



(12) 发明专利

(10) 授权公告号 CN 101200728 B

(45) 授权公告日 2011.02.09

(21) 申请号 200710199076.1
 (22) 申请日 2004.12.22
 (30) 优先权数据
 2003-425673 2003.12.22 JP
 (62) 分案原申请数据
 200480038471.6 2004.12.22
 (73) 专利权人 三得利控股株式会社
 地址 日本大阪府
 (72) 发明人 大山莞尔
 (74) 专利代理机构 北京信慧永光知识产权代理
 有限责任公司 11290
 代理人 梁兴龙 王维玉

Girkel T 等. Identification of a novel D6-acyl-group desaturase by targeted gene disruption in *Physcomitrella patens*. The Plant Journal 15 1. 2002, 15(1), 第 39-48 页.
 Kajikawa M 等. AY583465, *Marchantia polymorpha* polyunsaturated fatty acid delta-5-desaturase (DES5) gene, complete cds, 5822bp, DNA, linear. NCBI. 2004, 第 1-3 页.
 Kajikawa M 等. Isolation and characterization of D6-desaturase, an ELO-like enzyme and D5-desaturase from the liverwort *Marchantia polymorpha* and production of a rachidonican deicosapentaenoic acids in the methylotrophic yeast *Pichia pastoris*. Plant Molecular Biology 54 3. 2004, 54(3), 第 335-352 页.
 Michaelson LV 等. Isolation of a $\Delta 5$ -Fatty Acid Desaturase Gene from *Mortierella alpina*. Journal of Biological Chemistry 273 30. 1998, 273(30), 第 19055-19059 页.
 Kajikawa M 等. AY583464, *Marchantia polymorpha* polyunsaturated fatty acid elongase (EL01) gene, complete cds, 3875 bp, DNA, linear. NCBI. 2004, 第 1-4 页.

(51) Int. Cl.
C12N 15/52 (2006.01)
C12N 9/00 (2006.01)
C12N 15/82 (2006.01)
C07K 16/40 (2006.01)
C12P 7/64 (2006.01)
C12Q 1/68 (2006.01)
C12Q 1/25 (2006.01)
A01H 1/00 (2006.01)
 (56) 对比文件
 WO 00/75341 A, 2000.12.14, 全文.
 CN 1334340 A, 2002.02.06, 全文.
 JP 特表 2003-523746 A, 2003.08.12, 全文.
 JP 特表 2003-509050 A, 2003.11.03, 全文.
 CN 1108382 C, 2003.05.14, 全文.

审查员 朱宁

权利要求书 1 页 说明书 44 页 附图 1 页

(54) 发明名称

来自地钱的饱和脂肪酸合成系统酶基因及其利用

(57) 摘要

本发明从同种地钱中分离 $\Delta 5$ 脂肪酸去饱和酶基因、 $\Delta 6$ 脂肪酸去饱和酶基因及 $\Delta 6$ 脂肪酸碳链延长酶基因, 将这些基因导入高等植物, 从而获得能够生产花生四烯酸、二十碳五烯酸 (EPA) 的转基因植物体。

CN 101200728 B

1. 来自地钱目生物的基因,所述基因由序列号 3 所示的碱基序列构成,且编码具有 $\Delta 6$ 脂肪酸碳链延长活性的蛋白质。

2. 对具有 $\Delta 6$ 脂肪酸碳链延长活性且来自地钱目生物的蛋白质进行编码的基因,其中,所述基因对由序列号 4 所示的氨基酸序列构成的蛋白质进行编码。

3. 由权利要求 1 或 2 所述的基因所编码的蛋白质。

4. 含有权利要求 1 或 2 所述基因的重组表达载体。

5. 导入权利要求 1 或 2 所述基因而形成的转化体,其中,所述转化体为甲醇酵母转化体。

6. 一种生产导入了权利要求 1 或 2 所述基因的植物体、或与该植物体具有同一性状的该植物体的后代的方法,所述方法包括将权利要求 1 或 2 所述的基因导入宿主细胞。

7. 一种生产导入了权利要求 1 或 2 所述基因的植物体的组织的方法,所述方法包括将权利要求 1 或 2 所述的基因导入宿主细胞。

8. 脂肪酸生产方法,所述脂肪酸选自二均- γ -亚麻酸和二十碳四烯酸,其特征在于,所述方法利用通过权利要求 6 所述的方法制备的植物体。

来自地钱的不饱和脂肪酸合成系统酶基因及其利用

[0001] 本申请是中国专利申请第 200480038471.6 号的分案申请,第 200480038471.6 号的专利申请其申请日是 2004 年 12 月 22 日,发明名称是“来自地钱的不饱和脂肪酸合成系统酶基因及其利用”。

技术领域

[0002] 本发明涉及来自地钱 (*Marchantia polymorpha*) 的不饱和脂肪酸合成系统基因即 $\Delta 5$ 脂肪酸去饱和酶基因、 $\Delta 6$ 脂肪酸去饱和酶基因和 $\Delta 6$ 脂肪酸碳链延长酶基因及其利用。

背景技术

[0003] 花生四烯酸、二十碳五烯酸(以下简称“EPA”)等多不饱和脂肪酸(Polyunsaturated fatty acid;PUFA)在人体中以神经系统为主的细胞膜脂质中含量丰富。这些多不饱和脂肪酸作为前列腺素、白三烯等生理活性物质的前体而发挥作用,在药理学上非常重要。近年来,市场上出售含花生四烯酸、EPA 的健康食品。而且,脂肪酸还是洗涤剂、生物降解塑料等的原料,所以作为材料也备受注目。

[0004] 现在,通过从培养的微生物或鱼油中提取来生产多不饱和脂肪酸。因此存在以下问题:生产成本低,能源消耗量和废弃物量增加,特别是从鱼油提取制造多不饱和脂肪酸的方法,鱼资源有限。

[0005] 可认为花生四烯酸、EPA 分别以亚油酸、 α -亚麻酸为起点,通过 $\Delta 6$ 去饱和、脂肪酸碳链延长及 $\Delta 5$ 去饱和这三个连续的反应生物合成。这些反应分别由 $\Delta 6$ 脂肪酸去饱和酶(以下简称“ $\Delta 6$ 去饱和酶”)、 $\Delta 6$ 脂肪酸碳链延长酶(以下简称“ $\Delta 6$ 碳链延长酶”)及 $\Delta 5$ 脂肪酸去饱和酶(以下简称“ $\Delta 5$ 去饱和酶”)催化。

[0006] $\Delta 6$ 去饱和酶的基因在多种植物中被克隆。例如,在硅藻、小立碗藓(*Physcomitrella patens*)、角齿藓、琉璃苣、紫草、樱草及秋牡丹中克隆得到 $\Delta 6$ 去饱和酶的基因。除植物外,还在丝状菌、线虫、蓝藻、大鼠及人类中克隆得到 $\Delta 6$ 去饱和酶基因(参照非专利文献 1(Eur. J. Biochem. 269,p4105,2002)、非专利文献 2(Plant J. 15,p39,1998)、非专利文献 3(Eur. J. Biochem. ,267. p3801,2000)、非专利文献 4(Proc. Natl. Acad. Sci. USA 94, p4211,1997)、非专利文献 5(Lipids37,417,2002)、非专利文献 6(FEBS Lett. 542, p100,2003)、非专利文献 7(Whitney et al.,Planta Epub 2003)、非专利文献 8(Lipids34, p649,1999)、非专利文献 9(Gene,238, p445 1999)、非专利文献 10(Biochem J. 330, p611 1998)、非专利文献 11(Plant Mol. Biol. ,22, p293 1993)、非专利文献 12(Biochem. Biophys. res. Commun. 255,p575,1999)、非专利文献 13(J. Biol. Chem. 274,p471,1999))。而且,除蓝藻的 $\Delta 6$ 去饱和酶外,上述其他的 $\Delta 6$ 去饱和酶均在 N 末端存在细胞色素 b5 结构域。

[0007] $\Delta 6$ 碳链延长酶的基因最初在丝状菌和线虫中被克隆(参照非专利文献 14(Proc. Natl. Acad. Sci. USA 97, p8284,2000)、非专利文献 15(Proc. Natl. Acad. Sci. USA 97,

p6421,2000)。在植物中,唯一从 *Sollya heterophylla* 中被克隆(参照非专利文献 16(Plant J. 31, p255,2002))。

[0008] 在酵母 (*Saccharomyces cerevisiae*) 中,存在参与鞘脂长链饱和酰基链合成的 EL02 蛋白和 EL03 蛋白(参照非专利文献 17(J. Biol. Chem., 272, p17376,1997)), $\Delta 6$ 碳链延长酶的氨基酸序列与上述蛋白具有同源性。另一方面,在植物中存在另一类型的脂肪酸碳链延长酶,即 β -酮酯酰 CoA 合酶(KCS)。此酶催化长链饱和/单不饱和脂肪酸的链延长(参照非专利文献 15 和非专利文献 18(Plant Cell 7, p309,1995))。但是, $\Delta 6$ 碳链延长酶基因及酵母 EL02/EL03 基因在进化上与 KCS 基因没有直接关系(参照非专利文献 15 和 16)。

[0009] $\Delta 5$ 去饱和酶的基因最初从丝状菌被克隆(参照非专利文献 19(J. Biol. Chem. 273, p29360,1998)、非专利文献 20(J. Biol. Chem. 273, p19055))。 $\Delta 5$ 去饱和酶的结构与 $\Delta 6$ 去饱和酶相似,在 N 末端具有细胞色素 b5 结构域。在硅藻、线虫、大鼠、人类、小立碗藓 (*Physcomitrellapatens*) 等中克隆得到 $\Delta 5$ 去饱和酶基因(参照非专利文献 1、非专利文献 21(FEBS Lett. 439, p215,1998)、非专利文献 22(Arch. Biochem. Biophys. 391, p8,2001)、非专利文献 23(J. Biol. Chem. 274, p37335,1999)、非专利文献 24(J. Biol. Chem. 278, 35115,2003))。

[0010] 陆生植物由苔藓植物(苔藓植物门(Bryophyta))、蕨类植物、裸子植物及被子植物构成。苔藓植物是陆生植物中最古老的分类群,由藓类(藓纲(Bryosida))、苔类(苔纲(Hepaticopsida))及角藓类这三类构成。地钱与上述生物中的小立碗藓在分类上最近,但小立碗藓属于藓类,而地钱属于苔纲地钱亚纲(Marchantiidae)。上述三类约在 4 亿 3 千万年前就已经分化。因此虽然都称为苔藓,但小立碗藓与地钱在进化上大不相同,例如,基于 2 亿年前分化的拟南芥与稻子的差异(参照非专利文献 25(<http://www.nibb.ac.jp/~mhasebe/Physcomitrella.html>))。

[0011] 作为来自地钱的多不饱和脂肪酸合成系统酶基因,已经获得如上述 KCS 那样的链延长酶基因 MpFAE2 及 MpFAE3(参照非专利文献 26(Biosci. Biotechnol. Biochem. 67, p605,2003)、非专利文献 27(Biosci. Biotechnol. Biochem. 67, p1667,2003))。但是, MpFAE2 和 MpFAE3 不是 $\Delta 6$ 碳链延长酶基因。

[0012] 如上所述,从各种各样的生物中克隆得到多种多不饱和脂肪酸合成系统酶基因,但在植物中生产花生四烯酸、EPA 等碳原子数是 20 或大于 20 且不饱和度是 4 或大于 4 的多不饱和脂肪酸的例子少。上述例子的报道有:在亚麻中使来自硅藻的 $\Delta 6$ 去饱和酶基因及 $\Delta 5$ 去饱和酶基因、和来自小立碗藓的 $\Delta 6$ 碳链延长酶基因表达,生产花生四烯酸及 EPA,但其具体内容不详(参照非专利文献 24)。

[0013] 综上所述,因为花生四烯酸、EPA 等多不饱和脂肪酸是从培养的微生物或鱼油中提取来生产,所以存在以下问题:生产成本低、能源消耗量和废弃物量增加、鱼资源有限等。花生四烯酸、EPA 等多不饱和脂肪酸的分子内有多个双键,这一特性使其可在各种工业中得到应用(例如,薄膜、生物降解塑料、功能性纤维、润滑油、洗涤剂的原料等)。通过基因重组植物生产上述多不饱和脂肪酸,可以降低生产成本,同时还可以实现更加环保的生产过程。如果可以利用基因重组技术在油料作物中大量生产上述多不饱和脂肪酸,则作为廉价多用途原料将非常有用。

[0014] 另一方面,当在植物中使异种生物的基因表达时,因存在转录、翻译、翻译后修饰等过程,所以很难预测该基因在植物体内发挥良好功能的程度。特别是使多个异种生物的基因表达时,可预测:与上述非专利文献 24 那样使来自不同种生物的多个基因表达的情况相比,使来自同一种生物的多个基因在植物内表达时能良好地发挥功能。此外,最早的陆生植物即苔藓类的地钱作为高等植物的模式系统,受到人们关注,并期待其基因在植物内良好地发挥功能。因此,如果能够获得来自地钱的多不饱和脂肪酸合成酶基因即 $\Delta 5$ 去饱和酶基因、 $\Delta 6$ 去饱和酶基因及 $\Delta 6$ 碳链延长酶基因,可期待通过将把这些基因导入植物在植物体内高效地蓄积花生四烯酸、EPA。

[0015] 此外,虽然从同样是苔藓植物的小立碗藓中克隆得到 $\Delta 5$ 去饱和酶基因、 $\Delta 6$ 去饱和酶基因及 $\Delta 6$ 碳链延长酶基因,但地钱与小立碗藓在进化上相差很大,以现在的技术水平难以根据小立碗藓的基因获得地钱的基因。

发明内容

[0016] 鉴于上述以往的问题完成了本发明,本发明的目的是:提供能够在高等植物中生产花生四烯酸、EPA 的来自地钱 (*Marchantia polymorpha*) 的不饱和脂肪酸合成酶基因及其利用方法,所述来自地钱的不饱和脂肪酸合成酶基因即 $\Delta 5$ 去饱和酶基因、 $\Delta 6$ 去饱和酶基因、 $\Delta 6$ 碳链延长酶基因。

[0017] 本发明人等为了解决上述课题进行了潜心研究,从来自地钱 (*Marchantia polymorpha*) 的 cDNA 克隆中鉴定出编码上述 $\Delta 6$ 去饱和酶、 $\Delta 5$ 去饱和酶和 $\Delta 6$ 碳链延长酶的基因,还成功地将这些基因导入甲醇酵母 (*Pichia pastoris*) 并使其表达,发现表达这些基因的蛋白质分别具有 $\Delta 6$ 去饱和、 $\Delta 5$ 去饱和、 $\Delta 6$ 碳链延长的酶活性,从而完成了本发明。即,本发明包含以下的发明。

[0018] (1) 来自地钱目生物的基因,其与由序列号 1 所示碱基序列构成的 DNA 的全部或部分、或与由该 DNA 的互补碱基序列构成的 DNA 的全部或部分在严格条件下杂交,且编码具有 $\Delta 6$ 脂肪酸去饱和活性的蛋白质。

[0019] (2) 对具有 $\Delta 6$ 脂肪酸去饱和活性且来自地钱目生物的蛋白质进行编码的下述 (a) 或 (b) 中记载的基因。(a) 具有序列号 1 所示碱基序列的基因。(b) 与由序列号 1 所示碱基序列构成的 DNA、或与由该 DNA 的互补碱基序列构成的 DNA 在严格条件下杂交的基因。

[0020] (3) 对具有 $\Delta 6$ 脂肪酸去饱和活性且来自地钱目生物的蛋白质进行编码的下述 (a) 或 (b) 中记载的基因。(a) 具有序列号 1 所示碱基序列中第 253 至 1698 号的碱基序列的基因。(b) 与由序列号 1 所示碱基序列中第 253 至 1698 号的碱基序列构成的 DNA、或与由该 DNA 的互补碱基序列构成的 DNA 在严格条件下杂交的基因。

[0021] (4) 对具有 $\Delta 6$ 脂肪酸去饱和活性且来自地钱目生物的蛋白质进行编码的下述 (a) 或 (b) 中记载的基因。(a) 对由序列号 2 所示氨基酸序列构成的蛋白质进行编码的基因。(b) 对由序列号 2 所示氨基酸序列发生 1 个或大于 1 个氨基酸置换、缺失、插入和 / 或附加后的氨基酸序列构成的蛋白质进行编码的基因。

[0022] (5) 来自地钱目生物的基因,其与由序列号 3 所示碱基序列构成的 DNA 的全部或部分、或与由该 DNA 的互补碱基序列构成的 DNA 的全部或部分在严格条件下杂交,且编码具有 $\Delta 6$ 脂肪酸碳链延长活性的蛋白质。

[0023] (6) 对具有 $\Delta 6$ 脂肪酸碳链延长活性且来自地钱目生物的蛋白质进行编码的下述 (a) 或 (b) 中记载的基因。(a) 具有序列号 3 所示碱基序列的基因。(b) 与由序列号 3 所示碱基序列构成的 DNA、或与由该 DNA 的互补碱基序列构成的 DNA 在严格条件下杂交的基因。

[0024] (7) 对具有 $\Delta 6$ 脂肪酸碳链延长活性且来自地钱目生物的蛋白质进行编码的下述 (a) 或 (b) 中记载的基因。(a) 具有序列号 3 所示碱基序列中第 194 至 1066 号的碱基序列的基因。(b) 与由序列号 3 所示碱基序列中第 194 至 1066 号的碱基序列构成的 DNA、或与由该 DNA 的互补碱基序列构成的 DNA 在严格条件下杂交的基因。

[0025] (8) 对具有 $\Delta 6$ 脂肪酸碳链延长活性且来自地钱目生物的蛋白质进行编码的下述 (a 或 (b) 中记载的基因。其为对 (a) 由序列号 4 所示氨基酸序列构成的蛋白质、对 (b) 由序列号 4 所示氨基酸序列发生 1 个或大于 1 个氨基酸置换)、缺失、插入和 / 或附加后的氨基酸序列构成的蛋白质进行编码的基因。

[0026] (9) 来自地钱目生物的基因, 其与由序列号 5 所示碱基序列构成的 DNA 的全部或部分、或与由该 DNA 的互补碱基序列构成的 DNA 的全部或部分

[0027] 分在严格条件下杂交, 且编码具有 $\Delta 5$ 脂肪酸去饱和活性的蛋白质。

[0028] (10) 具有 $\Delta 5$ 脂肪酸去饱和活性且来自地钱目生物的蛋白质进行编码的下述 (a) 或 (b) 中记载的基因。(a) 具有序列号 5 所示碱基序列的基因。(b) 与由序列号 5 所示碱基序列构成的 DNA、或与由该 DNA 的互补碱基序列构成的 DNA 在严格条件下杂交的基因。

[0029] (11) 对具有 $\Delta 5$ 脂肪酸去饱和活性且来自地钱目生物的蛋白质进行编码的下述 (a) 或 (b) 中记载的基因。(a) 具有序列号 5 所示碱基序列中第 375 至 1829 号的碱基序列的基因。(b) 与由序列号 5 所示碱基序列中第 375 至 1829 号的碱基序列构成的 DNA、或与由该 DNA 的互补碱基序列构成的 DNA 在严格条件下杂交的基因。

[0030] (12) 对具有 $\Delta 5$ 脂肪酸去饱和活性且来自地钱目生物的蛋白质进行编码的下述 (a) 或 (b) 中记载的基因。其为对 (a) 由序列号 6 所示氨基酸序列构成的蛋白质、对 (b) 由序列号 6 所示氨基酸序列发生 1 个或大于 1 个氨基酸置换、缺失、插入和 / 或附加后的氨基酸序列构成的蛋白质进行编码的基因。

[0031] (13) 由上述 (1) ~ (12) 的任意一项中所述的基因所编码的蛋白质。

[0032] (14) 以下 (a) 或 (b) 所述的蛋白质。(a) 由序列号 2 所示氨基酸序列构成的蛋白质。(b) 由序列号 2 所示氨基酸序列发生 1 个或大于 1 个氨基酸置换、缺失、插入和 / 或附加后的氨基酸序列构成、且具有 $\Delta 6$ 脂肪酸去饱和活性的蛋白质。

[0033] (15) 以下 (a) 或 (b) 所述的蛋白质。(a) 由序列号 4 所示氨基酸序列构成的蛋白质。(b) 由序列号 4 所示氨基酸序列发生 1 个或大于 1 个氨基酸置换、缺失、插入和 / 或附加后的氨基酸序列构成、且具有 $\Delta 6$ 脂肪酸碳链延长活性的蛋白质。

[0034] (16) 以下 (a) 或 (b) 所述的蛋白质。(a) 由序列号 6 所示氨基酸序列构成的蛋白质。(b) 由序列号 6 所示氨基酸序列发生 1 个或大于 1 个氨基酸置换、缺失、插入和 / 或附加后的氨基酸序列构成、且具有 $\Delta 5$ 脂肪酸去饱和活性的蛋白质。

[0035] (17) 识别上述 (13) ~ (16) 任意一项中所述的蛋白质的抗体。

[0036] (18) 至少含有上述 (1) ~ (12) 任意一项中所述的基因的重组表达载体。

[0037] (19) 至少导入上述 (1) ~ (12) 任意一项中所述的基因而形成的转化体。

[0038] (20) 至少可表达地导入了上述 (1) ~ (12) 任意一项所述基因的植物体、或与该植

物体具有同一性状的该植物的后代、或该植物的组织。

[0039] (21) 至少可表达地导入了上述 (1) ~ (12) 任意一项所述的基因且脂肪酸组成被改变的植物体、或与该植物体具有同一性状的该植物的后代、或该植物的组织。

[0040] (22) 上述 (20) 或 (21) 中所述的植物的繁殖材料。

[0041] (23) 用上述 (21) 中所述的植物体或植物的组织生产脂肪酸的方法。

[0042] (24) 含有通过上述 (23) 所述的脂肪酸生产方法得到的 γ -亚麻酸、二均- γ -亚麻酸、花生四烯酸、十八碳四烯酸、二十碳四烯酸、二十碳五烯酸中至少一种脂肪酸的材料。

[0043] (25) 至少用 (1) ~ (12) 任意一项所述的基因改变脂肪酸组成的方法。

[0044] (26) 用上述 (1) ~ (12) 任意一项所述的基因的至少一部分碱基序列或其互补序列作为探针的基因检测工具。

[0045] (27) 用上述 (13) ~ (16) 任意一项中所述的蛋白质对调节该蛋白质的基因或调节该蛋白质的物质进行筛选的方法。

[0046] (28) 通过上述 (27) 中所述的筛选方法得到的基因或物质。

[0047] 另外,只要在本说明书中没有特别说明, A、C、G、T 分别表示腺嘌呤、胞嘧啶、鸟嘌呤、胸腺嘧啶这四个碱基。

[0048] 本发明的其他目的、特征及优点可通过以下的说明充分了解。而且,通过参照附图的以下说明,还可以明白本发明的有益效果。

附图说明

[0049] [图 1]

[0050] 表示将实施例 6 中使用的 MpDES6 基因、MpEL01 基因及 MpDES5 基因的各表达盒连接的结构构建顺序说明图。

具体实施方式

[0051] 以下说明本发明的实施形态之一。但是,本发明不限于此。

[0052] 下面,按照花生四烯酸及二十碳五烯酸 (EPA) 的合成途径、本发明涉及的基因、本发明涉及的蛋白质、本发明涉及的蛋白质及基因的获得方法、以及本发明涉及的基因及蛋白质的利用方法(有用性)这样的顺序,详细地说明本发明。

[0053] (1) 花生四烯酸及二十碳五烯酸 (EPA) 的合成途径

[0054] 可认为花生四烯酸、二十碳五烯酸 (EPA) 分别以亚油酸、 α -亚麻酸为起点,通过 $\Delta 6$ 去饱和、 $\Delta 6$ 碳链延长及 $\Delta 5$ 去饱和这三个连续的反应生物合成。这些反应分别由 $\Delta 6$ 去饱和酶、 $\Delta 6$ 碳链延长酶、 $\Delta 5$ 去饱和酶催化,分别称为 n-6 途径(花生四烯酸合成途径)和 n-3 途径(EPA 合成途径)。

[0055] 至今为止所报道的 $\Delta 6$ 去饱和酶、 $\Delta 6$ 碳链延长酶及 $\Delta 5$ 去饱和酶均参与 n-6 及 n-3 途径。即, $\Delta 6$ 去饱和酶在 n-6 途径中将亚油酸 ($18:2 \Delta^{9,12}$, 其中,18 表示碳原子数,2 表示双键数,9、12 表示双键的位置。以下同。) 转变成 γ -亚麻酸 (GLA ; $18:3 \Delta^{6,9,12}$), 在 n-3 途径中将 α -亚麻酸 (ALA ; $18:3 \Delta^{9,12,15}$) 转变成十八碳四烯酸 (STA ; $18:4 \Delta^{6,9,12,15}$)。 $\Delta 6$ 碳链延长酶在 n-6 途径中将 GLA 转变成二均- γ -亚麻酸 (DGLA ; $20:3 \Delta^{8,11,14}$), 在 n-3 途径中将 STA 转变成二十碳四烯酸 (ETA ; $20:4 \Delta^{8,11,14,17}$)。 $\Delta 5$ 去饱和酶在 n-6 途径中将 DGLA 转

变成花生四烯酸 ($20:4 \Delta^{5,8,11,14}$), 在 n-3 途径中将 ETA 转变成二十碳五烯酸 ($EPA; 20:5 \Delta^{5,8,11,14,17}$)。

[0056] (2) 本发明涉及的基因

[0057] [本发明涉及的 $\Delta 6$ 去饱和酶基因]

[0058] 本发明涉及的 $\Delta 6$ 去饱和酶基因是对具有 $\Delta 6$ 脂肪酸去饱和活性的蛋白质进行编码的来自地钱目生物的基因, 只要是符合以下条件的基因即可。

[0059] 1. 具有序列号 1 所示碱基序列的基因。

[0060] 2. 与由序列号 1 所示碱基序列构成的 DNA、或与由该 DNA 的互补碱基序列构成的 DNA 在严格条件下杂交的基因。

[0061] 3. 与由序列号 1 所示碱基序列构成的 DNA 的一部分、或与由该 DNA 的互补碱基序列构成的 DNA 的一部分在严格条件下杂交的基因。

[0062] 4. 具有序列号 1 所示碱基序列中第 253 至 1698 号的碱基序列的基因。另外, 序列号 1 所示碱基序列的第 253 至 1698 号的碱基序列是被翻译成由序列号 2 所示氨基酸序列构成的蛋白质的区域。

[0063] 5. 与由序列号 1 所示碱基序列中第 253 至 1698 号的碱基序列构成的 DNA、或与由该 DNA 的互补碱基序列构成的 DNA 在严格条件下杂交的基因。

[0064] 6. 对由序列号 2 所示氨基酸序列构成的蛋白质进行编码的基因。

[0065] 7. 对由序列号 2 所示氨基酸序列发生 1 个或大于 1 个氨基酸置换、缺失、插入、和 / 或附加后的氨基酸序列构成的蛋白质进行编码的基因。

[0066] [本发明涉及的 $\Delta 6$ 碳链延长酶基因]

[0067] 本发明涉及的 $\Delta 6$ 碳链延长酶基因是对具有 $\Delta 6$ 脂肪酸碳链延长酶活性的蛋白质进行编码的来自地钱目生物的基因, 只要是符合以下条件的基因即可。

[0068] 1. 具有序列号 3 所示碱基序列的基因。

[0069] 2. 与由序列号 3 所示碱基序列构成的 DNA、或与由该 DNA 的互补碱基序列构成的 DNA 在严格条件下杂交的基因。

[0070] 3. 与由序列号 3 所示碱基序列构成的 DNA 的一部分、或与由该 DNA 的互补碱基序列构成的 DNA 的一部分在严格条件下杂交的基因。

[0071] 4. 具有序列号 3 所示碱基序列中第 194 至 1066 号的碱基序列的基因。另外, 序列号 3 所示碱基序列的第 194 至 1066 号的碱基序列是被翻译成由序列号 4 所示氨基酸序列构成的蛋白质的区域。

[0072] 5. 与由序列号 3 所示碱基序列中第 194 至 1066 号的碱基序列构成的 DNA、或与由该 DNA 的互补碱基序列构成的 DNA 在严格条件下杂交的基因。

[0073] 6. 对由序列号 4 所示氨基酸序列构成的蛋白质进行编码的基因。

[0074] 7. 对由序列号 4 所示氨基酸序列发生 1 个或大于 1 个氨基酸置换、缺失、插入、和 / 或附加后的氨基酸序列构成的蛋白质进行编码的基因。

[0075] [本发明涉及的 $\Delta 5$ 去饱和酶基因]

[0076] 本发明涉及的 $\Delta 5$ 去饱和酶基因是对具有 $\Delta 5$ 脂肪酸去饱和活性的蛋白质进行编码的来自地钱目生物的基因, 只要是符合以下条件的基因即可。

[0077] 1. 具有序列号 5 所示碱基序列的基因。

[0078] 2. 与由序列号 5 所示碱基序列构成的 DNA、或与由该 DNA 的互补碱基序列构成的 DNA 在严格条件下杂交的基因。

[0079] 3. 与由序列号 5 所示碱基序列构成的 DNA 的一部分、或与由该 DNA 的互补碱基序列构成的 DNA 的一部分在严格条件下杂交的基因。

[0080] 4. 具有序列号 5 所示碱基序列中第 375 至 1829 号的碱基序列的基因。另外, 序列号 5 所示碱基序列的第 375 至 1829 号的碱基序列是被翻译成由序列号 6 所示氨基酸序列构成的蛋白质的区域。

[0081] 5. 与由序列号 5 所示碱基序列中第 375 至 1829 号的碱基序列构成的 DNA、或与由该 DNA 的互补碱基序列构成的 DNA 在严格条件下杂交的基因。

[0082] 6. 对由序列号 6 所示氨基酸序列构成的蛋白质进行编码的基因。

[0083] 7. 对由序列号 6 所示氨基酸序列发生 1 个或大于 1 个氨基酸置换、缺失、插入、和/或附加后的氨基酸序列构成的蛋白质进行编码的基因。

[0084] 另外, 上述“严格条件”意味着只有当序列间存在至少 90% 的同一性、优选至少 95% 的同一性、最优选至少 97% 的同一性时才发生杂交。

[0085] 上述杂交可以用 J. Sambrook et al. *Molecular Cloning, A Laboratory Manual*, 2d Ed., Cold Spring Harbor Laboratory (1989) 中所述的方法等以往公知的方法进行。通常, 温度越高, 盐浓度越低, 严格度越高 (越难杂交), 可以获得同源性更高的基因。作为杂交的条件, 可以适宜地采用以往公知的条件, 没有特殊限制, 例如, 42°C、6×SSPC、50% 甲酰胺、1% SDS、100 μg/ml salmon sperm DNA、5×Denhardt 溶液 (1×SSPE ; 0.18M 氯化钠、10mM 磷酸钠、pH 7.7、1mM EDTA)。

[0086] 上述“地钱目生物”不限于地钱 (*Marchantia polymorpha*), 包含属于地钱亚纲地钱目 (*Marchantiales*) 的生物。其中, 已知在 *Monocleaforsteri* (*Monocleales*)、*Corsinia coriandrina* (*Marchantiales*)、*Oximitra paleacea* (*Marchantiales*)、*Ricciocarpos natans* (*Marchantiales*)、*Ricca huebeneriana* (*Marchantiales*)、*Riccafluitans* (*Marchantiales*)、*Ricca duplex* (*Marchantiales*)、*Riccacaniculata* (*Marchantiales*)、*Ricca bifurca* (*Marchantiales*)、*Riccaciliifera* (*Marchantiales*)、*Ricca glauca* (*Marchantiales*)、*Riccasorocarpa* (*Marchantiales*)、*Ricca warnstorffii* (*Marchantiales*)、*Ricca michelii* (*Marchantiales*)、*Ricca papillosa* (*Marchantiales*) 以及 *Ricca zachariae* (*Marchantiales*) 中存在超长链多不饱和脂肪酸 (参照 *Prog. Lipid Res.* 32, p281, 1993)。以现有的技术水平可容易地从这些生物中获得 Δ6 去饱和酶、Δ6 碳链延长酶以及 Δ5 去饱和酶的基因。例如, 众所周知近缘生物中具有同一功能的酶的编码基因产生交叉杂交。

[0087] 本发明的基因不仅包含双链 DNA, 还包含构成双链 DNA 的被称为有意义链及反义链的各单链 DNA、RNA。反义链可用作探针或反义化合物。DNA 包括例如通过克隆、化学合成技术或这些技术的组合而得到的 cDNA、基因组 DNA 等。而且, 本发明的基因还可以是含有非翻译区 (UTR) 的序列、载体序列 (含有表达载体序列) 等序列的基因。

[0088] (3) 本发明涉及的蛋白质

[0089] [本发明涉及的 Δ6 去饱和酶蛋白质]

[0090] 本发明涉及的 Δ6 去饱和酶蛋白质, 只要是来自地钱目生物的生物蛋白质并且是具有

$\Delta 6$ 脂肪酸去饱和活性的蛋白质即可。更具体地说,只要是下述的蛋白质即可。

[0091] 1. 由上述 (2) 所述的本发明涉及的 $\Delta 6$ 去饱和酶基因所编码的蛋白质。

[0092] 2. 由序列号 2 所示氨基酸序列构成的蛋白质。

[0093] 3. 由序列号 2 所示氨基酸序列发生 1 个或大于 1 个氨基酸置换、缺失、插入、和 / 或附加后的氨基酸序列构成的蛋白质。

[0094] [本发明涉及的 $\Delta 6$ 碳链延长酶蛋白质]

[0095] 本发明涉及的 $\Delta 6$ 碳链延长酶蛋白质只要是来自地钱目生物的蛋白质并且是具有 $\Delta 6$ 脂肪酸碳链延长活性的蛋白质即可。更具体地说,只要是下述的蛋白质即可。

[0096] 1. 由上述 (2) 所述的本发明涉及的 $\Delta 6$ 碳链延长酶基因所编码的蛋白质。

[0097] 2. 由序列号 4 所示氨基酸序列构成的蛋白质。

[0098] 3. 由序列号 4 所示氨基酸序列发生 1 个或大于 1 个氨基酸置换、缺失、插入、和 / 或附加后的氨基酸序列构成的蛋白质。

[0099] [本发明涉及的 $\Delta 5$ 去饱和酶蛋白质]

[0100] 本发明涉及的 $\Delta 5$ 去饱和酶蛋白质只要是来自地钱目生物的蛋白质并且是具有 $\Delta 5$ 脂肪酸去饱和活性的蛋白质即可。更具体地说,只要是下述的蛋白质即可。

[0101] 1. 由上述 (2) 所述的本发明涉及的 $\Delta 5$ 去饱和酶基因所编码的蛋白质。

[0102] 2. 由序列号 6 所示氨基酸序列构成的蛋白质。

[0103] 3. 由序列号 6 所示氨基酸序列发生 1 个或大于 1 个氨基酸置换、缺失、插入、和 / 或附加后的氨基酸序列构成的蛋白质。

[0104] 上述“ $\Delta 6$ 脂肪酸去饱和和活性”意味着对亚油酸或 α -亚麻酸有底物特异性,分别将其转变成 γ -亚麻酸或十八碳四烯酸。上述“ $\Delta 6$ 脂肪酸碳链延长活性”意味着对 γ -亚麻酸或十八碳四烯酸有底物特异性,分别将其转变成二均- γ -亚麻酸或二十碳四烯酸。上述“ $\Delta 5$ 脂肪酸去饱和和活性”意味着对二均- γ -亚麻酸或二十碳四烯酸有底物特异性,分别将其转变成花生四烯酸或二十碳五烯酸 (EPA)。

[0105] 上述“1 个或大于 1 个氨基酸被置换、缺失、插入、和 / 或附加”意味着通过定点诱变法等公知的突变蛋白质制备方法所能够置换、缺失、插入、和 / 或附加的数目 (优选 10 个或小于 10 个、更优选 7 个或小于 7 个、进一步优选 5 个或小于 5 个) 的氨基酸被置换、缺失、插入、和 / 或附加。这样的突变蛋白质并不限于通过公知的突变蛋白质制备方法人工导入了突变的蛋白质,也可以是天然存在的同样的突变蛋白质经分离纯化后得到的蛋白质。

[0106] 另外,本发明的蛋白质只要是氨基酸通过肽键结合形成的多肽即可,但并不限于此,也可以是包含多肽以外的其他结构的复合蛋白质。这里所述的多肽以外的其他结构,可以是糖链、异戊间二烯基等,没有特殊限制。

[0107] 而且,本发明的蛋白质也可以含有附加多肽。被多肽附加的情况例如,本发明的蛋白质被 His、Myc、Flag 等表位标记等。

[0108] 此外,本发明的蛋白质可以是上述本发明的基因 (对本发明的蛋白质进行编码的基因) 导入宿主细胞使其蛋白质在细胞内表达的状态,也可以是从细胞、组织等分离纯化的状态。而且,根据上述宿主细胞中的表达条件,本发明的蛋白质还可以是与其他蛋白质相连的融合蛋白。此外,本发明的蛋白质也可以是化学合成的蛋白质。

[0109] (4) 本发明涉及的蛋白质及基因的获得方法

[0110] 对本发明涉及的蛋白质及基因的获得方法（生产方法）没有特殊限制，作为代表性方法，可以列举以下各种方法。

[0111] [蛋白质的获得方法]

[0112] 如上所述，对本发明的蛋白质的获得方法（生产方法）没有特殊限制，首先，可以列举这样的方法：从表达本发明蛋白质的细胞、组织等中简单纯化。对纯化方法没有特殊限制，用公知的方法从细胞、组织制备细胞提取液，将此细胞提取液用公知的方法如色谱柱等进行纯化即可。

[0113] 此外，本发明的蛋白质的获得方法还可以列举利用基因重组技术等的方法。例如，可以采用这样等方法：将本发明的基因整合进载体等，然后通过公知的方法可表达地导入宿主细胞，纯化细胞内翻译得到的上述蛋白质。

[0114] 另外，如上所述那样将外源基因导入宿主时，存在各种整合了使外源基因在宿主内表达的启动子的表达载体以及各种宿主，根据目的选择即可。生成的蛋白质的纯化方法根据使用的宿主、蛋白质的性质不同而不同，通过标签的利用等，可以较容易地纯化目的蛋白质。

[0115] 对突变蛋白质的制备方法也没有特殊限制。可以采用公知的突变蛋白质制备方法，例如，利用定点诱变法（Hashimoto-Gotoh, Gene152, 271-275(1995) 等）、PCR 法在碱基序列中导入点突变来制备突变蛋白质的方法、或者通过插入转位子制备突变株的方法等。制备突变蛋白质时，也可以利用市售的试剂盒。

[0116] 本发明蛋白质的获得方法并不限于上述方法，例如，还可以化学合成。例如，可以利用体外蛋白质合成液由本发明的基因合成本发明的蛋白质。

[0117] [基因的获得方法]

[0118] 对本发明的基因的获得方法（生产方法）也没有特殊限制，例如，利用差异筛选（subtractive cloning）的方法。在此方法中，按照公知的技术，在试管内反复进行直接杂交，浓缩目的 cDNA（本发明的基因）即可。

[0119] 在通常采用的条件下，实施上述差异筛选中的各个步骤即可。由此得到的克隆可以通过制作限制酶图谱及测定其碱基序列（sequencing）进行更详细地解析。根据上述解析，可以容易地判断是否获得了含有本发明基因序列的 DNA 片段。

[0120] 作为本发明的基因的获得方法，还可以是：通过公知的技术，将含有本发明的基因的 DNA 片段分离并克隆。例如，制备与本发明基因的碱基序列的一部分特异性杂交的探针，筛选基因组 DNA 文库、cDNA 文库即可。上述探针只要是与本发明基因的碱基序列或其互补序列的至少一部分特异性杂交的探针即可，可以采用任何序列和长度的探针。

[0121] 此外，从上述在地钱之间被良好保留的区域中选择上述探针的序列，对其他地钱的基因组 DNA（或 cDNA）文库进行筛选，可以分离并克隆对具有与上述蛋白质同样功能的同源分子、类似分子进行编码的基因。

[0122] 或者，本发明的基因的获得方法还可以是采用 PCR 等扩增手段的方法。例如，从本发明基因的 cDNA 序列的 5' 侧及 3' 侧的序列（或其互补序列）中分别制备引物，利用这些引物，以基因组 DNA（或 cDNA）等为模板进行 PCR 等，扩增两个引物之间的 DNA 区域，从而可以获得大量含有本发明的基因的 DNA 片段。

[0123] (5) 本发明涉及的基因及蛋白质的利用方法（有用性）

[0124] (5-1) 重组表达载体

[0125] 本发明涉及的重组表达载体只要含有前述 (2) 记载的本发明涉及的基因即可, 没有特殊限制。例如, 插入了 cDNA 的重组表达载体。制备重组表达载体时, 可以用质粒、噬菌体、或黏粒等, 没有特殊限制。而且, 制备方法采用公知的方法即可。

[0126] 对载体的具体种类没有特殊限制, 适宜地选择能在宿主细胞中表达的载体即可。即, 根据宿主细胞的种类, 选择确实能使基因表达的合适启动子序列, 将该启动子序列和本发明的基因整合进各种质粒等, 作为表达载体使用即可。

[0127] 还可以利用各种标记来确认本发明的基因是否已被导入宿主细胞, 而且是否已在宿主细胞中真实地表达。例如, 以宿主细胞中缺失的基因为标记, 以含有此标记和本发明的基因的质粒等为表达载体导入宿主细胞。根据标记基因的表达, 可以确认本发明的基因的导入。或者, 也可以使本发明的蛋白质作为融合蛋白表达, 例如, 以来自水母 *Aequoreacoerulescens* (Brandt) 的绿色荧光蛋白 GFP (Green Fluorescent Protein) 为标记, 使本发明的蛋白质作为 GFP 融合蛋白表达。

[0128] 上述宿主细胞没有特殊限制, 可以适宜地采用以往公知的各种细胞。具体可列举: 大肠杆菌 (*Escherichia coli*) 等细菌、酵母 (出芽酵母 *Saccharomyces cerevisiae*、分裂酵母 *Schizosaccharomyces pombe*)、线虫 (*Caenorhabditis elegans*)、非洲爪蟾 (*Xenopus laevis*) 的卵母细胞等, 没有特殊限制。

[0129] 对将上述表达载体导入宿主细胞的方法即转化方法没有特殊限制, 可以适宜地采用电穿孔法、磷酸钙法、核糖体法、DEAE 葡聚糖法等以往公知的方法。而且, 例如, 使本发明的蛋白质在昆虫中转移表达时, 可以采用杆状病毒表达系统。

[0130] (5-2) 转化体

[0131] 本发明涉及的转化体只要是导入了前述 (2) 记载的本发明涉及的基因的转化体即可, 没有特殊限制。这里所述的“转化体”不仅包含细菌、组织、器官, 还包含生物个体。

[0132] 对转化体的制备方法 (生产方法) 没有特殊限制, 例如, 可以是这样的方法: 将上述重组表达载体导入宿主细胞并转化。此外, 对成为转化对象的生物也没有特殊限制, 可以是上述宿主细胞中例举的各种微生物、动物等。

[0133] 本发明涉及的转化体优选: 可表达地导入了本发明涉及的基因的植物体、或与该植物体具有同一性状的该植物体的后代、或该植物体的组织。通过这样的转基因植物, 可以在成本低且环保的生产过程中生产花生四烯酸、EPA 等多不饱和脂肪酸。

[0134] 这里所述的“基因被可表达地导入”是指: 通过公知的基因工程手法 (基因操作技术), 将基因可表达地导入对象细胞 (宿主细胞)。

[0135] 植物体转化所使用的重组表达载体, 只要是能够使插入基因在该植物细胞内表达的重组表达载体即可, 没有特殊限制。例如, 可以利用以下载体: 具有在植物细胞内使基因恒定表达的启动子 (例如, 花椰菜花叶病毒 35S 启动子) 的载体、具有因外界刺激而诱导活化的启动子的载体。另外, 上述植物细胞包含各种形态的植物细胞, 例如: 悬浮培养的细胞、原生质体、叶的切片、愈伤组织等。

[0136] 向植物细胞导入重组表达载体时, 可以采用聚乙二醇法、电穿孔法 (Electroporation 法)、农杆菌介导法、微弹枪法 (particle gun 法) 等本领域技术人员公知的各种方法。可以根据植物细胞的种类不同, 采用本领域技术人员所公知的方法由转化

细胞再生植物体。

[0137] 例如,已建立多种在烟草中培育转基因植物体的方法,如:使转化后的农杆菌感染烟草叶盘的方法、利用聚乙二醇向原生质体导入基因并再生植物体的方法、通过电脉冲向原生质体导入基因并再生植物体的方法、通过微弹枪法向细胞直接导入基因并再生植物体的方法等。在本发明中,可以适宜地采用这些方法。

[0138] 而且,与拟南芥一样,烟草也是利用基因工程技术进行植物育种中的模式植物。如果说在上述烟草中可以获得花生四烯酸、EPA 含量增加了的转化体,则可以毫不夸张地说在所有植物中都可以获得转化体。另外,在本说明书中,如后述的实施例所示,不仅获得了转基因烟草,还获得了转基因稻子,由此证实:通过本发明可以获得所有种类的转基因植物体。

[0139] 例如,已建立多种在稻子中培育转基因植物体的方法,如:利用聚乙二醇向原生质体导入基因并再生植物体的方法、通过电脉冲向原生质体导入基因并再生植物体的方法、通过微弹枪法向细胞直接导入基因并再生植物体的方法等。在本发明中,可以适宜地采用这些方法。

[0140] 当上述转基因植物体是稻子时,因稻子中的花生四烯酸、EPA 含量增加,所以食用从该转基因植物体得到的种子即米,可以容易地向体内摄入花生四烯酸、EPA 等多不饱和脂肪酸。因此,转基因稻子作为粮食具有很高的价值,在食品产业、农业领域极其有用。此外,用目前较少得到利用的米糠、稻谷壳、藁等生产花生四烯酸、EPA,从中提取这些脂肪酸,从而可作为健康食品的原料得到有效利用。而且,还可以用作家畜的饲料。

[0141] 一旦获得基因组内导入了本发明的基因的转基因植物体,则可以通过有性生殖或无性生殖从该植物体获得后代。还可以从该植物体或其后代、或克隆中得到繁殖材料(例如,种子、果实、插条、块茎、块根、植株、愈伤组织、原生质体等),以繁殖材料为基础,大量生产该植物体。因此本发明包含:可表达地导入了本发明的基因的植物体、或、与该植物体具有同一性状的该植物体的后代、或、该植物体的组织、或、该植物体的繁殖材料。

[0142] 该植物体、与该植物体具有同一性状的该植物体的后代、及该植物体的组织也包含营养繁殖的植物体。营养繁殖也被称为营养生殖、克隆,一般通过插芽、插枝等进行繁殖,在试管内,可以由叶、茎、根等器官进行植物体的再分化、通过愈伤组织进行繁殖。根据植物种类不同,有时枝端形成特殊的冬芽、腋芽肉质化、花变成珠芽、形成块根、块茎等。

[0143] 本发明还包含:可表达地导入了本发明的基因且脂肪酸组成被改变了的植物体、或与该植物体具有同一性状的该植物体的后代、或该植物体的组织、该植物体的繁殖材料。“脂肪酸组成被改变了”是指:转化前植物体的脂肪酸组成与转化后植物体的脂肪酸组成不同。例如,原脂肪酸组成中不含花生四烯酸、EPA 的植物,通过用本发明涉及的基因进行转化,使转基因植物的脂肪酸组成中含有花生四烯酸、EPA。

[0144] (5-3) 脂肪酸生产方法

[0145] 本发明包含:利用经本发明涉及的基因转化且脂肪酸组成改变了的植物体或植物体的组织来生产脂肪酸的方法。

[0146] 例如,从上述花生四烯酸、EPA 含量增加了的本发明涉及的转基因植物制得的食用油,其花生四烯酸、EPA 的含量高,是高价值的食用油。上述转基因植物体的种子、果实、切穗、块茎、块根等作为含有花生四烯酸、EPA 的食物,也很有价值。

[0147] (5-4) 材料

[0148] 本发明还包含：含有由上述脂肪酸生产方法得到的物质即 γ -亚麻酸、二均- γ -亚麻酸、花生四烯酸、十八碳四烯酸、二十碳四烯酸、二十碳五烯酸中至少一种脂肪酸的材料。该“材料”指：除了作为上述食物的种子、果实、切穗、块茎、或块根外，还指能作为工业原料利用的所有材料。

[0149] 上述材料例如可以是含花生四烯酸、EPA 的健康食品，薄膜、生物降解塑料、功能性纤维、润滑油、洗涤剂的原料等。上述不饱和脂肪酸具有分子内存在多个双键这一特性。因此，通过本发明的转基因植物体生产花生四烯酸、EPA，可以降低生产成本。而且通过本发明，还可以实现环保的生产过程。

[0150] (5-5) 改变脂肪酸组成的方法

[0151] 本发明包含：用本发明涉及的基因改变脂肪酸组成的方法。例如，通过制备如上述那样导入了本发明涉及的基因的转化体，可以改变宿主细胞的脂肪酸组成。作为改变脂肪酸组成的对象，不受特殊限制，除植物以外还可以是动物、细菌、酵母等所有生物。

[0152] (5-6) 基因检测工具

[0153] 本发明涉及的基因检测工具用本发明涉及的基因的至少一部分碱基序列或其互补序列作为探针。基因检测工具可以在各种条件下用于本发明基因的表达模式的检出和测定等。

[0154] 本发明的基因检测工具例如可以是：将与本发明的基因特异性杂交的上述探针固定在支持物（载体）上的 DNA 芯片。这里所述的“DNA 芯片”主要指以合成的寡核苷酸为探针的合成型 DNA 芯片，但还包含以 PCR 产物等的 cDNA 为探针的粘贴型 DNA 微阵列。

[0155] 通过从 cDNA 序列中特定出特征序列的以往的公知方法，可以决定用作探针的序列。具体如 SAGE：Serial Analysis of Gene Expression 法（Science 276：1268，1997；Cell 88：243，1997；Science 270：484，1995；Nature 389：300，1997；美国专利第 5,695,937 号）等。

[0156] 另外，制造 DNA 芯片时，采用公知的方法即可。例如，当利用合成的寡核苷酸作为寡核苷酸时，通过照相平板印刷技术（Photolithography）与 DNA 固相合成技术相结合，在支持物上合成该寡核苷酸即可。而当利用 cDNA 作为寡核苷酸时，用阵列机粘贴在支持物上即可。

[0157] 此外，与一般的 DNA 芯片一样，也可以排布完全匹配探针（寡核苷酸）和该完全匹配探针中有一个碱基被置换的错配探针，进一步提高基因的检测精度。而且，为了平行检测不同的基因，还可以将多种寡核苷酸固定在同一支持物上构建 DNA 芯片。

[0158] 本发明涉及的基因检测工具不限于上述例举的 DNA 芯片，只要是用本发明涉及的基因的至少一部分碱基序列或其互补序列作探针的基因检测工具即可。

[0159] (5-7) 抗体

[0160] 本发明涉及的抗体是以本发明涉及的蛋白质、或其部分蛋白质·部分肽为抗原，通过公知的方法得到的多克隆抗体或单克隆抗体。公知的方法如文献（Harlow 等的“Antibodies：A laboratory manual”（ColdSpring Harbor Laboratory，New York（1988））、岩崎等的“单克隆抗体杂交瘤与 ELISA，讲谈社（1991）”）中记载的方法。按照上述方法得到的抗体，可用于本发明蛋白质的检出和测定等。

[0161] (5-8) 筛选方法

[0162] 本发明涉及的筛选方法是用本发明涉及的蛋白质对调节该蛋白质的基因或调节该蛋白质的物质进行筛选的方法。作为本发明的筛选方法,可以适宜地采用分析物质间有无结合、有无解离的以往公知的各种方法,没有特殊限制。例如,对促进本发明涉及的蛋白质的活性($\Delta 6$ 去饱和活性、 $\Delta 6$ 碳链延长活性和/或 $\Delta 5$ 去饱和活性)的物质进行筛选。

[0163] 而且,本发明还包含:通过上述筛选方法得到的基因或物质。

[0164] 下面,列举实施例,更详细地说明本发明,当然,本发明不限于以下的实施例,在细节上可以是各种形态。而且,本发明也不限于上述实施形态,在权利要求相所示的范围内可以有各种变更,适宜地组合已公开的技术手段的实施形态也包含在本发明的技术范围内。

[0165] [实施例]

[0166] 在本实施例中,只要没有特殊说明,实验方法遵照 MolecularCloning(Sambrook et. al. Cold Spring Harbour Laboratory Press,1989) 中所述的方法。

[0167] [实施例 1:来自地钱的 $\Delta 6$ 去饱和酶基因的分离]

[0168] 由至今为止克隆的 $\Delta 6$ 去饱和酶的氨基酸序列的比较可知:氨基酸序列 Trp-Trp-Lys-(Glu/Asp)-Lys-His-Asn(序列号 37) 和 Trp-Phe-Thr-Gly-Gly-Leu-Asn(序列号 38) 被保留。因此,为了分离来自地钱的 $\Delta 6$ 去饱和酶基因,采用对上述氨基酸序列进行编码的下述简并引物。

[0169] d 6DES-F 5' -TGGTGGAA(A/G)GA(A/G/T/C)AA(A/G)CA(T/C)AA-3' (序列号 7)

[0170] d 6DES-R 5' -(A/G)TTIA(A/G)ICCICIGT(A/G)AACCA-3' (序列号 8)

[0171] (I 表示肌苷, () 内表示多个碱基)

[0172] 试验材料用的是 E 系统地钱(参照 Transgenic Res. 9,p179,2000) 的叶状体。按照文献(Biosci. Biotechnol. Biochem. 67, p605,2003 ;Biosci. Biotechnol. Biochem. 67, p1667,2003) 中记载的方法,从叶状体中分离 poly(A)+RNA。将分离得到的 poly(A)+RNA 1.5 μ l 用 Ready-To-Go T-primed First Strand kit(Amersham 公司生产) 逆转录成 cDNA。以上述 cDNA 约 10ng 为模板,利用上述引物(d 6DES-F 和 d 6DES-R) 及酶(Takara Ex Taq、Takara 公司生产)0.5U,用制造者推荐的方法进行 PCR。反应液量为 20 μ l,采用 GeneAmp PCR system9700(PE Applied Biosystems 公司生产),在 94 $^{\circ}$ C 下保持 2 分钟后,进行 94 $^{\circ}$ C 下 1 分钟、45 $^{\circ}$ C 下 1.5 分钟、72 $^{\circ}$ C 下 2 分钟的反应,循环 35 次,然后冷却到 4 $^{\circ}$ C。

[0173] 将得到的 PCR 产物用 1%(w/v) 琼脂糖凝胶进行电泳,根据以往的 $\Delta 6$ 去饱和酶的氨基酸序列预测扩增片段的大小,将具有预测大小的扩增片段用 Prep-A Gene(Bio-rad 公司生产) 通过凝胶回收。将回收的扩增片段连接于 pT7Blue Vector(Takara 公司生产),转化入大肠杆菌 Electro-max DH10B cells(Invitrogen 公司生产, Carlsbad, CA)。

[0174] 利用 BigDye Terminator Cycle Sequencing kit(Applied Biosystems 公司生产) 及 automated sequencer ABI PRISM 377(Applied Biosystems 公司生产),测定得到的全克隆的碱基序列,寻找具有目的 cDNA 序列的克隆。

[0175] 为了获得全长 cDNA 序列,进行 5' -RACE 和 3' RACE。利用 5' -RACE System for Rapid Amplification of cDNA Ends Version 2.0(Invitrogen 公司生产)、Ready-To-Go T-primed First Strand kit(Amersham 公司生产) 以及下述引物(MpDES6-02R 和 MpDES6-01F),按照制造者推荐的方法进行。

[0176] MpDES6-02R :5' -AAGTTGCCTTCGATGTTTCTGG-3' (序列号 9)

[0177] MpDES6-01F :5' -GCTCGCCTGGAGCAAGGAAATC-3' (序列号 10)

[0178] 结果分离得到一种同源候选基因,将此基因作为 MpDES6 基因。分离的 MpDES6 基因的 cDNA 长度(除去 polyA 部分)为 2,522bp,其编码的推定氨基酸序列为 481 个残基。其碱基序列在序列号 1 表示,氨基酸序列在序列号 2 表示。

[0179] 将 MpDES6c DNA 的推定氨基酸序列与小立碗藓的 $\Delta 6$ 去饱和酶的氨基酸序列比较,结果显示:只有 47.5% 的同一性。

[0180] [实施例 2:来自地钱的 $\Delta 6$ 碳链延长酶基因的分离]

[0181] 由至今为止克隆的 $\Delta 6$ 碳链延长酶的氨基酸序列的比较可知:氨基酸序列 Val-Glu-Phe-Met-Asp-Thr-Val(序列号 39)及 Lys-Tyr-Leu-Phe-Trp-Gly-Arg(序列号 40)被保留。因此,为了分离来自地钱的 $\Delta 6$ 碳链延长酶基因,采用对上述氨基酸序列进行编码的下述简并引物。

[0182] d 6ELO-F 5' -GTIGA(A/G)TT(T/C)ATGGA(T/C)ACIGT-3' (序列号 11)

[0183] d 6ELO-R 5' -C(G/T)ICCCA(A/G)AAIA(A/G)(A/G)TA(T/C)TT-3' (序列号 12)

[0184] 用上述引物(d 6ELO-F 和 d 6ELO-R)进行 PCR,将获得的 DNA 片段进行亚克隆。测定得到的克隆的碱基序列,利用下述引物(MpELO1-02R 和 MpELO1-01F)从具有目的 cDNA 序列的克隆中获得全长 cDNA。另外,实验材料及方法与实施例 1 相同。

[0185] MpELO1-02R :5' -GCGAGCTTTCTCGTTCTTTCCC-3' (序列号 13)

[0186] MpELO1-01F :5' -TATGATTTTGAAGCGCAACACG-3' (序列号 14)

[0187] 结果分离得到一种同源候选基因,将此基因作为 MpELO1 基因。MpELO1 基因的 cDNA 长度(除去 polyA 部分)为 1,559bp,推定氨基酸序列为 290 个残基。其碱基序列在序列号 3 表示,氨基酸序列在序列号 4 表示。

[0188] 将 MpELO1c DNA 的推定氨基酸序列与小立碗藓的 $\Delta 6$ 碳链延长酶的氨基酸序列比较,结果显示:同一性为 62.7%。

[0189] [实施例 3:来自地钱的 $\Delta 5$ 去饱和酶基因的分离]

[0190] 其他生物种类的 $\Delta 5$ 去饱和酶在 N 末端有细胞色素 b5 结构域。由此预测:来自地钱的 $\Delta 5$ 去饱和酶基因与 $\Delta 6$ 去饱和酶基因一样,属于细胞色素 b5 结构域融合型去饱和酶基因家族。但在硅藻、真菌中, $\Delta 5$ 去饱和酶及 $\Delta 6$ 去饱和酶之间在氨基酸序列水平上的同源性非常低。因此,对丝状菌(*M. alpina*)的 $\Delta 5$ 去饱和酶及 $\Delta 6$ 去饱和酶间的氨基酸序列进行比较,结果发现:局部存在简并引物设计时最少需要的 4 至 5 个残基的连续保守序列。还惊奇地发现:这些保守序列在同种内的 $\Delta 5$ 去饱和酶和 $\Delta 6$ 去饱和酶之间比在异种的 $\Delta 5$ 去饱和酶的氨基酸序列之间被更好地保留。由此可认为:在细胞色素 b5 结构域融合型去饱和酶基因中,有时存在种特异性保守序列。于是,潜心比较并探讨了上述 MpDES6 和在 Genetics159, p981, 2001 中记载的功能未知的 MpDES 的碱基序列,结果发现:两处氨基酸序列(I(E/N)(G/D)KVYDV(序列号 41)、及 DPDI(Q/D)(Y/T)(M/V)P(序列号 42))被保留。与各氨基酸序列对应的简并引物序列如下所示。

[0191] d 5DES-F 5' -AT(A/T/C)(A/G)AIG(A/G)IAA(A/G)TITA(T/C)GA(T/C)GT-3' (序列号 15)

[0192] d 5DES-R 5' -GGIA(T/C)I(G/T)(A/T)IT(G/C)(A/G/T)AT(A/G)TCIGG(A/G)

TC-3' (序列号 16)

[0193] 用上述引物 (d 5DES-F 和 d 5DES-R) 进行 PCR, 将得到的 DNA 片段进行亚克隆。测定得到的克隆的碱基序列, 采用下述引物 (MpDES5-02R 及 MpDES5-01F) 从具有目的 cDNA 序列的克隆中获得全长 cDNA。另外, 实验材料及方法与实施例 1 相同。

[0194] MpDES5-02R :5' -GTGTGTACGATCCGTGGTTACC-3' (序列号 17)

[0195] MpDES5-01F :5' -AAGGCGGGACAGGATTCAACAC-3' (序列号 18)

[0196] 作为来自地钱的 $\Delta 5$ 去饱和酶的候选基因, 分离得到 2 种长度不同的克隆 c1 及 c2 (c1 :2, 427bp、c2 :2, 285bp)。比较 c1 和 c2 的碱基序列, 结果表明 :在 5' 非翻译区发生自动拼接 (选择性拼接)。自动拼接没有引起阅读框架的改变, 任一克隆均编码 484 个氨基酸 (序列号 6)。以下, 将 2, 427bp 长的克隆 c1 作为 MpDES5 基因 (序列号 5), 在下面的实施例中使用。

[0197] 将 MpDES5cDNA 的推定氨基酸序列与丝状菌 (*M. alpina*) 的 $\Delta 5$ 去饱和酶的氨基酸序列比较, 结果显示 :同一性为 31.4%。因为与地钱有近缘关系的小立碗藓的 $\Delta 5$ 去饱和酶的序列信息没有被公开, 所以在此没进行比较。

[0198] [实施例 4 :利用甲醇酵母 (*Pichia pastoris*) 的功能分析]

[0199] 为了分析 MpDES6、MpEL01 及 MpDES5 的 cDNA 的功能, 首先, 将各个 ORF 构建于甲醇诱导型启动子 AOX1 的下游。将制备的这些结构导入甲醇酵母 (*Pichia pastoris*), 分析其脂肪酸组成。利用下述引物, 通过 PCR 扩增 MpDES6、MpEL01 及 MpDES5 的 cDNA 碱基序列的 ORF 部分。

[0200] (MpDES6ORF 扩增用引物)

[0201] MpD6-17F :5' -GGAATTCGCGATGGCCTCGTCCACCACCAC-3' (序列号 19)

[0202] MpD6-18F :5' -GGAATTCTACTTTTCGCGCGTATGCTACC-3' (序列号 20)

[0203] (MpEL01ORF 扩增用引物)

[0204] MpD6EL01-15F :5' -GGAATTCGCGATGGAGGCGTACGAGATGG-3' (序列号 21)

[0205] MpD6EL01-16F :5' -GGAATTCTTCTGCCTTTTGTCTTGTATC-3' (序列号 22)

[0206] (MpDES5ORF 扩增用引物)

[0207] MpD5-11F :5' -GTTGAATTCGACAGTTATGCCGCCACACGC-3' (序列号 23)

[0208] MpD5-12R :5' -GTTGAATTCAGCCCAAAGCATGCTGTAC-3' (序列号 24)

[0209] 这些引物含有下划线所示的 EcoRI 识别序列, 在以下的克隆中利用。用 Pyrobest DNA polymerase (Takara 公司生产) 0.5U, 按照制造者推荐的方法, 反应液量为 20 μ l, 进行 PCR。反应条件为, 在 94°C 下保持 2 分钟后, 进行 94°C 下 1 分钟、57°C 下 1 分钟、72°C 下 1 分钟的反应, 循环 25 次, 然后冷却到 4°C。将得到的各 ORF 片段用 EcoRI 消化后, 用实施例 1 中记载的方法进行凝胶纯化。然后, 正向连接于甲醇酵母的表达载体 pPICZA (标记 :Zeocin 抗性基因, Invitrogen 公司生产) 内甲醇诱导型启动子 5' AOX1 下游的 EcoRI 部位。

[0210] 将各个表达结构以及作为对照的 pPICZA 载体用 *Pichia EasyCompkit* (Invitrogen 公司生产) 导入甲醇酵母的 PPY1 系统, 以 Zeocin 抗性为标记, 获得转化体。另外, 甲醇酵母可以合成 $\Delta 6$ 去饱和酶的底物即亚

[0211] 油酸和 α -亚麻酸, 但不能合成花生四烯酸、EPA 的其他的前体。

[0212] 为了使导入的基因表达, 利用 EasySelect *Pichia Expression Kit* (Invitrogen

公司生产),按照该试剂盒推荐的方法,将各转化体在以 1.0%甘油为唯一碳源的最低限度培养基中培养到 OD 值(600nm)到 0.5 后,在以 0.5%甲醇为唯一碳源的最低限度培养基中,在 30℃下培养 3 天到饱和状态。然后,用 GC-MS,通过公知的方法(Biosci. Biotechnol. Biochem. 67, p605, 2003),测定各转化体的脂肪酸组成。

[0213] 在表达 MpDES6 基因的转化体中,新检测出 $\Delta 6$ 去饱和酶的反应产物即 γ -亚麻酸、十八碳四烯酸,分别为总脂肪酸的 7.4%、0.7%。在导入了作为对照的 pPICZA 载体的酵母中,没有检测出上述物质。由此表明, MpDES6 编码 $\Delta 6$ 去饱和酶。

[0214] 在表达 MpELO1 基因的转化体中,当添加了 γ -亚麻酸时新检测出二均- γ -亚麻酸,为总脂肪酸的 14.1%,当添加了十八碳四烯酸时新检测出二十碳四烯酸,为 1.5%。在导入了作为对照的 pPICZA 载体的酵母中,没有检测出上述物质。由此表明, MpELO1 编码 $\Delta 6$ 碳链延长酶。

[0215] 在表达 MpDES5 基因的转化体中,当添加了二均- γ -亚麻酸时检测出花生四烯酸,为总脂肪酸的 1.1%,当添加了二十碳四烯酸时检测出二十碳五烯酸(EPA),为 0.1%。在导入了作为对照的 pPICZA 载体的酵母中,没有检测出上述物质。由此表明, MpDES5 编码 $\Delta 5$ 去饱和酶。

[0216] 如上所述,从地钱中获得了编码 $\Delta 6$ 去饱和酶、 $\Delta 6$ 碳链延长酶及 $\Delta 5$ 去饱和酶的基因,分别为 MpDES6、MpELO1 及 MpDES5。

[0217] [实施例 5 : 甲醇酵母 (*P. pastoris*) 中地钱多不饱和脂肪酸生物合成系统的再构建]

[0218] 为了使 MpDES6、MpELO1 及 MpDES5 共表达,将上述实施例 4 中制备的经 EcoRI 消化的 MpELO1 及 MpDES5 的 ORF 扩增片段,分别正向连接于不同的甲醇酵母表达载体 pPIC3K (标记 : HIS4 基因、Invitrogen 公司生产) 及 pPIC6A (标记 : 灭瘟素 (blasticidin) 抗性基因、Invitrogen 公司生产) 的 5' AOX1 启动子下游的 EcoRI 部位。此外, MpDES6 采用的是实施例 4 中制备的表达载体。以下各表达载体分别表示为 pPICZA-MpDES6、pPIC3K-MpELO1 及 pPIC6A-MpDES5。

[0219] 首先,将 pPICZA-MpDES6、或只将作为对照的 pPICZA 载体转化到甲醇酵母 PPY12 系统 (his4, arg4), 该甲醇酵母 PPY12 系统具有与上述实施例 4 中用的甲醇酵母 PPY1 系统相同的脂肪酸组成,并以 Zeocin 抗性为标记,获得转化体。然后,将 pPIC3K-MpELO1、或只将作为对照的 pPIC3K 载体分别导入基因组中整合了 pPICZA-MpDES6、或只整合了 pPICZA 的转化体,以组氨酸合成能力为标记,获得转化体。最后,将 pPIC6A-MpDES5、或只将作为对照的 pPIC6A 载体导入基因组中整合了 pPICZA-MpDES6 及 pPIC3K-MpELO1、或整合了 pPICZA 及 pPIC3K 的上述转化体中,以灭瘟素 (ブラストジン) 抗性为标记,获得转化体。

[0220] 用得到的导入了 2 种或 3 种基因的转化体,进行地钱的花生四烯酸 /EPA 生物合成系统的再构建实验。首先,利用上述导入了 2 种基因 (MpDES6 及 MpELO1) 的转化体,使 MpDES6 基因及 MpELO1 基因在甲醇酵母内共表达。其结果为,除了生成作为 $\Delta 6$ 去饱和反应产物的 γ -亚麻酸 (总脂肪酸的 2.9%) 及十八碳四烯酸 (总脂肪酸的 0.4%) 外,还生成其碳链延长反应产物即二均- γ -亚麻酸 (总脂肪酸的 2.8%) 及二十碳四烯酸 (总脂肪酸的 0.2%)。在作为对照的转化体中,没有检测出这些脂肪酸。在上述转化体中再导入 MpDES5 的转化体,除了生成 γ -亚麻酸 (总脂肪酸的 2.8%)、十八碳四烯酸 (总脂肪酸的

0.5%)、二均- γ -亚麻酸(总脂肪酸的1.5%)及二十碳四烯酸(总脂肪酸的0.1%)这4种脂肪酸外,还生成花生四烯酸(总脂肪酸的0.1%)及二十碳五烯酸(EPA、总脂肪酸的0.03%)。在作为对照的转化体中,没有检测出这些脂肪酸。该结果表明:通过使来自地钱的 $\Delta 6$ 去饱和酶基因、 $\Delta 6$ 碳链延长酶基因及 $\Delta 5$ 去饱和酶基因表达,即使在地钱以外的生物种类中也可以再构建多不饱和脂肪酸生物合成系统。

[0221] [实施例6:导入稻子的载体的构建及该载体向稻子的导入]

[0222] 在稻子中,为了使MpDES6、MpELO1及MpDES5基因表达,按照以下(i)~(iv)所示的顺序,构建表达结构。构建顺序在图1中显示。

[0223] (i) 利用在pBI221(TOYOCO公司生产)的花椰菜花叶病毒(CaMV)35S启动子和NOS终止子间设计的以下引物,进行PCR,从而制备除去了 β -Glucuronidase(GUS)基因部分的表达载体p35S-NOS。

[0224] MK001(F):5'-CGGGATCCTCTCCTGGCGCACCATCGTC-3'(序列号25)

[0225] MK002(R):5'-GGGGTACCAACGCGCTTTCCACCAACG-3'(序列号26)

[0226] 另外,引物MK001(F)含有下划线所示的BamHI识别序列,与GUS基因的3'末端退火,引物MK002(R)与GUS基因的5'末端退火(BamHI位点存在于退火部位的上游)。利用Pyrobest DNA polymerase(Takara公司生产)0.5U,按照制造者推荐的方法,反应液量为50 μ l,进行PCR。反应条件为,96 $^{\circ}$ C保持5分钟后,进行94 $^{\circ}$ C下30秒、68 $^{\circ}$ C下4分钟的反应,循环30次,然后冷却到4 $^{\circ}$ C。将得到的各ORF片段进行BamHI消化后,用实施例1中记载的方法进行凝胶纯化,然后使其自我连接。

[0227] (ii) 接着,分别将MpDES6基因、MpELO1基因及MpDES5基因的ORF连接到p35S-NOS的XbaI位点。ORF扩增采用含有下划线所示XbaI识别序列的以下引物。

[0228] (MpDES6ORF扩增用引物)

[0229] MpD6-21F:5'-GCTCTAGAGCGATGGCCTCGTCCACCACC-3'(序列号27)

[0230] MpD6-11R:5'-GCTCTAGACTATACTTTCGAGCGTATGC-3'(序列号28)

[0231] (MpELO1 ORF扩增用引物)

[0232] MpD6ELO1-18F:5'-GCTCTAGAGCGATGGAGGCGTACGAGATGG-3'

[0233] (序列号29)

[0234] MpD6ELO1-13R:5'-GCTCTAGATTATTCTGCCTTTTTGCTC-3'(序列号30)

[0235] (MpDES5 ORF扩增用引物)

[0236] MpD5-22F:5'-GCTCTAGAGACAGTTATGCCGCCACACGC-3'(序列号31)

[0237] MpD5-23R:5'-GCTCTAGAAGGCCCAAAGCATGCTGTAC-3'(序列号32)

[0238] 用Pyrobest DNA polymerase(Takara公司生产)0.5U,按照制造者推荐的方法,反应液量为20 μ l,进行PCR。反应条件为,在94 $^{\circ}$ C下保持2分钟后,进行94 $^{\circ}$ C下1分钟、57 $^{\circ}$ C下1分钟、72 $^{\circ}$ C下1分钟的反应,循环25次,然后冷却到4 $^{\circ}$ C。将得到的各ORF片段用XbaI消化后,用实施例1中记载的方法进行凝胶纯化,用于克隆。

[0239] (iii) 得到的各基因的表达结构(分别表示为p35S-MpDES6、p35S-MpELO1、p35S-MpDES5)均在CaMV35S启动子5'末端具有PstI位点、在NOS终止子3'末端具有EcoRI位点,利用此结构将上述3种基因的表达盒连接。首先,利用以下所示的引物,以p35S-MpDES5为模板,进行PCR,扩增MpDES5基因的表达盒部分,克隆到p35S-MpDES6的

CaMV35S 启动子 5' 末端的 PstI 位点。(参照图 1)

[0240] (MpDES5 基因表达盒扩增用引物)

[0241] M13R :5' -CAGGAAACAGCTATGACC-3' (序列号 33)

[0242] NOS-R4-PST :5' -AAACTGCAGATTCCCAGTCTAGTAACATAG-3' (序列号 34)

[0243] 另外, M13R 引物与 CaMV35S 启动子上游的载体序列退火。而 NOS-R4-PST 引物含有下划线所示的 PstI 识别序列,与 NOS 终止子 3' 末端退火。不包含同样存在于 NOS 终止子 3' 末端的 EcoRI 位点。

[0244] 用 Pyrobest DNA polymerase (Takara 公司生产) 0.5U, 按照制造者推荐的方法, 反应液量为 20 μ l, 进行 PCR。反应条件为, 在 94°C 下保持 2 分钟后, 进行 94°C 下 1 分钟、57°C 下 1 分钟、72°C 下 1 分钟的反应, 循环 25 次, 然后冷却到 4°C。将得到的 DNA 片段用 PstI 消化后, 用实施例 1 中记载的方法进行凝胶纯化, 克隆到含有 MpDES6 基因表达盒的质粒 (p35S-MpDES6) 的 PstI 位点。

[0245] (iv) 在上述得到的连有 MpDES5 及 MpDES6 基因的表达盒的结构 (表示为 p35S-MpDES5/35S-MpDES6) 上, 再连接 MpELO1 基因的表达盒。利用以下引物, 以 p35S-MpELO1 为模板进行 PCR, 扩增 MpELO1 基因的表达盒部分, 克隆到 MpDES6 基因表达盒内 NOS 终止子 3' 末端的 EcoRI 位点。

[0246] (MpELO1 基因表达盒扩增用引物)

[0247] 35S-F3-EI :5' -CCGGAATTCGCATGCCTGCAGGTCCCCAGA-3' (序列号 35)

[0248] M13F :5' -TGTA AACGACGGCCAGT-3' (序列号 36)

[0249] 另外, 35S-F3-EI 引物含有下划线所示的 EcoRI 识别序列, 与 CaMV35S 启动子的 5' 末端退火。M13F 引物与 NOS 终止子下游的载体序列退火。

[0250] 用 Pyrobest DNA polymerase (Takara 公司生产) 0.5U, 按照制造者推荐的方法, 反应液量为 20 μ l, 进行 PCR。反应条件为, 在 94°C 下保持 2 分钟后, 进行 94°C 下 1 分钟、57°C 下 1 分钟、72°C 下 1 分钟的反应, 循环 25 次, 然后冷却到 4°C。将得到的 DNA 片段用 EcoRI 消化后, 用实施例 1 中记载的方法进行凝胶纯化, 克隆到使 MpDES5 及 MpDES6 基因的表达盒连接的结构 (p35S-MpDES5/35S-MpDES6) 的 EcoRI 部位。

[0251] 通过以上操作, 制备 3 种基因的表达盒按照 MpDES5、MpDES6、MpELO1 的顺序连接的表达结构 (p35S-MpDES5/35S-MpDES6/p35S-MpELO1)。

[0252] 用公知的方法 (Genes Genet. Syst. 73, p219, 1998), 将按照上述方法得到的上述结构与以双丙氨酸为选择标记的质粒一起用微弹枪导入稻子, 获得转基因稻子。

[0253] [实施例 7 :烟草 (N. tabacum SR-1) 中地钱多不饱和脂肪酸合成系统的再构建]

[0254] 在本实施例中证实 :上述来自地钱的不饱和脂肪酸合成酶基因即 MpDES6 基因、MpDES5 基因、MpELO1 基因在植物体中良好地发挥功能。

[0255] 具体地说证实了 :在烟草中导入 MpDES6 基因、MpDES5 基因、MpELO1 基因, 在该烟草中可以生产花生四烯酸等。作为对照, 制备导入了来自丝状菌 (M. alpina) 的 $\Delta 6$ 去饱和酶基因 (MaDES6)、 $\Delta 5$ 去饱和酶基因 (MaDES5)、 $\Delta 6$ 脂肪酸碳链延长酶基因 (MaELO) 的烟草。

[0256] (i) 含有来自丝状菌基因的载体 (pSPB1519) 的构建

[0257] pE2113 (Mitsuhara et al. Plant Cell Physiol. 37, 45-59 1996) 具有花椰菜花叶病毒 35S (E1235S) 启动子及胭脂碱合成酶 (nos) 终止子, 所述花椰菜花叶病毒 35S (E1235S)

启动子具有反复的增强子序列。

[0258] 将 pE2113 用 SnaBI 消化后,与 XhoI 接头 (TAKARA 公司) 连接,从而制备质粒。将此质粒用 SacI 消化后,进行末端平滑化,再与 BamHI 接头 (TAKARA 公司) 连接,从而制备 pUE7。将 pUE7 经 HindIII 和 EcoRI 消化而得到的 DNA 片段中具有 E1235S 启动子的片段、与经 HindIII 和 EcoRI 消化的植物转化用双元载体 pBINPLUS (van Engelen et al. Transgenic research 4, p288, 1995) 连接,从而制备 pSPB505。

[0259] 另一方面,将含有 MaDES6 基因的质粒 pMLD101 用 XhoI 消化后,再用 BamHI 部分消化,从而得到约 1.6kb 的 DNA 片段。将此片段与 pSPB505 经 XhoI 和 BamHI 消化得到的双元载体部分的 DNA 片段连接,从而制备 pSPB559。

[0260] 将 pUCAP (van Engelen et al. Transgenic research 4, p288, 1995) 用 AscI 消化后,进行末端平滑化,与 PacI 接头连接,从而制备 pUCAPP。

[0261] 将 pE2113 用 SnaBI 消化后,与 BamHI 接头 (TAKARA 公司) 连接,制备 pUE6。将 pUE6 用 SacI 消化后,进行末端平滑化,与 SalI 接头 (TAKARA 公司) 连接,从而制备 pUE8。将 pUE8 经 HindIII 及 EcoRI 消化得到的 DNA 片段中具有 E1235S 启动子的片段,插入 pUCAPP 的 HindIII-EcoRI 间,从而制备质粒。将此质粒用 BamHI 和 SalI 消化得到的 DNA 片段、与 MaELO 基因的 cDNA 经 BamHI 和 XhoI 消化得到的 DNA 片段连接,制备 pSPB1130。将 pSPB1130 用 PacI 消化,将得到的约 2.3kb 的 DNA 片段插入 pBINPLUS 的 PacI 位点。选择 MaELO 基因的转录方向与 pBINPLUS 上的 nptII 基因的转录方向相同的质粒,将该质粒作为 pSPB1157P。

[0262] 将上述 pSPB559 用 PacI 消化后,进行末端平滑化,与 AscI 接头连接,从而制备 pSPB559A。然后,将 pSPB559A 经 AscI 消化得到的含有 MaDES6 基因的 DNA 片段,插入 pSPB1157P 的 AscI 位点,从而制备 pSPB1157。

[0263] 将 pCGP1364 (Plant Cell Physiol. 36, 1023, 1995) 经 HindIII 及 SacII 消化得到的约 1.3kb 的 DNA 片段、pCGP1364 经 PstI 消化并进行末端平滑化后又用 SacII 消化得到的约 2.9kb 的 DNA 片段、pUCAPA 经 SacI 消化并进行末端平滑化后又用 HindIII 消化得到的约 2.7kb 的 DNA 片段连接,从而得到 pSPB184。

[0264] 另一方面,从亚克隆了 MaDES5 基因的 pCRII 载体中,通过 XbaI 和 KpnI 的消化,切割含有 MaDES5 基因的 DNA 片段。将此 DNA 片段与上述 pSPB184 经 XbaI 和 KpnI 消化得到的 DNA 片段连接,从而制备 pSPB1519A。将 pSPB1519A 用 AscI 消化,将得到的 DNA 片段插入 pSPB1157 的 AscI 位点,制备 pSPB1519。在此 pSPB1519 上, MaDES6 基因、MaDES5 基因及 MaELO 基因在同一方向上被转录,受同一组成型启动子的控制。

[0265] (ii) 构建来自地钱的基因的载体 (pSPB2368)

[0266] 将 pUCAP (van Engelen et al. Transgenic Research 4, 288-290, 1995) 用 AscI 消化后与 SgfI 接头连接,再用 PacI 消化后与 FseI 接头连接,从而制备 pUCSAPF。此外,对 pBINPLUS 也实施同样的处理,制备 pBINSAPF。

[0267] 作为其他亚克隆用载体,将 pUC19 用 HindIII 消化后,与 PacI 接头连接,再用 EcoRI 消化后与 FseI 接头连接,从而制备 pUCPF。此外,将 pUC19 用 HindIII 消化后,与 SgfI 接头连接,再用 EcoRI 消化后与 AscI 接头连接,从而制备 pUCSA。将在 pUCSAPF 的 HindIII-XbaI 间插入 E1235S 且在 SacI-EcoRI 间插入甘露碱合成酶 (mas) 基因终止子的载体,用 XbaI 及 SacI 消化后,将末端平滑化,从而得到 pSPB2353A。在 pSPB2353A 的平滑末

端,连接从 p35S-MpDES6 中用 XbaI 切割并进行末端平滑化后的含有 MpDES6 基因的 DNA 片段,制备 pSPB2353。

[0268] 将在 pUCSA 的 HindIII-XbaI 间插入 E1235S 且在 SacI-EcoRI 间插入甘露碱合成酶 (mas) 基因终止子的载体,用 XbaI、SacI 消化,从而制备 pSPB2355A。

[0269] 另一方面,以 p35S-MpELO1 为模板,利用下述引物 XbaMpELOf 和 SacMpELOr,进行 PCR。

[0270] XbaMpELOf :5' -AGTCTCTAGAGCGATGGAGGCGTACG-3' (序列号 43)

[0271] SacMpELOr :5' -CAGTGAGCTCGGTGCTTATTCTGCC-3' (序列号 44)

[0272] PCR 的酶采用高精度的 KOD+DNA 聚合酶 (东洋纺),在 94°C 下保持 2 分钟后,94°C 下 15 秒、68°C 下 1-3 分钟,循环 25 次。将这样制备的 MpELO1DNA 片段经 XbaI 和 SacI 消化后的片段连接于上述 pSPB2355A,制备 pSPB2355。再将 pSPB2355 经 SgfI、AscI 消化得到的 DNA 片段,连接于经 SgfI 和 AscI 消化后的 pSPB2353,制备 pSPB2361。

[0273] 将在 pUCPF 的 HindIII-XbaI 位点插入 E1235S 且在 SacI-EcoRI 位点插入甘露碱合成酶 (mas) 基因终止子的载体,用 XbaI 及 SacI 消化,从而制备 pSPB2352A。

[0274] 另一方面,以 p35S-MpDES5 为模板,利用下述引物 XbaMpD5f 和 SacMpD5r,进行 PCR。PCR 的反应条件与上述 PCR 条件相同。

[0275] XbaMpD5f :5' -AGCTTCTAGAGCCATGCCGCCACACGCCC-3' (配列番号 45)

[0276] SacMpD5r :5' -CAGTGAGCTCTCAGCCATCCAGTCGT-3' (配列番号 46)

[0277] 将通过 PCR 制备的 MpD5DNA 片段用 XbaI、SacI 消化后,与 pSPB2352A 连接,制备 pSPB2352。

[0278] 在 pBINSAPF 经 PacI 和 FseI 消化得到的 DNA 片断上,连接从 pSPB2352 用 PacI 和 FseI 切出的含有 MpDES5 基因的 DNA 片断,制备 pSPB2368A。再将 pSPB2368A 用 SgfI 和 PacI 消化后,连接从 pSPB2361 用 SgfI 和 PacI 切出的含有 MpDES6 基因和 MpELO1 基因的 DNA 片断,从而得到 pSPB2368。在此质粒上, MpDES6 基因、MpDES5 基因、以及 MpELO1 基因在同一方向上被转录,受同一组成型启动子的控制。

[0279] (iii) 向烟草中导入基因

[0280] 接着,根据公知的方法 (Plant J. 5, 81, 1994),用 pSPB2368 或 pSPB1519 转化 *Agrobacterium tumefaciens* (菌株 :Ag10 (Lazo et al. 1991, Biotechnology 9 :963-967))。使该具有 pSPB2368 或 pSPB1519 的转化体农杆菌感染烟草叶盘。从这样得到的重组烟草叶片中用 RNeasyPlant miniKit (Qiagen) 提取 RNA,按照规定的方法,通过 RT-PCR,筛选表达了导入基因的系统。

[0281] 从按照上述方法得到的导入了由来自丝状菌的酶基因构成的 pSPB1519 的烟草 (pSPB1519 转基因烟草)、导入了由来自地钱的酶基因构成的 pSPB2368 的烟草 (pSPB2368 转基因烟草) 的叶片中,根据公知的方法 (藤野安彦编 (1978) 生物化学实验法 9 学会出版中心、山田晃弘编 (1989) 生物化学实验法 24 学会出版中心) 进行脂质的提取。将得到的脂质用气相色谱法 (Hewlett Packard 公司 HP-6800) 进行分析,分析结果在表 1 中显示。

[0282] 另外,作为对照,用没有导入基因的烟草叶片进行同样的分析。

[0283] [表 1]

[0284]

	对照 (%)	对照 (mg/gFW)	pSPB2368 (%)	pSPB2368 (mg/gFW)	pSPB1519 (%)
亚油酸	9.55	1.17	1.37	0.2	9.51
α -亚麻酸	49.99	6.12	17.83	2.58	39.24
γ -亚麻酸	0	0	4.06	0.59	3.37
二均 γ -亚麻酸	0	0	10.85	1.57	3.09
花生四烯酸	0	0	10.27	1.49	0
二十碳四烯酸	0	0	4.89	0.71	0
二十碳五烯酸	0	0	2.68	0.39	0
全脂质量	-	12.25	-	14.48	-

[0285] 本实施例的气相色谱法的分析条件如下所示。

[0286] (气相色谱法分析条件)

[0287] 色谱柱:Supelco SP-2330、Fused Silica Capillary Column、30m×0.32mm ID、0.2 μ m

[0288] 温度:Inj:240°C、Det:250°C、Oven:180°C 3分、180°C→220°C (2°C/min)

[0289] 柱流量:30cm/sec、压力 200kPa、检测器 FID

[0290] 根据标准脂肪酸甲酯的保留时间和 GC-MASS (Hewlett Packard 公司、HP-5973) 分析,决定色谱中的各个峰值,此外,由峰面积决定各脂肪酸的比例。

[0291] 在表 1 中,Control 表示对照,pSPB2368 表示 pSPB2368 转基因烟草,pSPB1519 表示 pSPB1519 转基因烟草。

[0292] 由表 1 所示的结果可知:在导入了由来自丝状菌的酶基因构成的 pSPB1519 的烟草 (pSPB1519 转基因烟草) 中,有二均- γ -亚麻酸的蓄积,但没有花生四烯酸的蓄积。而在导入了由来自地钱的酶基因构成的 pSPB2368 的烟草 (pSPB2368 转基因烟草) 中,不仅有花生四烯酸的蓄积,还有二十碳四烯酸和二十碳五烯酸的蓄积。由此可知,在高等植物中,来自地钱的酶在功能上优于来自丝状菌的酶,能够以亚油酸、 α -亚麻酸为底物合成以花生四烯酸为代表的多不饱和脂肪酸。

[0293] Abbadi 等报道:将硅藻 (Phaeodactylum tricornutum) 的 $\Delta 6$ 去饱和酶基因、 $\Delta 5$ 去饱和酶基因、以及小立碗藓 (Physcomitrella patens) 的碳链延长酶基因导入烟草及亚麻 (Linum usitatissimum) 中,使烟草种子中蓄积 1.5% 花生四烯酸,使亚麻种子中蓄积 1.0% 花生四烯酸 (Amine Abbadi et al. Plant Cell 16, 2734-2748, 2004)。

[0294] 在本实施例中,将来自地钱的 $\Delta 6$ 去饱和酶、 $\Delta 5$ 去饱和酶及碳链延长酶的基因导

入烟草后,在烟草叶中蓄积 10%或大于 10%花生四烯酸。此结果表明,与上述报道相比,本实施例的 pSPB2368 转基因烟草具有更高效地合成多不饱和脂肪酸的能力。

[0295] 此外,在用定鞭藻 (*Isochrysis galbana*) 的 $\Delta 9$ 碳链延长酶基因、小眼虫 (*Englena gracilis*) 的 $\Delta 8$ 去饱和酶基因、丝状菌的 $\Delta 5$ 去饱和酶基因在拟南芥 (*Arabidopsis thaliana*) 中进行脂质改变报道中,可知叶片中花生四烯酸占总脂质的 6.6mol%、碳原子数为 20 或大于 20 的脂质占 22.5mol% (Baoxiu Qi et al. Nature Biotechnology 22,739-745,2004)。在此报告中,经不同于采用 $\Delta 6$ 去饱和酶、 $\Delta 5$ 去饱和酶及碳链延长酶的修饰路径合成多不饱和脂肪酸,不能简单地作比较,但可知利用来自地钱的酶可蓄积更多的多不饱和脂肪酸。

[0296] 在总脂质的 30%或大于 30%是改变的脂肪酸的 pSPB2368 转基因烟草中,没有发现形态异常。而且,育性上也没问题,产生许多种子,从中可知异位多不饱和脂肪酸的增加对植物的生长影响小。

[0297] 根据至今为止的报道,在基因重组植物中生产的碳原子数 20 或大于 20 的多不饱和脂肪酸占总脂质的 20%左右已经是最高限,就花生四烯酸而言,最高为 6%左右。但是,按照本实施例的方法,通过利用来自地钱的脂肪酸合成酶,可以打破上述界限,在植物中生产更多的多不饱和脂肪酸。

[0298] 另外,基于具体实施方式的具体实施形态或实施例是对本发明技术内容的说明,不应该只限于上述具体例子被狭义地解释,在本发明的精神和所述权力要求的范围内,可以进行各种变更后实施。

[0299] 本发明涉及的基因是从同一种地钱中分离的 $\Delta 5$ 去饱和酶基因、 $\Delta 6$ 去饱和酶基因及 $\Delta 6$ 碳链延长酶基因。因此具有这样的效果:与使来自不同生物种类的多个基因在植物体内表达时相比,使上述 3 种基因在植物体内同时表达能发挥更好的功能。还具有这样的效果:因为地钱被认为是高等植物的模式系统,所以这些基因比来自植物以外的生物种类基因在植物体内能更好地发挥功能。

[0300] 而且,本发明涉及的转化体可以生产花生四烯酸、二十碳五烯酸 (EPA) 等多不饱和脂肪酸。特别是,本发明涉及的转基因植物可以在成本低且环保的生产过程中生产花生四烯酸、EPA 等多不饱和脂肪酸。按照上述方法得到的花生四烯酸、EPA 可作为廉价多用途的材料得到应用。而且,当上述转基因植物体作为食品时,因其花生四烯酸、EPA 的含量高,所以具有很高的价值。

[0301] 综上所述,本发明的基因·蛋白质在花生四烯酸、EPA 的生产中很有用。而且,可表达地导入了本发明的基因的转化体在制药产业、食品产业、各种材料产业等产业中生产花生四烯酸、EPA 时极其有用。尤其当上述转化体是植物体时,因植物体内花生四烯酸、EPA 的含量增加,所以在农业领域等领域上也非常有用。

[0302] 序列表

[0303]

序列表 (SEQUENCE LISTING)

<110> 三得利株式会社 (Suntory Limited)

<120> 来自地钱的不饱和脂肪酸合成系统酶基因及其利用 (MARCHANTIALES-DERIVED UNSATURATED FATTY ACID SYNTHETASE GENES AND USE OF THE SAME)

<130>

<150> JP 2003-425673

<151> 2003-12-22

<160> 46

<170> PatentIn Ver. 2.1

<210> 1

<211> 2519

<212> DNA

<213> Marchantia polymorpha

<220>

<221> CDS

<222> (253).. (1698)

<400> 1

```

atagatccaa tttcataagt cgacgagaaa ggcagaaggc gagaagcggc aggcagcgag 60
cgcgagcgcc agagctcttg ctcccctcgc tcatcgctcg cattgccgca ttttgtgagt 120
gtcggactga tcaactcagtc cgctcactgca aacgcgagcg agcgagagtg cgagtgagcg 180
agcgagcgag cgagagccgc ggtgtgtctg tgagatccaa tccttttct gctttgcgcg 240
ctgtggggcg cg atg gcc tcg tcc acc acc acc gcc gtg aag caa tct tcg 291
      Met Ala Ser Ser Thr Thr Thr Ala Val Lys Gln Ser Ser
          1             5             10

ggt ggg ctg tgg tcg aaa tgg ggc acc ggc agc aac ttg agc ttc gtg 339
Gly Gly Leu Trp Ser Lys Trp Gly Thr Gly Ser Asn Leu Ser Phe Val
      15             20             25

tcg cgc aag gag cag cag cag cag cag cag agc tct ccc gag gcg 387
Ser Arg Lys Glu Gln Gln Gln Gln Gln Gln Ser Ser Pro Glu Ala
      30             35             40             45

tcg act ccc gcg gcg cag cag gag aaa tcc atc agt aga gaa tcc atc 435
Ser Thr Pro Ala Ala Gln Gln Glu Lys Ser Ile Ser Arg Glu Ser Ile
          50             55             60

ccc gag ggc ttc ttg acc gtg gag gag gtg tcg aag cac gac aat ccg 483
Pro Glu Gly Phe Leu Thr Val Glu Val Ser Lys His Asp Asn Pro
      65             70             75

```

[0304]

agc gac tgc tgg atc gtc atc aac gac aag gtg tac gac gtg agc gca	531
Ser Asp Cys Trp Ile Val Ile Asn Asp Lys Val Tyr Asp Val Ser Ala	
80 85 90	
ttc ggg aag acg cat ccg ggc ggc cct gtg atc ttc acg cag gcc ggc	579
Phe Gly Lys Thr His Pro Gly Gly Pro Val Ile Phe Thr Gln Ala Gly	
95 100 105	
cgc gac gcc acg gat tct ttc aag gtt ttc cac tcc gcc aag gcg tgg	627
Arg Asp Ala Thr Asp Ser Phe Lys Val Phe His Ser Ala Lys Ala Trp	
110 115 120 125	
cag ttt ctc cag gac ctg tac atc gga gat ctg tac aat gcc gag cca	675
Gln Phe Leu Gln Asp Leu Tyr Ile Gly Asp Leu Tyr Asn Ala Glu Pro	
130 135 140	
gtg tcg gag ctg gtg aag gat tac cga gac ctg agg acg gcg ttc atg	723
Val Ser Glu Leu Val Lys Asp Tyr Arg Asp Leu Arg Thr Ala Phe Met	
145 150 155	
cgt tct cag cta ttc aag agc agt aaa atg tac tac gtg acc aag tgc	771
Arg Ser Gln Leu Phe Lys Ser Ser Lys Met Tyr Tyr Val Thr Lys Cys	
160 165 170	
gtc aca aat ttt gca att ctt gcc gcc agt ctc gca gtc atc gcg tgg	819
Val Thr Asn Phe Ala Ile Leu Ala Ala Ser Leu Ala Val Ile Ala Trp	
175 180 185	
agc cag acg tat ctg gcg gtt ttg tgc tcc agt ttc ctg ttg gct ctc	867
Ser Gln Thr Tyr Leu Ala Val Leu Cys Ser Ser Phe Leu Leu Ala Leu	
190 195 200 205	
ttc tgg cag caa tgt gga tgg tta tcg cac gat ttt ctc cac cac cag	915
Phe Trp Gln Gln Cys Gly Trp Leu Ser His Asp Phe Leu His His Gln	
210 215 220	
gtg acc gag aac cga tcg ctc aac acg tac ttc ggc ggc ctg ttc tgg	963
Val Thr Glu Asn Arg Ser Leu Asn Thr Tyr Phe Gly Gly Leu Phe Trp	
225 230 235	
ggt aac ttc gcc cag ggc tac agc gtg gga tgg tgg aag acc aag cac	1011
Gly Asn Phe Ala Gln Gly Tyr Ser Val Gly Trp Trp Lys Thr Lys His	
240 245 250	
aat gtg cac cac gcg gcc acg aac gaa tgc gac gac aag tat cag ccc	1059
Asn Val His His Ala Ala Thr Asn Glu Cys Asp Asp Lys Tyr Gln Pro	
255 260 265	
atc gat ccc gac atc gac acc gtg ccc ctg ctc gcc tgg agc aag gaa	1107
Ile Asp Pro Asp Ile Asp Thr Val Pro Leu Leu Ala Trp Ser Lys Glu	
270 275 280 285	
atc ttg gcc acc gtc gac gac caa ttc ttc cga tcg atc atc agc gtg	1155
Ile Leu Ala Thr Val Asp Asp Gln Phe Phe Arg Ser Ile Ile Ser Val	
290 295 300	

[0305]

cag cac ctt ctg ttc ttc ccg ctc ctc ttc ttg gca aga ttc agc tgg 1203
 Gln His Leu Leu Phe Phe Pro Leu Leu Phe Leu Ala Arg Phe Ser Trp
 305 310 315

ctg cat tcg agt tgg gcc cac gcc agc aac ttc gag atg cct cgg tac 1251
 Leu His Ser Ser Trp Ala His Ala Ser Asn Phe Glu Met Pro Arg Tyr
 320 325 330

atg aga tgg gcg gag aag gcc tcg ctc ctc ggg cac tac ggc gcc tca 1299
 Met Arg Trp Ala Glu Lys Ala Ser Leu Leu Gly His Tyr Gly Ala Ser
 335 340 345

atc ggc gcc gcc ttc tac att ttg ccc atc ccc cag gcc atc tgc tgg 1347
 Ile Gly Ala Ala Phe Tyr Ile Leu Pro Ile Pro Gln Ala Ile Cys Trp
 350 355 360 365

ctc ttc ttg tcg caa ctg ttt tgc ggc gct ctg ctc agc att gtc ttc 1395
 Leu Phe Leu Ser Gln Leu Phe Cys Gly Ala Leu Leu Ser Ile Val Phe
 370 375 380

gtg atc agc cac aat ggc atg gat gtg tac aac gac ccc cgg gac ttc 1443
 Val Ile Ser His Asn Gly Met Asp Val Tyr Asn Asp Pro Arg Asp Phe
 385 390 395

gtg acg gcc caa gtc acc tcg acc aga aac atc gaa ggc aac ttc ttc 1491
 Val Thr Ala Gln Val Thr Ser Thr Arg Asn Ile Glu Gly Asn Phe Phe
 400 405 410

aac gac tgg ttc acc gga ggc ctg aac agg cag att gag cac cat ctg 1539
 Asn Asp Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu His His Leu
 415 420 425

ttt ccg tct ctt ccg agg cac aac ctc gcc aag gtc gcg cca cac gtc 1587
 Phe Pro Ser Leu Pro Arg His Asn Leu Ala Lys Val Ala Pro His Val
 430 435 440 445

aag gcg ctc tgc gcc aag cac ggt ttg cat tac gaa gaa ttg agt ctg 1635
 Lys Ala Leu Cys Ala Lys His Gly Leu His Tyr Glu Glu Leu Ser Leu
 450 455 460

ggc acg gga gtc tgt cgt gtc ttc aat cgg cta gta gag gta gca tac 1683
 Gly Thr Gly Val Cys Arg Val Phe Asn Arg Leu Val Glu Val Ala Tyr
 465 470 475

gct gcg aaa gta tag atcgacgaga gtttcccacc aacacagtta gaacaaggga 1738
 Ala Ala Lys Val
 480

atagtacgag agaaggagac agcaacctgg actttttggt cctgatgttg catactttct 1798

cgaatatacg tctccacgcc ttaagtttc agcttcaact gattgtcttc agtaaccatc 1858

gcttgctcca actgggcgac ctgcagaatt gaagatcagt tttactgagt ttgtaccgag 1918

agtttcccaa attttgttgt aggctgatga cccaatccta gcatacactt taggaataag 1978

cagtctcaac ataattaggt ccatcattca gcaatttcga tacagcgctt gggattcgac 2038

[0306]

gagtttacac gatgagtatg gcttgtaact ggccttctca aggtagcctt ggatctcccc 2098
 gggcctcttg ccattcccatt cacccaatcg agattctgca gtctccaacc ttttctggaa 2158
 gttctcaatc tgtaacctct gttgtagaga tagcatacgc cacaagacaa ggtctttgtg 2218
 aacacagtcg tctaacaac agcaagttgt gtggattggc atctaaataa cgcctctgg 2278
 tcaagtaaca gcagggttc cgcagtttc aggaacatac tttgtttctg tcacagccag 2338
 gcggtgaata gtaaagccaa ttcaacacat acgggagaag atgggtcgat atttgtattt 2398
 ggcagggtgt ccagatttca cccatcagtc tctcacttgc ttgtatgtcc ctgacgtgct 2458
 tcaaaatttt gcgcggggaa tcatcaatat acttaccatt tgtaaaaaaa aaaaaaaaaa 2518
 a 2519

<210> 2

<211> 481

<212> PRT

<213> Marchantia polymorpha

<400> 2

Met	Ala	Ser	Ser	Thr	Thr	Thr	Ala	Val	Lys	Gln	Ser	Ser	Gly	Gly	Leu
1				5					10					15	
Trp	Ser	Lys	Trp	Gly	Thr	Gly	Ser	Asn	Leu	Ser	Phe	Val	Ser	Arg	Lys
		20						25					30		
Glu	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Ser	Ser	Pro	Glu	Ala	Ser	Thr	Pro
		35					40					45			
Ala	Ala	Gln	Gln	Glu	Lys	Ser	Ile	Ser	Arg	Glu	Ser	Ile	Pro	Glu	Gly
	50					55					60				
Phe	Leu	Thr	Val	Glu	Glu	Val	Ser	Lys	His	Asp	Asn	Pro	Ser	Asp	Cys
65				70						75					80
Trp	Ile	Val	Ile	Asn	Asp	Lys	Val	Tyr	Asp	Val	Ser	Ala	Phe	Gly	Lys
				85					90					95	
Thr	His	Pro	Gly	Gly	Pro	Val	Ile	Phe	Thr	Gln	Ala	Gly	Arg	Asp	Ala
			100					105					110		
Thr	Asp	Ser	Phe	Lys	Val	Phe	His	Ser	Ala	Lys	Ala	Trp	Gln	Phe	Leu
		115					120					125			
Gln	Asp	Leu	Tyr	Ile	Gly	Asp	Leu	Tyr	Asn	Ala	Glu	Pro	Val	Ser	Glu
		130				135					140				
Leu	Val	Lys	Asp	Tyr	Arg	Asp	Leu	Arg	Thr	Ala	Phe	Met	Arg	Ser	Gln
145				150						155					160
Leu	Phe	Lys	Ser	Ser	Lys	Met	Tyr	Tyr	Val	Thr	Lys	Cys	Val	Thr	Asn
				165					170					175	
Phe	Ala	Ile	Leu	Ala	Ala	Ser	Leu	Ala	Val	Ile	Ala	Trp	Ser	Gln	Thr
			180					185					190		
Tyr	Leu	Ala	Val	Leu	Cys	Ser	Ser	Phe	Leu	Leu	Ala	Leu	Phe	Trp	Gln
		195				200						205			
Gln	Cys	Gly	Trp	Leu	Ser	His	Asp	Phe	Leu	His	His	Gln	Val	Thr	Glu
		210				215					220				
Asn	Arg	Ser	Leu	Asn	Thr	Tyr	Phe	Gly	Gly	Leu	Phe	Trp	Gly	Asn	Phe
225				230						235					240
Ala	Gln	Gly	Tyr	Ser	Val	Gly	Trp	Trp	Lys	Thr	Lys	His	Asn	Val	His

[0307]

acg gaa tct gcg atc acc aaa ggt ttg cca tgc gtc gat agc ccg acg Thr Glu Ser Ala Ile Thr Lys Gly Leu Pro Cys Val Asp Ser Pro Thr 30 35 40	325
ccg atc gtt ctt ggg ttg tcg tcc tac ttg aca ttc gtg ttt ctc ggg Pro Ile Val Leu Gly Leu Ser Ser Tyr Leu Thr Phe Val Phe Leu Gly 45 50 55 60	373
ctc att gtc atc aag agc ctg gat ctt aag ccc cgc tcc aag gag ccc Leu Ile Val Ile Lys Ser Leu Asp Leu Lys Pro Arg Ser Lys Glu Pro 65 70 75	421
gcc att ttg aac ctg ttt gtg atc ttc cac aac ttc gtc tgc ttc gca Ala Ile Leu Asn Leu Phe Val Ile Phe His Asn Phe Val Cys Phe Ala 80 85 90	469
ctc agt ctg tac atg tgc gtg gga att gtc cgt caa gct atc ctc aac Leu Ser Leu Tyr Met Cys Val Gly Ile Val Arg Gln Ala Ile Leu Asn 95 100 105	517
agg tac tct ctg tgg ggc aat gcg tac aat ccc aaa gaa gtt caa atg Arg Tyr Ser Leu Trp Gly Asn Ala Tyr Asn Pro Lys Glu Val Gln Met 110 115 120	565
ggc cac ctg ctc tac att ttc tac atg tca aag tac atc gag ttt atg Gly His Leu Leu Tyr Ile Phe Tyr Met Ser Lys Tyr Ile Glu Phe Met 125 130 135 140	613
gac acg gtc att atg att ttg aag cgc aac acg cgc cag atc act gtg Asp Thr Val Ile Met Ile Leu Lys Arg Asn Thr Arg Gln Ile Thr Val 145 150 155	661
ttg cat gtg tac cac cac gca tcc atc tcc ttc atc tgg tgg atc atc Leu His Val Tyr His His Ala Ser Ile Ser Phe Ile Trp Trp Ile Ile 160 165 170	709
gcc tac cat gct cct ggc ggt gaa gct tat ttc tct gcc gca ttg aac Ala Tyr His Ala Pro Gly Gly Glu Ala Tyr Phe Ser Ala Ala Leu Asn 175 180 185	757
tcc gga gta cat gtg ctc atg tac ctc tac tac ctt ttg gca gca act Ser Gly Val His Val Leu Met Tyr Leu Tyr Tyr Leu Leu Ala Ala Thr 190 195 200	805
ctg gga aag aac gag aaa gct cgc cgc aag tac cta tgg tgg gga aaa Leu Gly Lys Asn Glu Lys Ala Arg Arg Lys Tyr Leu Trp Trp Gly Lys 205 210 215 220	853
tac ttg aca cag ctg cag atg ttc cag ttt gtc ctt aac atg att cag Tyr Leu Thr Gln Leu Gln Met Phe Gln Phe Val Leu Asn Met Ile Gln 225 230 235	901
gct tac tac gat att aag aac aac tcg cct tac cca caa ttt ttg atc Ala Tyr Tyr Asp Ile Lys Asn Asn Ser Pro Tyr Pro Gln Phe Leu Ile 240 245 250	949
cag att ttg ttc tac tac atg atc tcg ctt tta gcg cta ttt gga aac	997

[0309]

Gln Ile Leu Phe Tyr Tyr Met Ile Ser Leu Leu Ala Leu Phe Gly Asn
 255 260 265

ttt tac gtt cac aaa tac gta tca gcg ccc gca aaa cct gcg aag atc 1045
 Phe Tyr Val His Lys Tyr Val Ser Ala Pro Ala Lys Pro Ala Lys Ile
 270 275 280

aag agc aaa aag gca gaa taa gacaccacc tagtgacaag aagattttac 1096
 Lys Ser Lys Lys Ala Glu
 285 290

actaaactgt agtttttagca cccatcggtg acacgaatac attctgggtc tgctgtctt 1156

ggaagagtcg aagcattcag gagctctccc gttccatcga tcaaactcgg aacgaagtgc 1216

accttttagc tgcgatgaga gtctttactt cctgagccgt cgttcttgat gtggctctgta 1276

gctcagccat acgtgtagca tagctggaac atctggcttt tcaggaaagt cggcaaggca 1336

agaattcgac ccttgaacta gacaagggtc tgctgattca gcaaccatta gtgagtcact 1396

ggttaacaaa atcacagttt tgggccctta gttagtgaca accaacccta acactttgat 1456

acacagagta tcgttcgca gtggaagtgt aaaaatgtgc tttccaatc atcttgagtt 1516

ggttcctttt gaagtaaagg aaaattctat gattgttgag tccaaaaaaa aaaaaaaaaa 1576

a 1577

<210> 4

<211> 290

<212> PRT

<213> Marchantia polymorpha

<400> 4

Met Glu Ala Tyr Glu Met Val Asp Ser Phe Val Ser Lys Thr Val Phe
 1 5 10 15
 Glu Thr Leu Gln Arg Leu Arg Gly Gly Val Val Leu Thr Glu Ser Ala
 20 25 30
 Ile Thr Lys Gly Leu Pro Cys Val Asp Ser Pro Thr Pro Ile Val Leu
 35 40 45
 Gly Leu Ser Ser Tyr Leu Thr Phe Val Phe Leu Gly Leu Ile Val Ile
 50 55 60
 Lys Ser Leu Asp Leu Lys Pro Arg Ser Lys Glu Pro Ala Ile Leu Asn
 65 70 75 80
 Leu Phe Val Ile Phe His Asn Phe Val Cys Phe Ala Leu Ser Leu Tyr
 85 90 95
 Met Cys Val Gly Ile Val Arg Gln Ala Ile Leu Asn Arg Tyr Ser Leu
 100 105 110
 Trp Gly Asn Ala Tyr Asn Pro Lys Glu Val Gln Met Gly His Leu Leu
 115 120 125
 Tyr Ile Phe Tyr Met Ser Lys Tyr Ile Glu Phe Met Asp Thr Val Ile
 130 135 140
 Met Ile Leu Lys Arg Asn Thr Arg Gln Ile Thr Val Leu His Val Tyr
 145 150 155 160
 His His Ala Ser Ile Ser Phe Ile Trp Trp Ile Ile Ala Tyr His Ala

[0310]

act tcg tgg gtt aag gtc cat cca ggt gga agt ctc atc ttt gtg aag	602
Thr Ser Trp Val Lys Val His Pro Gly Gly Ser Leu Ile Phe Val Lys	
65 70 75	
gcg gga cag gat tca aca caa ctc ttt gat tct tat cac ccc ctc tat	650
Ala Gly Gln Asp Ser Thr Gln Leu Phe Asp Ser Tyr His Pro Leu Tyr	
80 85 90	
gtc aga aag cta ctt gca cag ttc tgc att ggt gaa ctc caa acg agt	698
Val Arg Lys Leu Leu Ala Gln Phe Cys Ile Gly Glu Leu Gln Thr Ser	
95 100 105	
gcg gga gat gag aag ttc aag tct tca acg ttg gag tat gct ggt gaa	746
Ala Gly Asp Glu Lys Phe Lys Ser Ser Thr Leu Glu Tyr Ala Gly Glu	
110 115 120	
gaa cat gaa gta ttt tac cac act ctc aag cag cgc gtg gaa acg tac	794
Glu His Glu Val Phe Tyr His Thr Leu Lys Gln Arg Val Glu Thr Tyr	
125 130 135 140	
ttc cgc aag cag aag ata aat cct cga tac cat ccg caa atg ctt gtg	842
Phe Arg Lys Gln Lys Ile Asn Pro Arg Tyr His Pro Gln Met Leu Val	
145 150 155	
aag tca gcc gtg atc att gga acc ctt ctt ctc tgt tac tat ttt ggc	890
Lys Ser Ala Val Ile Ile Gly Thr Leu Leu Leu Cys Tyr Tyr Phe Gly	
160 165 170	
ttc ttc tgg tct caa aat gta ctc ctc tcg atg ttt ctg gca agc atc	938
Phe Phe Trp Ser Gln Asn Val Leu Leu Ser Met Phe Leu Ala Ser Ile	
175 180 185	
atg ggg ttc tgc act gcg gag gtg ggc atg tcc atc atg cac gat ggt	986
Met Gly Phe Cys Thr Ala Glu Val Gly Met Ser Ile Met His Asp Gly	
190 195 200	
aac cac gga tcg tac aca caa tct acc ttg ctt ggt tac gtc atg ggc	1034
Asn His Gly Ser Tyr Thr Gln Ser Thr Leu Leu Gly Tyr Val Met Gly	
205 210 215 220	
gcc act ctt gat ctg gtg gga gct agc agt ttc atg tgg agg cag cag	1082
Ala Thr Leu Asp Leu Val Gly Ala Ser Ser Phe Met Trp Arg Gln Gln	
225 230 235	
cat gtg gcc ggg cac cac tcg ttc acc aac atc gac cat tac gat cca	1130
His Val Ala Gly His His Ser Phe Thr Asn Ile Asp His Tyr Asp Pro	
240 245 250	
gac att cgt gtg aag gat cct gat tta cga cgg gtt act tct caa caa	1178
Asp Ile Arg Val Lys Asp Pro Asp Leu Arg Arg Val Thr Ser Gln Gln	
255 260 265	
ccc cga aga tgg ttt cac gag tat cag cat atc tac tta gga gta ctc	1226
Pro Arg Arg Trp Phe His Glu Tyr Gln His Ile Tyr Leu Gly Val Leu	
270 275 280	

[0312]

tat ggc gtt ctt gcc tta aaa agt gtg ttg att gat gat ttc agc gcc	1274
Tyr Gly Val Leu Ala Leu Lys Ser Val Leu Ile Asp Asp Phe Ser Ala	
285 290 295 300	
ttc ttc agt ggt gct atc ggc cca gta aag ata gct caa atg aca cca	1322
Phe Phe Ser Gly Ala Ile Gly Pro Val Lys Ile Ala Gln Met Thr Pro	
305 310 315	
ctc gag atg ggc gtc ttc tgg gga ggg aag gtt gtg tac gca ctg tac	1370
Leu Glu Met Gly Val Phe Trp Gly Gly Lys Val Val Tyr Ala Leu Tyr	
320 325 330	
atg ttt ttg ctc cct atg atg tat ggt caa tac aac att ctt act ttc	1418
Met Phe Leu Leu Pro Met Met Tyr Gly Gln Tyr Asn Ile Leu Thr Phe	
335 340 345	
att ggt ctc tac att ctc tca cag tta gtt gca ggg tgg act ctt gcc	1466
Ile Gly Leu Tyr Ile Leu Ser Gln Leu Val Ala Gly Trp Thr Leu Ala	
350 355 360	
ctc ttc ttt caa gta gca cac gtt gtc gac gat gca gta ttt ccc gtt	1514
Leu Phe Phe Gln Val Ala His Val Val Asp Asp Ala Val Phe Pro Val	
365 370 375 380	
gcg gaa aca gat ggt gga aaa gca aag att cct tct ggt tgg gca gaa	1562
Ala Glu Thr Asp Gly Gly Lys Ala Lys Ile Pro Ser Gly Trp Ala Glu	
385 390 395	
atg cag gtc aga acc act acc aat ttc agc tca cga tca atg ttc tgg	1610
Met Gln Val Arg Thr Thr Thr Asn Phe Ser Ser Arg Ser Met Phe Trp	
400 405 410	
aca cat att agt ggc ggt ctg aac cat cag atc gag cac cat ctt ttc	1658
Thr His Ile Ser Gly Gly Leu Asn His Gln Ile Glu His His Leu Phe	
415 420 425	
ccg ggt gtc tgt cat gtt cac tac cca agc ata cag cca atc gtg aag	1706
Pro Gly Val Cys His Val His Tyr Pro Ser Ile Gln Pro Ile Val Lys	
430 435 440	
gct acc tgt gac gag ttc aac gtg cct tat act tcc tac ccc act ttc	1754
Ala Thr Cys Asp Glu Phe Asn Val Pro Tyr Thr Ser Tyr Pro Thr Phe	
445 450 455 460	
tgg gcg gcc ctt agg gca cat ttt caa cat ctg aaa aac gtc gga cta	1802
Trp Ala Ala Leu Arg Ala His Phe Gln His Leu Lys Asn Val Gly Leu	
465 470 475	
caa gat gga cta cga ctg gat ggc tga actgtgacag catgctttgg	1849
Gln Asp Gly Leu Arg Leu Asp Gly	
480 485	
gcctgcactt tcagatttcg gatcgaaggt gcgggcatg gaaataatca gataagagtt	1909
gtaagtaacg ttcaggagga gagcagaacg gattgatgag tgtccatttg tgaggcttcc	1969
acctttcagg aacagaagtt gattcgaatg cgaaacctcc aatgagcatt tcacagccgt	2029

[0313]

cttctccttg gccatcatgt gttcctccta gggagcttcg gtttttggaa gtttagtcagc 2089
 ttacttttca agatcgttca acgctcaagg ctagattttg tcgacactat ttagttaggt 2149
 ccgatagata ggtgataaga ttccgggtgcc ctacacatg tttcatcagt tgcgatgtaa 2209
 ttccagtaat ccacgtatgt ggctccagt tctgctgaaa tcagcacagg cagctatatac 2269
 atgctccttg atctctaaaa aaaaaaaaaa aaaaaaaaaa 2307

<210> 6

<211> 484

<212> PRT

<213> Marchantia polymorpha

<400> 6

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

[0314]

<223> i
 <220>
 <221> modified_base
 <222> (10)
 <223> i

 <220>
 <221> modified_base
 <222> (13)
 <223> i

 <400> 8
 rttnarncn ccngtraacc a 21

 <210> 9
 <211> 22
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence:Primer

 <400> 9
 aagttgcctt cgatgtttct gg 22

 <210> 10
 <211> 22
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence:Primer

 <400> 10
 gctcgcctgg agcaaggaaa tc 22

 <210> 11
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence:Primer

 <220>
 <221> modified_base
 <222> (3)
 <223> i

 <220>
 <221> modified_base
 <222> (18)
 <223> i

[0316]

<400> 11
gtngarttya tggayacngt 20

<210> 12
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<220>
<221> modified_base
<222> (3)
<223> i

<220>
<221> modified_base
<222> (12)
<223> i

<400> 12
cknccccara anarrtaytt 20

<210> 13
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 13
gcgagctttc tcgttctttc cc 22

<210> 14
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 14
tatgattttg aagcgcaaca cg 22

<210> 15
<211> 22
<212> DNA
<213> Artificial Sequence

<220>

[0317]

<223> Description of Artificial Sequence:Primer

<220>

<221> modified_base

<222> (6)

<223> i

<220>

<221> modified_base

<222> (9)

<223> i

<220>

<221> modified_base

<222> (14)

<223> i

<400> 15

athrangrna artntaygay gt

22

<210> 16

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Primer

<220>

<221> modified_base

<222> (3)

<223> i

<220>

<221> modified_base

<222> (6)

<223> i

<220>

<221> modified_base

<222> (9)

<223> i

<220>

<221> modified_base

<222> (18)

<223> i

<400> 16

ggnaynkwnt sdatrtcngg rtc

23

<210> 17

<211> 22

<212> DNA

<213> Artificial Sequence

[0318]

<220>
<223> Description of Artificial Sequence:Primer

<400> 17
gtgtgtacga tccgtggta cc 22

<210> 18
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 18
aaggcgggac aggattcaac ac 22

<210> 19
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 19
ggaattcgcg atggcctcgt ccaccaccac 30

<210> 20
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 20
ggaattctac tttcgagcg tatgctacc 29

<210> 21
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 21
ggaattcgcg atggaggcgt acgagatgg 29

<210> 22

[0319]

<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 22
ggaattcttc tgcctttttg ctcttgatc 29

<210> 23
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 23
gttgaattcg acagttatgc cgccacacgc 30

<210> 24
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 24
gttgaattca ggcccaaagc atgctgtcac 30

<210> 25
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 25
cgggatcctc tcctggcgca ccacgctc 28

<210> 26
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 26
gggtaccaa cgcgctttcc caccaacg 28

[0320]

<210> 27
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 27
gctctagagc gatggcctcg tccaccacc 29

<210> 28
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 28
gctctagact atactttcgc agcgtatgc 29

<210> 29
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 29
gctctagagc gatggaggcg tacgagatgg 30

<210> 30
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

<400> 30
gctctagatt attctgcctt tttgctc 27

<210> 31
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Primer

[0321]

<400> 31 gctctagaga cagttatgcc gccacacgc	29
<210> 32 <211> 29 <212> DNA <213> Artificial Sequence	
<220> <223> Description of Artificial Sequence:Primer	
<400> 32 gctctagaag gcccaaagca tgctgtcac	29
<210> 33 <211> 18 <212> DNA <213> Artificial Sequence	
<220> <223> Description of Artificial Sequence:Primer	
<400> 33 caggaaacag ctatgacc	18
<210> 34 <211> 30 <212> DNA <213> Artificial Sequence	
<220> <223> Description of Artificial Sequence:Primer	
<400> 34 aaactgcaga ttcccgatct agtaacatag	30
<210> 35 <211> 30 <212> DNA <213> Artificial Sequence	
<220> <223> Description of Artificial Sequence:Primer	
<400> 35 ccggaattcg catgcctgca ggtccccaga	30
<210> 36 <211> 18 <212> DNA <213> Artificial Sequence	

[0322]

<220>
<223> Description of Artificial Sequence:Primer

<400> 36
tgtaaaacga cggccagt 18

<210> 37
<211> 7
<212> PRT
<213> Unknown Organism

<220>
<221> Xaa
<222> (4)
<223> Description of Unknown Organism:conserved amino acid sequence

<400> 37
Trp Trp Lys Xaa Lys His Asn
1 5

<210> 38
<211> 7
<212> PRT
<213> Unknown Organism

<220>
<223> Description of Unknown Organism:conserved amino acid sequence

<400> 38
Trp Phe Thr Gly Gly Leu Asn
1 5

<210> 39
<211> 7
<212> PRT
<213> Unknown Organism

<220>
<223> Description of Unknown Organism:conserved amino acid sequence

<400> 39
Val Glu Phe Met Asp Thr Val
1 5

<210> 40
<211> 7
<212> PRT
<213> Unknown Organism

[0323]

<220>
<223> Description of Unknown Organism:conserved amino
acid sequence

<400> 40
Lys Tyr Leu Phe Trp Gly Arg
1 5

<210> 41
<211> 8
<212> PRT
<213> Unknown Organism

<220>
<223> Description of Unknown Organism:conserved amino
acid sequence

<220>
<221> Xaa
<222> (2)
<223> Glu or Asn

<220>
<221> Xaa
<222> (3)
<223> Gly or Asp

<400> 41
Ile Xaa Xaa Lys Val Tyr Asp Val
1 5

<210> 42
<211> 8
<212> PRT
<213> Unknown Organism

<220>
<223> Description of Unknown Organism:conserved amino
acid sequence

<220>
<221> Xaa
<222> (5)
<223> Gln or Asp

<220>
<221> Xaa
<222> (6)
<223> Tyr or Thr

<220>
<221> Xaa
<222> (7)
<223> Met or Val

[0324]

<400> 42
Asp Pro Asp Ile Xaa Xaa Xaa Pro
1 5

<210> 43
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer
XbaMpELOf

<400> 43
agtctctaga gcgatggagg cgtacg 26

<210> 44
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer
SacMpELOr

<400> 44
cagtgagctc ggtgtcttat tctgcc 26

<210> 45
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer XbaMpD5f

<400> 45
agcttctaga gccatgccgc cacacgccc 29

<210> 46
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer SacMpD5r

<400> 46
cagtgagctc tcagccatcc agtcgt 26

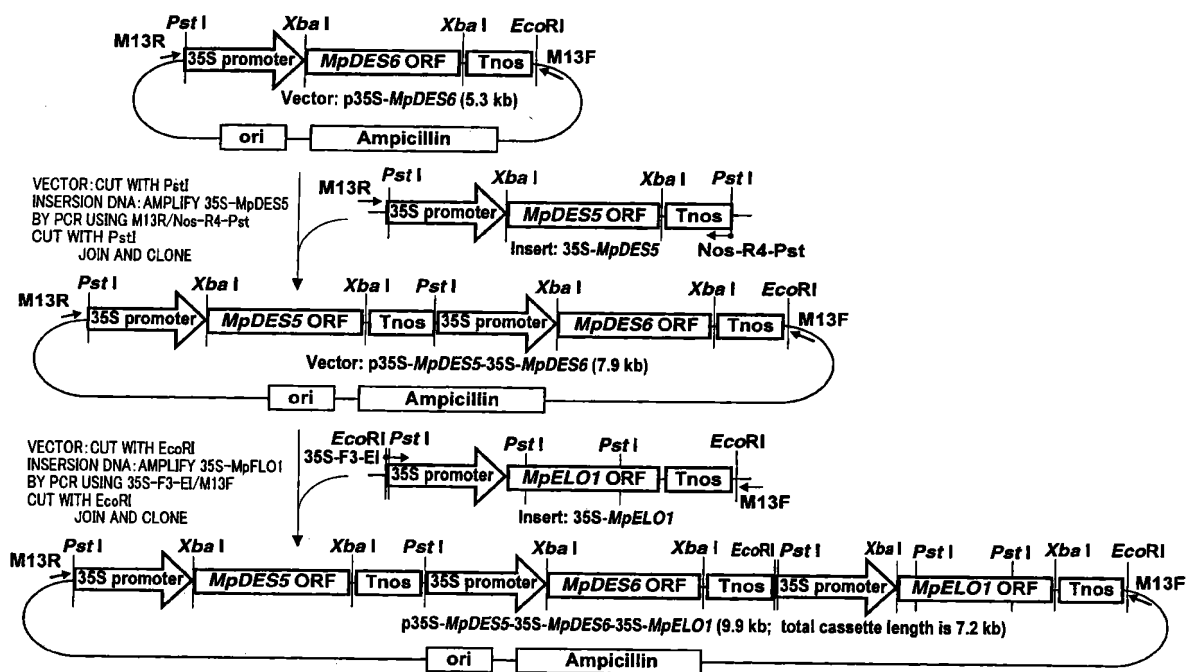


图 1