1) MGAGGGMTEKEREKQE---QLGRAGGGAQFORSPTKPPFGLGQKKKAIP

SEQ ID NO:34(1) MGAGGRMTEKEREKQE---QLGRANGGAAYQRSPTKPPFGLGQKKKAIP

SEQ ID NO:22(1) MGAGGRMTEKEREEOQK--LLGRAGNGAAVQRSPTKPPFGLGQKKKAIP

51 100

SEQ ID NO:2 (42) PHCFKRSIPRSFSY LISDIIASC FYVATNYFSLLPQPLSYLAWPLYWA

SEQ ID NO:6 (1) -----MSYWRDVIIVFGLAVAAAYLN-----NWWVWPLYWA

SEQ ID NO:10(42) PHCFORSIFRSFSYVYDEITIAFCLYYVATHYFHL LPSPLSLAWPLYWA

SEQ ID NO:14 (1) -----

SEQ ID NO:26(31) PHCFKRSIPRSFYVAYDEITIAIFYYIATTYFHL LPSPLNYLAWPLYWA

SEQ ID NO:18(51) PHCFERSVLRFSYVAHDSIAALLYLAVAVIPALPCPLRYAAWPLYWV

SEQ ID NO:30(48) PHCFORSIIKSFSYVAHDSIIALLYAALVVIPTLPTVLQLGAWPLYWI

SEQ ID NO:34(48) PHCFORSIIKSFSYVAHDSIIVIALLYAALVVIPTLPTVLQLGAWPLYWI

SEQ ID NO:22(49) PHCFORSVIKSFSYVAHDSIIVIALLYFALVMIPVLPSGMEFAAWPLYWI

101 150

SEQ ID NO:2 (92) CQGCVL TGIWVIAHECGHHAFSDYQWLDDTVGLIFHSFLLVPYFSWKYSH

SEQ ID NO:6 (33) AQGTMFWALEVVGHCCHGCSFNNPKLNSVAGHLHSSLLVPYHGWKISH

SEQ ID NO:10 (92) VQGCILTGWVIAHECGHHAFSDYQLLDDIVGLVLHSGLLVPYFSWKYSH

SEQ ID NO:14 (1) -----

SEQ ID NO:26(81) CQGCILTGWVIAHECGHHAFSDYQWLDDMVGFVLHSSLLVPYFSWKISH

SEQ ID NO:18(101) AQGCVCTGWWVIAHECGHHAFSDYALLDDAVGLALHSALLVPYFSWKYSH

SEQ ID NO:30(98) VQGCVM TGVWVIAHECGHHAFSDYSLDDTVGLVLHSSLLVPYFSWKYSH

SEQ ID NO:34(98) VQGCVM TGVWVIAHECGHHAFSDYSLDDTVGLVLHSSLLVPYFSWKYSH

SEQ ID NO:22(99) AQGCVLTGVWVIAHECGHHAFSDYSVLDDIVGLVLHSSLLVPYFSWKYSH

151 200

SEQ ID NO:2 (142) RRHHSNTGSLERDEVFVPKQKSAIKWYGKYL N-NPLGRIMMLTVQFVGLGW

SEQ ID NO:6 (83) RTHHQNHGHVENDESHPLPEKIFKSLDNVTR--ILR---FTLPFPLAY

SEQ ID NO:10(142) RRHHSNTGSLERDEVFVPKQKSCIKWYSKYLN-NPPGRVETLAVTLTGLW

SEQ ID NO:14 (1) -----MWYSKYLN-NPPGRVETLAVTLTGLW

图 16

SEQ ID NO:20(151)RRHHSNTGSLERDEVFVPKQKEAIGWYSKYLN-NPPGRVTLAVTLTGW
SEQ ID NO:18(151)RRHHSNTGSLERDEVFVPKKEAPWYAPYHGSPAGRAHVAVOLTTGW
SEQ ID NO:30(148)RRHHSNTGSLERDEVFVPKQKEALAWYTPYVYNNPIGREVHVVOLTTGW
SEQ ID NO:34(148)RRHHSNTGSLERDEVFVPKQKEALAWYTPYVYNNPIGREVHVVOLTTGW
SEQ ID NO:22(149)RRHHSNTGSLERDEVFVPKQKSALAWYTPYVYHNPVIGREVHVVOLTTGW

201

250

SEQ ID NO:2(191)ELVAFNVSERPYDGFACFFENAPIYNDREELQIYLEDAGILAVCFGY
SEQ ID NO:6(128)ELVWSRSPK----TSEFMDSDDEVPSKDKVITSTVCWTAMAALV
SEQ ID NO:10(191)ELVAFNVSERPYDRFACFDEYCPYSDREELQIYLEDAGILAVCFGY
SEQ ID NO:14(26)ELVAFNVSERPYDRFACFDEYCPYSDREELQIYLEDAGILAVCFGY
SEQ ID NO:26(180)ELVAFNVSERPYDRFACFDEYCPYSDREELQIYLEDAGIFTVCMY
SEQ ID NO:18(201)ELVAFNVSERPYPRFACFDEYCPYSDREELQIYLEDAGILAVCFGY
SEQ ID NO:30(198)ELVAFNVSERPYPRFACFDEYCPYSDREELQIYLEDAGILAVCFGY
SEQ ID NO:34(198)ELVAFNVSERPYPRFACFDEYCPYSDREELQIYLEDAGILAVCFGY
SEQ ID NO:22(199)ELVAFNVSERPYPRFACFDEYCPYSDREELQIYLEDAGILAVCFGY

251

300

SEQ ID NO:2(241)RYAAAOEMASMICLGVELLIVNAFVLIITYOFT--HPSLPHDSSEWD
SEQ ID NO:6(174)GLGFVMEFVQLLKLGLGYAIFVMEDVYHGHGHEDKKEWGEES
SEQ ID NO:10(241)RLAMAKELAWVVCVGVLLVNCFFVLIITYOFT--HPALPHDSSEWD
SEQ ID NO:14(76)CLAMAKELAWVVCVGVLLVNCFFVLIITYOFT--HPALPHDSSEWD
SEQ ID NO:26(230)RLVLTKEFVWVSVGVLLVNCFFVLIITYOFT--HPSLPHDKVLRMG
SEQ ID NO:18(251)KLAAAFGVWVVRVGVLLVNAFVLIITYOFT--HPALPHDSGEWD
SEQ ID NO:30(248)KLVSSEFFWVVRVGVLLVNAFVLIITYOFT--HPALPHDSGEWD
SEQ ID NO:34(248)KLVSSEFFWVVRVGVLLVNAFVLIITYOFT--HPALPHDSGEWD
SEQ ID NO:22(249)KLSSAFEFWVVRVGVLLVNAFVLIITYOFT--HPALPHDSSEWD

301

350

SEQ ID NO:2(289)WLRGALATVDRDYGILNKVFHNITDTHVAHFLSTMPHYNAMEAKAIKE
SEQ ID NO:6(224)WLRGLTTDRDYGILNKVHHDIGTHVHFLFPOPHYHLAEAEAAKE
SEQ ID NO:10(289)WLRGALATVDRDYGILNKVFHNITDTHVAHFLSTMPHYNAMEAKAIKE
SEQ ID NO:14(124)WLRGALATVDRDYGILNKVFHNITDTHVAHFLSTMPHYNAMEAKAIKE
SEQ ID NO:26(278)TGCEAPSRP---WIETGCTSPCSTSPHMSRTSSPRCLITERWRL
SEQ ID NO:18(299)WLRGALATVDRDYGILNKVFHNITDTHVAHFLSTMPHYNAMEAKAIKE
SEQ ID NO:30(296)WLRGALATVDRDYGILNKVFHNITDTHVAHFLSNMPHYNAMEAKAIKE
SEQ ID NO:34(296)WLRGALATVDRDYGILNKVFHNITDTHVAHFLSTMPHYNAMEAKAIKE
SEQ ID NO:22(297)WLRGALATVDRDYGILNKVFHNITDTHVAHFLSTMPHYNAMEAKAIKE

351

400

SEQ ID NO:2 (339) ILGDYYQFDGTPWYVAMYREAKECIYVEPDREGDKKGVYWYNNKL-----
 SEQ ID NO:6 (273) VLGGYYREPKKSSPPIYLIGELLRSKKDHFVSDSGDIVYYGTDPTLSS
 SEQ ID NO:10(339) ILGGYYRFDETPEVKAMWREARECIYVEPDQSTESKGVWYNNKL-----
 SEQ ID NO:14(174) ILGGYYRFDGTPPEVKAMWREARECIYVEPDQSTQSKGVWYNNKL-----
 SEQ ID NO:26(324) RRSSRESGSITSSMGLPL-----
 SEQ ID NO:18(349) VLGGYYQFDPTPVAKATWREARECIYVEPENRN-RKGVWYNSKF-----
 SEQ ID NO:30(346) ILGGYYQFDGTPVAKATWREAKECIYVEPENR---KGVWYNSKF-----
 SEQ ID NO:34(346) ILGGYYQFDPTPVAKATWREAKECIYVEPENR---KGVWYNSKF-----
 SEQ ID NO:22(347) ILGGYYQFDPTPVAKATWREAKECIYVEPENR---KGVWYNNKF-----

401

SEQ ID NO:2 (384)-----
 SEQ ID NO:6 (323)SSTSQ
 SEQ ID NO:10(384)-----
 SEQ ID NO:14(219)-----
 SEQ ID NO:26(342)-----
 SEQ ID NO:18(393)-----
 SEQ ID NO:30(388)-----
 SEQ ID NO:34(388)-----
 SEQ ID NO:22(389)-----