



(12) 发明专利

(10) 授权公告号 CN 107858204 B

(45) 授权公告日 2021.10.29

(21) 申请号 201710929893.1

(22) 申请日 2009.11.17

(65) 同一申请的已公布的文献号

申请公布号 CN 107858204 A

(43) 申请公布日 2018.03.30

(30) 优先权数据

61/199,669 2008.11.18 US

61/270,710 2009.07.09 US

(62) 分案原申请数据

200980154876.9 2009.11.17

(73) 专利权人 联邦科学技术研究组织

地址 澳大利亚澳大利亚首都地区

专利权人 粮食研究发展公司

(72) 发明人 J·R·皮特里 A·M·麦肯齐

柳青 P·什雷斯塔

P·D·尼科尔斯

S·I·E·布莱克本

M·P·曼苏尔 S·S·罗伯特

D·M·F·弗兰普顿 周雪荣

S·P·辛格 C·C·伍德

(74) 专利代理机构 永新专利商标代理有限公司

72002

代理人 左路 林晓红

(51) Int.Cl.

C12N 9/00 (2006.01)

C12N 15/52 (2006.01)

C12N 5/10 (2006.01)

C12P 7/64 (2006.01)

(56) 对比文件

CN 101018862 A, 2007.08.15

WO 2007096387 A1, 2007.08.30

审查员 赵建民

权利要求书3页 说明书118页

序列表63页 附图13页

(54) 发明名称

产生 ω -3脂肪酸的酶和方法

(57) 摘要

本发明涉及在重组细胞(例如酵母或植物细胞)中合成长链多不饱和脂肪酸(特别是二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸)的方法。本发明还提供产生长链多不饱和脂肪酸的重组细胞或植物。此外,本发明涉及一组新的酶,它们具有去饱和酶或延长酶活性,可用于合成长链多不饱和脂肪酸的方法。具体地,本发明提供具有新活性的 ω 3去饱和酶、 Δ 5延长酶和 Δ 6去饱和酶。本发明还提供用于以多种基因瞬时和/或稳定转化细胞(特别是植物细胞)的方法和DNA构建体。

1. 从包含二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸的种子提取油的方法, 所述方法包括以下步骤:

(i) 获得欧洲油菜 (*Brassica napus*) 种子, 所述种子包含编码微藻 $\Delta 6$ 去饱和酶、 $\Delta 6$ 延长酶、 $\Delta 5$ 去饱和酶、微藻 $\Delta 5$ 延长酶和 $\Delta 4$ 去饱和酶的外源性多核苷酸, 其中各多核苷酸与一或多个在种子发育期间指导所述多核苷酸表达的启动子可操作连接, 以产生包含二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸的种子, 所述二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸在所述种子发育期间产生自 α -亚麻酸 (ALA), 并且以酯化形式作为甘油三酯的部分以基于将 α -亚麻酸转换为二十二碳五烯酸或二十二碳六烯酸的效率是至少 15.4% 的水平存在, 其中所述微藻 $\Delta 6$ 去饱和酶对作为脂肪酸底物的 ALA-CoA 的 $\Delta 6$ 去饱和酶活性高于对作为脂肪酸底物的连接于卵磷脂的 sn-2 位置的 ALA 的 $\Delta 6$ 去饱和酶活性, 其中所述 $\Delta 4$ 去饱和酶对酰基-CoA 底物的活性高于对相应的酰基-卵磷脂底物的活性, 其中所述 $\Delta 6$ 去饱和酶是 *Ostreococcus tauri* 或细小微胞藻 (*Micromonas pusilla*) $\Delta 6$ -去饱和酶, 且其中所述 $\Delta 4$ 去饱和酶是路氏巴夫藻 (*Pavlova lutheri*) 或盐生巴夫藻 (*Pavlova salina*) $\Delta 4$ 去饱和酶, 和

(ii) 从所述种子提取油。

2. 权利要求 1 的方法, 其中所述 $\Delta 5$ 去饱和酶是微藻 $\Delta 5$ 去饱和酶。

3. 权利要求 1 的方法, 其中花生四烯酸、二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸总计包含种子内总脂肪酸的 15%。

4. 权利要求 1 的方法, 其中所述 $\Delta 5$ 延长酶具有由登记号 AAV67798 所提供的氨基酸序列。

5. 权利要求 1 的方法, 其中所述 $\Delta 4$ 去饱和酶具有由登记号 AY332747 所提供的氨基酸序列。

6. 权利要求 1 的方法, 其中各外源性多核苷酸与种子特异性启动子可操作连接。

7. 权利要求 1 的方法, 其包括纯化所提取的脂质的步骤。

8. 权利要求 1 的方法, 其中 $\Delta 6$ 延长酶具有由登记号 AF428243 所提供的氨基酸序列。

9. 产生包含二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸的欧洲油菜种子的方法, 所述方法包括

i) 向欧洲油菜细胞中引入编码微藻 $\Delta 6$ 去饱和酶、 $\Delta 6$ 延长酶、 $\Delta 5$ 去饱和酶、微藻 $\Delta 5$ 延长酶和 $\Delta 4$ 去饱和酶的外源性多核苷酸, 其中各多核苷酸与一或多个在欧洲油菜种子发育期间指导所述多核苷酸表达的启动子可操作连接, 其中所述微藻 $\Delta 6$ 去饱和酶对作为脂肪酸底物的 ALA-CoA 的 $\Delta 6$ 去饱和酶活性高于对作为脂肪酸底物的连接于卵磷脂的 sn-2 位置的 ALA 的 $\Delta 6$ 去饱和酶活性, 其中所述 $\Delta 4$ 去饱和酶对酰基-CoA 底物的活性高于对相应的酰基-卵磷脂底物的活性, 其中所述 $\Delta 6$ 去饱和酶是 *Ostreococcus tauri* 或细小微胞藻 $\Delta 6$ -去饱和酶, 且其中所述 $\Delta 4$ 去饱和酶是路氏巴夫藻或盐生巴夫藻 $\Delta 4$ 去饱和酶, 和

(ii) 产生包含种子的转基因植物, 所述种子包含二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸, 所述二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸在所述种子发育期间产生自 α -亚麻酸 (ALA), 并且以酯化形式作为甘油三酯的部分以基于将 α -亚麻酸转换为二十二碳五烯酸或二十二碳六烯酸的效率是至少 15.4% 的水平存在。

10. 权利要求 9 的方法, 其进一步包括产生所述转基因植物的祖先。

11. 获得包含二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸的遗传学修饰的欧洲油菜细胞的方法,所述方法包括:

a) 向欧洲油菜细胞中引入编码微藻 $\Delta 6$ 去饱和酶、 $\Delta 6$ 延长酶、 $\Delta 5$ 去饱和酶、微藻 $\Delta 5$ 延长酶和 $\Delta 4$ 去饱和酶的外源性多核苷酸,其中各多核苷酸与一或多个在欧洲油菜种子发育期间指导所述多核苷酸表达的启动子可操作连接,其中所述微藻 $\Delta 6$ 去饱和酶对作为脂肪酸底物的ALA-CoA的 $\Delta 6$ 去饱和酶活性高于对作为脂肪酸底物的连接于卵磷脂的sn-2位置的ALA的 $\Delta 6$ 去饱和酶活性,其中所述 $\Delta 4$ 去饱和酶对酰基-CoA底物的活性高于对相应的酰基-卵磷脂底物的活性,其中所述 $\Delta 6$ 去饱和酶是*Ostreococcus tauri*或细小微胞藻 $\Delta 6$ -去饱和酶,且其中所述 $\Delta 4$ 去饱和酶是路氏巴夫藻或盐生巴夫藻 $\Delta 4$ 去饱和酶,

b) 在所述细胞或其子代细胞中表达所述外源性多核苷酸,

c) 分析所述细胞或其子代细胞的脂肪酸组成,以及

d) 选择以至少15.4%的效率将 α -亚麻酸转化为二十二碳六烯酸的细胞。

12. 权利要求11的方法,其中所述外源性多核苷酸稳定整合入所述细胞的基因组中。

13. 来自欧洲油菜种子的长链多不饱和脂肪酸脂质提取物,所述种子包含编码微藻 $\Delta 6$ 去饱和酶、 $\Delta 6$ 延长酶、 $\Delta 5$ 去饱和酶、微藻 $\Delta 5$ 延长酶和 $\Delta 4$ 去饱和酶的外源性多核苷酸,其中各多核苷酸与一或多个在种子发育期间指导所述多核苷酸表达的启动子可操作连接,从而所述种子包含包含二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸,所述二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸在所述种子发育期间产生自 α -亚麻酸(ALA),并且以酯化形式作为甘油三酯的部分以基于将 α -亚麻酸转换为二十二碳五烯酸或二十二碳六烯酸的效率是至少15.4%的水平存在,其中所述微藻 $\Delta 6$ 去饱和酶对作为脂肪酸底物的ALA-CoA的 $\Delta 6$ 去饱和酶活性高于对作为脂肪酸底物的连接于卵磷脂的sn-2位置的ALA的 $\Delta 6$ 去饱和酶活性,其中所述 $\Delta 4$ 去饱和酶对酰基-CoA底物的活性高于对相应的酰基-卵磷脂底物的活性,其中所述 $\Delta 6$ 去饱和酶是*Ostreococcus tauri*或细小微胞藻 $\Delta 6$ -去饱和酶,其中所述 $\Delta 4$ 去饱和酶是路氏巴夫藻或盐生巴夫藻 $\Delta 4$ 去饱和酶,且其中所述提取物包含二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸。

14. 饲料,包含压碎的欧洲油菜种子或权利要求13的提取物,所述种子包含编码微藻 $\Delta 6$ 去饱和酶、 $\Delta 6$ 延长酶、 $\Delta 5$ 去饱和酶、微藻 $\Delta 5$ 延长酶和 $\Delta 4$ 去饱和酶的外源性多核苷酸,其中各多核苷酸与一或多个在种子发育期间指导所述多核苷酸表达的启动子可操作连接,从而所述种子包含包含二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸,所述二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸在所述种子发育期间产生自 α -亚麻酸(ALA),并且以酯化形式作为甘油三酯的部分以基于将 α -亚麻酸转换为二十二碳五烯酸或二十二碳六烯酸的效率是至少15.4%的水平存在,其中所述微藻 $\Delta 6$ 去饱和酶对作为脂肪酸底物的ALA-CoA的 $\Delta 6$ 去饱和酶活性高于对作为脂肪酸底物的连接于卵磷脂的sn-2位置的ALA的 $\Delta 6$ 去饱和酶活性,其中所述 $\Delta 4$ 去饱和酶对酰基-CoA底物的活性高于对相应的酰基-卵磷脂底物的活性,其中所述 $\Delta 6$ 去饱和酶是*Ostreococcus tauri*或细小微胞藻 $\Delta 6$ -去饱和酶,其中所述 $\Delta 4$ 去饱和酶是路氏巴夫藻或盐生巴夫藻 $\Delta 4$ 去饱和酶,其中所述压碎的欧洲油菜种子不能萌发,并且其中所述饲料不是植物品种。

15. 制备饲料的方法,所述方法包括将欧洲油菜种子或压碎的欧洲油菜种子或权利要求13的提取物与合适的营养物和/或载体混合,所述种子包含编码微藻 $\Delta 6$ 去饱和酶、 $\Delta 6$ 延

长酶、 $\Delta 5$ 去饱和酶、微藻 $\Delta 5$ 延长酶和 $\Delta 4$ 去饱和酶的的外源性多核苷酸,其中各多核苷酸与一或多个在种子发育期间指导所述多核苷酸表达的启动子可操作连接,从而所述种子包含包含二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸,所述二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸在所述种子发育期间产生自 α -亚麻酸(ALA),并且以酯化形式作为甘油三酯的部分以基于将 α -亚麻酸转换为二十二碳五烯酸或二十二碳六烯酸的效率是至少15.4%的水平存在,其中所述微藻 $\Delta 6$ 去饱和酶对作为脂肪酸底物的ALA-CoA的 $\Delta 6$ 去饱和酶活性高于对作为脂肪酸底物的连接于卵磷脂的sn-2位置的ALA的 $\Delta 6$ 去饱和酶活性,其中所述 $\Delta 4$ 去饱和酶对酰基-CoA底物的活性高于对相应的酰基-卵磷脂底物的活性,其中所述 $\Delta 6$ 去饱和酶是*Ostreococcus tauri*或细小微胞藻 $\Delta 6$ -去饱和酶,其中所述 $\Delta 4$ 去饱和酶是路氏巴夫藻或盐生巴夫藻 $\Delta 4$ 去饱和酶。

产生 ω -3 脂肪酸的酶和方法

[0001] 本申请是申请日为2009年11月17日,申请号为200980154876.9,标题为“产生 ω -3 脂肪酸的酶和方法”的专利申请的分案申请。

技术领域

[0002] 本发明涉及在重组细胞(例如酵母或植物细胞)中合成长链多不饱和脂肪酸(特别是二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸)的方法。本发明还提供产生长链多不饱和脂肪酸的重组细胞或植物。此外,本发明涉及一组新的酶,它们具有去饱和酶或延长酶活性,可用于合成长链多不饱和脂肪酸的方法。具体地,本发明提供具有新活性的 ω 3 去饱和酶、 Δ 5 延长酶和 Δ 6 去饱和酶。本发明还提供用于以多种基因瞬时和/或稳定转化细胞(特别是植物细胞)的方法和DNA构建体。

背景技术

[0003] 现在广泛认为 ω -3 长链多不饱和脂肪酸(LC-PUFA和VLC-PUFA)是关乎人和动物健康的重要化合物。可从膳食来源或通过转换亚油酸(LA, 18:2 ω 6)或 α -亚麻酸(ALA, 18:3 ω 3)而获得这些脂肪酸,LA和ALA被认为是人类膳食中的必需脂肪酸。尽管人和许多其他脊椎动物能够将植物来源的LA或ALA转换为VLC-PUFA,但这种转换的速率很慢。此外,大多数现代社会的饮食不平衡,其中至少90%的多不饱和脂肪酸(PUFA)是 ω 6 脂肪酸,而非理想的4:1或更低的 ω 6: ω 3 脂肪酸的比例(Trautwein, 2001)。对于人类,VLC-PUFA例如二十碳五烯酸(EPA, 20:5 ω 3)和二十二碳六烯酸(DHA, 22:6 ω 3)的直接来源主要是来自鱼类或鱼油。因此,健康专家们推荐将含有高水平VLC-PUFA的鱼类常规纳入人类饮食。例如,来源于鱼类的VLC-PUFA油被越来越多地添加到食品和婴儿配方奶中。不过,鉴于全球和国家鱼业的下滑,需要找到这些有益健康的油类的替代来源。

[0004] 与动物不同,高等植物不能合成链长超过18个碳原子的多不饱和脂肪酸。具体地,农作物和园艺植物以及其他被子植物不具有合成长链 ω 3 脂肪酸例如衍生自ALA的EPA、二十碳五烯酸(DPA, 22:5 ω 3)和DHA所需的酶。因此,植物生物技术的一个重要目的是通过遗传工程学方法改造出产生大量VLC-PUFA的农作物,由此提供这些化合物的替代来源。

[0005] VLC-PUFA生物合成途径

[0006] 生物体(例如微藻类、苔藓和真菌)中VLC-PUFA的生物合成通常是一系列的依赖氧的去饱和作用和延长反应(图1)。这些生物体中产生EPA的最常见的途径包括 Δ 6-去饱和作用、 Δ 6-延长作用和 Δ 5-去饱和作用(称为 Δ 6-去饱和作用途径),较不常见的途径则使用 Δ 9-延长作用、 Δ 8-去饱和作用和 Δ 5-去饱和作用(称为 Δ 9-去饱和作用途径)。这些连续的去饱和作用和延长反应可以从 ω 6 脂肪酸底物LA开始,见图1左上部分(ω 6),或从 ω 3 底物ALA开始,见图1右下部分(ω 3)。如果起始的 Δ 6-去饱和作用在 ω 6 底物LA上进行,该系类三种酶的VLC-PUFA产物会是 ω 6 脂肪酸ARA。合成VLC-PUFA的生物体可利用 ω 3-去饱和酶将 ω 6 脂肪酸转换为 ω 3 脂肪酸,如图1所示的将花生四烯酸(ARA, 20:4 ω 6)转换为EPA的 Δ 17-去饱和酶步骤。 ω 3-去饱和酶家族的一些成员可作用于从LA到ARA的大量底物。植物 ω 3-去

饱和酶通常特异性催化由LA到ALA的 $\Delta 15$ -去饱和作用,而真菌和酵母 $\omega 3$ -去饱和酶对由ARA到EPA的 $\Delta 17$ -去饱和作用具有特异性(Pereira等,2004a;Zank等,2005)。一些报道提示可能存在非特异性的 $\omega 3$ -去饱和酶,可将多种 $\omega 6$ 底物转换为相应的 $\omega 3$ 产物(Zhang等,2007)。其他 $\omega 3$ -去饱和酶可对 $\omega 3$ 底物具有偏向性(Sayanova等,2003)。

[0007] 在这些生物体中,EPA转换为DHA相对简单,包括通过 $\Delta 5$ -延长作用由EPA产生DPA,随后通过 $\Delta 4$ -去饱和作用产生DHA(图1)。相反,哺乳动物则利用所谓的“Sprecher”途径,通过不依赖于 $\Delta 4$ 去饱和酶的3个分开的反应将DPA转换为DHA(Sprecher等,1995)。

[0008] 前端去饱和酶通常存在于植物、苔藓、微藻类和低等动物(例如秀丽隐杆线虫(*Caenorhabditis elegans*))中,主要接受与磷脂酰胆碱(PC)底物的sn-2位置发生酯化的脂肪酸底物。这些去饱和酶因此称为酰基-PC脂质连接的前端去饱和酶(Domergue等,2003)。相反,高等动物前端去饱和酶通常接受酰基辅酶A底物,其中脂肪酸底物连接于CoA而非PC(Domergue等,2005)。

[0009] PUFA和VLC-PUFA延长反应各自包括由多组分蛋白复合体催化的4个步骤:首先,缩合反应给脂肪酸加上来自丙二酰-CoA的2C单元,导致形成 β -酮酰基中间体,然后通过NADPH将其还原,随后脱水产生烯醇基中间体,该中间体最后被再次还原,产生延长的脂肪酸。通常认为这4个反应中的缩合步骤具有底物特异性,而其他步骤则不具有。在实践中,这意味着,只要引入对VLC-PUFA具有特异性的缩合酶(典型地称为“延长酶”),植物中的天然延长机制便能够延长VLC-PUFA,只不过植物中的天然延长机制延长非天然VLC-PUFA底物的效率可能是低的。2007年鉴定并表征了酵母延长作用循环脱水酶(Denic and Weissman,2007)。

[0010] 在植物、苔藓和微藻类中,VLC-PUFA去饱和作用天然发生在主要属于酰基-PC库的脂肪酸底物,而延长作用则发生在酰基-CoA库的底物。脂肪酸从酰基-PC分子转移至CoA载体是由磷脂酶类(PLA)进行的,而酰基-CoA脂肪酸转移至PC载体是由溶血卵磷脂酰基转移酶类(LPCAT)进行的(图2)(Singh等,2005)。在去饱和作用可以跟上延长作用之前必需发生酰基交换(反之亦然),由此引起的流量下降可通过使用对酰基辅酶A底物具有特异性的去饱和酶而克服(Hoffmann等,2008)。

[0011] 遗传工程化产生VLC-PUFA

[0012] 绝大多数VLC-PUFA代谢工程一直采用需氧 $\Delta 6$ -去饱和作用/延长作用途径进行。1996年首次报道了使用来自蓝细菌集胞藻(*Synechocystis*)的 $\Delta 6$ -去饱和酶在烟草中生物合成 γ -亚麻酸(GLA,18:3 $\omega 6$)(Reddy and Thomas,1996)。最近,已经在农作物例如向日葵(73%GLA;Knauf等,2006)和大豆(28%GLA;Sato等,2004)中产生GLA。由于涉及更多的去饱和作用和延长作用步骤,因此产生VLC-PUFA例如EPA和DHA涉及更复杂的工程化。Qi等(2004)最早报道了在陆生植物中产生EPA,其将编码来自球等鞭金藻(*Isochrysis galbana*)的 $\Delta 9$ -延长酶的基因、编码来自小眼虫(*Euglena gracilis*)的 $\Delta 8$ -去饱和酶的基因、和编码来自高山被孢霉(*Mortierella alpina*)的 $\Delta 5$ -去饱和酶的基因引入拟南芥,产生高达3%的EPA。随后Abbadì等(2004)报道了使用编码来自小立碗藓(*Physcomitrella patens*)的 $\Delta 6$ -去饱和酶和 $\Delta 6$ -延长酶的基因和编码来自三角褐指藻(*Phaeodactylum tricornutum*)的 $\Delta 5$ -去饱和酶的基因在亚麻子中产生了高达0.8%的EPA。

[0013] WO 04/017467首次报道了产生DHA,并迄今为止其报道的产生VLC-PUFA的水平是最高的,其描述通过引入编码异丝水霉(*Saprolegnia diclina*) $\Delta 6$ -去饱和酶、高山被孢霉

$\Delta 6$ -去饱和酶、高山被孢霉 $\Delta 5$ -去饱和酶、异丝水霉 $\Delta 4$ -去饱和酶、异丝水霉 $\Delta 17$ -去饱和酶、高山被孢霉 $\Delta 6$ -延长酶和路氏巴夫藻 (*Pavlova lutheri*) $\Delta 5$ -延长酶的基因,在大豆胚(而非种子)中产生了3%的DHA。在也产生DHA的胚中,最高EPA水平是19.6%,说明EPA转换为DHA的效率低下(WO 2004/071467)。这一发现与Robert等(2005)发表的结果类似,后者的从EPA到DHA的流量低下,其使用鲇(*Danio rerio*) $\Delta 5/6$ -去饱和酶、秀丽隐杆线虫 $\Delta 6$ -延长酶、和盐生巴夫藻 (*Pavlova salina*) $\Delta 5$ -延长酶和 $\Delta 4$ -去饱和酶,在拟南芥中产生了3% EPA和0.5%DHA。也在2005,Wu et al.报道使用畸雌腐霉 (*Pythium irregulare*) $\Delta 6$ -去饱和酶、破囊壶菌 (*Thraustochytrid*) $\Delta 5$ -去饱和酶、小立碗藓 $\Delta 6$ -延长酶、金盏花 (*Calendula officianalis*) $\Delta 12$ -去饱和酶、破囊壶菌 $\Delta 5$ -延长酶、致病疫霉 (*Phytophthora infestans*) $\Delta 17$ -去饱和酶、虹鳟 (*Oncorhynchus mykiss*) VLC-PUFA延长酶、破囊壶菌 $\Delta 4$ -去饱和酶和破囊壶菌LPCAT,在芥菜 (*Brassica juncea*) 中产生了25% ARA、15%EPA和1.5%DHA (Wu等,2005)。

[0014] 因此仍需要在重组细胞中,特别是在油料种子植物的种子中更加高效地产生LC-PUFA。

发明内容

[0015] 本发明人首次鉴定到了能够在重组细胞中将EPA高效转换为DPA的 $\Delta 5$ 延长酶。

[0016] 因此,本发明还提供重组细胞,优选地是植物细胞,更优选地是植物种子细胞,所述细胞包含编码具有 $\Delta 5$ 延长酶活性的脂肪酸延长酶的外源性多核苷酸,其中当所述延长酶由所述细胞、优选植物细胞中的所述外源性多核苷酸表达时,所述延长酶对EPA具有活性,以至少60%、至少65%、至少70%或至少75%的效率产生DPA。

[0017] 在一个实施方式中,所述延长酶包含SEQ ID NO:6所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:6具有至少47%相同性的氨基酸序列。

[0018] 在另一个实施方式中,所述细胞还包含编码以下各项的外源性多核苷酸:

[0019] i) $\Delta 8$ 去饱和酶和/或 $\Delta 6$ 去饱和酶,

[0020] ii) $\Delta 9$ 延长酶和/或 $\Delta 6$ 延长酶,

[0021] iii) $\Delta 5$ 去饱和酶,和

[0022] iv) 任选存在的 $\Delta 4$ 去饱和酶和/或 $\omega 3$ 去饱和酶,

[0023] 其中每一多核苷酸可操纵地连接于一或多种能够在所述细胞中指导所述多核苷酸表达的启动子。

[0024] 本发明人还鉴定到了具有新特性的 $\omega 3$ 去饱和酶。所述 $\omega 3$ 去饱和酶可用于为产生EPA、下游脂肪酸DPA和DHA、以及其他 $\omega 3$ VLC-PUFA而设计的重组途径中。

[0025] 因此,本发明提供重组细胞,优选地是植物细胞,更优选地是植物种子细胞,所述细胞包含编码具有 $\omega 3$ 去饱和酶活性的脂肪酸去饱和酶的外源性多核苷酸,其中当所述去饱和酶由所述细胞中的所述外源性多核苷酸表达时,所述去饱和酶能够进行以下至少一种去饱和作用:将ARA去饱和化为EPA、将DGLA去饱和化为ETA、将GLA去饱和化为SDA、将ARA去饱和化为EPA并将DGLA去饱和化为ETA、将ARA去饱和化为EPA并将GLA去饱和化为SDA、或所有这三种去饱和作用。

[0026] 所述去饱和酶优选地是前端去饱和酶。

[0027] 在另一个实施方式中,所述去饱和酶对酰基链中具有至少三个碳碳双键的C20脂肪酸、优选ARA具有 $\Delta 17$ 去饱和酶活性。

[0028] 在另一个实施方式中,所述去饱和酶对酰基链中具有三个碳碳双键的C18脂肪酸、优选GLA具有 $\Delta 15$ 去饱和酶活性。

[0029] 所述去饱和酶优选地对酰基辅酶A底物的活性高于对相应的酰基-PC底物的活性。

[0030] 在一个实施方式中,所述酰基辅酶A底物是ARA-CoA,而所述酰基-PC底物包含位于的PC的sn-2位置的ARA。

[0031] 在另一个实施方式中,所述细胞是植物细胞,且当所述去饱和酶由所述细胞中的所述外源性多核苷酸表达时,所述去饱和酶对ARA具有活性,以至少40%的效率产生EPA。

[0032] 在一个具体的实施方式中,所述去饱和酶包含SEQ ID NO:15、17或20所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:15具有至少35%相同性的氨基酸序列、与SEQ ID NO:17具有至少60%相同性的氨基酸序列和/或与SEQ ID NO:20具有至少60%相同性的氨基酸序列。

[0033] 此外,本发明人鉴定到了编码一种 $\Delta 6$ 去饱和酶的基因,所述 $\Delta 6$ 去饱和酶在植物和/或酵母中对 $\omega 3$ 脂肪酸底物的转换效率高于对 $\omega 6$ 脂肪酸底物的转换效率。该 $\Delta 6$ 去饱和酶还表现出 $\Delta 8$ 去饱和酶活性。与对 $\omega 3$ 去饱和的脂肪酸底物不具有偏向性的去饱和酶相比,在植物的重组LC-PUFA途径中使用该 $\Delta 6$ 去饱和酶或其他对 $\omega 3$ 去饱和的脂肪酸底物具有高特异性的去饱和酶提高了EPA、DPA和DHA的水平。

[0034] 因此,本发明还提供重组细胞,优选地是植物细胞,更优选地是植物种子细胞,所述细胞包含编码具有 $\Delta 6$ 去饱和酶活性的脂肪酸去饱和酶的外源性多核苷酸,其中所述去饱和酶进一步具有至少两种、优选全部三种以下特性:

[0035] i) 对作为脂肪酸底物的ALA的 $\Delta 6$ 去饱和酶活性高于对LA的 $\Delta 6$ 去饱和酶活性,优选地在植物细胞中是这样,

[0036] ii) 对作为脂肪酸底物的ALA-CoA的 $\Delta 6$ 去饱和酶活性高于对作为脂肪酸底物的连接于PC的sn-2位置的ALA的 $\Delta 6$ 去饱和酶活性,优选地在植物细胞中是这样,和

[0037] iii) 对ETrA具有 $\Delta 8$ 去饱和酶活性,优选地在植物细胞中是这样。

[0038] 本发明还提供重组细胞,优选地是植物细胞,更优选地是植物种子细胞,所述细胞包含编码具有 $\Delta 6$ 去饱和酶活性的脂肪酸去饱和酶的外源性多核苷酸,其中所述去饱和酶对 $\omega 3$ 底物的活性高于对相应的 $\omega 6$ 底物的活性,且其中当所述去饱和酶由所述细胞中的所述多核苷酸表达时,所述去饱和酶对ALA具有活性,以至少5%、至少7.5%、或至少10%的效率产生SDA,或当所述去饱和酶在酵母细胞中表达时,其以至少35%的效率产生SDA。

[0039] 在一个实施方式中,所述去饱和酶对脂肪酸底物ALA的 $\Delta 6$ 去饱和酶活性高于对LA的 $\Delta 6$ 去饱和酶活性,优选地在植物细胞中是这样。

[0040] 所述 $\Delta 6$ 去饱和酶优选地对ALA底物的 $\Delta 6$ 去饱和酶活性是对LA的至少大约2倍、至少3倍、至少4倍、或至少5倍,优选地在植物细胞中是这样。

[0041] 在另一个实施方式中,所述 $\Delta 6$ 去饱和酶对作为脂肪酸底物的ALA-CoA的活性高于对作为脂肪酸底物的连接至PC的sn-2位置的ALA的活性,优选地在植物细胞中是这样。

[0042] 所述 $\Delta 6$ 去饱和酶优选地对作为脂肪酸底物的ALA-CoA的 $\Delta 6$ 去饱和酶活性是对作为脂肪酸底物的连接至PC的sn-2位置的ALA的 $\Delta 6$ 去饱和酶活性的至少大约5倍或至少10-

倍,优选地在植物细胞中是这样。

[0043] 所述 $\Delta 6$ 去饱和酶优选地是前端去饱和酶。

[0044] 在另一个实施方式中,本发明的细胞还包含编码以下各项的外源性多核苷酸:

[0045] i) $\Delta 6$ 延长酶,

[0046] ii) $\Delta 5$ 去饱和酶,

[0047] iii) $\Delta 5$ 延长酶,和

[0048] iv) 任选存在的 $\Delta 4$ 去饱和酶和/或 $\omega 3$ 去饱和酶,

[0049] 其中每一多核苷酸可操纵地连接于一或多种能够在所述细胞中指导所述多核苷酸表达的启动子。

[0050] 本发明细胞中的所述 $\Delta 6$ 去饱和酶优选地对ETA不具有可检测到的 $\Delta 5$ 去饱和酶活性。

[0051] 所述 $\Delta 6$ 去饱和酶优选地包含SEQ ID NO:10所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:10具有至少77%相同性的氨基酸序列。

[0052] 在另一个实施方式中,所述 $\Delta 6$ 去饱和酶包含SEQ ID NO:8所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:8具有至少67%相同性的氨基酸序列并具有 $\Delta 8$ 去饱和酶活性。

[0053] 本发明人还发现,表达 $\Delta 9$ 延长酶、 $\Delta 8$ 去饱和酶和 $\Delta 5$ 去饱和酶的重组细胞能够更高效地将脂肪酸底物转换为EPA、DPA和DHA。

[0054] 因此,本发明还提供重组细胞,优选地是植物细胞,更优选地是植物种子细胞,所述细胞包含编码以下各项的外源性多核苷酸:

[0055] i) $\Delta 9$ 延长酶,

[0056] ii) $\Delta 8$ 去饱和酶,

[0057] iii) $\Delta 5$ 去饱和酶,

[0058] iv) 任选存在的 $\Delta 5$ 延长酶,和

[0059] v) 如果存在 $\Delta 5$ 延长酶,则任选存在的 $\Delta 4$ 去饱和酶,

[0060] 其中每一多核苷酸可操纵地连接于一或多种能够在所述细胞中指导所述多核苷酸表达的启动子,且其中所述细胞内的总脂肪酸有至少15%、至少20%、或至少25%在它们的酰基链中包含至少20个碳和至少3个碳碳双键。

[0061] 在一个实施方式中,本发明细胞内脂肪酸中的ARA、EPA、DPA和DHA总计占所述细胞内的总脂肪酸的至少15%、至少20%、或至少25%。

[0062] 在进一步的实施方式中,本发明的细胞将油酸转换为二十碳烯酸(C20:1)的能力与野生型植物相比是降低的,和/或在所述细胞中有低于5%的油酸被转换为二十碳烯酸。

[0063] 在一个具体的实施方式中,所述细胞内的总脂肪酸具有低于1%的C20:1。

[0064] 在进一步的实施方式中,本发明的细胞与野生型细胞相比具有降低的内源性 $\Delta 15$ 去饱和酶活性,和/或所述细胞中有低于10%的LA被转换为ALA。

[0065] 在一个具体的实施方式中,所述内源性 $\Delta 15$ 去饱和酶对酰基-PC底物的活性高于对相应的酰基辅酶A底物的活性,优选地其中所述酰基是LA。

[0066] 在进一步的实施方式中,本发明的细胞与缺乏所述外源性多核苷酸的相应细胞相比,其将GLA转换为SDA和/或将ARA转换为EPA的转换是增加的。

[0067] 在一个实施方式中,本发明的细胞内脂肪酸中的DHA的量占所述细胞内的总脂肪酸的至少3%、至少5%、或至少10%。

[0068] 在另一个实施方式中,在本发明的细胞内,LA转换为ARA和/或ALA转换为EPA的转换效率是至少80%或至少90%。

[0069] 在一个实施方式中,所述 $\Delta 9$ 延长酶包含SEQ ID NO:22所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:22具有至少80%相同性的氨基酸序列。

[0070] 在进一步的实施方式中,所述 $\Delta 8$ 去饱和酶包含SEQ ID NO:24所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:24具有至少80%相同性的氨基酸序列。

[0071] 在另一个实施方式中,所述 $\Delta 5$ 去饱和酶包含SEQ ID NO:26或SEQ ID NO:13所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:26和/或SEQ ID NO:13具有至少80%相同性的氨基酸序列。

[0072] 本发明人已经获得的结果表明,可以利用表达所述 $\Delta 6$ -去饱和酶、 $\Delta 6$ 延长酶、 $\Delta 5$ 去饱和酶、 $\Delta 5$ 延长酶、和 $\Delta 4$ 去饱和酶的一组基因、或类似的(特别是其中所述去饱和酶对酰基辅酶A底物具有活性的)一组基因,来合成显著水平的EPA、DPA和DHA。

[0073] 因此,本发明还提供重组细胞,优选地是植物细胞,更优选地是植物种子细胞,所述细胞包含编码以下各项的外源性多核苷酸:

[0074] i) $\Delta 6$ 延长酶和/或 $\Delta 9$ 延长酶,

[0075] ii) $\Delta 6$ 去饱和酶和/或 $\Delta 8$ 去饱和酶,

[0076] iii) $\Delta 5$ 去饱和酶,

[0077] iv) $\Delta 5$ 延长酶,

[0078] v) $\Delta 4$ 去饱和酶,和

[0079] vi) 任选存在的二酰基甘油酰基转移酶,

[0080] 其中每一多核苷酸可操纵地连接于一或多种能够在所述细胞中指导所述多核苷酸表达的启动子,特征在于具有一种或多种或全部以下特性:

[0081] a) 将ALA转换为EPA、DPA或DHA的转换效率是至少17.3%、或至少23%,

[0082] b) 将ALA转换为DPA或DHA的转换效率是至少15.4%、或至少21%,

[0083] c) 将ALA转换为DHA的效率是至少9.5%、或至少10.8%,和

[0084] d) 将EPA转换为DHA的效率是至少45%、或至少50%,

[0085] 且优选地进一步特征在于所述细胞内总脂肪酸中有至少4%是DHA。

[0086] 优选地,掺入到细胞内的三酰基甘油中的总脂肪酸中有至少6%、至少11%或至少15%是DHA。

[0087] 在一个实施方式中,DHA构成所述细胞内SDA、ETA、EPA、DPA和DHA的总和的20-65%,优选地为40-65%。

[0088] 优选地,细胞内的 $\omega 3$ 脂肪酸中的0.1-25%是SDA、0.1-10%是ETA、0.1-60%是EPA、0.1-50%是DPA、且30-95%是DHA,更优选地,所述细胞内的 $\omega 3$ 脂肪酸中的0.1-25%是SDA、0.1-10%是ETA、0.1-50%是EPA、0.1-50%是DPA和40-95%是DHA。

[0089] 所述 $\Delta 4$ 去饱和酶优选地包含SEQ ID NO:73所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:73具有至少80%相同性的氨基酸序列。

[0090] 在另一个方面,本发明提供重组细胞,优选地是植物细胞,更优选地是植物种子细

胞,所述细胞包含编码以下各项的外源性多核苷酸:

[0091] i) $\Delta 6$ 延长酶和/或 $\Delta 9$ 延长酶,

[0092] ii) $\Delta 6$ 去饱和酶和/或 $\Delta 8$ 去饱和酶,

[0093] iii) $\Delta 5$ 去饱和酶,

[0094] iv) $\Delta 5$ 延长酶,和

[0095] v) 任选存在的二酰基甘油酰基转移酶

[0096] 其中每一多核苷酸可操纵地连接于一或多种能够在所述细胞中指导所述多核苷酸表达的启动子,特征在于具有一种或多种或全部以下特性:

[0097] a) 将ALA转换为EPA或DPA的转换效率是至少17.3%、或至少23%,和

[0098] b) 将ALA转换为DPA的效率是至少15.4%、或至少21%,

[0099] 且优选地进一步特征在于所述细胞内的总脂肪酸中有至少4%是DPA。

[0100] 优选地,掺入到细胞内的三酰基甘油中的总脂肪酸中有至少6%、至少11%或至少15%是DPA。

[0101] 优选地,DPA构成细胞内SDA、ETA、EPA和DPA的总量的20-65%,更优选地40-65%。

[0102] 优选地,细胞内的 $\omega 3$ 脂肪酸中的0.1-35%是SDA、0.1-15%是ETA、0.1-60%是EPA、而30-75%是DPA,更优选地,细胞内的 $\omega 3$ 脂肪酸中的0.1-35%是SDA、0.1-15%是ETA、0.1-50%是EPA而40-75%是DPA。

[0103] 在一个实施方式中,所述 $\Delta 6$ 延长酶包含SEQ ID NO:4所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:4具有至少55%相同性的氨基酸序列。

[0104] 在另一个实施方式中,所述 $\Delta 6$ 去饱和酶包含SEQ ID NO:8所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:8具有至少67%相同性的氨基酸序列。

[0105] 在一个实施方式中,所述 $\Delta 5$ 去饱和酶包含SEQ ID NO:26所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:26具有至少80%相同性的氨基酸序列。

[0106] 在另一个实施方式中,所述 $\Delta 5$ 延长酶包含SEQ ID NO:6所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:6具有至少47%相同性的氨基酸序列。

[0107] 在一个实施方式中,所述二酰基甘油酰基转移酶包含SEQ ID NO:75或SEQ ID NO:108所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:75和/或SEQ ID NO:108具有至少80%相同性的氨基酸序列。

[0108] 本发明显然涵盖任何两种、三种、四种或全部上述酶的组合。

[0109] 在另一个实施方式中,本发明的细胞、优选地植物细胞、且更优选地植物种子细胞,还包含编码以下各项的外源性多核苷酸:

[0110] i) $\Delta 17$ 去饱和酶,

[0111] ii) $\Delta 15$ 去饱和酶,和/或

[0112] iii) $\Delta 12$ 去饱和酶

[0113] 其中每一多核苷酸可操纵地连接于一或多种能够在所述细胞中指导所述多核苷酸表达的启动子,

[0114] 在另一个实施方式中,本发明的细胞中由外源性多核苷酸表达的一种或多种或全部所述去饱和酶对酰基辅酶A底物的活性高于对相应的酰基-PC底物的活性。在具体的实施方式中,所述 $\Delta 6$ 去饱和酶和 $\Delta 5$ 去饱和酶、 $\Delta 5$ 去饱和酶和 $\Delta 4$ 去饱和酶、 $\Delta 6$ 去饱和酶和 $\Delta 4$

去饱和酶、或全部三种所述 $\Delta 6$ 去饱和酶、 $\Delta 5$ 去饱和酶和 $\Delta 4$ 去饱和酶、或在每一上述这些组合的基础上额外加上任何 $\Delta 17$ 去饱和酶、 $\Delta 15$ 去饱和酶和/或 $\Delta 12$ 去饱和酶,均对它们的酰基辅酶A底物的活性高于对相应的酰基-PC底物的活性。在这一实施方式中,细胞中由外源性多核苷酸表达的其他去饱和酶对酰基辅酶A底物的活性可高于或不高于对相应的酰基-PC底物的活性。可以理解,每种酶的优选的酰基辅酶A底物是不同的。

[0115] 本发明还提供重组细胞,优选地是植物细胞,更优选地是植物种子细胞,所述细胞包含编码具有 $\Delta 6$ 延长酶和 $\Delta 9$ 延长酶活性的脂肪酸延长酶的外源性多核苷酸,其中所述延长酶的 $\Delta 6$ 延长酶活性高于 $\Delta 9$ 延长酶活性。

[0116] 在一个实施方式中,所述延长酶转换SDA以产生ETA的效率是至少50%或至少60%,和/或转换ALA以产生ETrA的效率是至少6%或至少9%。

[0117] 优选地,所述延长酶具有的 $\Delta 6$ 延长酶活性是 $\Delta 9$ 延长酶活性的至少大约6.5倍。

[0118] 在另一个实施方式中,所述延长酶不具有可检测到的 $\Delta 5$ 延长酶活性。

[0119] 延长酶优选地包含SEQ ID NO:4所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:4具有至少55%相同性的氨基酸序列。

[0120] 在另一个实施方式中,所述细胞还包含编码以下各项的外源性多核苷酸:

[0121] i) $\Delta 8$ 去饱和酶和/或 $\Delta 6$ 去饱和酶,

[0122] ii) $\Delta 5$ 去饱和酶,

[0123] iii) $\Delta 5$ 延长酶,和

[0124] iv) 任选存在的 $\Delta 4$ 去饱和酶和/或 $\omega 3$ 去饱和酶,

[0125] 其中每一多核苷酸可操纵地连接于一或多种能够在所述细胞中指导所述多核苷酸表达的启动子。

[0126] 本发明还提供重组细胞,优选地是植物细胞,更优选地是植物种子细胞,所述细胞包含编码具有 $\Delta 5$ 去饱和酶活性的脂肪酸去饱和酶的外源性多核苷酸,其中所述去饱和酶包含SEQ ID NO:13所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:13具有至少53%相同性的氨基酸序列。

[0127] 本发明还提供重组细胞,优选地是植物细胞,更优选地是植物种子细胞,所述细胞包含编码具有 $\Delta 9$ 延长酶活性的脂肪酸延长酶的外源性多核苷酸,其中所述延长酶包含SEQ ID NO:28、94和96中任一序列所示的氨基酸序列、其生物学活性片段、与SEQ ID NO:28具有至少81%相同性的氨基酸序列、或与SEQ ID NO:94和/或SEQ ID NO:96具有至少50%相同性的氨基酸序列。

[0128] 在一个实施方式中,所述 $\Delta 9$ 延长酶包含SEQ ID NO:94或SEQ ID NO:96所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:94和/或SEQ ID NO:96具有至少50%相同性的氨基酸序列,且其中所述延长酶对 $\omega 6$ 底物的活性高于对相应的 $\omega 3$ 底物的活性。更优选地,所述 $\Delta 9$ 延长酶对 $\omega 6$ 底物(例如LA)的活性是对相应的 $\omega 3$ 底物(例如ALA)的活性的至少2倍,更优选地至少4倍。

[0129] 在另一个方面,本发明提供一种重组细胞,其包含编码二酰基甘油酰基转移酶的外源性多核苷酸,其中所述二酰基甘油酰基转移酶包含SEQ ID NO:108所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:108具有至少54%相同性的氨基酸序列。

[0130] 在本发明细胞的一个实施方式中,所述去饱和酶和/或延长酶、或多种去饱和酶

和/或延长酶,可纯化自微藻类。优选的微藻类是巴夫藻(Pavlova spp)、塔胞藻(Pyramimonas spp)和微胞藻(Micromonas spp)。

[0131] 在优选的实施方式中,本发明的细胞是真核细胞。例如所述细胞可以是植物细胞、哺乳动物细胞、昆虫细胞、真菌细胞或酵母细胞。所述细胞可以是组织培养物中的细胞、体外的细胞和/或分离的细胞。

[0132] 在一个实施方式中,所述细胞在植物中和/或是成熟的植物种子细胞。所述植物可以是田间的植物或作为植物部分收获的植物,或所述种子可以是收获的种子。

[0133] 在一个具体的实施方式中,所述植物或植物种子分别是油料种子植物或油料种子。

[0134] 本领域人员可以理解,可以在同一细胞中共表达一或多种所定义的延长酶和/或去饱和酶。

[0135] 在进一步的实施方式中,本发明的细胞能够合成长链多不饱和脂肪酸(LC-PUFA),其中所述衍生自不能合成所述LC-PUFA的细胞。

[0136] 本发明人还发现,共表达沉默阻抑物(silencing suppressor)可升高植物细胞内脂肪酸生物合成酶类的水平,特别是在从最初的转化植物重复繁殖数代之后。因此,在优选的实施方式中,本发明的细胞,优选地植物细胞,更优选地植物贮藏器官细胞或种子细胞,包含编码沉默阻抑物的外源性多核苷酸。

[0137] 优选地,编码沉默阻抑物的外源性多核苷酸可操纵地连接于植物贮藏器官特异性启动子。在一个实施方式中,所述植物贮藏器官特异性启动子是种子特异性启动子,或是优先在发育的种子中表达的子叶-特异性启动子或胚乳-特异性启动子。

[0138] 本发明还提供获得细胞、优选地植物细胞、更优选地植物种子细胞的方法,所述的细胞能够合成一或多种长链多不饱和脂肪酸(LC-PUFA),所述方法包括

[0139] a) 将编码脂肪酸 ω 3去饱和酶活性的外源性多核苷酸引入细胞,优选地是不能合成所述LC-PUFA的细胞,其中所述多核苷酸可操纵地连接于能够在所述细胞中指导所述多核苷酸表达的启动子,

[0140] b) 在所述细胞中表达所述外源性多核苷酸,

[0141] c) 分析所述细胞的脂肪酸组成,和

[0142] d) 选择能够进行以下至少一种去饱和作用的细胞:将ARA去饱和化为EPA、将DGLA去饱和化为ETA、将GLA去饱和化为SDA、将ARA去饱和化为EPA并将DGLA去饱和化为ETA、将ARA去饱和化为EPA并将GLA去饱和化为SDA、或所有这三种去饱和作用。

[0143] 在一个实施方式中,所选的细胞是本发明的细胞。具体地,所述细胞可进一步包含本申请所描述的去饱和酶和延长酶的组合。

[0144] 本发明还提供获得细胞、优选地植物细胞、更优选地植物种子细胞的方法,所述的细胞能够合成一或多种长链多不饱和脂肪酸(LC-PUFA),所述方法包括

[0145] a) 将编码脂肪酸 Δ 5延长酶的外源性多核苷酸引入细胞,优选地是不能合成所述LC-PUFA的细胞,其中所述多核苷酸可操纵地连接于能够在所述细胞中指导所述多核苷酸表达的启动子,

[0146] b) 在所述细胞中表达所述外源性多核苷酸,

[0147] c) 分析所述细胞的脂肪酸组成,和

[0148] d) 选择其中所述 $\Delta 5$ 延长酶对 EPA 具有活性并以至少 60%、至少 65%、至少 70% 或至少 75% 的效率产生 DPA 的细胞。

[0149] 在本发明方法的一个实施方式中,所选的细胞是本发明的细胞。具体地,所述细胞可进一步包含本申请所描述的去饱和酶和延长酶的组合。

[0150] 本发明还提供获得细胞、优选地植物细胞、更优选地植物种子细胞的方法,所述的细胞能够合成一或多种长链多不饱和脂肪酸 (LC-PUFA),所述方法包括

[0151] a) 将编码脂肪酸 $\Delta 6$ 去饱和酶的外源性多核苷酸引入细胞,优选地是不能合成所述 LC-PUFA 的细胞,其中所述多核苷酸可操纵地连接于能够在所述细胞中指导所述多核苷酸表达的启动子,

[0152] b) 在所述细胞中表达所述外源性多核苷酸,

[0153] c) 分析所述细胞的脂肪酸组成,和

[0154] d) 选择具有至少两种、优选全部三种以下特性的细胞:

[0155] i) 对作为脂肪酸底物的 ALA 的 $\Delta 6$ 去饱和酶活性高于对 LA 的 $\Delta 6$ 去饱和酶活性,优选地在植物细胞中是这样,

[0156] ii) 对作为脂肪酸底物的 ALA-CoA 的 $\Delta 6$ 去饱和酶活性高于对作为脂肪酸底物的连接于 PC 的 sn-2 位置的 ALA 的 $\Delta 6$ 去饱和酶活性,优选地在植物细胞中是这样,和

[0157] iii) 对 ALA 具有 $\Delta 6$ 去饱和酶活性并对 EtrA 具有 $\Delta 8$ 去饱和酶活性,优选地在植物细胞中是这样。

[0158] 本发明还提供获得细胞、优选地植物细胞、更优选地植物种子细胞的方法,所述的细胞能够合成一或多种长链多不饱和脂肪酸 (LC-PUFA),所述方法包括

[0159] a) 将编码脂肪酸 $\Delta 6$ 去饱和酶的外源性多核苷酸引入细胞,优选地是不能合成所述 LC-PUFA 的细胞,其中所述多核苷酸可操纵地连接于能够在所述细胞中指导所述多核苷酸表达的启动子,

[0160] b) 在所述细胞中表达所述外源性多核苷酸,

[0161] c) 分析所述细胞的脂肪酸组成,和

[0162] d) 选择具有 $\Delta 6$ 去饱和酶活性的细胞,所述 $\Delta 6$ 去饱和酶对 $\omega 3$ 底物的活性高于对相应的 $\omega 6$ 底物的活性,且所述 $\Delta 6$ 去饱和酶对 ALA 具有活性并以至少 5%、至少 7.5%、或至少 10% 的效率产生 SDA,或当其在酵母细胞中表达时以至少 35% 的效率产生 SDA。

[0163] 在一个实施方式中,所选的细胞是本发明的细胞。具体地,所述细胞可进一步包含本申请所描述的去饱和酶和延长酶的组合。

[0164] 本发明还提供获得细胞、优选地植物细胞、更优选地植物种子细胞的方法,所述的细胞能够合成一或多种长链多不饱和脂肪酸 (LC-PUFA),所述方法包括

[0165] a) 将编码以下各项的外源性多核苷酸引入细胞,优选地是不能合成所述 LC-PUFA 的细胞:

[0166] i) $\Delta 9$ 延长酶,

[0167] ii) $\Delta 8$ 去饱和酶,

[0168] iii) $\Delta 5$ 去饱和酶,

[0169] iv) 任选存在的 $\Delta 5$ 延长酶,和

[0170] v) 如果存在 $\Delta 5$ 延长酶,则任选存在的 $\Delta 4$ 去饱和酶,

[0171] 其中每一多核苷酸可操纵地连接于一或多种能够在所述细胞中指导所述多核苷酸表达的启动子，

[0172] b) 在所述细胞中表达所述外源性多核苷酸，

[0173] c) 分析所述细胞的脂肪酸组成，和

[0174] d) 选择其中总脂肪酸有至少15%、至少20%或至少25%在它们的酰基链中包含至少20个碳和至少3个碳碳双键的细胞。

[0175] 在一个实施方式中，所选的细胞是本发明的细胞。

[0176] 本发明还提供获得细胞、优选地植物细胞、更优选地植物种子细胞的方法，所述的细胞能够合成一或多种长链多不饱和脂肪酸 (LC-PUFA)，所述方法包括

[0177] a) 将编码以下各项的外源性多核苷酸引入细胞，优选地是不能合成所述LC-PUFA的细胞：

[0178] i) $\Delta 6$ 延长酶和/或 $\Delta 9$ 延长酶，

[0179] ii) $\Delta 6$ 去饱和酶和/或 $\Delta 8$ 去饱和酶，

[0180] iii) $\Delta 5$ 去饱和酶，

[0181] iv) $\Delta 5$ 延长酶，

[0182] v) $\Delta 4$ 去饱和酶，和

[0183] vi) 任选存在的二酰基甘油酰基转移酶，

[0184] 其中每一多核苷酸可操纵地连接于一或多种能够在所述细胞中指导所述多核苷酸表达的启动子，

[0185] b) 在所述细胞中表达所述外源性多核苷酸，

[0186] c) 分析所述细胞的脂肪酸组成，和

[0187] d) 选择具有一种或多种或全部以下特性的细胞：

[0188] 1) 将ALA转换为EPA、DPA或DHA的转换效率是至少17.3%、或至少23%；

[0189] 2) 将ALA转换为DPA或DHA的转换效率是至少15.4%、或至少21%；

[0190] 3) 将ALA转换为DHA的效率是至少9.5%、或至少10.8%；和

[0191] 4) 将EPA转换为DHA的效率是至少45%、或至少50%；

[0192] 且优选地进一步特征在于所述细胞内总脂肪酸中有至少4%是DHA。

[0193] 在另一方面，本发明提供获得细胞、优选地植物细胞、更优选地植物种子细胞的方法，所述的细胞能够合成一或多种长链多不饱和脂肪酸 (LC-PUFA)，所述方法包括

[0194] a) 将编码以下各项的外源性多核苷酸引入细胞，优选地是不能合成所述LC-PUFA的细胞：

[0195] i) $\Delta 6$ 延长酶和/或 $\Delta 9$ 延长酶，

[0196] ii) $\Delta 6$ 去饱和酶和/或 $\Delta 8$ 去饱和酶，

[0197] iii) $\Delta 5$ 去饱和酶，

[0198] iv) $\Delta 5$ 延长酶，和

[0199] v) 任选存在的二酰基甘油酰基转移酶

[0200] 其中每一多核苷酸可操纵地连接于一或多种能够在所述细胞中指导所述多核苷酸表达的启动子，

[0201] b) 在所述细胞中表达所述外源性多核苷酸，

- [0202] c) 分析所述细胞的脂肪酸组成,和
- [0203] d) 选择具有一种或多种或全部以下特性的细胞:
- [0204] a) 将ALA转换为EPA或DPA的转换效率是至少17.3%、或至少23%,和
- [0205] b) 将ALA转换为DPA的效率是至少15.4%、或至少21%,
- [0206] 且优选地进一步特征在于所述细胞内的总脂肪酸中有至少4%是DPA。
- [0207] 在本发明的方法的一个实施方式中,所述外源性多核苷酸稳定整合到所述细胞的基因组中。
- [0208] 在另一个实施方式中,所述方法还包括从步骤a)的细胞再生转化植物的步骤。
- [0209] 在进一步的实施方式中,所述外源性多核苷酸在所述细胞中瞬时表达。
- [0210] 在一个实施方式中,所述细胞是植物的叶细胞。
- [0211] 本发明还提供选择参与脂肪酸去饱和作用的核酸分子的方法,包括:
- [0212] i) 获得可操纵地连接于启动子的核酸分子,所述核酸分子编码多肽,所述多肽可能是脂肪酸去饱和酶;
- [0213] ii) 将所述核酸分子引入细胞,其中所述启动子在所述细胞中具有活性;
- [0214] iii) 在所述细胞中表达所述核酸分子;
- [0215] iv) 分析所述细胞中的脂肪酸组成;和
- [0216] v) 基于以下事实选择参与脂肪酸去饱和作用的核酸分子:所述多肽具有 ω 3去饱和酶活性并能够进行以下至少一种去饱和作用:将ARA去饱和化为EPA、将DGLA去饱和化为ETA、将GLA去饱和化为SDA、将ARA去饱和化为EPA并将DGLA去饱和化为ETA、将ARA去饱和化为EPA并将GLA去饱和化为SDA、或所有这三种去饱和作用。
- [0217] 在所述方法的一个实施方式中,所述多肽的氨基酸序列与SEQ ID NO:15具有至少35%的相同性、与SEQ ID NO:17具有至少60%的相同性、和/或与SEQ ID NO:20具有至少60%的相同性。
- [0218] 本发明还提供选择参与脂肪酸延长作用的核酸分子的方法,包括:
- [0219] i) 获得可操纵地连接于启动子的核酸分子,所述核酸分子编码多肽,所述多肽可能是脂肪酸延长酶,
- [0220] ii) 将所述核酸分子引入细胞,其中所述启动子在所述细胞中具有活性,
- [0221] iii) 在所述细胞中表达所述核酸分子;
- [0222] iv) 分析所述细胞中的脂肪酸组成;和
- [0223] v) 基于以下事实选择参与脂肪酸延长作用的核酸分子:所述多肽具有 Δ 5延长酶活性且其转换EPA产生DPA的效率是至少60%、至少65%、至少70%或至少75%。
- [0224] 在一个实施方式中,所述多肽的氨基酸序列与SEQ ID NO:6具有至少47%的相同性。
- [0225] 本发明还提供选择参与脂肪酸去饱和作用的核酸分子的方法,包括:
- [0226] i) 获得可操纵地连接于启动子的核酸分子,所述核酸分子编码多肽,所述多肽可能是脂肪酸去饱和酶,
- [0227] ii) 将所述核酸分子引入细胞,其中所述启动子在所述细胞中具有活性,
- [0228] iii) 在所述细胞中表达所述核酸分子;
- [0229] iv) 分析所述细胞中的脂肪酸组成;和

[0230] v) 基于以下事实选择参与脂肪酸去饱和作用的核酸分子:所述多肽具有 $\Delta 6$ 去饱和酶活性并具有至少两种、优选全部三种以下特性:

[0231] a) 对作为脂肪酸底物的ALA的 $\Delta 6$ 去饱和酶活性高于对LA的 $\Delta 6$ 去饱和酶活性,优选地在植物细胞中是这样,

[0232] b) 对作为脂肪酸底物的ALA-CoA的 $\Delta 6$ 去饱和酶活性高于对作为脂肪酸底物的连接于PC的sn-2位置的ALA的 $\Delta 6$ 去饱和酶活性,优选地在植物细胞中是这样,和

[0233] c) 对ALA具有 $\Delta 8$ 去饱和酶活性,优选地在植物细胞中是这样。

[0234] 在所述方法的一个实施方式中,所述多肽的氨基酸序列与SEQ ID NO:10具有至少77%的相同性和/或与SEQ ID NO:8具有至少67%的相同性。

[0235] 本发明还提供选择参与脂肪酸去饱和作用的核酸分子的方法,包括:

[0236] i) 获得可操纵地连接于启动子的核酸分子,所述核酸分子编码多肽,所述多肽可能是脂肪酸去饱和酶,

[0237] ii) 将所述核酸分子引入细胞,其中所述启动子在所述细胞中具有活性,

[0238] iii) 在所述细胞中表达所述核酸分子;

[0239] iv) 分析所述细胞中的脂肪酸组成;和

[0240] v) 基于以下事实选择参与脂肪酸去饱和作用的核酸分子:所述多肽具有 $\Delta 6$ 去饱和酶活性和 $\Delta 8$ 去饱和酶活性。

[0241] 在一个实施方式中,所述多肽的氨基酸序列与SEQ ID NO:8具有至少67%的相同性。

[0242] 在另一个实施方式中,本发明的方法的步骤(v) 包括选择编码对酰基辅酶A底物具有活性的去饱和酶的核酸分子或编码前端去饱和酶的核酸分子。

[0243] 本发明还提供本文中所定义的外源性多核苷酸的组合,当它们用于产生重组细胞时,它们在重组细胞中表达至少两种脂肪酸去饱和酶和两种脂肪酸延长酶的组合、和/或在重组细胞中产生LC-PUFA。

[0244] 本发明还提供基本上纯化的和/或重组脂肪酸 $\Delta 5$ 延长酶,其中当由外源性多核苷酸在细胞中表达时,所述延长酶对EPA具有活性,以至少60%、至少65%、至少70%或至少75%的效率产生DPA。

[0245] 在一个实施方式中,所述 $\Delta 5$ 延长酶的特征在于具有本文所定义的任何一或多种特性。

[0246] 本发明还提供基本上纯化的和/或重组脂肪酸 $\omega 3$ 去饱和酶,当由外源性多核苷酸在细胞中表达时,其能够进行以下至少一种去饱和作用:将ARA去饱和化为EPA、将DGLA去饱和化为ETA、将GLA去饱和化为SDA、将ARA去饱和化为EPA并将DGLA去饱和化为ETA、将ARA去饱和化为EPA并将GLA去饱和化为SDA、或所有这三种去饱和作用。

[0247] 在一个实施方式中,本发明的 $\omega 3$ 去饱和酶的特征在于具有本文所定义的任何一或多种特性。

[0248] 本发明还提供基本上纯化的和/或重组脂肪酸 $\Delta 6$ 去饱和酶,其中所述去饱和酶进一步具有至少两种、优选全部三种以下特性:

[0249] i) 对作为脂肪酸底物的ALA的 $\Delta 6$ 去饱和酶活性高于对LA的 $\Delta 6$ 去饱和酶活性,优选地在植物细胞中是这样,

[0250] ii) 对作为脂肪酸底物的ALA-CoA的 $\Delta 6$ 去饱和酶活性高于对作为脂肪酸底物的连接于PC的sn-2位置的ALA的 $\Delta 6$ 去饱和酶活性,优选地在植物细胞中是这样,和

[0251] iii) 对ETrA具有 $\Delta 8$ 去饱和酶活性,优选地在植物细胞中是这样。

[0252] 本发明还提供基本上纯化的和/或重组脂肪酸 $\Delta 6$ 去饱和酶,其中所述去饱和酶对 $\omega 3$ 底物的活性高于对相应的 $\omega 6$ 底物的活性,且其中所述去饱和酶对ALA具有活性,当所述去饱和酶由外源性多核苷酸在细胞中表达时以至少5%、至少7.5%、或至少10%的效率产生SDA,或当其在酵母细胞中表达时以至少35%的效率产生SDA。

[0253] 在一个实施方式中,本发明的 $\Delta 6$ 去饱和酶的特征在于具有本文所定义的任何一或多种特性。

[0254] 本发明还提供基本上纯化的和/或重组脂肪酸 $\Delta 6$ 延长酶和 $\Delta 9$ 延长酶,其中所述延长酶的 $\Delta 6$ 延长酶活性高于 $\Delta 9$ 延长酶活性。

[0255] 在一个实施方式中,本发明的 $\Delta 6$ 延长酶和 $\Delta 9$ 延长酶的特征在于具有本文所定义的任何一或多种特性。

[0256] 本发明还提供基本上纯化的和/或重组脂肪酸 $\Delta 5$ 去饱和酶,其包含SEQ ID NO:13所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:13具有至少53%相同性的氨基酸序列。

[0257] 本发明还提供基本上纯化的和/或重组脂肪酸 $\Delta 9$ 延长酶,其包含SEQ ID NO:28、94和96任一所示的氨基酸序列、其生物学活性片段、与SEQ ID NO:28具有至少81%相同性的氨基酸序列,或与SEQ ID NO:94和/或SEQ ID NO:96具有至少50%相同性的氨基酸序列。

[0258] 在一个实施方式中,所述 $\Delta 9$ 延长酶包含SEQ ID NO:94或SEQ ID NO:96所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:94和/或SEQ ID NO:96具有至少50%相同性的氨基酸序列,且其中所述延长酶对 $\omega 6$ 底物的活性高于对相应的 $\omega 3$ 底物的活性。

[0259] 在另一个方面,本发明提供基本上纯化的和/或重组二酰基甘油酰基转移酶,其包含SEQ ID NO:108所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:108具有至少54%相同性的氨基酸序列。

[0260] 在一个实施方式中,本发明的去饱和酶或延长酶可纯化自微藻类。优选的微藻类是巴夫藻、塔胞藻和微胞藻。

[0261] 本发明还提供分离的和/或外源性多核苷酸,其包含:

[0262] i) 选自SEQ ID NO:3、5、7、9、11、12、14、16、18、19、27、29、93、95、107或125至129中任一序列的核苷酸序列,

[0263] ii) 编码本发明的去饱和酶或延长酶的核苷酸序列,

[0264] iii) 与SEQ ID NO:3、5、7、9、11、12、14、16、18、19、27、29、93、95、107或125至129所示的一或多种序列具有至少50%相同性的核苷酸序列,和/或

[0265] iv) 在严格性条件下与i)至iii)中任一序列杂交的序列。

[0266] 在一个实施方式中,所述分离的和/或外源性多核苷酸包含与SEQ ID NO:3和/或SEQ ID NO:126具有至少57%相同性的核苷酸序列,并编码 $\Delta 6$ 延长酶。

[0267] 在另一个实施方式中,所述分离的和/或外源性多核苷酸包含与SEQ ID NO:14、SEQ ID NO:16、SEQ ID NO:18和/或SEQ ID NO:19具有至少50%相同性的核苷酸序列,并编码 $\omega 3$ 去饱和酶。

[0268] 在另一个实施方式中,所述分离的和/或外源性多核苷酸包含与SEQ ID NO:5和/或SEQ ID NO:128具有至少50%相同性的核苷酸序列,并编码 $\Delta 5$ 延长酶。

[0269] 在一个实施方式中,所述分离的和/或外源性多核苷酸包含与SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11和/或SEQ ID NO:125具有至少75%相同性的核苷酸序列,并编码 $\Delta 6$ 去饱和酶。

[0270] 在另一个实施方式中,所述分离的和/或外源性多核苷酸包含与SEQ ID NO:12具有至少60%相同性的核苷酸序列,并编码 $\Delta 5$ 去饱和酶。

[0271] 在另一个实施方式中,所述分离的和/或外源性多核苷酸包含与SEQ ID NO:27、SEQ ID NO:29、SEQ ID NO:93和/或SEQ ID NO:96具有至少50%相同性的核苷酸序列,并编码 $\Delta 9$ 延长酶。

[0272] 在另一个实施方式中,所述分离的和/或外源性多核苷酸包含与SEQ ID NO:107具有至少60%相同性的核苷酸序列,并编码二酰基甘油酰基转移酶。

[0273] 在一个具体的实施方式中,所述分离的和/或外源性多核苷酸与SEQ ID NO:3、5、7、9、11、12、14、16、18、19、27、29、93、95、107或125至129所示序列之一具有至少80%、或至少90%、或至少95%、或至少99%的相同性。

[0274] 本发明还提供用于整合到和/或已整合到植物细胞基因组中的DNA构建体,所述构建体包含一簇至少三个编码调控植物细胞内脂肪酸合成的蛋白质的可读框,优选地每一蛋白质是脂肪酸去饱和酶或脂肪酸延长酶,其中具有相同转录方向的各可读框间隔至少750bp、至少1,000bp或至少1,250bp,且至少有两个可读框具有不同的转录方向,其中每一可读框可操纵地连接于在植物细胞中具有活性的启动子,且每一启动子可独立地是相同的或不同的。

[0275] 优选地,所述DNA构建体中至少有两个启动子是不同的。

[0276] 一或多个或每一所述可读框优选地可操纵地连接于异源性5'前导序列,每一前导序列可独立地是相同的或不同的,其中与特定可读框的天然存在的5'前导序列相比,每一异源性5'前导序列使翻译效率提高。

[0277] 在本发明的DNA构建体中,所述异源性5'前导序列优选地是烟草花叶病毒(TMV)5'前导序列。

[0278] 在本发明的DNA构建体中,所述蛋白质优选地是延长酶和/或去饱和酶,更优选地是本申请所述的组合。

[0279] 在另一个实施方式中,本发明的DNA构建体仅具有三个或四个被翻译成蛋白质的可读框。

[0280] 本发明还提供载体,其包含本发明的多核苷酸和/或本发明的DNA构建体。

[0281] 优选地,所述多核苷酸可操纵地连接于启动子。

[0282] 本发明还提供产生本发明的去饱和酶或延长酶的方法,所述方法包括在细胞或无细胞表达系统中表达本发明的多核苷酸、本发明的DNA构建体和/或本发明的载体。

[0283] 本发明人还令人惊异地发现,可将至少三种独立的包含不同外源性多核苷酸的染色体外转移核酸瞬时转染至真核细胞内并可组合检测细胞内每一外源性多核苷酸的活性。因此,在另一方面,本发明还提供以至少三种不同的外源性多核苷酸瞬时转染真核细胞的方法,所述方法包括:

[0284] i) 获得至少

[0285] a) 第一细菌,其包含包含第一外源性多核苷酸的染色体外转移核酸,

[0286] b) 第二细菌,其包含包含第二外源性多核苷酸的染色体外转移核酸,和

[0287] c) 第三细菌,其包含包含第三外源性多核苷酸的染色体外转移核酸,和

[0288] ii) 将所述细胞与步骤i)的细菌接触,

[0289] 其中每一染色体外转移核酸从细菌转移至细胞中以产生瞬时转染的细胞,其中每一外源性多核苷酸包含在所述细胞中具有活性的启动子,其中每一启动子可以独立地是相同的或不同的,且其中至少有一种外源性多核苷酸编码沉默阻抑物。

[0290] 可以一或多种所述细菌依次或同时进行步骤ii)。例如,细胞可以接触第一细菌,然后接触第二细菌,等等。在另一实例中,细菌同时接触各种细菌,优选地为所述细菌的混合物。不同细菌的浓度可彼此不同或可以是相同的或相近的。所述细菌可以是混合的分离物(pooled isolates),例如包含来自菌株文库的多种分离物。

[0291] 在一个实施方式中,所述方法还包括获得所述细胞然后使所述细胞接触一或多种额外的细菌,每一细菌包含包含不同外源性多核苷酸的染色体外转移核酸。例如,在一个实施方式中,所述方法包括获得所述细胞然后使所述细胞接触第四细菌,所述第四细菌包含包含第四外源性多核苷酸的染色体外转移核酸。在另外一实施方式中,所述方法包括获得所述细胞然后使所述细胞接触第五细菌,所述第五细菌包含包含第五外源性多核苷酸的染色体外转移核酸。在另外一实施方式中,所述方法包括获得所述细胞然后使所述细胞接触第六细菌,所述第六细菌包含包含第六外源性多核苷酸的染色体外转移核酸。在另外一实施方式中,所述方法包括获得所述细胞然后使所述细胞接触第七细菌,所述第七细菌包含包含第七外源性多核苷酸的染色体外转移核酸。在又一实施方式中,所述方法包括获得所述细胞然后使所述细胞接触第八细菌,所述第八细菌包含包含第八外源性多核苷酸的染色体外转移核酸。

[0292] 优选地,所述的不同外源性多核苷酸编码不同RNA分子和/或多肽。

[0293] 在一个实施方式中,每一外源性多核苷酸编码形成酶途径的一部分的酶或是此种酶的候选物。

[0294] 上述方面特别可用于研究形成大的和/或复杂的生物学途径的多核苷酸和/或多肽。因此,在优选的实施方式中,每一外源性多核苷酸编码酶或是此种酶的候选物,所述酶参与脂肪酸合成、脂肪酸修饰、二酰基甘油组装、三酰基甘油组装、或其中两种或更多种的组合。

[0295] 在一个实施方式中,一或多种所述细菌的形式是原生质体。可用于本发明的细菌的实例包括,但不限于,土壤杆菌、根瘤菌、苜蓿中华根瘤菌(*Sinorhizobium meliloti*)、百脉根中慢生根瘤菌(*Mezorhizobium loti*)、弗氏志贺菌(*Shigella flexneri*)、鼠伤寒沙门氏菌(*Salmonella typhimurium*)、猪霍乱沙门氏菌(*Salmonella choleraesuis*)、单核细胞增多性李斯特菌(*Listeria monocytogenes*)、大肠杆菌(*Escherichia coli*)、假结核病耶尔森菌(*Yersinia pseudotuberculosis*)和小肠结肠炎耶尔森菌(*Yersinia enterocolitica*)。

[0296] 可用于本发明的染色体外转移核酸的实例包括,但不限于,P-DNA、土壤杆菌、T-DNA、或它们的组合。

[0297] 优选地,上述方面的细胞是植物细胞或哺乳动物细胞。在一个实施方式中,所述细胞是组织或器官的一部分。在另一个实施方式中,所述细胞是植物细胞,所述组织或器官是叶、茎、根、分生组织、愈伤组织、或胚珠。

[0298] 本发明人还确定了,如果启动子是种子特异性启动子,外源性多核苷酸在叶细胞中的表达可共表达种子特异性转录因子(例如多叶子叶2(leafy cotyledon 2)、视褐质3(fusca3)或脱落酸敏感3(abscisic acid-senstive3))而增强。多叶子叶2蛋白的实例包括,但不限于,WO 01/70777中所述的那些。因此,在优选的实施方式中,所述植物细胞是植物的叶细胞,至少一种所述启动子是种子特异性启动子,且至少一种所述外源性多核苷酸编码种子特异性转录因子例如多叶子叶2。

[0299] 在一个实施方式中,所述外源性多核苷酸无一是病毒基因。在一个实施方式中,一种或多种所述外源性多核苷酸在所述染色体外转移核酸中仅以所述核酸序列的单拷贝形式存在,而不是所述核酸序列的多聚体或部分多聚体。在另一个实施方式中,至少一种所述染色体外转移核酸不包括在所述细胞中具有功能的复制起始点,优选地不是病毒复制起始点,更优选地不是FBNYV复制起始点。在另一个实施方式中,所述外源性多核苷酸无一编码病毒复制酶或病毒移动蛋白,例如WO 2007/137788和Marillonnet等(2005)所述的那些。

[0300] 本发明还提供筛选瞬时转染的细胞的所需活性的方法,所述方法包括以至少三种本发明的外源性多核苷酸瞬时转染真核细胞,并测试所述细胞的所需活性。

[0301] 本发明人还发现,通过以不同的染色体外转移核酸来提供6种以上不同基因,可增强所述6种以上不同基因对细胞、特别是植物细胞的转染。因此,在另一方面,本发明提供以至少六种不同的外源性多核苷酸转化真核细胞的方法,所述方法包括

[0302] i) 获得至少

[0303] a) 第一细菌,其包含第一染色体外转移核酸,所述第一染色体外转移核酸包含三种、四种、五种或六种不同的外源性多核苷酸,和

[0304] b) 第二细菌,其包含不同于所述第一染色体外转移核酸的第二染色体外转移核酸,所述第二染色体外转移核酸包含三种、四种、五种或六种不同的外源性多核苷酸,

[0305] ii) 将所述细胞与步骤i)的细菌接触,和

[0306] iii) 任选地选择被所述第一和第二染色体外转移核酸的外源性多核苷酸稳定转化的细胞,

[0307] 其中每一所述第一和第二染色体外转移核酸的外源性多核苷酸由所述细菌转移至所述细胞中以产生转化的细胞,其中每一所述外源性多核苷酸包含启动子,所述启动子在所述细胞中或可由其衍生的细胞中具有活性,且其中每一启动子可以独立地是相同的或不同的。

[0308] 可以两种细菌依次或同时进行步骤i) a) 和i) b)。例如,所述细胞可先接触第一细菌,然后接触第二细菌。与第二细菌接触的细胞可以是与第一细菌接触后的细胞的子代细胞或衍生自与第一细菌接触后的细胞。在另一实例中,所述细胞同时与这两种细菌接触。

[0309] 在一个实施方式中,所述i) 第一染色体外转移核酸只有3-6种、3-5种、3-4种、4-6种、4-5种或5-6种不同的外源性多核苷酸,和ii) 第二染色体外转移核酸只有3-6种、3-5种、3-4种、4-6种、4-5种或5-6种不同的外源性多核苷酸。

[0310] 优选地,每一外源性多核苷酸编码多肽,且其中各多肽是不同的。

[0311] 在另一个实施方式中，

[0312] i) 第一染色体外转移核酸包含两种外源性多核苷酸，它们独立地编码选自如下一组的多肽： $\Delta 6$ 去饱和酶、 $\Delta 12$ 去饱和酶和 $\Delta 15$ 去饱和酶，和

[0313] ii) 第二染色体外转移核酸包含编码多肽的外源性多核苷酸，所述多肽是该组中的第三种酶。

[0314] 优选地，所述细胞是植物细胞，且所述方法还包括从所述稳定转化的细胞再生转化植物的步骤。

[0315] 本发明还提供细胞，所述细胞是通过所述以至少三种本发明的外源性多核苷酸瞬时转染真核细胞的方法而产生的，或是通过所述以至少6种不同的本发明外源性多核苷酸转化真核细胞的方法而产生的。

[0316] 在又一方面中，本发明提供产生以至少六种不同的外源性多核苷酸稳定转化的植物的方法，所述方法包括

[0317] i) 获得包含第一外源性基因组区的第一种稳定转化植物，所述第一外源性基因组区包含3、4、5、或6种不同的外源性多核苷酸，

[0318] ii) 获得与所述第一种稳定转化植物具有性亲和性并包含第二外源性基因组区的第二种稳定转化植物，所述第二外源性基因组区不同于所述第一外源性基因组区并包含3、4、5、或6种不同的外源性多核苷酸，

[0319] iii) 使所述第一种稳定转化植物与所述第二种稳定转化植物杂交，和

[0320] iv) 选择由步骤iii) 产生的包含所述第一和第二基因组区的植物或其子代，由此产生稳定转化植物，

[0321] 其中每一所述外源性多核苷酸包含启动子，所述启动子在植物中具有活性，且其中每一启动子可以独立地是相同的或不同的。

[0322] 在优选的实施方式中，按照前述本发明的DNA构建体中所述来定向和间隔所述第一和/或第二外源性基因组区的外源性多核苷酸。

[0323] 在所述第一和第二外源性基因组区中，任一启动子序列可出现多次或仅出现一次，或者，在所述第一和第二外源性基因组区中，一或多种启动子可出现多次而一或多种其他启动子仅出现一次。每一植物启动子可独立地优先在所述植物的组织或器官中具有活性，例如，相对于其他组织或器官而言，在叶或种子中具有活性。这可使得所有引入的蛋白质编码区在所述植物器官或组织中同时表达或重叠表达。在并列的实施方式中，一或多种启动子在所述植物中组成型表达，而一或多种其他启动子优先在植物器官或组织中表达。

[0324] 在一个实施方式中，步骤i) 包括通过以下步骤产生所述第一种稳定转化植物

[0325] a) 将植物细胞与包含第一染色体外转移核酸的第一细菌接触，所述第一染色体外转移核酸包含三种、四种、五种或六种不同的外源性多核苷酸，

[0326] b) 从步骤a) 的植物细胞产生稳定转化植物，和任选地

[0327] c) 产生来自步骤b) 的稳定转化植物的子代植物；和/或步骤ii) 包括通过以下步骤产生所述第二种稳定转化植物，

[0328] d) 将植物细胞与包含第二染色体外转移核酸的第二细菌接触，所述第二染色体外转移核酸包含三种、四种、五种或六种不同的外源性多核苷酸，

[0329] e) 从步骤d) 的植物细胞产生稳定转化植物，和任选地

[0330] f) 产生来自步骤e)的稳定转化植物的子代植物。

[0331] 在另一个方面,本发明提供产生以至少六种不同的外源性多核苷酸稳定转化的植物的方法,所述方法包括

[0332] i) 获得包含第一外源性基因组区的第一种稳定转化植物或植物部分,所述第一外源性基因组区包含3、4、5、或6种不同的外源性多核苷酸,

[0333] ii) 使所述第一种稳定转化植物或植物部分的细胞接触包含染色体外转移核酸的细菌,所述染色体外转移核酸包含三种、四种、五种或六种不同的外源性多核苷酸,

[0334] iii) 从所述细胞产生植物,和

[0335] iv) 任选地,选择由步骤iii)产生的包含所述至少六种不同的外源性多核苷酸的植物。

[0336] 对于上述各方面的“将细胞与步骤i)的细菌接触”这一步骤,正如本领域人员能够理解的那样,这一步骤要在合适的条件下进行合适的时间,以便染色体外转移核酸从细菌转移至细胞。

[0337] 在另一方面,本发明提供真核细胞至少包含

[0338] a) 包含第一外源性多核苷酸的第一染色体外转移核酸,

[0339] b) 包含第二外源性多核苷酸的第二染色体外转移核酸,和

[0340] c) 包含第三外源性多核苷酸的第三染色体外转移核酸。

[0341] 在一个实施方式中,所述细菌进一步一或多种额外的细菌,每一细菌包含包含外源性多核苷酸的染色体外转移核酸。

[0342] 本发明还提供植物或其子代或种子,其包含包含3、4、5、或6种不同的外源性多核苷酸的第一外源性基因组区,和包含3、4、5、或6种不同的外源性多核苷酸的第二外源性基因组区。所述外源性基因组区的外源性多核苷酸优选地按照前述DNA构建体中所述来定向和间隔。

[0343] 本发明还提供包含本发明的细胞的转基因非人类生物体。在一个实施方式中,所述生物体的每一细胞均是本发明的细胞。

[0344] 优选地,所述转基因非人类生物体是转基因植物,更优选地是转基因油料种子植物,其产生下文所列的油。在进一步的实施方式中,所述转基因植物包含至少一种额外的编码沉默阻抑物的外源性多核苷酸,其可操纵地连接于植物贮藏器官特异性启动子,其中所述植物具有正常的表型。

[0345] 本发明还提供种子,其包含本发明的细胞或得自本发明的转基因植物。

[0346] 本发明还提供油,其由本发明的细胞、本发明的转基因非人类生物体、或本发明的种子产生,或得自本发明的细胞、本发明的转基因非人类生物体、或本发明的种子。

[0347] 在一个实施方式中,所述油是通过来自油料种子的油进行提取而获得的。

[0348] 在一个实施方式中,所述油介花油(欧洲油菜(*Brassica napus*)、芜菁(*Brassica rapa*))、芥子油(芥菜(*Brassica juncea*))、其他芸苔油、葵花子油(向日葵(*Helianthus annuus*))、亚麻籽油(栽培亚麻(*Linum usitatissimum*))、大豆油(大豆(*Glycine max*))、红花油(红花(*Carthamus tinctorius*))、玉米油(玉蜀黍(*Zea mays*))、烟草油(烟草(*Nicotiana tabacum*))、花生油(花生(*Arachis hypogaea*))、棕榈油、棉籽油(陆地棉(*Gossypium hirsutum*))、椰子油(椰子(*Cocos nucifera*))、鳄梨油(鳄梨(*Persea*

americana))、橄榄油(橄榄(*Olea europaea*))、腰果油(腰果(*Anacardium occidentale*))、夏威夷果油(夏威夷果(*Macadamia intergrifolia*))、扁桃仁油(扁桃(*Prunus amygdalus*))或拟南芥籽油(拟南芥(*Arabidopsis thaliana*))。

[0349] 本发明还提供脂肪酸,其由本发明的细胞、本发明的转基因非人类生物体、或本发明的种子产生,或得自本发明的细胞、本发明的转基因非人类生物体、或本发明的种子。

[0350] 本发明还提供产生含不饱和脂肪酸的油的方法,所述方法包括从本发明的细胞、本发明的转基因非人类生物体、或本发明的种子提取油。

[0351] 本发明还提供组合物,其包含本发明的细胞、本发明的去饱和酶或延长酶、本发明的多核苷酸、本发明的DNA构建体、本发明的载体、本发明的油或本发明的脂肪酸。

[0352] 本发明还提供饲料、化妆品或化学制品,其包含本发明的细胞、本发明的转基因非人类生物体、本发明的种子、本发明的油和/或本发明的脂肪酸。

[0353] 本发明还提供进行去饱和酶反应的方法,所述方法包括使与CoA酯化的多不饱和脂肪酸接触本发明的去饱和酶。

[0354] 本发明还提供基本上纯化的抗体或其片段,其特异性结合本发明的去饱和酶或延长酶。

[0355] 本发明还提供治疗或预防可能从PUFA获益的病症的方法,所述方法包括给对象施用本发明的细胞、本发明的去饱和酶或延长酶、本发明的多核苷酸、本发明的DNA构建体、本发明的载体、本发明的转基因非人类生物体、本发明的种子、本发明的油或本发明的脂肪酸和/或本发明的饲料。

[0356] 在一个实施方式中,所述病症是心律失常、血管成形术、炎症、哮喘、银屑病、骨质疏松症、肾脏结石、AIDS、多发性硬化、类风湿关节炎、克隆病、精神分裂症、癌症、胎儿酒精综合征、注意力缺陷多动症、囊性纤维化、苯丙酮尿症、单相抑郁、攻击敌意、肾上腺脑白质营养不良、冠心病、高血压、糖尿病、肥胖、阿尔茨海默病、慢性阻塞性肺疾病、溃疡性结肠炎、血管成形术后狭窄、湿疹、高血压、血小板聚集、胃肠道出血、子宫内膜异位症、经前期综合征、肌痛脑脊髓炎、病毒感染后慢性疲劳或眼部疾病。

[0357] 本发明还提供本发明的细胞、本发明的去饱和酶或延长酶、本发明的多核苷酸、本发明的DNA构建体、本发明的载体、本发明的转基因非人类生物体、本发明的种子、本发明的油或本发明的脂肪酸和/或本发明的饲料在制备用于治疗或预防可能从PUFA获益的病症的药物中的用途。

[0358] 本发明人已经令人惊异地发现,沉默阻抑物可优先地表达于植物贮藏器官以增强转基因在植物中的表达水平,而不明显影响植物的发育。

[0359] 因此,本发明提供植物细胞,其包含:

[0360] i) 编码沉默阻抑物的第一外源性多核苷酸,其可操纵地连接于植物贮藏器官特异性启动子,和

[0361] ii) 编码RNA分子的第二外源性多核苷酸,其可操纵地连接于在植物贮藏器官中指导基因转录的启动子。

[0362] 优选地,所述植物贮藏器官特异性启动子是种子特异性启动子,例如子叶特异性启动子或胚乳特异性启动子。

[0363] 在一个实施方式中,所述沉默阻抑物是病毒阻抑蛋白,例如,但不限于,P1、P19、

V2、P38、P15、Pe-Po和RPV-P0。

[0364] 典型地,如果植物组成型地表达病毒阻抑蛋白,则植物具有异常的表型,但如果沉默阻抑物特异性表达于贮藏器官,则植物具有正常的表型。此类病毒阻抑蛋白的实例包括,但不限于,P1、P19和P15。

[0365] 在进一步的实施方式中,所述病毒阻抑蛋白降低microRNA的积累和/或降低microRNA的导向型裂解。

[0366] RNA分子本身可以是功能性的,例如,但不限于,反义多核苷酸、催化性多核苷酸、dsRNA和/或microRNA。或者,所述RNA分子可编码具有所需功能的多肽,例如,但不限于,参与脂肪酸合成或修饰的酶、种子贮藏蛋白(例如,谷类麦谷蛋白或麦醇溶蛋白)、参与碳水化合物的合成或修饰、次级代谢的酶或药物。药用蛋白质的实例包括,但不限于,抗体以及抗体相关分子及其片段、抗原性多肽,它们能够例如产生针对癌症、感染因子的免疫保护,细胞因子例如,粒细胞巨噬细胞集落刺激因子、干扰素- α 、人血清白蛋白和促红细胞生成素。

[0367] 在进一步的实施方式中,所述细胞包含至少一种、至少两种、至少三种、至少四种或至少五种或更多额外的不同的外源性多核苷酸,每一外源性多核苷酸编码RNA分子并可操纵地连接于在贮藏器官中指导基因转录的启动子。每一外源性多核苷酸可以可操纵地连接于同一启动子、不同的启动子或它们的组合。

[0368] 在一个实施方式中,所述外源性多核苷酸是DNA。

[0369] 在进一步的实施方式中,所述细胞存在于植物贮藏器官例如种子中。

[0370] 在优选的实施方式中,与缺乏所述第一外源性多核苷酸的同源细胞相比,所述RNA分子以升高的水平存在。优选地,所述水平增加至少10%、至少20%、更优选地至少30%。

[0371] 在另一实施方式中,与缺乏所述第一外源性多核苷酸的同源细胞相比,由至少所述额外的外源性多核苷酸编码的至少一种RNA分子以升高的水平存在。

[0372] 本发明还提供转基因植物,其包含以上方面的细胞。在一个实施方式中,所述植物的每一细胞均是以上方面所定义的细胞。在特别优选的实施方式中,与缺乏所述细胞的植物相比,所述植物具有正常的表型。

[0373] 在另一方面,本发明提供植物贮藏器官,其包含以上方面的细胞和/或其得自前面定义的转基因植物。

[0374] 在一个实施方式中,所述植物贮藏器官是种子。

[0375] 在另一方面,本发明提供获得在贮藏器官中具有升高水平的RNA分子的表型正常植物的方法,包括

[0376] a) 向植物细胞中引入

[0377] i) 编码沉默阻抑物的第一外源性多核苷酸,其可操纵地连接于植物贮藏器官特异性启动子,和

[0378] ii) 编码RNA分子的第二外源性多核苷酸,其可操纵地连接于在植物贮藏器官中指导基因转录的启动子,

[0379] b) 从步骤a)的细胞再生转化植物,

[0380] c) 生长所述转化植物直至其产生贮藏器官,

[0381] d) 确定所述贮藏器官中的RNA分子水平,和

[0382] e) 选择表型正常的植物,但其中所述贮藏器官内的RNA分子的水平与缺乏所述第

一外源性多核苷酸的相应的贮藏器官相比是升高的。

[0383] 在又一方面,本发明提供获得在贮藏器官中RNA分子稳定表达的表型正常植物的方法,包括

[0384] a) 向植物细胞中引入

[0385] i) 编码沉默阻抑物的第一外源性多核苷酸,其可操纵地连接于植物贮藏器官特异性启动子,和

[0386] ii) 编码RNA分子的第二外源性多核苷酸,其可操纵地连接于在植物贮藏器官中指导基因转录的启动子,

[0387] b) 从步骤a)的细胞再生转化植物,

[0388] c) 从步骤b)的植物产生包含贮藏器官的第三代子代植物,和

[0389] d) 选择第三代子代植物,其中所述贮藏器官中的所述RNA分子的水平是前代植物的贮藏器官中水平的至少90%。

[0390] 优选地,以上方面的外源性多核苷酸稳定整合到所述细胞的基因组中。

[0391] 在又一方面,本发明提供在转基因植物的贮藏器官中稳定表达RNA分子的方法,包括

[0392] i) 表达编码沉默阻抑物的、可操纵地连接于植物贮藏器官特异性启动子的第一外源性多核苷酸,和

[0393] ii) 表达编码RNA分子的、可操纵地连接于在植物贮藏器官中指导基因转录的启动子的第二外源性多核苷酸,

[0394] 其中所述转基因植物是从以所述外源性多核苷酸转化的母本植物获得的至少第三代子代植物,且其中所述植物的贮藏器官中的所述RNA分子的水平是前代植物的贮藏器官中水平的至少90%。

[0395] 在一个实施方式中,所述植物生长在田间。

[0396] 显然,本发明一个方面的优选的特征也适用于本发明的许多其他方面。

[0397] 在本说明书中,“包含”一词或其变化形式例如“包括”应理解为旨在纳入所提及的元件、整数或步骤或是元件、整数或步骤的组,但不排除任何其他的元件、整数或步骤或是元件、整数或步骤的组。

[0398] 下面通过以下非限制性实施例并结合附图进一步阐述本发明。

附图说明

[0399] 图1.需氧DHA生物合成途径。

[0400] 图2.在PC、CoA库和TAG库之间转移脂肪酸的各种酰基交换酶(acyl exchange enzymes)。改编自Singh等(2005)。

[0401] 图3.微胞藻CS-0170 Δ 6-延长酶和相关基因之间的多重比对。AAV33630, C20-多不饱和脂肪酸延长酶[巴夫藻CCMP459]; AAY15135, 延长酶1[盐生巴夫藻]; ABR67690, C20延长酶[绿色巴夫藻(*Pavlova viridis*)]; AAV67797, 多不饱和脂肪酸延长酶1[*Ostreococcus tauri*]; CAL55414, 多不饱和脂肪酸延长酶2(ISS) [Ostreococcus tauri]; XP_001419791, 推定的蛋白质[Ostreococcus lucimarinus CCE9901]; MicCS0170-d6E, 微胞藻CS-0170 Δ 6-延长酶(本发明); AAV67800, 多不饱和脂肪酸延长酶2[假微型海链藻(*Thalassiosira*

pseudonana)];XP_001416454,推定的蛋白质[*Ostreococcus lucimarinus* CCE9901];ABC18313,多不饱和脂肪酸延长酶1[破囊壶菌(*Thraustochytrium* sp.)FJN-10];AAV67799,多不饱和脂肪酸延长酶1[假微型海链藻];AAW70157, Δ -6-延长酶[三角褐指藻];ABC18314,多不饱和脂肪酸延长酶2[破囊壶菌FJN-10];CAD58540,未命名蛋白质产物[球等鞭金藻]。

[0402] 图4.塔胞藻CS-0140 Δ 6-延长酶和相关基因之间的多重比对。AAL84174,多不饱和脂肪酸特异性延长酶1[小立碗藓];AAT85662,多不饱和脂肪酸延长酶[地钱(*Marchantia polymorpha*)];AAV67797,多不饱和脂肪酸延长酶1[*Ostreococcus tauri*];AB094747,推定的蛋白质[*Ostreococcus lucimarinus* CCE9901];Pyrco-d6E,塔胞藻CS-0140 Δ 6-延长酶(本发明);ABC18313,多不饱和脂肪酸延长酶1[破囊壶菌FJN-10];BAF97073,多不饱和脂肪酸延长酶[高山被孢霉];XP_001567488,长链多不饱和脂肪酸延长酶样蛋白[巴西利什曼原虫(*Leishmania braziliensis*)MHOM/BR/75/M2904];BAE71129, Δ 5-延长酶[地钱];XP_001779105;推定的蛋白质[小立碗藓]。

[0403] 图5.塔胞藻CS-0140 Δ 5-延长酶和相关基因之间的多重比对。AAI52204,Elov14蛋白[鲐];CAG01780,未命名蛋白质产物[墨绿凹鼻鲀(*Tetraodon nigroviridis*)];AAV67800,多不饱和脂肪酸延长酶2[假微型海链藻];AAV33630,C20-多不饱和脂肪酸延长酶[巴夫藻CCMP459];ABR67690,C20延长酶[绿色巴夫藻];AAY15135,延长酶1[盐生巴夫藻];AAV67798,多不饱和脂肪酸延长酶2[*Ostreococcus tauri*];AB098084,推定的蛋白质[*Ostreococcus lucimarinus* CCE9901];Pyrco-d5E,塔胞藻CS-0140 Δ 5-延长酶(本发明)。

[0404] 图6.微胞藻CCMP1545 Δ 6-去饱和酶和相关基因之间的多重比对。AAM09687, Δ 5-脂肪酸去饱和酶[破囊壶菌ATCC21685];AAV33631, Δ 4-去饱和酶[球等鞭金藻];AAW70159, Δ 6-去饱和酶[*Ostreococcus tauri*];AB099366,推定的蛋白质[*Ostreococcus lucimarinus* CCE9901];Mic-d6D,微胞藻CCMP1545 Δ 6-去饱和酶(本发明);ABF58685, Δ 5-去饱和酶[*Perkinsus marinus*];ABL96295, Δ 5-去饱和酶[盐生巴夫藻]。

[0405] 图7.*Ostreococcus lucimarinus* Δ 6-去饱和酶和相关基因之间的多重比对。AAM09687, Δ 5-脂肪酸去饱和酶[破囊壶菌ATCC21685];AAR27297, Δ 6-去饱和酶[淀粉菌(*Amylomyces rouxii*)];AAS93682, Δ 6-脂肪酸去饱和酶[稻根霉菌(*Rhizopus oryzae*)];AAV33631, Δ 4-去饱和酶[球等鞭金藻];AAW70159, Δ 6-去饱和酶[*Ostreococcus tauri*];Ostlu-d6D,*Ostreococcus lucimarinus* Δ 6-去饱和酶(本发明);ABF58685, Δ 5-去饱和酶[*Perkinsus marinus*];ABL96295, Δ 5-去饱和酶[盐生巴夫藻];EDQ92231,推定的蛋白质[领鞭毛虫(*Monosiga brevicollis*)MX1];CAM41728,脂肪酸去饱和酶,推定的[巴西利什曼原虫];CAM65683,脂肪酸去饱和酶,推定的[婴儿利什曼原虫(*Leishmania infantum*)]。

[0406] 图8.塔胞藻CS-0140 Δ 5-去饱和酶和相关基因之间的多重比对。AAM09687, Δ 5-脂肪酸去饱和酶[破囊壶菌ATCC21685];Pyrco-d5D,塔胞藻CS-0140 Δ 5-去饱和酶(本发明);AAT85661, Δ 6-脂肪酸去饱和酶[地钱];AAX14505, Δ 6-脂肪酸去饱和酶[假微型海链藻];ABP49078, Δ 6-脂肪酸去饱和酶[三角褐指藻];AAW70159, Δ 6-去饱和酶[*Ostreococcus tauri*];XP_001421073,推定的蛋白质[*Ostreococcus lucimarinus* CCE9901];ABF58685, Δ 5-去饱和酶[*Perkinsus marinus*];CAJ07076,脂肪酸去饱和酶,推定的[硕大利什曼原虫(*Leishmania major*)];ABL96295, Δ 5-去饱和酶[盐生巴夫藻];EDQ92231,推定的蛋白质

[领鞭毛虫MX1]。

[0407] 图9.Ostreococcus tauri (Ot) (SEQ ID NO:30)、Ostreococcus lucimarinus (Ol) (SEQ ID NO:10) 和微胞藻 (M) CCMP1545 (SEQ ID NO:8) 的 $\Delta 6$ -去饱和酶蛋白序列的多重比对。

[0408] 图10.显示各种去饱和酶之间的关系的系统发育树。Pavsa-d5D=盐生巴夫藻 $\Delta 5$ -去饱和酶 (ABL96295); Thr21685-d5D=破囊壶菌 ATCC21685 $\Delta 5$ -去饱和酶 (AAM09687); Mic1545-d6D=微胞藻 CCMP1545 $\Delta 6$ -去饱和酶 (本发明); Ostta-d6D=Ostreococcus tauri $\Delta 6$ -去饱和酶 (AAW70159); Ostlu-d6D=Ostreococcus lucimarinus $\Delta 6$ -去饱和酶 (本发明); Galga-d6D=原鸡 (Gallus gallus) $\Delta 6$ -去饱和酶 (XP_421053); Homsa-d6D=智人 $\Delta 6$ -去饱和酶 (AAG23121); Musmu-d6D=小家鼠 $\Delta 6$ -去饱和酶 (NP_062673); Danre-d5/6D=鲱 $\Delta 5$ -/ $\Delta 6$ -去饱和酶 (AAG25710); Spaau-d6D=金黄色石斑鱼 (Sparus aurata) 推定的 $\Delta 6$ -去饱和酶 (AAL17639); Scoma-d6D=大菱鲆 (Scophthalmus maximus) $\Delta 6$ -去饱和酶 (AAS49163); Oncmy-d6D=虹鳟 $\Delta 6$ -去饱和酶 (AAK26745); Salsa-d5D=Salmo salar $\Delta 5$ -去饱和酶 (AAL82631); Prifa-d6D=Primula farinosa $\Delta 6$ -去饱和酶 (AAP23034); Euggr-d6D=小眼虫 $\Delta 6$ -去饱和酶; Borof-d6D=琉璃苣 (Borago officianalis) $\Delta 6$ -去饱和酶 (AAC49700); Caeel-d6D=秀丽隐杆线虫 $\Delta 6$ -去饱和酶 (AAC15586); Rhior-d6D=稻根霉菌 $\Delta 6$ -去饱和酶 (AAS93682); Moral-d6D=高山被孢霉 $\Delta 6$ -去饱和酶 (AAF08685); Moris-d6D=深黄被孢霉 (Mortierella isabellina) $\Delta 6$ -去饱和酶 (AAL73948); Marpo-d6D=地钱 $\Delta 6$ -去饱和酶 (AAT85661); Cerpu-d6D=角齿藓 (Ceratodon purpureus) $\Delta 6$ -去饱和酶 (CAB94993); Phatr-d6D=三角褐指藻 $\Delta 6$ -去饱和酶 (AAL92563); Thaps-d6D=假微型海链藻 $\Delta 6$ -去饱和酶 (AAX14505); Pavsa-d8D=盐生巴夫藻 $\Delta 8$ -去饱和酶 (ABL96296); Phatr-d5D=三角褐指藻 $\Delta 5$ -去饱和酶 (AAL92562); Marpo-d5D=地钱 $\Delta 5$ -去饱和酶 (AAT85663); Moral-d5D=高山被孢霉 $\Delta 5$ -去饱和酶 (AAR28035); Dicdi-d5D=盘基网柄菌 (Dictyostelium discoideum) $\Delta 5$ -去饱和酶 (BAA37090)。

[0409] 图11.来自以 linP-mic1545-d6D-linT 构建体转化的 T2 拟南芥种子的 GC 结果。显示了单个事件 1-19 的 SDA 和 GLA 水平,各柱状图上方显示了 $\omega 3$ 到 $\omega 6$ 的转换效率的比。细小微胞藻 $\Delta 6$ -去饱和酶对 $\omega 3$ 底物显示出明显的偏向性。

[0410] 图12.构成渗入到本氏烟草 (Nicotiana benthamiana) 中的 EPA 途径的酶的转换效率。EPA 途径含有 $\Delta 6$ -去饱和酶 (车前叶蓝蓟 (Echium plantagineum) $\Delta 6$ -去饱和酶、Ostreococcus tauri $\Delta 6$ -去饱和酶或细小微胞藻 (Micromonas pusilla) $\Delta 6$ -去饱和酶)、Pyramimonas cordata $\Delta 6$ -延长酶和盐生巴夫藻 $\Delta 5$ -去饱和酶。a 显示各途径的 $\omega 3$ 库转换效率; b. 含有车前叶蓝蓟途径、细小微胞藻途径和含有这两种去饱和酶的途径之间的直接比较; c. 含有 O. tauri 途径、细小微胞藻途径和含有这两种酰基-CoA 去饱和酶的途径之间的直接比较。

[0411] 图13. GC 和 GC-MS 确认通过瞬时表达的微胞藻 RCC299 $\omega 3$ 去饱和酶在本氏烟草中产生 EPA。

[0412] 图14. 二元载体 pJP101acq 的图显示了该二元载体的关键特征,包括启动子方向、TMV 前导序列定位、间隔区域定位和基因插入的独特克隆位点。NosT=NOS 终止子,FP1=截短的 napin 终止子,LininT=Linin 终止子。

[0413] 图15.二元载体pJP107的图。

[0414] 图16.确定在基于叶的分析法中实现近最大基因活性所需的土壤杆菌的浓度。以不同培养密度的含有二元表达构建体Ig Δ 9elo的土壤杆菌AGL1进行渗入后的本氏烟草中的球等鞭金藻 Δ 9-延长酶 (Ig Δ 9elo) 活性。将共渗入的P19设定在OD_{600nm}为0.4的浓度。y-轴显示的是产生EDA和ETRA的Ig Δ 9elo活性的总和。

[0415] 图17.采用在叶子中瞬时表达或稳定表达的方法比较LC-PUFA途径的转基因表达。转换效率基于总脂肪酸特征谱。^a结果来自 (Qi等,2004); ^b结果来自 (Robert等,2005)。

[0416] 图18.本氏烟草中的代谢适应。a.是对金枪鱼油中产生的脂肪酸甲基酯 (FAME) 的气相色谱 (GC) 示踪,其仅含有少量的EPA,但含有大量的DHA。b.和c.是对来自以单基因CaMV 35S二元构建体 (含有细小微胞藻 Δ 6-去饱和酶、P.cordata Δ 6-延长酶、盐生巴夫藻 Δ 5-去饱和酶、P.cordata Δ 5-延长酶 (b.) 或盐生巴夫藻 Δ 5-延长酶 (c.) 和盐生巴夫藻 Δ 4-去饱和酶) 瞬时转化的本氏烟草叶组织中的TAG成分的FAME的GC示踪。EPA在使用盐生巴夫藻 Δ 5-延长酶的样品中的累积证实了如何通过仔细选择途径中的单个基因而使得代谢途径发生适应。

[0417] 图19.载体pJP3075中包含转基因的区域的图。

[0418] 图20.载体pJP3059中包含转基因的区域的图。

[0419] 图21.载体pJP3060中包含转基因的区域的图。

[0420] 图22.载体pJP3115中包含转基因的区域的图。

[0421] 图23.载体pJP3116中包含转基因的区域的图。

[0422] 序列表要点

[0423] SEQ ID NO:1-编码微胞藻CS-0170 Δ 6-延长酶的可读框。

[0424] SEQ ID NO:2-微胞藻CS-0170 Δ 6-延长酶。

[0425] SEQ ID NO:3-编码塔胞藻CS-0140 Δ 6-延长酶/ Δ 9-延长酶的可读框。

[0426] SEQ ID NO:4-塔胞藻CS-0140 Δ 6-延长酶/ Δ 9-延长酶。

[0427] SEQ ID NO:5-编码塔胞藻CS-0140 Δ 5-延长酶的可读框。

[0428] SEQ ID NO:6-塔胞藻CS-0140 Δ 5-延长酶。

[0429] SEQ ID NO:7-编码微胞藻CCMP1545 Δ 6-去饱和酶/ Δ 8-去饱和酶的可读框。

[0430] SEQ ID NO:8-微胞藻CCMP1545 Δ 6-去饱和酶/ Δ 8-去饱和酶。

[0431] SEQ ID NO:9-编码Ostreococcus lucimarinus Δ 6-去饱和酶的可读框。

[0432] SEQ ID NO:10-Ostreococcus lucimarinus Δ 6-去饱和酶。

[0433] SEQ ID NO:11-用于在植物中表达Ostreococcus lucimarinus Δ 6-去饱和酶的经密码子优化的可读框。

[0434] SEQ ID NO:12-编码塔胞藻CS-0140 Δ 5-去饱和酶的可读框。

[0435] SEQ ID NO:13-塔胞藻CS-0140 Δ 5-去饱和酶。

[0436] SEQ ID NO:14-编码微胞藻CS-0170 ω 3-去饱和酶的部分可读框。

[0437] SEQ ID NO:15-部分微胞藻CS-0170 ω 3-去饱和酶。

[0438] SEQ ID NO:16-编码微胞藻RCC299 ω 3-去饱和酶的可读框。

[0439] SEQ ID NO:17-微胞藻RCC299 ω 3-去饱和酶。

[0440] SEQ ID NO:18-用于在植物中表达微胞藻RCC299 ω 3-去饱和酶的经密码子优化的

可读框。

- [0441] SEQ ID NO:19-编码微胞藻CCMP1545 ω 3-去饱和酶的可读框。
- [0442] SEQ ID NO:20-微胞藻CCMP1545 ω 3-去饱和酶。
- [0443] SEQ ID NO:21-编码球等鞭金藻 Δ 9-延长酶的可读框。
- [0444] SEQ ID NO:22-球等鞭金藻 Δ 9-延长酶。
- [0445] SEQ ID NO:23-编码盐生巴夫藻 Δ 8-去饱和酶的可读框。
- [0446] SEQ ID NO:24-盐生巴夫藻 Δ 8-去饱和酶。
- [0447] SEQ ID NO:25-编码盐生巴夫藻 Δ 5-去饱和酶的可读框。
- [0448] SEQ ID NO:26-盐生巴夫藻 Δ 5-去饱和酶。
- [0449] SEQ ID NO:27-编码Emiliana huxleyi CCMP1516 Δ 9延长酶的可读框。
- [0450] SEQ ID NO:28-Emiliana huxleyi CCMP1516 Δ 9延长酶。
- [0451] SEQ ID NO:29-用于在植物中表达Emiliana huxleyi Δ 9延长酶的经密码子优化的可读框。
- [0452] SEQ ID NO:30-Ostreococcus tauri Δ 6-去饱和酶。
- [0453] SEQ ID NO:31-延长酶共有结构域1。
- [0454] SEQ ID NO:32-延长酶共有结构域2。
- [0455] SEQ ID NO:33-延长酶共有结构域3。
- [0456] SEQ ID NO:34-延长酶共有结构域4。
- [0457] SEQ ID NO:35-延长酶共有结构域5。
- [0458] SEQ ID NO:36-延长酶共有结构域6。
- [0459] SEQ ID NO:37-去饱和酶共有结构域1。
- [0460] SEQ ID NO:38-去饱和酶共有结构域2。
- [0461] SEQ ID NO:39-去饱和酶共有结构域3。
- [0462] SEQ ID NO:40-去饱和酶共有结构域4。
- [0463] SEQ ID NO:41-71和78-92-寡核苷酸引物。
- [0464] SEQ ID NO:72-编码盐生巴夫藻 Δ 4-去饱和酶的可读框。
- [0465] SEQ ID NO:73-盐生巴夫藻 Δ 4-去饱和酶。
- [0466] SEQ ID NO:74-编码拟南芥二酰基甘油酰基转移酶1的可读框。
- [0467] SEQ ID NO:75-拟南芥二酰基甘油酰基转移酶1。
- [0468] SEQ ID NO:76-延长酶共有结构域7。
- [0469] SEQ ID NO:77-延长酶共有结构域8。
- [0470] SEQ ID NO:93-编码Pavlova pinguis Δ 9-延长酶的可读框。
- [0471] SEQ ID NO:94-Pavlova pinguis Δ 9-延长酶。
- [0472] SEQ ID NO:95-编码盐生巴夫藻 Δ 9-延长酶的可读框。
- [0473] SEQ ID NO:96-盐生巴夫藻 Δ 9-延长酶。
- [0474] SEQ ID NO:97-P19病毒阻抑物。
- [0475] SEQ ID NO:98-V2病毒阻抑物。
- [0476] SEQ ID NO:99-P38病毒阻抑物。
- [0477] SEQ ID NO:100-Pe-P0病毒阻抑物。

- [0478] SEQ ID NO:101-RPV-P0病毒抑制物。
- [0479] SEQ ID NO:102-编码P19病毒抑制物的可读框。
- [0480] SEQ ID NO:103-编码V2病毒抑制物的可读框。
- [0481] SEQ ID NO:104-编码P38病毒抑制物的可读框。
- [0482] SEQ ID NO:105-编码Pe-P0病毒抑制物的可读框。
- [0483] SEQ ID NO:106-编码RPV-P0病毒抑制物的可读框。
- [0484] SEQ ID NO:107-密码子优化的编码微胞藻CCMP1545二酰基甘油酰基转移酶2的可读框。
- [0485] SEQ ID NO:108-微胞藻CCMP1545二酰基甘油酰基转移酶2。
- [0486] SEQ ID NO:109-124-转移核酸边界序列。
- [0487] SEQ ID NO:125-用于在植物中表达微胞藻CCMP1545 $\Delta 6$ 去饱和酶/ $\Delta 8$ 去饱和酶的经密码子优化的可读框。
- [0488] SEQ ID NO:126-用于在植物中表达塔胞藻CS-0140 $\Delta 6$ 延长酶/ $\Delta 9$ 延长酶的经密码子优化的可读框(在3'端被截短并编码功能性延长酶)。
- [0489] SEQ ID NO:127-用于在植物中表达盐生巴夫藻 $\Delta 5$ 去饱和酶的经密码子优化的可读框。
- [0490] SEQ ID NO:128-用于在植物中表达塔胞藻CS-0140 $\Delta 5$ 延长酶的经密码子优化的可读框。
- [0491] SEQ ID NO:129-用于在植物中表达盐生巴夫藻 $\Delta 4$ 去饱和酶的经密码子优化的可读框。
- [0492] 发明详述
- [0493] 通用技术和定义
- [0494] 除非另有说明,否则本说明书中所使用的科技术语均具有与本领域(例如细胞培养、分子遗传学、脂肪酸合成、转基因植物、蛋白质化学和生物化学等领域)普通技术人员通常所理解的含义相同的含义。
- [0495] 除非另有说明,本发明所采用的重组蛋白质、细胞培养和免疫学技术均为本领域人员所熟知的标准技术。此类技术在文献中有充分的描述和解释,例如参见J.Perbal, A Practical Guide to Molecular Cloning, John Wiley and Sons (1984); J.Sambrook等, Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory Press (1989); T.A.Brown (编辑), Essential Molecular Biology: A Practical Approach, Volumes 1 and 2, IRL Press (1991); D.M.Glover and B.D.Hames (编辑), DNA Cloning: A Practical Approach, Volumes 1-4, IRL Press (1995 and 1996); 和F.M.Ausubel等 (编辑), Current Protocols in Molecular Biology, Greene Pub.Associates and Wiley-Interscience (1988, 包括迄今所有的更新版); Ed Harlow and David Lane (编辑) Antibodies: A Laboratory Manual, Cold Spring Harbour Laboratory, (1988); 和J.E.Coligan等 (编辑) Current Protocols in Immunology, John Wiley&Sons (包括迄今所有的更新版), 通过引用将上述文献并入本申请。
- [0496] 选定的定义
- [0497] 在本文中, 术语“脂肪酸”指的是羧酸(有机酸), 通常具有长的脂肪族尾部, 可以是

饱和的或不饱和的。典型地,脂肪酸具有碳碳键连接的链,其长度为至少8个碳原子,更优选地长度为至少12个碳。绝大多数天然存在的脂肪酸具有偶数个碳原子,因为它们的生物合成涉及具有2个碳原子的乙酸酯。脂肪酸可以是游离状态的(非酯化的)或者是酯化的形式,例如作为甘油三酯、甘油二酯、单酰基甘油酯的部分、酰基辅酶A(硫酯)结合型或其他结合形式。脂肪酸可酯化为磷脂,例如磷脂酰胆碱、磷脂酰乙醇胺、磷脂酰丝氨酸、磷脂酰甘油、磷脂酰肌醇或二磷脂酰甘油的形式。

[0498] “饱和脂肪酸”在其链上不含有任何双键或其他官能团。术语“饱和”指的是氢,其中所有的碳(除羧基 $[-COOH]$ 之外)均含有尽可能多的氢。换言之, ω 端含有3个氢(CH_3-)而链内的各个碳含有2个氢($-CH_2-$)。

[0499] “不饱和脂肪酸”与饱和脂肪酸形式相似,不同之处为沿着链存在一或多个烯烃官能团,每一烯烃将链中的单键的 $-CH_2-CH_2-$ 部分置换为双键的 $-CH=CH-$ 部分(即,与另一碳双键连接的碳)。链中连接于该双键任一侧的2个相邻碳原子可以顺式或反式构型存在。

[0500] 在本文中,术语“单不饱和脂肪酸”指的是这样的脂肪酸,其碳链中包含至少12个碳原子且链中仅有一个烯烃基(碳碳双键)。在本文中,术语“多不饱和脂肪酸”或“PUFA”指的是这样的脂肪酸,其碳链中包含至少12个碳原子和至少2个烯烃基(碳碳双键)。

[0501] 在本文中,术语“长链多不饱和脂肪酸”或“LC-PUFA”指的是这样的脂肪酸,其碳链中包含至少20个碳原子和至少2个碳碳双键,因此也包括VLC-PUFA。在本文中,术语“极长链多不饱和脂肪酸”和“VLC-PUFA”指的是这样的脂肪酸,其碳链中包含至少22个碳原子和至少3个碳碳双键。通常,脂肪酸碳链的碳原子数涉及的是非分支碳链。如果碳链是分支的,则碳原子数不包括侧基碳原子。在一个实施方式中,所述长链多不饱和脂肪酸 ω 3脂肪酸,即脂肪酸自甲基端开始的第三个碳碳键发生去饱和(碳碳双键)。在另一个实施方式中,所述长链多不饱和脂肪酸 ω 6脂肪酸,即脂肪酸自甲基端开始的第六个碳碳键发生去饱和(碳碳双键)。在进一步的实施方式中,所述长链多不饱和脂肪酸选自:花生四烯酸(ARA, $20:4 \Delta 5, 8, 11, 14; \omega 6$)、二十碳四烯酸(ETA, $20:4 \Delta 8, 11, 14, 17, \omega 3$)、二十碳五烯酸(EPA, $20:5 \Delta 5, 8, 11, 14, 17; \omega 3$)、二十二碳五烯酸(DPA, $22:5 \Delta 7, 10, 13, 16, 19, \omega 3$)、或二十二碳六烯酸(DHA, $22:6 \Delta 4, 7, 10, 13, 16, 19, \omega 3$)。LC-PUFA也可以是双高- γ -亚油酸(DGLA)或二十碳三烯酸(ETrA, $20:3 \Delta 11, 14, 17, \omega 3$)。显然,本发明所产生的LC-PUFA可以是任何或全部上述各项的混合物,并可包括其他LC-PUFA或任何这些LC-PUFA的衍生物。在优选的实施方式中, ω 3脂肪酸是EPA、DPA、和/或DHA,优选地,EPA和/或DPA,或优选地,DPA和/或DHA。

[0502] 此外,在本文中,术语“长链多不饱和脂肪酸”和“极长链多不饱和脂肪酸”指的是游离状态的(非酯化的)或者是酯化的形式的脂肪酸,例如作为甘油三酯、甘油二酯、单酰基甘油酯的部分、酰基辅酶A(硫酯)结合型或其他结合形式。脂肪酸可酯化为磷脂,例如磷脂酰胆碱、磷脂酰乙醇胺、磷脂酰丝氨酸、磷脂酰甘油、磷脂酰肌醇或二磷脂酰甘油的形式。因此,LC-PUFA可以混合物的形式存在于细胞内或提取自细胞、组织或生物体的纯化的油或脂质中。在优选的实施方式中,本发明提供的油包含至少75%或85%的三酰基甘油,其余部分以其他形式的脂质(例如本文中所提到的那些)存在,至少所述的三酰基甘油包含LC-PUFA。所述油可被进一步纯化或处理,例如以强碱水解去除游离脂肪酸或被分级、蒸馏等等。

[0503] 可用于本发明的去饱和酶、延长酶和酰基转移酶蛋白及其编码基因是本领域已知

的或是其同源物或衍生物。表1列出了这些基因的实例及其编码的蛋白质的大小。已被发现参与LC-PUFA生物合成的去饱和酶均属于所谓的“前端”去饱和酶。

[0504]

表 1. 参与 LC-PUFA 生物合成的克隆的基因

酶	生物体 类型	物种	登录号	蛋白质大小 (氨基酸)	参考文献
Δ4-去饱和酶	原生生物	小眼虫	AY278558	541	Meyer 等, 2003
	藻类	路氏巴夫藻	AY332747	445	Tonon 等, 2003
		球等鞭金藻	AAV33631	433	Pereira 等, 2004b
		盐生巴夫藻	AAV15136	447	Zhou 等, 2007
	破囊壶菌	金黄色破囊壶菌 (<i>Thraustochytrium aureum</i>)	AAN75707 AAN75708 AAN75709 AAN75710	515	N/A
Δ5-去饱和酶		破囊壶菌 ATCC21685	AAM09688	519	Qiu 等 2001
	哺乳动物类	智人	AF199596	444	Cho 等, 1999b Leonard 等, 2000b
	线虫	秀丽隐杆线虫	AF11440, NM_069350	447	Michaelson 等, 1998b; Watts and Browse, 1999b
	真菌	高山被孢霉	AF067654	446	Michaelson 等, 1998a; Knutzon 等, 1998
		畸雌腐霉	AF419297	456	Hong 等, 2002a
		盘基网柄菌	AB022097	467	Saito 等, 2000
		异丝水霉		470	WO02081668
	硅藻	三角褐指藻	AY082392	469	Domergue 等, 2002
	藻类	破囊壶菌	AF489588	439	Qiu 等, 2001
		金黄色破囊壶菌		439	WO02081668
		球等鞭金藻		442	WO02081668
	Moss	地钱	AY583465	484	Kajikawa 等, 2004

[0505]

酶	生物体 类型	物种	登录号	蛋白质大小 (氨基酸)	参考文献
$\Delta 6$ -去饱和酶	哺乳动物类	智人	NM_013402	444	Cho 等, 1999a; Leonard 等, 2000
		小家鼠	NM_019699	444	Cho 等, 1999a
	线虫	秀丽隐杆线虫	Z70271	443	Napier 等, 1998
	植物类	玻璃苣	U79010	448	Sayanova 等, 1997
		蓝蓟	AY055117 AY055118		Garcia-Maroto 等, 2002
		<i>Primula vialii</i>	AY234127	453	Sayanova 等, 2003
		<i>Anemone leveillei</i>	AF536525	446	Whitney 等, 2003
	苔藓	角齿藓	AJ250735	520	Sperling 等, 2000
		地钱	AY583463	481	Kajikawa 等, 2004
		小立碗藓	CAA11033	525	Girke 等, 1998
	真菌	高山被孢霉	AF110510 AB020032	457	Huang 等, 1999; Sakuradani 等, 1999
		畸雌腐霉	AF419296	459	Hong 等, 2002a
		卷枝毛霉菌 (<i>Mucor circinelloides</i>)	AB052086	467	NCBI*
		根霉菌	AY320288	458	Zhang 等, 2004
		异丝水霉		453	WO02081668
	硅藻	三角褐指藻	AY082393	477	Domergue 等, 2002
	细菌	集胞藻	L11421	359	Reddy 等, 1993
	藻类	金黄色破囊壶菌		456	WO02081668
双功能 $\Delta 5/\Delta 6$ 去饱和酶	鱼类	鲇	AF309556	444	Hastings 等, 2001
C20 $\Delta 8$ -去饱和酶	藻类	小眼虫	AF139720	419	Wallis and Browse, 1999
	植物类	玻璃苣	AAG43277	446	Sperling 等, 2001

[0506]

酶	生物体 类型	物种	登录号	蛋白质大小 (氨基酸)	参考文献
Δ6-延长酶	线虫	秀丽隐杆线虫	NM_069288	288	Beaudoin 等, 2000
	苔藓	小立碗藓	AF428243	290	Zank 等, 2002
		地钱	AY583464	290	Kajikawa 等, 2004
	真菌	高山被孢霉	AF206662	318	Parker-Barnes 等, 2000
	藻类	路氏巴夫藻**		501	WO 03078639
		破囊壶菌	AX951565	271	WO 03093482
		破囊壶菌**	AX214454	271	WO 0159128
PUFA-延长酶	哺乳动物类	智人	AF231981	299	Leonard 等, 2000b; Leonard 等, 2002
		褐家鼠	AB071985	299	Inagaki 等, 2002
		褐家鼠**	AB071986	267	Inagaki 等, 2002
		小家鼠	AF170907	279	Tvrđik 等, 2000
		小家鼠	AF170908	292	Tvrđik 等, 2000
	鱼类	鲇	AF532782	291 (282)	Agaba 等, 2004
		鲇**	NM_199532	266	Lo 等, 2003
	蠕虫	秀丽隐杆线虫	Z68749	309	Abbott et al 1998 Beaudoin et al 2000
	藻类	金黄色破囊壶菌**	AX464802	272	WO 0208401-A2
		路氏巴夫藻**		320	WO 03078639
Δ9-延长酶	藻类	球等鞭金藻	AF390174	263	Qi 等, 2002
		小眼虫		258	WO 08/128241
Δ5-延长酶	藻类	O. tauri	AAV67798	300	Meyer 等, 2004
		P. cordata		268	this work
		巴夫藻 CCMP459	AAV33630	277	Pereira 等, 2004b
		盐生巴夫藻	AAY15135	302	Robert 等, 2009
	硅藻	假微型海链藻	AAV67800	358	Meyer 等, 2004
	鱼类	虹鳟	CAM55862	295	WO 06/008099
	苔藓	地钱	BAE71129	348	Kajikawa 等, 2006

* <http://www.ncbi.nlm.nih.gov/> ** 功能未证实/未确定

[0507] 在本文中,术语“前端去饱和酶”指的是许多这样的一类酶,这些酶在脂质的酰基链的羧基与已经存在的不饱和部分之间引入双键,它们在结构上的特征为存在N-端细胞色素b5结构域以及典型的脂肪酸去饱和酶结构域,后者包括三个高度保守性的组氨酸盒(Napier等,1997)。

[0508] 可如下测试可用于本发明的任何延长酶或去饱和酶的活性:在细胞例如酵母细胞或植物细胞中表达编码所述酶的基因,并确定所述细胞产生LC-PUFA的能力与具有可比性

的不表达所述酶的细胞相比是否升高。

[0509] 在一个实施方式中,可用于本发明的去饱和酶和/或延长酶可纯化自微藻类。

[0510] 尽管某些酶在本文中被具体描述为是“双功能性的”,但不使用该词并不一定暗示一种具体的酶除具体提及的活性以外便不具有其他活性。

[0511] 去饱和酶类

[0512] 在本文中,术语“去饱和酶”指的是这样的酶,其能够在脂肪酸底物的酰基中引入碳碳双键,所述脂肪酸底物典型地是酯化的形式,例如是脂肪酸辅酶A酯。酰基可与磷脂例如磷脂酰胆碱(PC)发生酯化,或与酰基载体蛋白(ACP)发生酯化,或在优选的实施方式中与辅酶A发生酯化。相应地,去饱和酶通常可分为3类。在一个实施方式中,所述去饱和酶是前端去饱和酶。

[0513] 在本申请中,术语“ $\Delta 4$ 去饱和酶”指的是这样的蛋白质,其所进行去饱和酶反应导致在脂肪酸底物的自羧基端起的第4位处引入碳碳双键。“ $\Delta 4$ 去饱和酶”至少能够将DPA转换为DHA。在除哺乳动物之外的生物体中,所述从DPA产生DHA的去饱和步骤由 $\Delta 4$ 去饱和酶催化,已经从淡水原生生物小眼虫和海洋生物破囊壶菌中分离了编码此酶的基因(Qiu等,2001;Meyer等,2003)。在一个实施方式中,所述 $\Delta 4$ 去饱和酶包含SEQ ID NO:73所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:73具有至少80%相同性的氨基酸序列。

[0514] 在本申请中,“ $\Delta 5$ 去饱和酶”指的是这样的蛋白质,其所进行去饱和酶反应导致在脂肪酸底物的自羧基端起的第5位处引入碳碳双键。表1列出了 $\Delta 5$ 去饱和酶的实例。在一个实施方式中,所述 $\Delta 5$ 去饱和酶包含SEQ ID NO:26所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:26具有至少80%相同性的氨基酸序列。在另一个实施方式中,所述 $\Delta 5$ 去饱和酶包含SEQ ID NO:13所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:13具有至少53%相同性的氨基酸序列。

[0515] 在本申请中,“ $\Delta 6$ 去饱和酶”指的是这样的蛋白质,其所进行去饱和酶反应导致在脂肪酸底物的自羧基端起的第6位处引入碳碳双键。表1列出了 $\Delta 6$ 去饱和酶的实例。

[0516] 在一个实施方式中,所述 $\Delta 6$ 去饱和酶的特征为还具有至少两种、优选地全部三种以下特性,优选地在植物细胞中是这样:i)对 α -亚麻酸(ALA,18:3 $\Delta 9,12,15,\omega 3$)脂肪酸底物的 $\Delta 6$ 去饱和酶活性高于对亚油酸(LA,18:2 $\Delta 9,12,\omega 6$)的活性;ii)对作为脂肪酸底物的ALA-CoA的 $\Delta 6$ 去饱和酶活性高于对作为脂肪酸底物的连接至PC的sn-2位置的ALA的活性;和iii)对ETra具有 $\Delta 8$ 去饱和酶活性。

[0517] 在另一实施方式,所述 $\Delta 6$ 去饱和酶对 $\omega 3$ 底物的活性高于对相应的 $\omega 6$ 底物的活性,并对ALA具有活性,当由重组细胞中的外源性多核苷酸表达时以至少5%,更优选地至少7.5%,或最优选地至少10%的效率产生octadecatetraenoic acid(十八碳四烯酸,SDA,18:4 $\Delta 6,9,12,15,\omega 3$),当在酵母细胞中表达时该效率是至少35%。在一个实施方式中,所述 $\Delta 6$ 去饱和酶对作为脂肪酸底物的ALA的活性高于对LA的活性,例如,对ALA的 $\Delta 6$ 去饱和酶活性是对LA的至少两倍。在另一个实施方式中,所述 $\Delta 6$ 去饱和酶对作为脂肪酸底物的ALA-CoA的活性高于对作为脂肪酸底物的连接至PC的sn-2位置的ALA的活性,例如,对前者的 $\Delta 6$ 去饱和酶活性是对后者的至少5倍或至少10倍。

[0518] 在一个实施方式中,所述 $\Delta 6$ 去饱和酶对ETA不具有可检测到的 $\Delta 5$ 去饱和酶活性。在另一个实施方式中,所述 $\Delta 6$ 去饱和酶包含SEQ ID NO:10所示的氨基酸序列、其生物学活

性片段、或与SEQ ID NO:10具有至少77%相同性的氨基酸序列。在另一个实施方式中,所述 $\Delta 6$ 去饱和酶包含SEQ ID NO:8所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:8具有至少67%相同性的氨基酸序列。所述 $\Delta 6$ 去饱和酶还可具有 $\Delta 8$ 去饱和酶活性。

[0519] 在本申请中,“ $\Delta 8$ 去饱和酶”指的是这样的蛋白质,其所进行的去饱和酶反应导致在脂肪酸底物的自羧基端起的第8位处引入碳碳双键。所述 $\Delta 8$ 去饱和酶至少能够将EtrA转换为ETA。表1列出了 $\Delta 8$ 去饱和酶的实例。在一个实施方式中,所述 $\Delta 8$ 去饱和酶包含SEQ ID NO:24所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:24具有至少80%相同性的氨基酸序列。

[0520] 在本申请中,“ $\omega 3$ 去饱和酶”指的是这样的蛋白质,其所进行的去饱和酶反应导致在脂肪酸底物的自甲基端起的第3位处引入碳碳双键。 $\omega 3$ 去饱和酶的实例包括Pereira等(2004a)、Horiguchi等(1998)、Berberich等(1998)和Spychalla等(1997)所描述的那些。

[0521] 在一个实施方式中,所述 $\omega 3$ 去饱和酶至少能够进行将ARA转换为EPA、将dGLA转换为ETA、将 γ -亚麻酸(GLA, 18:3 $\Delta 6, 9, 12, \omega 6$)转换为SDA中的一种转换、将ARA转换为EPA和将dGLA转换为ETA这两种转换、将ARA转换为EPA和将GLA转换为SDA这两种转换、或全部这三种转换。

[0522] 在一个实施方式中,所述 $\omega 3$ 去饱和酶对具有至少三个碳碳双键的C20脂肪酸、优选地ARA,具有 $\Delta 17$ 去饱和酶活性。在另一个实施方式中,所述 $\omega 3$ 去饱和酶对具有三个碳碳双键的C18脂肪酸、优选地GLA,具有 $\Delta 15$ 去饱和酶活性。

[0523] 在本申请中,“ $\Delta 15$ 去饱和酶”指的是这样的蛋白质,其所进行的去饱和酶反应导致在脂肪酸底物的自羧基端起的第15位处引入碳碳双键。

[0524] 在本申请中,“ $\Delta 17$ 去饱和酶”指的是这样的蛋白质,其所进行的去饱和酶反应导致在脂肪酸底物的自羧基端起的第17位处引入碳碳双键。

[0525] 在另一个实施方式中,所述 $\omega 3$ 去饱和酶对酰基辅酶A底物(例如,ARA-CoA)的活性高于对相应的酰基-PC底物的活性。在本申请中,“相应的酰基-PC底物”指的是在磷脂酰胆碱(PC)的sn-2位置发生酯化的脂肪酸,其中所述脂肪酸与所述酰基辅酶A底物中的脂肪酸相同。在一个实施方式中,所述活性是至少两倍高。

[0526] 在进一步的实施方式中,所述 $\omega 3$ 去饱和酶包含SEQ ID NO:15、17或20所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:15具有至少35%相同性的氨基酸序列、与SEQ ID NO:17具有至少60%相同性的氨基酸序列和/或与SEQ ID NO:20具有至少60%相同性的氨基酸序列。

[0527] 在另一实施方式中,用于本发明的去饱和酶对酰基辅酶A底物的活性高于对相应的酰基-PC底物的活性。如上所述,“相应的酰基-PC底物”指的是在磷脂酰胆碱(PC)的sn-2位置发生酯化的脂肪酸,其中所述脂肪酸与所述酰基辅酶A底物中的脂肪酸相同。在一个实施方式中,所述活性是至少两倍高。在一个实施方式中,所述去饱和酶是 $\Delta 5$ 或 $\Delta 6$ 去饱和酶,表2给出了它们的非限制性实例。

[0528] 表2. 对酰基辅酶A底物的活性高于对相应的酰基-PC底物的活性的去饱和酶

[0529]

酶	生物体类型	物种	登录号	蛋白质大小 (氨基酸)	参考文献
$\Delta 6$ -去饱和酶	藻类	<i>Mantoniella squamata</i>	CAQ30479	449	Hoffmann 等 2008
		<i>Ostreococcus tauri</i>	AAW70159	456	Domergue 等 2005
$\Delta 5$ -去饱和酶	藻类	<i>Mantoniella squamata</i>	CAQ30478	482	Hoffmann 等 2008
	植物	<i>Anemone leveillei</i>	N/A		Sayanova 等 2007

[0530] 延长酶类

[0531] 生化证据提示脂肪酸延长作用包括4个步骤:缩合、还原、脱水以及第二步还原。就本发明而言,“延长酶”指的是这样的多肽,其在合适的生理学条件下,在存在延长作用复合体的其他成员的情况下,催化缩合步骤。已经发现,相应的酰基链的延长仅需要在细胞中异源或同源表达延长作用蛋白复合体的缩合组份(“延长酶”)。因此,所引入的延长酶能够成功募集转基因宿主的还原和脱水活性,以进行成功的酰基延长作用。就脂肪酸底物的链长及其去饱和程度而言,认为延长反应的特异性在于所述缩合组份。该组份也被认为在延长反应中具有限速作用。

[0532] 在本申请中,“ $\Delta 5$ 延长酶”至少能够将EPA转换为DPA。 $\Delta 5$ 延长酶的实例包括W02005/103253中所描述的那些。在一个实施方式中,所述 $\Delta 5$ 延长酶对EPA具有活性并以至少60%,更优选地至少65%,更优选地至少70%或最优选地至少75%的效率产生DPA。在进一步的实施方式中,所述 $\Delta 5$ 延长酶包含SEQ ID NO:6所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:6具有至少47%相同性的氨基酸序列。

[0533] 在本申请中,“ $\Delta 6$ 延长酶”至少能够将SDA转换为ETA。 $\Delta 6$ 延长酶的实例包括表1所列的那些。在一个实施方式中,所述延长酶包含SEQ ID NO:4所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:4具有至少55%相同性的氨基酸序列。

[0534] 在本申请中,“ $\Delta 9$ 延长酶”至少能够将ALA转换为EtrA。 $\Delta 9$ 延长酶的实例包括表1所列的那些。在一个实施方式中,所述 $\Delta 9$ 延长酶包含SEQ ID NO:22所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:22具有至少80%相同性的氨基酸序列。在另一个实施方式中,所述 $\Delta 9$ 延长酶包含SEQ ID NO:28所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:28具有至少81%相同性的氨基酸序列。在另一个实施方式中,所述 $\Delta 9$ 延长酶包含SEQ ID NO:94所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:94具有至少50%相同性的氨基酸序列。在另一个实施方式中,所述 $\Delta 9$ 延长酶包含SEQ ID NO:96所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:96具有至少50%相同性的氨基酸序列。在进一步的实施方式中,所述 $\Delta 9$ 延长酶包含SEQ ID NO:94或SEQ ID NO:96所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:94和/或SEQ ID NO:96具有至少50%相同性的氨基酸序列,且其中所述延长酶对 $\omega 6$ 底物的活性高于对相应的 $\omega 3$ 底物的活性。

[0535] 在本文中,术语“对 $\omega 6$ 底物的活性高于对相应的 $\omega 3$ 底物的活性”指的是所述酶对底物的相对活性不同于 $\omega 3$ 去饱和酶的作用。优选地, $\omega 6$ 底物是LA而 $\omega 3$ 底物是ALA。

[0536] 在本申请中,“具有 $\Delta 6$ 延长酶和 $\Delta 9$ 延长酶活性的延长酶”至少能够(i)将SDA转换

为ETA和(ii)将ALA转换为EtrA,且其 $\Delta 6$ 延长酶活性高于 $\Delta 9$ 延长酶活性。在一个实施方式中,所述延长酶转换SDA以产生ETA的效率是至少50%,更优选地至少60%,和/或转换ALA以产生EtrA的效率是至少6%,或更优选地至少9%。在另一个实施方式中,所述延长酶的 $\Delta 6$ 延长酶活性是 $\Delta 9$ 延长酶活性的至少大约6.5倍。在另一个实施方式中,所述延长酶不具有可检测到的 $\Delta 5$ 延长酶活性。在另一实施方式中,所述延长酶包含SEQ ID NO:4所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO:4具有至少55%相同性的氨基酸序列。

[0537] 其他酶

[0538] 在本文中,术语“二酰基甘油酰基转移酶”(EC 2.3.1.20;DGAT)指的是这样的蛋白质,其将脂肪酰基从酰基-CoA转移至二酰基甘油底物,以产生三酰基甘油。因此,术语“二酰基甘油酰基转移酶活性”指的是将酰基-CoA转移至二酰基甘油以产生三酰基甘油。已知有三种类型的DGAT,分别称为DGAT1、DGAT2和DGAT3。DGAT1多肽典型地具有10个跨膜结构域,DGAT2典型地具有2个跨膜结构域,而DGAT3典型地是可溶性的。DGAT1多肽的实例包括由来自以下物种的DGAT1基因编码的多肽:烟曲霉(*Aspergillus fumigatus*) (登录号XP_755172)、拟南芥(CAB44774)、蓖麻(*Ricinus communis*) (AAR11479)、油桐(*Vernicia fordii*) (ABC94472)、咸虾花(*Vernonia galamensis*) (ABV21945,ABV21946)、鬼箭羽(*Euonymus alatus*) (AAV31083)、秀丽隐杆线虫(AAF82410)、褐家鼠(*Rattus norvegicus*) (NP_445889)、智人(NP_036211),以及它们的变体和/或突变体。DGAT2多肽的实例包括由来自以下物种的DGAT2基因编码的多肽:拟南芥(登录号NP_566952)、蓖麻(AAY16324)、油桐(ABC94474)、拉曼被孢霉(*Mortierella ramanniana*) (AAK84179)、智人(Q96PD7,Q58HT5)、欧洲牛(*Bos taurus*) (Q70VD8)、小家鼠(*Mus musculus*) (AAK84175)、微胞藻CCMP1545,以及它们的变体和/或突变体。DGAT3多肽的实例包括由来自花生(花生,Saha,等,2006)的DGAT3基因编码的多肽,及其变体和/或突变体。

[0539] 多肽/肽

[0540] 本发明还提供多肽,它们可以是纯化的或重组的。“基本上纯化的多肽”或“纯化的多肽”指的是这样的多肽,其通常已经与在产生所述多肽的细胞中或在天然状态下与之结合的那些脂质、核酸、其他的肽以及其他污染分子分离。优选地,基本上纯化的多肽至少60%、更优选地至少75%、且更优选地至少90%不含在产生所述多肽的细胞中或在天然状态下与之天然结合的其他成分。

[0541] 术语“重组”当涉及多肽时指的是,当所述多肽由细胞或者无细胞表达系统产生时,与其天然状态相比,所述多肽产生的量或速度发生改变。在一个实施方式中,细胞在天然状态下不产生所述多肽。不过,细胞可包含使得产生的所述多肽的量发生改变的非内源性基因。本发明的重组多肽包括细胞、组织、器官或无细胞表达系统中的多肽,即尚未自产生所述多肽的转基因(重组)细胞中的其他组分分离或纯化的多肽,也包括在此类细胞或无细胞系统中产生并随后自至少某些其他成分中纯化出来的多肽。

[0542] 术语“多肽”和“蛋白质”通常可互换使用。

[0543] 一种多肽或一类多肽可通过其氨基酸序列与参比氨基酸序列的相同性(%相同性)程度来定义,或通过其与一种参比氨基酸序列的%相同性大于其与另一参比氨基酸序列的%相同性来定义。多肽与参比氨基酸序列的%相同性通常通过GAP分析(Needleman and Wunsch,1970;GCG程序)确定,使用的参数为:缺口创建罚分=5,而缺口延伸罚分=

0.3。查询序列长度为至少15个氨基酸,且GAP分析在至少15个氨基酸的区域比对两个序列。更优选地,查询序列长度为至少50个氨基酸,且GAP分析在至少50个氨基酸的区域比对两个序列。更优选地,查询序列长度为至少100个氨基酸,且GAP分析在至少100个氨基酸的区域比对两个序列。甚至更优选地,查询序列长度为至少250个氨基酸,且GAP分析在至少250个氨基酸的区域比对两个序列。甚至更优选地,GAP分析比对两个序列的全长。所述多肽或一类多肽可与所述参比多肽具有相同的酶活性,或具有不同于所述参比多肽的活性,或缺乏所述参比多肽的活性。优选地,所述多肽的酶活性是所述参比多肽的至少10%。

[0544] 在本申请中,“生物学活性”片段是本发明的多肽的一部分,其保持全长参比多肽的所定义的活性,例如具有去饱和酶和/或延长酶活性或其他酶活性。本申请中所使用的生物学活性片段不包括全长多肽。生物学活性片段可以是任何长度,条件是它们保持所定义的活性。优选地,生物学活性片段保持全长蛋白质活性的至少10%。

[0545] 就所定义的多肽或酶而言,可以理解高于上述数值的百分比相同性是优选的实施方式。因此,在适用的情况下,与最小的百分比相同性数值相比,优选地所述多肽/酶包含这样的氨基酸序列,其与相应的指定SEQ ID NO具有至少60%、更优选地至少65%、更优选地至少70%、更优选地至少75%、更优选地至少76%、更优选地至少80%、更优选地至少85%、更优选地至少90%、更优选地至少91%、更优选地至少92%、更优选地至少93%、更优选地至少94%、更优选地至少95%、更优选地至少96%、更优选地至少97%、更优选地至少98%、更优选地至少99%、更优选地至少99.1%、更优选地至少99.2%、更优选地至少99.3%、更优选地至少99.4%、更优选地至少99.5%、更优选地至少99.6%、更优选地至少99.7%、更优选地至少99.8%、且甚至更优选地至少99.9%的相同性。

[0546] 在一个实施方式中,本发明的基本上纯化的和/或重组 $\Delta 6$ 去饱和酶不包含登录号EEH58637.1或XP_001421073.1提供的序列。在另一个实施方式中,本发明的基本上纯化的和/或重组 $\omega 3$ 去饱和酶不包含登录号XP_002505536.1提供的序列。在另一个实施方式中,本发明的基本上纯化的和/或重组DGAT不包含登录号EEH54819.1提供的序列。

[0547] 本发明的多肽的氨基酸序列突变体可通过在本发明的核酸内引入合适的核苷酸改变或通过体外合成所需多肽而制备。此类突变包括,例如,氨基酸序列内残基的缺失、插入或取代。可通过缺失、插入和取代的组合而获得最终的构建体,条件是最终的肽产物具有所需的特征。

[0548] 可通过本领域所知的任何技术制备突变的(改变的)肽。例如,可对本发明的多核苷酸进行体外诱变。此类体外诱变技术包括将多核苷酸亚克隆到合适的载体内,将载体转化到“增变株(mutator)”菌株如大肠杆菌(E.coli)XL-1red(Stratagene)内并将转化的细菌增殖合适的代数。在另一个实例中,对本发明的多核苷酸进行DNA重排,参见Harayama(1998)。可采用在此所述的技术容易地筛选源自突变的/改变的DNA的产物以确定它们是否具有去饱和酶和/或延长酶活性。

[0549] 在设计氨基酸序列突变体时,突变的位置和突变的性质取决于要修饰的特征。可单个地或系列地修饰突变位点,例如通过(1)先选择保守性氨基酸进行取代,然后根据所得结果进行更加激进的选择,(2)缺失靶残基,或(3)在邻近被定位的位点处插入其他残基。

[0550] 氨基酸序列缺失通常为大约1至15个残基,更优选地大约1至10个残基,且典型地为大约1至5个连续的残基。

[0551] 取代突变是去除多肽分子内的至少一个氨基酸残基并在该处插入一个不同的残基。取代诱变的最感兴趣的位点包括那些被鉴定为是活性位点的位点。感兴趣的其他位点是那些在各种菌株或物种之间均相同的特定残基。这些位置可能对生物学活性十分重要。这些位点,特别是那些具有至少3个其他同等保守的位点的序列内的位点,优选地以相对保守的方式进行取代。表3在“示例性取代”一栏给出了此类保守性取代。

[0552] 在优选的实施方式中,与天然存在的多肽相比,突变体/变体多肽具有1或2或3或4个保守性氨基酸改变。表1给出了保守性氨基酸改变的详情。表1给出了保守性氨基酸改变的详情。本领域人员能够理解,可以合理地预期此类少数的改变不会改变在重组细胞内表达的所述多肽的活性。

[0553] 表3. 示例性取代

[0554]

原残基	示例性取代
Ala (A)	val;leu;ile;gly
Arg (R)	lys
Asn (N)	gln;his
Asp (D)	glu
Cys (C)	ser
Gln (Q)	asn;his
Glu (E)	asp
Gly (G)	pro,ala
His (H)	asn;gln
Ile (I)	leu;val;ala
Leu (L)	ile;val;met;ala;phe
Lys (K)	arg
Met (M)	leu;phe
Phe (F)	leu;val;ala
Pro (P)	gly
Ser (S)	thr
Thr (T)	ser
Trp (W)	tyr
Tyr (Y)	trp;phe
Val (V)	ile;leu;met;phe;ala

[0555] 本发明也包括在合成过程中或合成之后被差异性修饰的本发明的多肽,例如通过生物素化、苄化作用、糖基化、乙酰化、磷酸化、酰胺化、通过已知的保护/阻断基团进行衍生作用、蛋白裂解、连接于抗体分子或其他细胞配体等等。这些修饰可提高本发明的多肽的稳定性和/或生物活性。

[0556] 可以多种形式产生本发明的多肽,包括产生和回收天然多肽、产生和回收重组多肽以及化学合成所述多肽。在一个实施方式中,如下产生重组多肽,在有效产生所述多肽的条件下培养能够表达所述多肽的细胞。重组多肽可随后从细胞分泌并被回收,或从细胞中提取并回收,并优选地自污染分子中纯化出来。其可以被进一步地以化学或酶学方法修饰

或者不进行这样的修饰。用于培养的优选的细胞是本发明的重组细胞。有效的培养条件包括,但不限于,允许多肽产生的有效的培养基、生物反应器、温度、pH和氧条件。有效的培养基指的是可将细胞培养于其中以产生本发明的多肽的任何培养基。该培养基典型地包含具有可同化的碳源、氮源和磷酸源以及合适的盐、矿物质、金属和其他营养物如维生素的水性培养基。本发明的细胞可培养于常规的发酵生物反应器、摇瓶、试管、微滴定板和培养皿中。可在适合于重组细胞的温度、pH和氧含量进行培养。此类培养条件是本领域人员熟知的。产生所述多肽的更优选的细胞是植物中的细胞,特别是植物种子中的细胞。

[0557] 在本发明中,除非另有说明,术语“抗体”包括完整抗体的片段,所述片段仍保留与靶分析物的结合活性,也包括包含所述片段的化合物。此类片段包括Fv、F(ab')和F(ab')₂片段,以及单链抗体(scFv)。本发明的抗体可以是多克隆或单克隆的,并可采用本领域的标准工艺制备。

[0558] 多核苷酸

[0559] 本发明还提供多核苷酸,其可以是,例如,基因、分离的多核苷酸或嵌合DNA。其可以是DNA或RNA,可以是基因组来源的或合成的,双链的或单链的,并与碳水化合物、脂质、蛋白质或其他材料组合,以进行在此所定义的具体活性。术语“多核苷酸”在此与术语“核酸分子”互换使用。“分离的多核苷酸”指的是这样的多核苷酸,如果所述多核苷酸是从自然来源获得的,其已经与天然状态下与之结合或连接的那些多核苷酸序列分离,或是非天然存在的多核苷酸。优选地,分离的多核苷酸至少60%、更优选地至少75%、且更优选地至少90%不含与之天然结合的其他成分。

[0560] 在本文中,术语“基因”取其最广泛的含义,并包括脱氧核糖核苷酸序列,后者包含转录区和(如果翻译的话)结构基因的蛋白质编码区、并包括邻近所述编码区5'和3'端且位于距所述编码区5'和3'端之任一端至少大约2kb距离处的序列,所述序列参与该基因的表达。就此而言,基因包括调控信号,例如与特定基因天然相关的启动子、增强子、终止信号和/或多聚腺苷酸信号,或异源性调控信号,在后一种情况中,该基因称为“嵌合基因”。位于蛋白质编码区5'端且存在于mRNA上的序列称为5'非翻译序列。位于蛋白质编码区3'端或下游且存在于mRNA上的序列称为3'非翻译序列。术语“基因”涵盖cDNA和基因组形式的基因。基因的基因组形式或克隆含有被非编码序列中断的编码区,这些非编码序列称为“内含子”或“干扰区”或“干扰区”或“干扰序列”。内含子是被转录至核RNA(hnRNA)内的基因区段。内含子可含有调控元件例如增强子。内含子要自核转录产物或初级转录产物去除或“剪切掉”,因此信使RNA(mRNA)转录产物内没有内含子。mRNA的功能是在翻译过程中确定新生多肽内的氨基酸序列或次序。术语“基因”包括编码全部或部分在此所述的蛋白质的合成分子或融合分子以及上述任一的互补核苷酸序列。

[0561] 在本申请中,“嵌合DNA”指的是任何这样的DNA分子,其不是天然位置上的天然DNA分子,在此也称为“DNA构建体”。典型地,嵌合DNA或嵌合基因包含在天然情况下并不在一起的调控序列、转录的序列或蛋白编码序列。因此,嵌合DNA或嵌合基因可包含不同来源的调控序列和编码序列,或包含同一来源但以不同于天然情况下所见的方式排列的调控序列和编码序列。

[0562] 术语“内源性”在此指的是在未修饰植物与所研究的植物中在相同的发育阶段均正常存在或产生的物质。“内源性基因”指的是生物体基因组中存在于其天然位置上的天然

基因。在本申请中，“重组核酸分子”、“重组多核苷酸”或其变化形式指的是通过重组DNA技术而构建或修饰的核酸分子。术语“外来多核苷酸”或“外源性多核苷酸”或“异源性多核苷酸”等指的是通过实验操控引入细胞基因组内的任何核酸。外来或外源基因可以是插入至非天然生物体内的基因，引入至天然宿主细胞内的新位置上的天然基因，或嵌合基因。“转基因”是已经通过转化程序而引入至基因组中的基因。术语“遗传学修饰”、“转基因”及其变化形式包括通过转化或转导将基因引入至细胞内，突变细胞内的基因，并改变或调节已进行上述这些步骤的细胞或生物体或它们的子代内的基因调控。“基因组区”在此指的是基因组内的位置，其中已经将转基因或一组转基因（在此也称为簇）插入至细胞内或其祖先内。此类区域仅包括通过人工干预例如通过在此所述的方法而插入的那些核苷酸。

[0563] 术语“外源性”当涉及多核苷酸时指的是，当所述多核苷酸存在于细胞中时，与其天然状态相比，所述多核苷酸的量发生改变。在一个实施方式中，所述细胞天然条件下不包含所述多核苷酸。不过，所述细胞可包含导致所编码的多肽产生的量发生改变的非内源性多核苷酸。本发明的外源多核苷酸包括尚未与它们存在于其中的转基因（重组）细胞或无细胞表达系统中的其他成分分离的多核苷酸，也包括在此类细胞或无细胞系统中产生并随后自至少某些其他成分中纯化出来的多核苷酸。外源多核苷酸（核酸）可以是天然存在的一段连续的核苷酸，或是包含由来自不同来源的（天然存在的和/或合成的）两或多段连续的核苷酸连接形成的单一多核苷酸。典型地，此类嵌合多核苷酸至少包含编码本发明的多肽的可读框，该可读框可操纵地连接于适合在感兴趣的细胞中驱动所述可读框转录的启动子。

[0564] 在本文中，术语“不同的外源性多核苷酸”或其变化形式指的是各种多核苷酸的核苷酸序列具有至少一个、优选地多个核苷酸的不同。多核苷酸编码的RNA可以在细胞内翻译成蛋白质或者不翻译。在一个实例中，优选的是每一多核苷酸编码具有不同活性的蛋白质。在另一实例中，每一外源性多核苷酸与另一外源性多核苷酸的相同性少于95%、少于90%、或少于80%。优选地，外源性多核苷酸编码功能性蛋白质/酶。此外，优选的是不同的外源性多核苷酸是不重叠的，即每一多核苷酸是例如染色体外转移核酸的不同区域，其与另一外源性多核苷酸不重叠。至少，每一外源性多核苷酸具有转录起始和终止位点以及所指定的启动子。一个外源性多核苷酸可包含或不包含内含子。

[0565] 就所定义的多核苷酸而言，可以理解高于上述数值的百分比相同性是优选的实施方式。因此，在适用的情况下，就最小的百分比相同性数值而言，优选地所述多核苷酸包含这样的多核苷酸序列，其与相应的指定SEQ ID NO具有至少60%、更优选地至少65%、更优选地至少70%、更优选地至少75%、更优选地至少80%、更优选地至少85%、更优选地至少90%、更优选地至少91%、更优选地至少92%、更优选地至少93%、更优选地至少94%、更优选地至少95%、更优选地至少96%、更优选地至少97%、更优选地至少98%、更优选地至少99%、更优选地至少99.1%、更优选地至少99.2%、更优选地至少99.3%、更优选地至少99.4%、更优选地至少99.5%、更优选地至少99.6%、更优选地至少99.7%、更优选地至少99.8%、且甚至更优选地至少99.9%的相同性。

[0566] 在一个实施方式中，本发明的编码 $\Delta 6$ 去饱和酶的分离的和/或外源性多核苷酸不包含来自微胞藻或*Ostreococcus*基因组的推定为分别编码登录号EEH58637.1或XP_001421073.1中提供的氨基酸序列的序列。在另一个实施方式中，本发明的编码 $\omega 3$ 去饱和酶的分离的和/或外源性多核苷酸不包含来自微胞藻基因组的推定为编码登录号XP_

002505536.1中提供的氨基酸序列的序列。在另一个实施方式中,本发明的编码DGAT的分离的和/或外源性多核苷酸不包含来自微胞藻基因组的推定为编码登录号EEH54819.1中提供的氨基酸序列的序列。

[0567] 本发明的多核苷酸可在严格性条件下与编码本发明的多肽的多核苷酸选择性杂交。在本文中,在严格性条件下指的是(1)杂交过程中采用变性剂例如甲酰胺,例如,50% (v/v) 甲酰胺含0.1% (w/v) 牛血清白蛋白,0.1% Ficol1,0.1% 聚乙烯吡咯烷酮,50mM磷酸钠缓冲液 (pH 6.5) 含750mM NaCl,75mM 柠檬酸钠,42℃;或(2)采用50% 甲酰胺,5x SSC (0.75M NaCl,0.075M 柠檬酸钠),50mM磷酸钠 (pH 6.8),0.1% 焦磷酸钠,5x 登哈特溶液,超声处理的鲑精DNA (50g/ml),0.1% SDS和10% 硫酸葡聚糖,42℃,在0.2x SSC和0.1% SDS中;和/或(3)采用低离子强度和高洗涤温度,例如,0.015M NaCl/0.0015M 柠檬酸钠/0.1% SDS,50℃。

[0568] 与天然存在的分子相比,本发明的多核苷酸可具有一或多个突变,其可以是核苷酸残基的缺失、插入、或取代。相对于参比序列,具有突变的多核苷酸可以是天然产生的(即分离自天然来源)或合成的(例如,通过在上述核酸上进行定点诱变或DNA重排)。因此,本发明的多核苷酸显然可以是天然存在的或重组的。

[0569] 重组载体

[0570] 本发明的一个实施方式包括重组载体,其包含至少一种本发明的核苷酸分子,所述核苷酸分子插入至能够将所述多核苷酸分子输送至宿主细胞内的任何载体内。重组载体包括表达载体。重组载体含有异源多核苷酸序列,即所述多核苷酸序列天然情况下不与本发明的核苷酸分子邻近,且优选地其所来源的物种不同于所述多核苷酸分子所来源的物种。载体可以是RNA或DNA,真核细胞的或原核细胞的,且典型地是病毒载体、来源于病毒或质粒。质粒载体典型地包括额外的核酸序列,以便于筛选、扩增和将表达盒转化至原核细胞,例如pUC-来源的载体、pSK-来源的载体、pGEM-来源的载体、pSP-来源的载体、pBS-来源的载体、或含有一或多个T-DNA区域的二元载体。额外的核酸序列包括复制起点(用于载体自主复制)、选择标记物基因(优选地编码抗生素抗性或除草剂抗性)、独特的多克隆位点(用于提供多个位点以便将编码的核酸序列或基因插入核酸构建体中)、以及增强原核细胞和真核细胞(特别是植物细胞)的转化的序列。重组载体包含的本发明的多核苷酸可多于1种,例如3、4、5或6种本发明的多核苷酸的组合,每一个均可操纵地连接于在感兴趣的细胞中起作用的表达调控序列。优选地,所述多于1种的多核苷酸,例如3、4、5或6种多核苷酸,在单独一个重组载体内共价连接在一起,然后可将它们作为一个单独的分子引入细胞内,形成本发明的重组细胞,且优选地整合至该重组细胞的基因组内,例如转基因植物的重组细胞。由此,连接在一起的这些多核苷酸可作为单独一个基因座而被重组细胞或植物的子代整体继承。重组载体或植物可包含2个或更多个此类重组载体,每一个均含有多种多核苷酸,例如其中每一重组载体包含3、4、5或6种多核苷酸。

[0571] “可操纵地连接”在本文中指的是两个或多个核酸(例如DNA)区段之间的功能性关系。典型地,其指的是转录调控元件(启动子)与被转录序列的功能性关系。例如,启动子可操纵地连接于编码序列,例如在此所述的多核苷酸,条件是其在合适的细胞内刺激或调变编码序列的转录。通常,可操纵地连接于被转录序列的启动子转录调控元件与被转录序列在物理上是连续的,即,它们具有顺式作用。不过,某些转录调控元件,例如增强子,不需要

与被增强转录的编码序列在物理上连续或位于与之十分接近的位置。

[0572] 当有多个启动子时,每一启动子可独立地是相同的或不同的。

[0573] 重组分子,例如嵌合DNA,还可含有(a)一或多种编码信号肽序列的分泌信号,以使得所表达的本发明的多肽被产生所述多肽的细胞分泌,或为所表达的多肽提供定位,例如将多肽保留在细胞的内质网(ER)中或转移至质体内,和/或(b)含有融合序列,这导致核酸分子表达为融合蛋白。合适的信号区段的实例包括任何能够指导本发明的多肽分泌或定位的信号区段。优选的信号区段包括,但不限于,Nicotiana nectarin信号肽(US 5,939,288)、烟草延伸信号或大豆油质蛋白油体(soy oleosin oil body)结合蛋白信号。重组分子还可包含位于本发明的核酸分子的核酸序列周围和/或内部的干扰和/或非翻译序列。

[0574] 为了有助于鉴定转化体,希望核酸构建体在外来或外源性多核苷酸以外还包含选择标记物基因或筛选标记物基因,或包含选择标记物基因或筛选标记物基因作为外来或外源性多核苷酸。“标记物基因”指的是赋予表达该标记物基因的细胞以不同表型的基因,这使得此类被转化的细胞能够与不含有该标记物基因的细胞区别开。选择标记物基因赋予一种性状,以便基于对选择剂(例如,除草剂、抗生素、照射、热或其他对未转化细胞造成损害的处理)的抗性进行“选择”。筛选标记物基因(或报告基因)赋予一种性状,以便通过观察或测试,例如通过“筛选”(例如, β -葡糖醛酸糖苷酶、萤光素酶、GFP或未转化细胞中不存在的其他酶活性),而鉴定。标记物基因和感兴趣的核苷酸序列不一定要连接。标记物的实际选择并非至关重要,只要其与所选择的细胞(例如植物细胞)相组合具有功能(即,选择性)即可。标记物基因和感兴趣的外来或外源性多核苷酸不一定要连接,因为共转化不连接的基因(例如,见US 4,399,216)也是有效的植物转化方法。

[0575] 细菌选择标记物的实例是赋予抗生素抗性的标记物,例如氨苄西林、红霉素、氯霉素或四环素抗性,优选地卡那霉素抗性。用于选择植物转化体的示例性选择标记物包括,但不限于,hyg基因,其编码潮霉素B抗性;新霉素磷酸转移酶(nptII)基因,赋予对卡那霉素、巴龙霉素、G418的抗性;大鼠肝脏的谷胱甘肽-S-转移酶基因,赋予对谷胱甘肽衍生的除草剂的抗性,例如,见EP 256223;谷氨酰胺合成酶基因,过表达后赋予对谷氨酰胺合成酶抑制剂例如草胺膦(phosphinothricin)的抗性,例如,见W0 87/05327;来自产绿色链霉菌(*Streptomyces viridochromogenes*)的乙酰基转移酶基因,赋予对选择剂草胺膦的抗性,例如,见EP 275957;编码5-烯醇莽草酸-3-磷酸酯合酶(EPSPS)的基因,赋予对的耐受性N-膦酰基甲基甘氨酸,例如,见Hinchee等(1988);bar基因,赋予对双丙氨膦的抗性,例如,见W091/02071;腈水解酶基因,例如来自臭鼻克雷伯菌(*Klebsiella ozaenae*)的bxn,其赋予对溴苯腈的抗性(Stalker等,1988);二氢叶酸还原酶(DHFR)基因,赋予对氨甲蝶呤的抗性(Thillet等,1988);突变的乙酰乳酸合酶基因(ALS),其赋予对咪唑啉酮、磺酰脲或其他ALS-抑制性化合物的抗性(EP 154,204);突变的邻氨基苯甲酸合酶基因,其赋予对5-甲基色氨酸的抗性;或茅草枯脱卤素酶基因,其赋予对除草剂的抗性。

[0576] 优选的筛选标记物包括,但不限于,uidA基因,其编码 β -葡糖醛酸糖苷酶(GUS),该酶的各种显色底物是已知的; β -半乳糖苷酶基因,其所编码的酶的显色底物是已知的;水母萤光素基因(Prasher等,1985),其可用于钙敏感型生物发光检测;绿色荧光蛋白基因(Niedz等,1995)或其衍生物;萤光素酶(luc)基因(Ow等,1986),其可用于生物发光检测;和其他本领域已知的那些。“报道分子”在本说明书中指的是这样的分子,依据其化学性质,提

供可分析鉴定的信号,所述信号有助于通过参照蛋白质产物而确定启动子活性。

[0577] 优选地,核酸构建体稳定整合至细胞例如植物细胞的基因组内。因此,所述核酸可包含合适的元件,使得分子能够整合至基因组内,或者所述构建体被置于合适的载体内,以整合至细胞的染色体内。

[0578] 表达

[0579] 在本申请中,表达载体是DNA或RNA载体,其能够转化宿主细胞并实现一或多种特定多核苷酸分子的表达。优选地,表达载体还能够在宿主细胞复制。表达载体可以是原核细胞的或真核细胞的,且典型地是病毒或质粒。本发明的表达载体包括在本发明的宿主细胞内有功能(即,指导基因表达)的任何载体,宿主细胞包括细菌、真菌、内寄生物、节肢动物细胞、动物细胞和植物细胞。本发明的特别优选的表达载体可指导基因在酵母和/或植物细胞内表达。

[0580] 本发明的表达载体含有调控序列,例如转录控制序列、翻译控制序列、复制起点和与重组细胞具有相容性并控制本发明的核苷酸分子表达的其他调控序列。具体而言,本发明的多核苷酸或载体包括转录控制序列。转录控制序列是控制转录起始、延长和终止的序列。特别重要的转录控制序列是那些控制转录起始的序列,例如启动子、增强子、操纵子和阻遏物序列。合适的转录控制序列包括能够在至少一种本发明的重组细胞中起作用的转录控制序列。所使用的调控序列的选择取决于靶生物体,例如感兴趣的植物和/或靶器官或组织。此类调控序列可获自任何真核生物体,例如植物或植物病毒,或可化学合成。各种各样的此类转录控制序列是本领域人员已知的。特别优选的转录控制序列是那些在植物中具有转录指导活性的启动子,根据所使用的植物或其部分,可以是组成型的或具有阶段和/或组织特异性。

[0581] 适用于稳定转染植物细胞或建立转基因植物的多种载体已被公开,例如参见Pouwels等,Cloning Vectors:A Laboratory Manual,1985,supp.1987;Weissbach and Weissbach,Methods for Plant Molecular Biology,Academic Press,1989;和Gelvin等,Plant Molecular Biology Manual,Kluwer Academic Publishers,1990。典型地,植物表达载体包括,例如,受5' 和3' 调控序列转录调控的一或多个克隆的植物基因和显性选择标记物。此类植物表达载体还可含有启动子调控区(例如,调控区控制诱导型或组成型表达、环境调控型或发育调控型表达、或细胞或组织特异性表达)、转录起始位点、核糖体结合位点、RNA加工信号、转录终止位点、和/或多腺苷酸化信号。

[0582] 已经公开了在植物细胞内具有活性的多种组成型启动子。用于在植物中进行组成型表达的合适的启动子包括,但不限于,花椰菜花叶病病毒(CaMV) 35S启动子、玄参花叶病病毒(FMV) 35S、甘蔗杆状病毒启动子、鸭跖草黄色斑点病毒启动子、来自核酮糖-1,5-二磷酸羧化酶小亚基的光诱导型启动子、稻的细胞质磷酸丙糖异构酶启动子、拟南芥的腺嘌呤磷酸核糖基转移酶启动子、稻肌动蛋白1基因启动子、甘露氨酸合酶和章鱼碱合酶启动子、Adh启动子、蔗糖合酶启动子、R基因复合体启动子和叶绿素 α/β 结合蛋白基因启动子。这些启动子已经用于构建在植物中表达的DNA载体;参见,例如WO 84/02913。所有这些启动子均已用于构建可在植物中表达的各类重组DNA载体。

[0583] 为了在植物的源组织例如叶、种子、根或茎中进行表达,优选地,本发明所用的启动子在这些特定组织中具有相对高的表达。为此,可为基因选择多种具有组织或细胞特异

性或增强型的启动子。文献报导的此类启动子的实例包括豌豆的叶绿体谷氨酰胺合成酶GS2启动子、小麦的叶绿体果糖-1,6-二磷酸酶启动子、马铃薯的核光合ST-LS1启动子、拟南芥的丝氨酸/苏氨酸激酶启动子和葡糖淀粉酶(CHS)启动子。以下启动子也被报导在具有光合作用活性的组织中具有活性,东方落叶松(*Larix laricina*)的核酮糖-1,5-二磷酸羧化酶启动子、松Cab基因Cab6的启动子、小麦Cab-1基因启动子、菠菜Cab-1基因启动子、稻Cab1R基因启动子、玉蜀黍的丙酮酸-正磷酸双激酶(PPDK)启动子、烟草Lhcb1*2基因的启动子、拟南芥Suc2蔗糖-H³⁰协同转运子启动子、菠菜的类囊体膜蛋白基因启动子(PsaD、PsaF、PsaE、PC、FNR、AtpC、AtpD、Cab、RbcS)。

[0584] 叶绿素 α/β -结合蛋白的其他启动子也可用于本发明,例如白芥菜(*Sinapis alba*)LhcB基因和PsbP基因的启动子。应答于环境、激素、化学品和/或发育信号而被调控的各种植物基因启动子也可用于在植物细胞中表达RNA-结合蛋白基因,包括受到以下调控的启动子:(1)热;(2)光(例如豌豆RbcS-3A启动子、玉蜀黍RbcS启动子);(3)激素,例如脱落酸;(4)致伤(wounding)(例如WunI);或(5)化学制品,例如茉莉酮酸甲酯、水杨酸、类固醇激素、醇、Safeners(WO 9706269);或者使用(6)器官特异性启动子也是有利的。

[0585] 在本文中,术语“植物贮藏器官特异性启动子”指的是这样的启动子,与其他植物组织相比,其优先在植物的贮藏器官中指导基因转录。优选地,所述启动子仅在贮藏器官中指导感兴趣的基因的表达,和/或在植物的其他部分例如叶中,通过Northern印迹分析和/或RT-PCR检测不到感兴趣的基因的表达。典型地,启动子在贮藏器官的生长和发育过程中驱动基因表达,特别是在贮藏器官中合成并积聚贮藏化合物的阶段中。此类启动子可在整个植物贮藏器官中驱动基因表达,或仅在其部分中,例如双子叶植物种子的种皮、胚或子叶中或单子叶植物种子的胚乳或糊粉层中驱动基因表达。

[0586] 为了在植物的库组织(sink tissues)中表达,例如在马铃薯的块茎、番茄的果实、或大豆、芸苔、棉花、玉蜀黍、小麦、稻和大麦的种子中表达,本发明所用的启动子优选地在这些特异性组织中具有相对高的表达。已知多种具有块茎特异性或增强型表达的基因启动子,包括I类patatin启动子、马铃薯块茎ADPGPP基因启动子(大和小亚单位)、蔗糖合酶启动子、包括22kD蛋白质复合物和蛋白酶抑制剂的主要块茎蛋白的启动子、颗粒结合型淀粉合酶基因(GBSS)的启动子以及其他I和II类patatin启动子。其他启动子也可用于在特定组织例如种子或果实中表达蛋白质。可使用 β -伴大豆球蛋白启动子或其他种子特异性启动子例如napin、玉米醇溶蛋白、核丝(linin)和菜豆球蛋白启动子。还可使用根特异性启动子。所述启动子的实例是酸性壳多糖酶基因启动子。还可使用已经得到鉴定的CaMV 35S启动子的根特异性亚结构域来实现在根组织中的表达。

[0587] 在一个特别优选的实施方式中,启动子在脂肪酸和油进行生物合成的组织和器官中指导表达。在种子发育过程中,此类启动子在合适的时机起作用,改变种子的油的组成。

[0588] 在进一步特别优选的实施方式中,且在本发明的一些方面,启动子是植物贮藏器官特异性启动子。在一个实施方式中,所述植物贮藏器官特异性启动子是种子特异性启动子。在更加优选的实施方式中,相对于在种子的胚或植物的其他器官例如叶中的表达而言,所述启动子优先在双子叶植物的子叶或单子叶植物的胚乳中指导表达。用于种子特异性表达的优选的启动子包括:i)来自编码参与种子中脂肪酸生物合成和积聚的酶(例如去饱和酶和延长酶)的基因的启动子,ii)来自编码种子贮藏蛋白的基因的启动子,和iii)来自编

码参与种子中碳水化合物生物合成和积聚的酶的基因的启动子。合适的种子特异性启动子是油料种子油菜napin基因启动子(US 5,608,152)、蚕豆(*Vicia faba*)USP启动子(Baumlein等,1991)、拟南芥油质蛋白启动子(WO 98/45461)、菜豆(*Phaseolus vulgaris*)的菜豆球蛋白启动子(美国5,504,200)、芸苔属Bce4启动子(WO 91/13980)或豆球蛋白B4启动子(Baumlein等,1992),以及在单子叶植物例如玉蜀黍、大麦、小麦、黑麦、稻等中引起种子特异性的表达启动子。合适的著名启动子是大麦lpt2或lpt1基因启动子(WO 95/15389和WO 95/23230)或WO 99/16890所述的启动子(大麦的大麦醇溶蛋白基因的启动子、稻的谷蛋白基因的启动子、稻的oryzin基因的启动子、稻的醇溶谷蛋白基因的启动子、小麦的麦醇溶蛋白基因的启动子、小麦的谷蛋白基因的启动子、玉蜀黍的玉米醇溶蛋白基因的启动子、燕麦的谷蛋白基因的启动子、蜀黍kasirin基因的启动子、黑麦的黑麦醇溶蛋白基因的启动子)。其他启动子包括Broun等(1998)、Potenza等(2004)、US 20070192902和US 20030159173所述的启动子。在一个实施方式中,所述种子特异性启动子优先表达于种子的特定部分,例如子叶或胚乳。子叶特异性启动子的实例包括,但不限于,FP1启动子(Ellerstrom等,1996)、豌豆豆球蛋白启动子(Perrin等,2000)和豆植物凝血素启动子(Perrin等,2000)。胚乳特异性启动子的实例包括,但不限于,玉米的玉米醇溶蛋白-1启动子(Chikwamba等,2003)、稻谷蛋白-1启动子(Yang等,2003)、大麦D-大麦醇溶蛋白启动子(Horvath等,2000)和小麦HMW麦谷蛋白启动子(Alvarez等,2000)。在进一步的实施方式中,所述种子特异性启动子在胚中和/或在种子萌发后不表达,或是仅以低水平表达。

[0589] 在另一个实施方式中,所述植物贮藏器官特异性启动子是块茎特异性启动子。实例包括,但不限于,马铃薯patatin B33、PAT21和GBSS启动子,以及甜马铃薯sporamin启动子(综述见Potenza等,2004)。在优选的实施方式中,相对于块茎的外层(皮)或块茎的胚而言,所述启动子优先在块茎的木髓中指导表达。

[0590] 在另一个实施方式中,所述植物贮藏器官特异性启动子是果实特异性启动子。实例包括,但不限于,番茄多聚半乳糖醛酸酶-E8和Pds启动子以及苹果ACC氧化酶启动子(综述见Potenza等,2004)。在优选的实施方式中,相对于果实的皮或果实内的种子而言,所述启动子优先在果实的可食用部分例如果实的木髓中指导表达。

[0591] 可从选定的启动子衍生出5'非翻译前导序列以表达本发明的多核苷酸的异源基因序列,或者,5'非翻译前导序列可以与待产生的酶的编码区是异源的,需要的话还可对其进行特异性修饰以便增加mRNA的翻译。如何优化转基因的表达可参见Kozziel等(1996)的综述。5'非翻译区也可得自植物病毒RNA(烟草花叶病病毒、烟草刻蚀病毒、玉蜀黍矮花叶病病毒、紫花苜蓿花叶病病毒等等)、合适的真核细胞基因、植物基因(小麦和玉蜀黍叶绿素a/b结合蛋白基因前导区)、或合成的基因序列。本发明不限于其中所述非翻译区衍生自伴随启动子序列的5'非翻译序列的构建体。前导序列也可衍生自无关的启动子或编码序列。可用于本发明的前导序列包括玉蜀黍Hsp70前导区(美国5,362,865和美国5,859,347)和实施例8中例证的TMV ω 元件。

[0592] 转录的终止是通过嵌合载体内与感兴趣的多核苷酸可操纵地连接的3'非翻译DNA序列而实现的。重组DNA分子的3'非翻译区含有多腺苷酸化信号,其在植物中的作用是导致在RNA的3'端添加腺嘌呤核苷酸。3'非翻译区可得自在植物细胞中表达的各种基因。就此能力而言,常用的是胭脂碱合酶3'非翻译区、豌豆的小亚单位核酮糖二磷酸羧化酶(Rubisco)

基因的3'非翻译区、大豆的7S种子贮藏蛋白基因的3'非翻译区。含有聚腺苷酸化信号的土壤杆菌 (*Agrobacterium*) 肿瘤诱导型 (Ti) 质粒基因的3'转录的非翻译区也是合适的。

[0593] 可采用重组DNA技术来提高转化的多核苷酸分子的表达,例如提高操控宿主细胞内多核苷酸分子的拷贝数、所述多核苷酸分子的转录效率、所得转录产物的翻译效率、以及转录后修饰的效率。可用于增加本发明的核苷酸分子表达的重组技术包括,但不限于,将多核苷酸分子可操纵地连接于高拷贝数质粒,将多核苷酸分子整合至一或多个宿主细胞染色体中,给质粒添加载体稳定序列,置换或改进转录控制信号(例如启动子、操纵子、增强子),置换或改进翻译控制信号(例如核糖体结合位点、Shine-Dalgarno序列),改进本发明的核苷酸分子以符合宿主细胞内的密码子使用情况,以及去除使转录产物不稳定的序列。

[0594] 转移核酸

[0595] 本发明的转移核酸至少包含一个、优选地两个边界序列,以及外源性多核苷酸。转移核酸可编码或不编码选择标记物。优选地,转移核酸形成细菌中的二元载体的一部分,其中所述二元载体还包含使得载体在细菌中复制的元件或能够选择或维持含有所述载体的细菌细胞的元件。转移至真核细胞后,二元载体的转移核酸成分能够整合至真核细胞的基因组中。

[0596] 在本文中,术语“染色体外转移核酸”指的是这样的核酸分子,其能够从细菌例如土壤杆菌转移至真核细胞例如植物的叶细胞。染色体外转移核酸是熟知的遗传学元件,能够被转移并随后将其边界内所含的核苷酸序列整合至受者细胞的基因组内。就此而言,转移核酸典型地在其两侧有两个“边界”序列,不过在一些情况中仅在一侧有单个边界,并在转移过程中随机产生所转移的核酸的另一端。合意的外源性多核苷酸典型地位于转移核酸的左侧边界样序列与右侧边界样序列之间。转移核酸内所含的合意的多核苷酸可以可操纵地连于多种不同的启动子和终止子调控元件,以有助于其表达,即多核苷酸的转录和/或翻译。来自土壤杆菌例如根癌土壤杆菌 (*Agrobacterium tumefaciens*) 或毛根土壤杆菌 (*Agrobacterium rhizogenes*) 的T-DNA及其人工变体/突变体可能是转移核酸的表征的最清楚的实例。另一实例是P-DNA (“植物-DNA”),其包含来自植物的T-DNA边界样序列。

[0597] 在本申请中,“T-DNA”指的是,例如,来自根癌土壤杆菌Ti质粒或毛根土壤杆菌Ri质粒的T-DNA,或是其具有T-DNA (转移的-DNA) 功能的人工变体。T-DNA可包含整个T-DNA,包括右侧和左侧边界序列,但仅需要包含用于转移的顺式所需的最少序列,即右侧和T-DNA边界序列。本发明的T-DNA在其右侧和左侧边界序列(如果有的话)之间的任何位置插入了外源性多核苷酸,后者的两侧为位点特异性重组酶的靶位点。编码用于将T-DNA转移至植物细胞的反式所需的因子的序列,例如vir基因,可插入至T-DNA,或可与T-DNA存在于同一复制子中,或优选地在土壤杆菌宿主中的相适应的复制子中式反式的。此类“二元载体系统”是本领域已知的。

[0598] 在本申请中,“P-DNA”指的是从植物基因组中分离的转移核酸,或是其人工的变体/突变体,并在每一端或至少一端包含T-DNA边界样序列。所述边界样序列优选地与来自土壤杆菌例如根癌土壤杆菌或毛根土壤杆菌的T-DNA边界序列具有至少50%、至少60%、至少70%、至少75%、至少80%、至少90%或至少95%,但低于100%的序列相同性。因此,P-DNA可替代T-DNA,用于将P-DNA中所含的核苷酸序列从例如土壤杆菌转移至另一种细胞。在插入待转移的外源性多核苷酸之前,可修饰P-DNA以便克隆,P-DNA优选地应该不编码任何

蛋白质。P-DNA的特征在于其含有至少右侧边界序列,优选地也含有左侧边界序列。

[0599] 在本申请中,转移核酸的“边界”序列可分离自选定的生物体例如植物或细菌,或是其人工变体/突变体。边界序列促使并有助于转移与之连接的外源性多核苷酸,并有助于其整合至受者细胞基因组内。

[0600] 在一个实施方式中,边界序列的长度为5-100bp、10-80bp、15-75bp、15-60bp、15-50bp、15-40bp、15-30bp、16-30bp、20-30bp、21-30bp、22-30bp、23-30bp、24-30bp、25-30bp、或26-30bp。

[0601] 来自土壤杆菌的T-DNA的边界序列是本领域熟知的,见Lacroix等(2008)、Tzfira和Citovsky(2006)和Glevin(2003)。P-DNA的边界序列可分离自任何植物,例如马铃薯和小麦。在一个实施方式中,所述P-DNA具有核酸序列ANGATNTATN6GT (SEQ ID NO:109),其中“N”是任何核苷酸,例如“A,”“G,”“C,”或“T”。可用于本发明的其他边界序列的实例包括,但不限于,

[0602] TGACAGGATATATTGGCGGGTAAAC (SEQ ID NO:110);

[0603] TGGCAGGATATATTGTGGTGTAAC (SEQ ID NO:111);

[0604] TGGCAGGATATATACCGTTGTAATT (SEQ ID NO:112);

[0605] CGGCAGGATATATTCAATTGTAATT (SEQ ID NO:113);

[0606] TGGTAGGATATATACCGTTGTAATT (SEQ ID NO:114);

[0607] TGGCAGGATATATGGTACTGTAATT (SEQ ID NO:115);

[0608] YGRYAGGATATATWSNVBKGTAAWY (SEQ ID NO:116);

[0609] CGGCAGGATATATCCTGATGTAAAT (SEQ ID NO:117);

[0610] TGGCAGGAGTTATTCGAGGGTAAAC (SEQ ID NO:118);

[0611] TGACAGGATATATCGTGATGTCAAC (SEQ ID NO:119);

[0612] GGGAAGTACATATTGGCGGGTAAAC (SEQ ID NO:120);

[0613] TTACAGGATATATTAATATGTATGA (SEQ ID NO:121);

[0614] TAACATGATATATTCCTTGTAAT (SEQ ID NO:122);

[0615] TGACAGGATATATGGTAATGTAAAC (SEQ ID NO:123);和

[0616] TGGCAGGATATATACCGATGTAAAC (SEQ ID NO:124),

[0617] 其中*Y=C或T;R=A或G;K=G或T;W=A或T;S=C或G;V=A、C、或G;B=C、G、或T。

[0618] 尽管传统上仅采用土壤杆菌将基因转移至植物细胞,但现在已经鉴定/开发了许多与土壤杆菌类似的方式起作用的系统。一些非土壤杆菌物种已经被遗传改造为适合进行基因转移(Chung等,2006;Broothaerts等,2005)。这些包括根瘤菌NGR234、苜蓿中华根瘤菌和百脉根中慢生根瘤菌。通过给细菌提供转化过程所需的机构(即,由土壤杆菌Ti-质粒编码的一组致病基因和位于一单独的小二元质粒上的T-DNA区段)而使得细菌适合于基因转移。这样的工程化细菌能够转化不同的植物组织(叶盘(leafdisks)、愈伤组织和卵形组织(oval tissue))、单子叶植物或双子叶植物、以及多种不同植物(例如,烟草、稻)。

[0619] 几十年前,通过将哺乳动物细胞与携带质粒的大肠杆菌的原生质体融合而首次实现了将真核表达质粒从细菌直接转移至真核宿主(Schaffner,1980)。其后,能够将基因输送至哺乳动物细胞的细菌的数量稳定增加(Weiss,2003),这是由四组人员独立发现的(Sizemore等1995;Courvalin等,1995;Powell等,1996;Darji等,1997)。

[0620] 已经发现,弱化的弗氏志贺菌、鼠伤寒沙门氏菌或经由弗氏志贺菌的致病质粒(pWR100)而具有侵袭性的大肠杆菌,在侵入宿主细胞后能够转移表达质粒,并因代谢的弱化而发生细胞内死亡。通过鼻部或口腔粘膜施用此类重组志贺菌或沙门氏菌,诱导了针对该表达质粒所编码的抗原的免疫应答。同时,已发现的能够在体外和体内将表达质粒转移至哺乳动物宿主细胞的细菌的数量已经增加了一倍以上,已经报道的伤寒沙门氏菌(*S. typhi*)、猪霍乱沙门氏菌、单核细胞增多性李司忒菌、假结核病耶尔森菌和小肠结肠炎耶尔森菌(Fennelly等,1999;Shiau等,2001;Dietrich等,1998,2001;Hense等,2001;Al-Mariri等,2002)。

[0621] 一般而言,可以假设所有能够进入宿主细胞的细胞质并在其中裂解的细菌(如弗氏志贺菌或单核细胞增多性李司忒菌)都应该可以转移DNA。这种情况称为“流产”或“自杀”侵入,因为细菌要裂解才能实现DNA转移(Grillot-Courvalin等,1999)。此外,甚至停留在吞噬泡中的许多细菌(如鼠伤寒沙门氏菌)也能做到这一点。因此,已经被工程化为具有侵袭性但不能发生吞噬体逃逸的大肠杆菌重组实验株,仍旧能够将其质粒负载输送至被感染的哺乳动物细胞的核中(Grillot-Courvalin等,1998)。此外,最近也发现根癌土壤杆菌可将转基因引入哺乳动物细胞(Kunik等,2001)。

[0622] 采用染色体外转移元件的转移方法典型地将多个拷贝的该元件转移至受者细胞内。在本文中,术语“瞬时转染”指的是尽管有些外源性多核苷酸可以稳定整合至细胞的基因组内,但并不筛选出稳定整合的细胞。因此,细胞内的许多转移核酸仍然处于染色体外,例如,转移至受者细胞内的外源性多核苷酸中有超过90%的拷贝没有整合至基因组内。

[0623] 在本文中,术语“转染”、“转化”及其变化形式通常可互换使用。“转染的”或“转化的”细胞是经操作引入了外源性多核苷酸的细胞,或是由所述细胞产生的子代细胞。

[0624] 重组细胞

[0625] 本发明还提供重组细胞,优选地重组植物细胞,其是以一或多种重组分子转化的宿主细胞,所述重组分子例如为本发明的多核苷酸、嵌合DNAs或重组载体。重组细胞可包含它们的任意组合,例如两种或三种重组载体,或重组载体和一或多种额外的多核苷酸或嵌合DNA。本发明的合适的细胞包括任何能够被本发明的多核苷酸、嵌合DNA或重组载体转化的细胞,例如,编码本发明的多肽或酶的分子。所述细胞优选地是由此能够用于产生LC-PUFA的细胞。重组细胞可以是培养的细胞、体外的细胞、或生物体内的细胞,例如植物或器官例如种子或叶内的细胞。优选地,所述细胞处于植物内,更优选地处于植物的种子内。

[0626] 被引入所述多核苷酸的宿主细胞可以是未转化的细胞,或是已经以一或多种核酸分子转化的细胞。此类核酸分子可与LC-PUFA合成有关或无关。本发明的宿主细胞可以内源性(即,天然地)具有产生本发明的蛋白质的能力,在这种情况下,由其产生的所述重组细胞具有增强的产生所述多肽的能力,或者,所述宿主细胞仅在以一或多种本发明的多核苷酸转化之后才能产生此类蛋白质。在一个实施方式中,本发明的重组细胞具有增强的合成长链多不饱和脂肪酸的能力。在本文中,术语“具有增强的合成长链多不饱和脂肪酸的能力的细胞”是一个将本发明的重组细胞与缺乏本发明的多核苷酸的宿主细胞相比的比较性的术语,其中与天然细胞相比,所述重组细胞产生更多的长链多不饱和脂肪酸,或更高浓度的LC-PUFA例如EPA、DPA或DHA(相对于其他脂肪酸)。具有增强的合成另一种产物(例如另一种脂肪酸、脂质、碳水化合物如淀粉、RNA分子、多肽、药物或其他产物)的能力的细胞具有相应

的含义。

[0627] 本发明的宿主细胞可以是任何能够产生至少一种在此所述的蛋白质的细胞,包括细菌、真菌(包括酵母)、寄生虫、节肢动物、底物和植物细胞。细胞可以是原核细胞或真核细胞。优选的宿主细胞是酵母和植物细胞。在优选的实施方式中,所述植物细胞是种子细胞,特别是种子的子叶或胚乳内的细胞。在一个实施方式中,所述细胞是动物细胞或海藻细胞。动物细胞可以是任何类型的动物细胞,例如,非人类动物细胞、非人类脊椎动物细胞、非人类哺乳动物细胞,或水生动物如鱼类或甲壳类动物、无脊椎动物、昆虫等的细胞。节肢动物细胞的非限制性实例包括昆虫细胞,例如草地夜蛾(*Spodoptera frugiperda*) (Sf) 细胞,如 Sf9、Sf21、粉纹夜蛾(*Trichoplusia ni*) 细胞和果蝇 S2 细胞。可用作本发明的宿主细胞的细菌细胞的实例是聚球藻(*Synechococcus*) (也称集胞藻(*Synechocystis*)), 例如细长聚球藻(*Synechococcus elongatus*)。

[0628] 所述细胞可来自适合发酵方法的生物体。在本文中,术语“发酵方法”指的是任何发酵方法或任何包括发酵步骤的方法。发酵方法包括,但不限于,用于产生醇类(例如,乙醇、甲醇、丁醇);有机酸类(例如,柠檬酸、乙酸、甲叉丁二酸、乳酸、葡糖酸);酮类(例如,丙酮);氨基酸(例如,谷氨酸);气体(例如, H_2 和 CO_2);抗生素(例如,青霉素和四环素);酶;维生素(例如,核黄素、 β -胡萝卜素);和激素的发酵方法。发酵方法还包括消费性酿酒业(例如啤酒和葡萄酒)、乳品业(例如发酵乳产品)、皮革业和烟草业所使用的发酵方法。优选的发酵方法包括醇类发酵方法,这是本领域人员所熟知的。优选的发酵方法是厌氧发酵方法,这是本领域人员所熟知的。合适的发酵细胞,典型地是微生物,能够将糖例如葡萄糖或麦芽糖直接或间接发酵为(即转换为)所需的发酵产物。发酵微生物的实例包括真菌生物体,例如酵母。在本文中,“酵母”包括酵母属(*Saccharomyces* spp.), 酿酒酵母(*Saccharomyces cerevisiae*)、卡尔酵母(*Saccharomyces carlbergensis*)、假丝酵母、克鲁维酵母、毕赤酵母属物种、汉逊酵母、木霉、斯氏油脂酵母(*Lipomyces starkey*) 和解脂耶氏酵母(*Yarrowia lipolytica*)。优选的酵母包括酵母属物种的菌株,特别是酿酒酵母。

[0629] 转基因植物

[0630] 本发明还提供包含本发明的细胞的植物,例如包含一或多种本发明的多核苷酸的转基因植物。术语“植物”在本文中作为名词指的是整个植物,但作为形容词使用时指的是存在于植物中的、自植物获得的、来源于植物的或与植物有关的任何物质,例如,植物器官(例如,叶、茎、根、花)、单个的细胞(例如,花粉)、种子、植物细胞等等。术语“植物部分”指的是所有包括植物 DNA 的植物部分,包括营养性结构(vegetative structures) 例如,叶或茎、根、开花器官或结构、花粉、种子、种子部分例如胚、胚乳、盾片或种皮、植物组织例如维管组织、以及它们的细胞和子代。

[0631] “转基因植物”、“遗传学修饰的植物”或其变化形式指的是这样的植物,其含有基因构建体(“转基因”),该基因构建体在同样物种、品种或栽培变种的野生型植物中是不存在的。就本发明而言,在此所述的转基因植物包括采用重组技术进行遗传学改造以便在所需植物或植物器官中产生至少一种本文所述的多肽的植物及其子代。转基因植物细胞和转基因植物部分也具有相应的含义。“转基因”在此具有生物技术领域的通常含义,并包括通过重组 DNA 或 RNA 技术而产生或改变并被引入至本发明的细胞、优选地植物细胞内的遗传学序列。转基因包括的遗传序列可来源于与引入该转基因的植物细胞属于同样物种、品种或

栽培变种的植物细胞,或者来源于不同的物种、品种或栽培变种的植物细胞,或是来源于非植物细胞。典型地,转基因已经通过人为操作而引入细胞例如植物,例如通过转化,但也可采样本领域人员所知的任何方法。

[0632] 术语“种子”和“籽粒”在此可互换使用。“籽粒”指的是成熟的籽粒例如收获的籽粒或虽然仍在植物上但已适合收获的籽粒,不过,根据上下文也可指吸涨或萌芽后的籽粒。成熟籽粒具有的含水量通常低于大约18-20%。“发育中的种子”在本文中指的是成熟之前的种子,典型地见于授粉后或开花期后的植物繁殖结构,不过也可指与植物分离的成熟之前的此类种子。

[0633] 在本文中,术语“植物贮藏器官”指的是植物专门用于以例如蛋白质、碳水化合物、脂肪酸和/或油的形式储藏能量的部分。植物贮藏器官的实例是种子、果实、块状根和块茎。本发明优选的植物贮藏器官是种子。

[0634] 在本文中,术语“具有正常的表型”指的是,本发明的遗传学改造的植物或植物器官,特别是贮藏器官例如种子、块茎或果实,与未改造的植物或植物器官相比,其生长和繁殖能力没有显著降低。在一个实施方式中,所述具有正常的表型的遗传学改造的植物或植物器官包含可操纵地连接于植物贮藏器官特异性启动子的编码沉默阻抑物的外源性多核苷酸,并与不包含所述多核苷酸的同基因植物或器官具有基本上相同的生长或繁殖能力。优选地,如果生长在相同的条件下,所产生的生物量、生长速度、萌发率、贮藏器官大小、种子大小和/或活种子数量,与缺乏所述外源性多核苷酸的植物相比,不低于后者的90%。该术语不包括植物的如下特征,所述特征可能不同于野生型植物,但并不影响植物的商业目的用途,例如,幼苗叶子的芭蕾舞短裙表型(ballerina phenotype)。

[0635] 本发明所提供的或者预期可用于本发明的植物包括单子叶植物和双子叶植物。在优选的实施方式中,本发明的植物是农作物(例如,谷类和豆类、玉蜀黍、小麦、马铃薯、木薯(tapioca)、稻、蜀黍、粟、木薯、大麦、或豌豆)或其他豆科植物。可种植植物用于产生可食用的根、块茎、叶、茎、花或果实。植物可以是蔬菜或观赏植物。本发明的植物可以是:玉米(*Zea mays*)、芸苔(甘蓝型油菜(*Brassica napus*))、芜菁(*Brassica rapa ssp.*)、亚麻(*Linum usitatissimum*)、紫花苜蓿(*Medicago sativa*)、稻(*Oryza sativa*)、黑麦(*Secale cereale*)、蜀黍(*Sorghum bicolor*、*Sorghum vulgare*)、向日葵(*Helianthus annuus*)、小麦(*Triticum aestivum*)、大豆(*Glycine max*)、烟草(*Nicotiana tabacum*)、马铃薯(*Solanum tuberosum*)、花生(*Arachis hypogaea*)、棉花(*Gossypium hirsutum*)、蕃薯(*Lopmoeba batatus*)、木薯(*Manihot esculenta*)、咖啡(*Cofea spp.*)、椰子(*Cocos nucifera*)、凤梨(*Anana comosus*)、柑橘(*Citrus spp.*)、可可(*Theobroma cacao*)、茶(*Camellia senensis*)、香蕉(*Musa spp.*)、鳄梨(*Persea americana*)、无花果(*Ficus casica*)、番石榴(*Psidium guajava*)、芒果(*Mangifer indica*)、橄榄(*Olea europaea*)、木瓜(*Carica papaya*)、腰果(*Anacardium occidentale*)、夏威夷果(*Macadamia intergrifolia*)、杏仁(*Prunus amygdalus*)、糖用甜菜(*Beta vulgaris*)、燕麦、或大麦。

[0636] 在优选的实施方式中,植物是被子植物。

[0637] 在一个实施方式中,植物是油料种子植物,优选地是油料种子农作物植物。在本文中,“油料种子植物”是用于从植物的种子中商品化生产油的植物物种。油料种子植物可以是油料种子油菜(例如芸苔)、玉蜀黍、向日葵、大豆、蜀黍、亚麻(亚麻子)或糖用甜菜。此外,

油料种子植物可以是其他的芸苔属 (Brassicaceae)、棉花、花生、罌粟、芥菜、蓖麻子、芝麻、红花、或产生坚果的植物。植物在其果实中可产生高水平的油,例如橄榄、油椰子或椰子。本发明可使用的园艺植物是莴苣、菊苣、或芸苔属蔬菜,包括甘蓝、椰菜、或花椰菜。本发明可用于烟草、葫芦、胡萝卜、草莓、番茄或胡椒。

[0638] 在另一个优选的实施方式中,用于产生本发明的转基因植物的非转基因植物 (特别是在其种子中) 产生的油 i) 具有低于 20%、低于 10% 或低于 5% 的 18:2 脂肪酸和/或 ii) 具有低于 10% 或低于 5% 的 18:3 脂肪酸。

[0639] 在优选的实施方式中,转基因植物对于已经被引入的每个基因 (转基因) 均是纯合子,这样其子代的所需表型不会出现分离。转基因植物对于引入的转基因也可以是杂合子,优选地对于所述转基因均一地是杂合子,例如,在从杂交种子生长而来的 F1 子代中即是如此。此类植物可具有一些优点,例如本领域熟知的杂种优势。

[0640] 在相关的情况中,转基因植物还可包含额外的转基因,所述额外的转基因编码参与产生 LC-PUFA 的酶,例如,但不限于,Δ 6 去饱和酶、Δ 9 延长酶、Δ 8 去饱和酶、Δ 6 延长酶、Δ 5 去饱和酶、ω 3 去饱和酶、Δ 4 去饱和酶、Δ 5 延长酶、二酰基甘油酰基转移酶、Δ 17 去饱和酶、Δ 15 去饱和酶和/或 Δ 12 去饱和酶。具有一或多种这些活性的此类酶的实例是本领域已知的,包括本文和 WO 05/103253 中所述那些 (见,例如,WO 05/103253 的表 1)。在具体的实例中,转基因植物至少包含编码以下各项的外源性多核苷酸:

[0641] a) Δ 4 去饱和酶、Δ 5 去饱和酶、Δ 6 去饱和酶、Δ 5 延长酶和 Δ 6 延长酶,

[0642] b) Δ 4 去饱和酶、Δ 5 去饱和酶、Δ 8 去饱和酶、Δ 5 延长酶和 Δ 9 延长酶,

[0643] c) Δ 4 去饱和酶、Δ 5 去饱和酶、Δ 6 去饱和酶、Δ 5 延长酶、Δ 6 延长酶和 Δ 15 去饱和酶,

[0644] d) Δ 4 去饱和酶、Δ 5 去饱和酶、Δ 8 去饱和酶、Δ 5 延长酶、Δ 9 延长酶和 Δ 15 去饱和酶,

[0645] e) Δ 4 去饱和酶、Δ 5 去饱和酶、Δ 6 去饱和酶、Δ 5 延长酶、Δ 6 延长酶和 Δ 17 去饱和酶,或

[0646] f) Δ 4 去饱和酶、Δ 5 去饱和酶、Δ 8 去饱和酶、Δ 5 延长酶、Δ 9 延长酶和 Δ 17 去饱和酶。

[0647] 植物的转化

[0648] 可通过本领域已知的技术产生转基因植物,这些技术例如可参见 A.Slater 等, Plant Biotechnology-The Genetic Manipulation of Plants, Oxford University Press (2003) 和 P.Christou and H.Klee, Handbook of Plant Biotechnology, John Wiley and Sons (2004)。

[0649] 在本文中,术语“稳定转化”、“稳定转化的”及其变化形式指的是外源核酸分子整合至细胞的基因组内,由此在细胞分裂过程中传递给子代,而无需对其进行阳性选择。可通过本领域已知的任何手段选择稳定的转化体或其子代,例如在染色体 DNA 上进行 Southern 印迹或基因组 DNA 的原位杂交。

[0650] 土壤杆菌介导的转移是将基因引入植物细胞的广泛适用的系统,这是因为 DNA 可被引入整个植物组织或植物器官或培养的外植体的细胞内,以便瞬时表达或 DNA 稳定整合至植物细胞基因组内。使用土壤杆菌介导的植物整合载体将 DNA 引入植物细胞是本领域熟

知的(参见,例如,US5177010、US 5104310、US 5004863或US 5159135)。待转移的DNA区域由边界序列限定,干预DNA(T-DNA)通常被插入至植物基因组内。此外,T-DNA的整合是相对精确的过程,其极少产生重排。鉴于其基因转移的简便和确定性,在有效发生土壤杆菌介导的转化的植物品种中成为首选的方法。优选的土壤杆菌转化载体能够在大肠杆菌以及土壤杆菌中复制,因此操作简便(Klee等,In:Plant DNA Infectious Agents,Hohn and Schell, eds.,Springer-Verlag,New York,pp.179-203(1985)。

[0651] 可使用的加速方法包括,例如,微粒轰击等等。将转化核酸分子输送至植物细胞内的方法的一个实例是微粒轰击。该方法的综述见Yang等,Particle轰击Technology for Gene Transfer,Oxford Press,Oxford,England(1994)。非生物学颗粒(微粒)可包被上核酸并通过推进力输送至细胞内。示例性的颗粒包括含钨、金、铂等的颗粒。其除了是可重复性地转化单子叶植物的有效手段之外,微粒轰击的特别优势还在于,既不需要分离原生质体,也不需要土壤杆菌感染的易感性。通过加速而将DNA输送至玉蜀黍细胞内的方法的一个示例性实施方式是生物弹道学 α -颗粒输送系统,其能够驱动包被有DNA的颗粒通过一个屏,例如不锈钢或Nytex屏,到达以悬浮培养的玉黍蜀细胞所覆盖的滤膜表面上。适合用于本发明的颗粒输送系统是氦加速枪PDS-1000/He,可自Bio-Rad Laboratories获得。

[0652] 为进行轰击,可将悬浮的细胞浓缩于滤膜上。将含有待轰击细胞的滤膜放置在距微粒停止板下方合适的距离处。需要的话,还可在枪和待轰击细胞之间放置一或多个屏。

[0653] 或者,可将不成熟的胚或其他靶细胞设置在固相培养基上。待轰击细胞放置在距微粒停止板下方适当距离处。如果需要,还可在加速装置和待轰击细胞之间放置一或多个屏。通过在此所述的技术可获得1000或更多个瞬时表达标记物基因的细胞焦点。轰击后48小时焦点内表达外源基因产物的细胞数量在1到10个,平均为1到3个。

[0654] 在轰击转化中,可优化轰击前培养条件和轰击参数,以产生最大数量的稳定转化体。轰击的物理学和生物学参数对于该技术均十分重要。物理因素涉及操控DNA/微粒沉淀物或影响大颗粒或微粒的飞行和速率的因素。生物学因素包括涉及在轰击之前和轰击后即刻对细胞的操控步骤、调节靶细胞的渗透压以便减轻与轰击相关的损伤、以及转化DNA的性质,例如是线性化DNA还是完好的超螺旋质粒。轰击前操控被认为对于成功转化未成熟胚而言是特别重要的。

[0655] 在另一个替代性实施方式中,可稳定转化质体。在高等植物中转化质体的已知方法包括用颗粒枪输送含有选择标记物的DNA并通过同源重组将DNA导向质体基因组(US 5,451,513、US 5,545,818、US 5,877,402、US 5,932,479和WO 99/05265)。

[0656] 因此,预期可以的是,可通过在小规模的研究中调整轰击参数的各个方面以便充分优化条件。特别希望调整的是物理参数,例如间隙距离、飞行距离、组织距离和氦气压力。还可通过改变影响受者细胞的生理学状态并因此可影响转化和整合效率的条件以使得降低损伤因素最小化。例如,可调节渗透压状态、组织水合和受者细胞的亚培养阶段或细胞周期以达到最佳转化。根据在此公开的内容,本领域人员知晓如何进行其他常规的调整。

[0657] 可采用基于磷酸钙沉淀、聚乙二醇处理、电穿孔以及这些处理的组合的方法实现植物原生质体的转化。如何将这些系统应用于不同植物品种取决于自原生质体再生出特定植株的能力。自原生质体再生谷类的示例性方法已被公开(Fujimura等,1985;Toriyama等,1986;Abdullah等,1986)。

[0658] 还可采用其他的细胞转化方法,包括但不限于通过以下方法将DNA引入植物:直接将DNA转移至花粉,直接将DNA注射至植物的繁殖性器官,或直接将DNA注射至未成熟胚的细胞内随后对干燥的胚进行再水化。

[0659] 从单个植物原生质体转化体或各种转化的外植体再生、发育和栽培出植物是本领域熟知的(Weissbach等,In:Methods for Plant Molecular Biology,Academic Press, San Diego,Calif.,(1988)。这种再生和生长方法典型地包括以下步骤:选择转化的细胞,培养个体化细胞历经从胚发育至有根的小植株阶段的通常阶段。转基因胚和种子的再生类似。得到的转基因的有根的苗随后栽植在合适的植物生长培养基例如土壤中。

[0660] 含有外来的外源基因的植物的发育或再生是本领域熟知的。优选地,再生的植物自花传粉以产生纯合子转基因植物。或者,将得自再生的植物的花粉与由具有重要农艺价值的株系的种子长成的植物进行杂交。反过来,可使用来自这些重要株系的植物的花粉给再生的植物授粉。可采用本领域人员熟知的方法栽培含有所需外源核酸的本发明的转基因植物。

[0661] 主要采用根癌土壤杆菌转化双子叶植物以及获得转基因植物的方法已经用于棉花(美国5,004,863、美国5,159,135、美国5,518,908);大豆(美国5,569,834、美国5,416,011);芸苔属(美国5,463,174);花生(Cheng等,1996);和豌豆(Grant等,1995)。

[0662] 转化谷类植物例如小麦和大麦以便通过引入外源核酸而在植物中引入遗传变异以及用于从原生质体或未成熟的植物胚再生植物的方法是本领域熟知的,参见例如,CA 2,092,588、AU 61781/94、AU 667939、US 6,100,447、PCT/US97/10621、US 5,589,617、US 6,541,257,其他方法参见专利说明书W099/14314。优选地,通过根癌土壤杆菌介导的转化方法产生转基因小麦或大麦植物。携带所需核酸构建体的载体可被引入组织培养的植物或外植体的可再生的小麦细胞内、或合适的植物系统例如原生质体中。

[0663] 可再生的小麦细胞优选地来自未成熟胚、成熟胚、来自它们的愈伤组织、或分生组织的盾片。

[0664] 为了证实转基因细胞和植物内存在转基因,可采用本领域人员已知的方法进行聚合酶链反应(PCR)扩增或Southern印迹分析。根据产物的性质,可通过多种途径检测转基因表达产物,包括免疫印迹和酶分析。在不同植物组织中定量蛋白质表达和检测复制的一种特别有用的方法是使用报告基因,例如GUS。一旦获得转基因植物,可使之生长产生具有所需表型的植物组织或部分。可收获植物组织或植物部分和/或收集种子。种子可用作来源用于生长出具有所需特性的额外的植物、组织或部分。

[0665] 采用土壤杆菌或其他转化方法形成的转基因植物典型地含有位于一条染色体上的单个基因座。此类转基因植物可被称为所添加的基因的半合子。更优选的转基因植物是所添加的基因的纯合子;即,转基因植物含有两个所添加的基因,一对染色体的每一条的相同基因座上具有一个基因。可通过半合子转基因植物自花授粉,使所得种子发芽,并分析所得植物中的感兴趣的基因,从而获得纯合子转基因植物。

[0666] 还要理解,含有两种独立地分离外源基因或基因座的两种不同的转基因植物可杂交,以产生含有两套基因或基因座的子代。合适F1子代的自花授粉可产生两种外源基因或基因座均为纯合子的植物。也预期了与亲代植物进行回交以及与非转基因植物进行异系杂交,以及营养体繁殖。常规采用的对不同性状和作物的其他育种方法可参见Fehr,In:

Breeding Methods for Cultivar Development, Wilcox J. ed., American Society of Agronomy, Madison Wis. (1987)。

[0667] 转基因非人类动物

[0668] “转基因非人类动物”指的是动物而非人类,所述动物含有同一物种或品种的野生型动物所不具有的基因构建体(“转基因”)。在此所述的“转基因”具有生物技术领域内的通常含义,且包括这样的遗传学序列,所述序列是经重组DNA或RNA技术产生或改变的且已经被引入至动物细胞内。转基因可包括来源于动物细胞的遗传学序列,所述动物细胞与待引入转基因的细胞可以是相同的物种或不同的物种。典型地,转基因已经通过人为操作引入至动物内,例如,通过转化,但也可通过本领域人员所知的任何方法。

[0669] 产生转基因动物的技术是本领域熟知的。该方面的通用教科书是Houdebine, Transgenic animals-Generation and Use (Harwood Academic, 1997)。可通过任何方法将多核苷酸分子转化至细胞内,所述方法可将多核苷酸分子插入至细胞内。转化技术包括,但不限于,转染、电穿孔、微注射、脂染、吸附和原生质体融合。重组细胞可保持单细胞状态或可生长为组织、器官或多细胞生物体。转化的多核苷酸分子可保持在染色体外或可整合至被转化(即,重组)细胞的染色体内的一或多个位置,使得其能够被表达。异源DNA可被引入至例如哺乳动物的受精卵内。例如,可通过微注射、磷酸钙介导的沉淀、脂质体融合、反转录病毒感染或其他手段转化全能或多能干细胞,然后将转化的细胞引入胚胎,并使胚胎发育为转基因动物。在高度优选的方法中,使用含有所需DNA的反转录病毒感染发育的胚胎,并自感染的胚胎产生转基因动物。不过,在最优选的方法中,合适的DNA被共注射至前核或胚胎的细胞质内,优选地在单个细胞阶段,并使胚胎发育为成熟的转基因动物。

[0670] 产生转基因动物的另一方法涉及通过标准的方法将核酸微注射至原核期卵内。然后培养被注射的卵,随后转移至假孕受者的输卵管内。

[0671] 还可通过核移植技术产生转基因动物。采用这种方法,用质粒稳定转染来自供者动物的成纤维细胞,该质粒整合有置于调控序列控制下的感兴趣的结合结构域或结合区的编码序列。然后将稳定的转染体与去核的卵母细胞融合,培养并转移至雌性受者内。

[0672] 提高外源RNA水平和稳定的表达

[0673] 沉默阻抑物

[0674] 转录后基因沉默(PTGS)是核苷酸序列特异性防卫机制,能够将细胞的和病毒的mRNA靶向降解。PTGS出现在以外来(外源性)或内源性DNA稳定或瞬时转化植物或真菌,导致与所引入的核酸具有序列相似性的RNA分子的积聚被降低。

[0675] 普遍认为共表达沉默阻抑物和感兴趣的转基因可升高细胞内由转基因转录而来的RNA的水平。尽管这在体外证明是正确的,但在许多整体植物共表达研究中已经发现其具有显著的副作用。更具体地,如Mallory等(2002)、Chapman等(2004)、Chen等(2004)、Dunoyer等(2004)、Zhang等(2006)、Lewsey等(2007)和Meng等(2008)所述,表达沉默阻抑物(通常置于组成型启动子控制下)的植物通常具有异常的表型,其程度导致它们无法用于商业生产。

[0676] 如上所述,本发明人发现,可通过将沉默阻抑物的表达局限于植物的贮藏器官或其部分而提高RNA分子水平和/或使得RNA分子水平在许多世代内保持稳定。在本申请中,“沉默阻抑物”是任何如下所述的多核苷酸或多肽,其能够在植物细胞中表达并提高来自植

物细胞内的不同转基因的表达产物的水平,特别是经过从最初转化的植物反复传代之后。在一个实施方式中,所述沉默阻抑物是病毒沉默阻抑物或其突变体。本领域已知大量病毒沉默阻抑物,包括,但不限于,P19、V2、P38、Pe-Po和RPV-P0。在一个实施方式中,所述病毒沉默阻抑物包含SEQ ID NO 97至101任一所示的氨基酸序列、其生物学活性片段、或与SEQ ID NO 97至101中的任何一或多种序列具有至少50%相同性的氨基酸序列,并具有沉默阻抑物的活性。

[0677] 在本文中,术语“使表达稳定”、“稳定表达的”、“稳定表达”及其变化形式指的是,经反复传代后(例如至少3代、至少5代或至少10代)的子代植物中的RNA分子的水平与缺乏编码所述沉默阻抑物的外源性多核苷酸的同基因植物相比,基本上是相同的或更高。但是,该术语不排除经反复传代后RNA分子的水平与前一代相比有所下降这一可能性,例如每一代损失不低于10%。

[0678] 阻抑物可选自任何来源,例如植物、病毒、哺乳动物等。阻抑物可以是,例如:

- [0679] ● 兽棚病毒B2;
- [0680] ● 绿萝潜隐病毒P14;
- [0681] ● 绿萝潜隐病毒AC2;
- [0682] ● 非洲木薯花叶病毒AC4;
- [0683] ● 秋葵黄脉花叶病C2;
- [0684] ● 秋葵黄脉花叶病C4;
- [0685] ● 秋葵黄脉花叶病BC1;
- [0686] ● 番茄褪绿病毒p22;
- [0687] ● 番茄褪绿病毒CP;
- [0688] ● 番茄褪绿病毒CPm;
- [0689] ● 番茄金色花叶病毒AL2;
- [0690] ● tomato leaf curl Java virusBC1
- [0691] ● 番茄黄曲叶病毒V2;
- [0692] ● 番茄黄曲叶病毒-中国C2
- [0693] ● 番茄黄曲叶中国病毒Y10分离株BC1;
- [0694] ● 番茄黄曲叶病毒以色列分离株V2;
- [0695] ● 绿豆黄花叶病毒-豇豆AC2;
- [0696] ● 木槿褪绿环斑病毒CP;
- [0697] ● 芜菁皱病毒P38;
- [0698] ● 芜菁皱病毒CP;
- [0699] ● 花椰菜花叶病病毒P6;
- [0700] ● 甜菜黄化病病毒p21;
- [0701] ● 柑橘腐根病病毒p20;
- [0702] ● 柑橘腐根病病毒p23;
- [0703] ● 柑橘腐根病病毒CP;
- [0704] ● 豇豆花叶病毒SCP;
- [0705] ● 甘薯褪绿矮缩病毒p22

- [0706] ●黄瓜花叶病毒2b;
- [0707] ●番茄不孕病毒HC-Pro
- [0708] ●甜菜卷顶病毒L2;
- [0709] ●土传小麦花叶病毒19K;
- [0710] ●大麦条斑花叶病毒Gammab;
- [0711] ●poa semilatifolius virus Gammab;
- [0712] ●花生丛簇病毒P15;
- [0713] ●稻米矮小病毒Pns10;
- [0714] ●curubiti aphid borne yellows virus P0;
- [0715] ●甜菜西部黄化病毒P0;
- [0716] ●马铃薯病毒X P25;
- [0717] ●黄瓜脉黄化病毒P1b;
- [0718] ●李痘病毒HC-Pro;
- [0719] ●甘蔗花叶病毒HC-Pro
- [0720] ●马铃薯病毒Y strain HC-Pro;
- [0721] ●烟草蚀纹病毒P1/HC-Pro;
- [0722] ●芜菁花叶病毒P1/HC-Pro;
- [0723] ●鸭茅斑点病毒P1;
- [0724] ●鸭茅斑点病毒-挪威分离株P1
- [0725] ●水稻黄斑点病毒P1;
- [0726] ●水稻黄斑点病毒-尼日利亚分离株P1;
- [0727] ●水稻白叶病毒NS3
- [0728] ●稻条纹病毒NS3
- [0729] ●十字花科感染烟草花叶病毒126K;
- [0730] ●十字花科感染烟草花叶病毒p122;
- [0731] ●烟草花叶病毒p122;
- [0732] ●烟草花叶病毒126
- [0733] ●烟草花叶病毒130K;
- [0734] ●烟草脆裂病毒16K;
- [0735] ●番茄丛矮病毒P19;
- [0736] ●蕃茄斑萎病毒NSs;
- [0737] ●苹果褪绿叶斑病毒P50;
- [0738] ●葡萄黄脉病毒A p10;
- [0739] ●葡萄卷叶相关病毒-2BYV p21的同源物,
- [0740] 及其变体/突变体。以上所列的提供了可由其获得阻抑物的病毒以及来自每一具体病毒的阻抑物的蛋白质(例如,B2、P14等)或编码区命名。
- [0741] 可以使用多个拷贝的阻抑物。不同阻抑物可以一起使用(例如,串联)。
- [0742] RNA分子
- [0743] 基本上任何适合在植物贮藏器官中表达的RNA分子均可与沉默阻抑物共表达。RNA

分子可影响农业性状、抗虫性、抗病性、除草剂抗性、不育性、籽粒特征等等。所编码的多肽可参与油、淀粉、碳水化合物、营养物等的代谢,或负责合成蛋白质、肽、脂肪酸、脂质、蜡、油、淀粉、糖、碳水化合物、香料、气味(odor)、毒素、类胡萝卜素、激素、聚合物、类黄酮、贮藏蛋白、酚酸、生物碱、木质素、丹宁酸、纤维素、糖蛋白、糖脂等等。

[0744] 在具体的实例中,所述植物中以下的酶或蛋白质的水平是升高的:在植物(例如为芸苔、例如油料种子油菜或向日葵、红花、亚麻、棉、大豆或玉米)中产生油的酶;参与植物(例如马铃薯、玉米和谷物例如小麦、大麦或稻)中淀粉合成的酶;合成药物的酶或其本身即是天然药物的蛋白质,例如药物或兽医产品。

[0745] 预期以本发明的方法产生的多肽的类型包括:用于包括人在内的哺乳动物类的药用蛋白质,例如胰岛素、前胰岛素原、胰岛素原、胰高血糖素、干扰素例如 α -干扰素和 γ -干扰素、凝血因子例如因子VII、VIII、IX、X、XI和XII、生育激素例如黄体生成素、卵泡刺激素、生长因子例如表皮生长因子、血小板衍生的生长因子、粒细胞集落刺激因子、泌乳刺激素、催产素、促甲状腺激素、促肾上腺皮质激素、降钙素、甲状旁腺激素、生长激素抑制素、促红细胞生成素(EPO)、酶例如 β -葡糖脑苷脂酶、血红蛋白、血清白蛋白、胶原、生长激素、人血清白蛋白、人分泌的碱性磷酸酶、抑肽酶、 α 1-抗胰蛋白酶、IgG1(磷酸酯)、IgM(神经肽半抗原)、SIgA/G(变异链球菌(*Streptococcus mutans*)外源凝集素)、scFv-bryodin 1免疫毒素(CD 40)、IgG(HSV)、LSC(HSV)等等。

[0746] 此外,本发明的方法也可用于产生特异性抗体,包括抗体相关性分子或其活性片段,它们结合,例如,骨形态生成蛋白受体IB型;E16;STEAP1;MPF;Napi3b;Sema 5b;PSCA;B型内皮缩血管肽受体;MSG783;STEAP2;TrpM4;CRIPT0;CD21;CD79b;FcRH2;HER2;NCA;MDP;IL20R α ;短蛋白聚糖(Brevican);EphB2R;ASLG659;PSCA;GEDA;B细胞活化因子受体;CD22;CD79a;CXCR5;HLA-DOB;P2X5;CD72;LY64;FcRH1;IRTA2;TENB2;CD20;VEGF包括VEGF_A、B、C或D;p53;EGFR;孕激素受体;组织蛋白酶D;Bcl-2;E钙粘蛋白;CEA;Lewis X;Ki67;PCNA;CD3;CD4;CD5;CD7;CD11c;CD11d;c-Myc;tau;PrPSC;或A β 。

[0747] 此外,本发明的方法也可用于产生抗原,所述抗原可通过或不通过摄入所述贮藏器官而输送,抗原的实例包括乙型肝炎病毒包膜蛋白、狂犬病毒糖蛋白、大肠杆菌热不稳定肠毒素、Norwalk病毒衣壳蛋白、糖尿病自身抗原、霍乱毒素B亚单位、霍乱毒素B和A2亚单位、轮状病毒肠毒素和产肠毒素大肠杆菌菌毛抗原融合物、猪传播性胃肠炎病毒糖蛋白S、人类鼻病毒15(HRV-14)和1型人免疫缺陷病毒(HIV-1)表位、貂肠炎病毒表位、口蹄疫病毒VP1结构蛋白、人巨细胞病毒糖蛋白B、龋齿(*S.mutans*)抗原和呼吸道合胞病毒抗原。

[0748] 产生的LC-PUFA的水平

[0749] 重组细胞中产生的LC-PUFA或LC-PUFA的组合的水平十分重要。所述水平可以表示为特定LC-PUFA或一组相关的LC-PUFA(例如 ω 3LC-PUFA或 ω 6LC-PUFA)、或VLC-PUFA、或其他占总脂肪酸的组成(百分比),这可通过本领域已知的方法确定。所述水平还可表示为LC-PUFA的含量,例如LC-PUFA占包含所述重组细胞的材料干重的百分比,例如LC-PUFA占种子干重的百分比。可以理解,就LC-PUFA的含量而言,油料种子中产生的LC-PUFA可以显著高于不为产油而生长的蔬菜或籽粒,但两者可具有相似的LC-PUFA组成,且两者均可用作人或动物摄入LC-PUFA的来源。

[0750] 可通过本领域已知的任何方法确定LC-PUFA的水平。在优选的方法中,自细胞、组

织或生物体提取总脂质,并将脂肪酸转换为甲基酯,随后通过气相色谱(GC)进行分析。此类技术见实施例1。色谱的峰位置可用来鉴定各特定脂肪酸,而各峰下的面积求积分后应用于确定量。在本申请中,除非另有说明,样品中的特定脂肪酸的百分比是按照该脂肪酸的峰下的面积占色谱中的脂肪酸总面积的百分比确定的。这基本上对应于重量百分比(w/w)。脂肪酸的鉴定可通过GC-MS加以确认。总脂质可以通过本领域已知的纯化组份(例如TAG组份)的技术进行分离。例如,可进行分析尺度的薄层色谱(TLC)从而将TAG与其他脂质组份例如DAG、酰基-CoA或磷脂分离,以便确定TAG具体的脂肪酸组成。

[0751] 在一个实施方式中,所述细胞内脂肪酸中的ARA、EPA、DPA和DHA总计占所述细胞内的总脂肪酸的至少15%,更优选地至少20%或至少25%。在更加优选的实施方式中,这些脂肪酸的总和占所述细胞内的总脂肪酸的至少29%、至少30%或至少31%。在另一个实施方式中,所述细胞内的总脂肪酸具有低于1%的C20:1。在另一个实施方式中,所述细胞内脂肪酸中的DHA的量占所述细胞内的总脂肪酸的至少3%,更优选地至少4%,更优选地至少5%或至少7%,或最优选地至少10%。在优选的实施方式中,细胞内可提取的TAG包含本段所提及的水平的脂肪酸。还预期了这些特征的每一种可能的组合。例如,所述细胞内脂肪酸中的ARA、EPA、DPA和DHA总计可占细胞内总脂肪酸的至少15%、至少20%、至少25%、至少29%、至少30%或至少31%,其中至少3%、至少4%、至少5%、至少7%或至少10%的细胞内总脂肪酸是DHA,而C20:1的水平可低于1%。

[0752] 在这些实施方式的每一实施方式中,重组细胞可以是适合发酵的生物体的细胞,例如,单细胞微生物,其可以是原核生物或真核生物,例如酵母或植物细胞。在优选的实施方式中,细胞是被子植物(高等植物)细胞。在进一步优选的实施方式中,细胞是种子内的细胞,例如,油料种子或籽粒或谷类。

[0753] 重组细胞内LC-PUFA产生的水平还可以转换率表示,即所形成的LC-PUFA的量占一或多种底物PUFA或LC-PUFA的百分比。例如,就EPA而言,可以表示为EPA的水平(占总脂肪酸的百分比)与底物脂肪酸(ALA、SDA、ETA或ETrA)水平的比率。

[0754] 在一个实施方式中,ALA转换为EPA的转换效率是至少80%或更优选地至少90%。在另一个实施方式中,ALA转换为EPA、DPA或DHA(以EPA、DPA和DHA的百分比的总和/ALA和ALA经 $\Delta 6$ -去饱和的所有脂肪酸产物的百分比的总和计算)的转换效率是至少17.3%、或至少23%。在另一个实施方式中,ALA转换为DPA或DHA(以DPA和DHA的百分比的总和/ALA和ALA经 $\Delta 6$ -去饱和的所有脂肪酸产物的百分比的总和计算)的转换效率是至少15.4%、或至少21%。在另一个实施方式中,ALA转换为DHA(以DHA的百分比/ALA和ALA经 $\Delta 6$ -去饱和的所有脂肪酸产物的百分比的总和计算)的转换效率是至少9.5%、或至少10.8%。在另一个实施方式中,EPA转换为DHA(以DHA的百分比/EPA和EPA经 $\Delta 5$ -延长的所有脂肪酸产物的百分比的总和计算)的转换效率是至少45%、或至少50%。在另一个实施方式中,转换SDA产生ETA(以ETA和ETA经 $\Delta 5$ -去饱和的所有脂肪酸产物的百分比的总和/SDA和SDA经 $\Delta 6$ -延长的所有脂肪酸产物的百分比的总和计算)的转换效率是至少50%,更优选地至少60%。在另一个实施方式中,ALA转换为ETrA的转换效率是至少6%,更优选地至少9%。在另一个实施方式中,通过 $\Delta 5$ 延长酶步骤将EPA转换为DPA(以DPA和DHA的百分比的总和/EPA、DPA和DHA的百分比的总和计算)的转换效率是至少60%,更优选地至少65%,更优选地至少70%或最优选地至少75%。

[0755] 如果用于引入这些基因的亲代细胞被选择为使得所产生的或外源提供的脂肪酸底物的水平达到最佳,则可以使得重组细胞中的LC-PUFA含量最大化。可以通过将细胞在最佳条件下生长或温育而使得LC-PUFA水平最大化,例如在略低于该细胞的标准温度的温度下生长,认为这样有利于多不饱和脂肪酸的积聚。不过,目前的证据提示,在酵母或植物中外源表达的某些去饱和酶与某些延长酶组合时具有相对低的活性。可通过提供在LC-PUFA合成过程中能够利用酰基-CoA形式的脂肪酸作为底物的去饱和酶来改善这一点,这在酵母以外的其他重组细胞例如植物细胞中被认为是有利的。

[0756] 油的产生

[0757] 可采用本领域常规使用的技术来提取、加工和分析由本发明的细胞、植物、种子等产生的油。典型地,植物种子经熟化、压榨和提取以产生粗制油,然后进行脱胶、精制、脱色和除味。通常,种子压碎技术是本领域已知的。例如,给油料种子喷水使之变软,并使之含水量升高至,例如8.5%,使用间隙设置为0.23至0.27mm的平滑滚筒将其制成薄片。根据种子的类型,在压碎之前可以不加水。加热可使酶失活,且有助于细胞进一步破坏,使油滴凝结,并使蛋白质凝聚,所有这些均有助于提取加工。

[0758] 通过螺旋压榨机可使得大部分种子油释放出来。从螺旋压榨机排出的饼随后进行溶剂提取,例如以己烷,使用伴热柱(heat traced column)。或者,压榨工序所产生的粗制油可通过沉降罐,其上部具有槽状金属丝网引流构造,以去除压榨工序过程中随油而挤出的固形物。净化的油再通过板框式滤器,以去除任何残余的微小固形颗粒。如果需要,自提取工序所回收的油可与净化的油混合以产生掺合的粗制油。

[0759] 将溶剂从粗制油中去除后,将压榨部分和提取部分混合饼对其进行常规的油加工过程(即,脱胶、碱法净化、脱色和除味)。可通过向粗制油中添加浓磷酸进行脱胶,将不可水合的磷脂转化为可水合形式,并螯合存在的少量金属。通过离心将胶与油分离。可通过添加足量的氢氧化钠溶液滴定全部脂肪酸并移除所形成脂肪酸盐,以精制油。

[0760] 可通过在真空条件下将油加热至260℃并以0.1ml/分钟/100ml油的速度缓慢向油中引入蒸汽而进行除味。喷射30分钟后,将油在真空中冷却。典型地,油被转移至玻璃容器并充入氩气,然后冷藏。如果油量有限,则可置于真空条件下,例如Parr反应器内,并加热至260℃,加热时间与用于除味所需的时间一样。这一处理可改善油的颜色并去除绝大部分的挥发性物质。

[0761] 饲料(Feedstuffs)

[0762] 本发明包括可用作饲料的组合物。在本发明中,“饲料”包括供人类或动物摄入(包括用于胃肠道和/或胃肠外摄入)的任何食物或制品,当其进入肌体后,(a)用于营养或强健组织或提供能量;和/或(b)保持、恢复或支持足够的营养状态或代谢功能。本发明的饲料包括婴儿和/或幼儿的营养组合物。

[0763] 本发明的饲料包含,例如,本发明的细胞、本发明的植物、本发明的植物部分、本发明的种子、本发明的提取物、本发明的方法的产物、本发明的发酵方法的产物、或本发明的组合物,以及合适的载剂。术语“载剂”取其最为广泛的含义,涵盖具有或不具有营养价值的任何成分。本领域人员能够理解,载剂必须适合于用于饲料(或以足够低的浓度使用),以便其不会对摄入所述饲料的生物体产生有害作用。

[0764] 本发明的饲料包含使用本发明的方法、细胞或植物直接或间接产生的油、脂肪酸

酯或脂肪酸。组合物可以是固体或液体的形式。此外,组合物可包括可食用的常量营养物、维生素、和/或矿物质,它们的量符合特定的需要。这些成分的量可以变化,这取决于组合物是用于正常个体还是用于具有特殊需求的个体,例如患有代谢性疾病和类似情况的个体。

[0765] 具有营养价值的合适的载剂的实例包括,但不限于,常量营养物例如可食用的脂肪、糖类和蛋白质。此类可食用的脂肪的实例包括,但不限于,椰子油、琉璃苣油、真菌油、黑醋栗油、大豆油和甘油一酸酯和甘油二酯。此类糖类的实例包括(但不限于):葡萄糖、食用乳糖、和水解淀粉。此外,可用于本发明的营养组合物中的蛋白质的实例包括(但不限于):大豆蛋白、电渗析乳清、电渗析脱脂乳、乳乳清、或这些蛋白质的水解产物。

[0766] 至于维生素和矿物质,以下物质可添加至本发明的饲料组合物中:钙、磷、钾、钠、氯、镁、锰、铁、铜、锌、硒、碘和维生素A、E、D、C、和复合B。其他的此类维生素和矿物质也可添加。

[0767] 用于本发明的饲料组合物的成分可以是半纯化来源的或纯化来源的。半纯化来源的或纯化来源的指的是通过对天然材料进行纯化而制备的物质或通过从头合成而制备的物质。

[0768] 本发明的饲料组合物也可添加至食物中,即使是在不需要饮食增补的情况下也可以。例如,组合物可添加至任何类型的食物中,包括(但不限于):人造黄油、改性黄油、奶酪、乳、酸奶、巧克力、糖果、点心、色拉油、烹饪用油、烹饪用脂肪、肉、鱼和饮料。

[0769] 酵母属物种用于酿造啤酒和葡萄酒以及烘焙,特别是烘制面包。酵母是植物提取物的主要成分。酵母也用作动物饲料的添加剂。显然,可产生遗传工程化的酵母菌株使之适合于合成在此所述的LC-PUFA。这些酵母菌株随后可用于食品以及葡萄酒和啤酒酿造,以便提供具有增加的脂肪酸含量的产品。

[0770] 此外,根据本发明产生的脂肪酸或经转化后含有并表达对象基因的宿主细胞也可用作动物食物增补剂,以改变动物组织或乳的脂肪酸组成,使之更适合人类或动物摄入。此类动物的实例包括羊、牛、马等等。

[0771] 此外,本发明的饲料还可用于水产业以提高供人类或动物摄入的鱼的脂肪酸水平。

[0772] 本发明的优选的饲料是植物、种子、或其他植物部分,例如叶和茎,它们可以直接用作人或其他动物的食物或饲料。例如,动物可直接在生长此类植物的田间放牧或以更加定量的可控饲养方式饲喂。本发明包括使用此类植物和植物部分作为饲料以提高人和其他动物的LC-PUFA水平。

[0773] 组合物

[0774] 本发明还包括组合物,特别是药物组合物,所述组合物包含一或多种采用本发明的方法产生的脂肪酸和/或所得的油。

[0775] 药物组合物可包含一或多种脂肪酸和/或油,连同标准的、熟知的、非毒性药用可接受的载剂、佐剂或载体,例如磷酸缓冲盐溶液、水、乙醇、多元醇、植物油、湿润剂或乳状液,例如水/油乳状液。组合物可以是液体或固体形式。例如,组合物的形式可以是片剂、胶囊、可摄入的液体或粉末、注射形式、或局部用药膏或乳剂。可通过例如保持分散体系所需要的粒径和通过使用表面活性物质而保持合适的流动性。还可加入等渗剂,例如,糖、氯化钠等等。除了此类惰性修饰剂之外,组合物还可包括佐剂,例如湿润剂、乳化和混悬剂、甜味

剂、调味剂和芳香剂。

[0776] 混悬液除了含有活性化合物之外还包含混悬剂,例如乙氧基异硬脂醇(ethoxylated isostearyl alcohols)、聚氧乙烯山梨醇酯和聚氧乙烯山梨糖酯、微晶纤维素、偏氢氧化铝(aluminum metahydroxide)、班脱土、琼脂、和黄芪胶或这些物质的混合物。

[0777] 可通过采用本领域熟知的技术制备固体剂量形式例如片剂和胶囊。例如,根据本发明产生的脂肪酸可压成片剂,其中使用常规的片基例如乳糖、蔗糖和玉米淀粉,联合使用的有粘合剂例如阿拉伯树胶、玉米淀粉或白明胶,崩解剂例如马铃薯淀粉或褐藻酸,以及润滑剂例如硬脂酸或硬脂酸镁。可通过将这些赋形剂连同抗氧化剂和相关的脂肪酸装入白明胶胶囊而制备胶囊。

[0778] 对于静脉施用,可根据本发明产生的脂肪酸或其衍生物掺入至商品化的配制品中。

[0779] 特定脂肪酸的典型剂量是0.1mg至20g,每天摄入1至5次(高达每日100g),且优选地在每天大约10mg至大约1、2、5、或10g的范围内(一次或分多次摄入)。本领域已知,每天需要最少大约300mg的脂肪酸,特别是LC-PUFA。不过,应该理解任何量的脂肪酸对对象均可能是有益的。

[0780] 本发明的药物组合物的可能的施用途包括,例如,胃肠道(例如口服和直肠)和胃肠外。例如,可口服或直肠施用液体制备物。此外,可将均质的混合物完全分散至水中,在无菌条件下与生理学可接受的稀释剂、防腐剂、缓冲剂或推进物混合物,形成喷剂或吸入剂。

[0781] 待施用于患者的组合物的剂量可由本领域人员确定,这取决于多种因素,例如患者的体重、患者的年龄、患者的总体健康状况、患者的既往病史、患者的免疫状态等等。

[0782] 此外,本发明的组合物可用于化妆品。其可添加至已有的化妆品组合物中一般形成混合物,或是将根据本发明产生的脂肪酸用作化妆品组合物中的唯一的“活性”成分。

实施例

[0783] 实施例1:材料和方法

[0784] 培养微藻类

[0785] 在标准培养条件下培养来自CSIRO活微藻保藏中心(CSIRO Collection of Living Microalgae (<http://www.marine.csiro.au/microalgae>))的微胞藻CS-0170和塔胞藻CS-0140分离物。将来自保藏中心的原培养物在1LErlenmeyer瓶中按照1:10稀释进行次代培养并连续传代扩增,然后转移至10L聚碳酸酯大瓶。培养基为f/2,是改良的Guillard和Ryther's (1962) f培养基,含有半量的营养物,生长温度为 $20 \pm 1^\circ\text{C}$ 。其他培养条件包括,光密度为 $100 \mu\text{mol photons PAR} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$, 12:12小时的亮:暗光周期,含1%CO₂的空气充入气泡,速度为 $200 \text{ mL} \cdot \text{L}^{-1} \cdot \text{min}^{-1}$ 。

[0786] 分离微藻基因组DNA

[0787] 采用DNeasy Plant Mini Kit系统分离微胞藻CS-0170和塔胞藻CS-0140的基因组DNA,方法见所附说明书(QIAGEN, 目录#69106)。

[0788] 分离微藻总RNA

[0789] 采用以下方法从微胞藻CS-0170和塔胞藻CS-0140细胞分离总RNA。用研钵和杵将

2g (湿重) 细胞在液氮中粉碎, 缓慢撒进大口杯中, 其中含有持续搅拌的22mL提取缓冲液。向其中加入5%不溶性聚乙烯吡咯烷酮, 90mM 2-巯基乙醇, 和10mM二硫苏糖醇, 混合物继续搅拌10分钟, 然后转移至Corex™管。加入18.4mL的3M醋酸铵, 充分混合。然后将样品6000×g离心20分钟, 4℃。将上清液转移至新的管中并加入0.1体积的3M NaAc (pH 5.2) 和0.5体积的冷异丙醇沉淀核酸。−20℃孵育1小时, 之后将样品在6000×g旋转离心30分钟。将沉淀团垂悬于1mL水中, 以酚/氯仿抽提。将水相转移至新的管中, 再次加入0.1体积的3M NaAc (pH 5.2) 和2.5体积的冰冷乙醇沉淀核酸。将沉淀团垂悬于水中并通过分光光度计确定核酸浓度。

[0790] 载体和菌株

[0791] 质粒pYES2和酵母株INVSC1来自Invitrogen, 质粒载体pGEMT-Easy来自Promega, 质粒载体pBluescript II KS-来自Stratagene。Lazo等(1991)提到的根癌土壤杆菌菌株AGL1, Coutu等(2007)提到的pORE二元载体系列。

[0792] PCR条件

[0793] 除非另有说明, 否则采用标准条件聚合酶链反应(PCR)扩增DNA片段。通过改变扩增循环的数量、引物退火温度、Mg²⁺浓度和其他参数的条件最佳化, 这在本领域是常规的。缓冲液如聚合酶供货商所述。典型地, 反应条件如下。94℃初始变性2-3分钟, 然后将反应混合物进行20-40个变性/退火/延伸循环, 其中在94℃变性30-60秒, 40-60℃引物退火30秒, 30-60秒聚合酶延伸70-72℃, 随后70-72℃再延伸3分钟。

[0794] 典型地, 采用Superscript III One-Step RT-PCR系统(Invitrogen)进行反转录-PCR(RT-PCR)扩增, 体积为25μL, 使用10pmol的正向引物和30pmol的反向引物, MgSO₄终浓度为2.5mM, 400ng的总RNA, 按照生产商的说明使用缓冲液和核苷酸。典型的温度方案为: 1个循环的45℃ 30分钟进行反转录; 然后1个循环的94℃ 2分钟, 随后40个循环的94℃ 30秒、52℃ 30秒、70℃ 1分钟; 然后1个循环的72℃ 2分钟, 然后将反应混合物冷却至5℃。

[0795] 5' 和3' -RACE

[0796] 为了获得相应于部分长度基因片段的全长cDNA, 通过5' - 和3' -RACE (Rapid Amplification of cDNA Ends) 方法获得cDNA的5' 和/或3' 末端。使用实施例中所所述的基因特异性正向引物和oligo-dT反向引物5'-ATTTAGGTGACACTATAGTTTTTTTTTTTTTTTTTTT-3' (SEQ ID NO:41) 分离cDNA的3' 末端, 其中V代表A、G或C, 其在所有3' -RACE反应中均是相同的。采用Superscript III One-Step RT-PCR系统(Invitrogen)进行RT-PCR扩增, 体积为25μL, 使用10pmol的正向引物和30pmol的反向引物, MgSO₄终浓度为2.5mM, 400ng的总RNA作为合成cDNA的模板, 按照生产商的说明使用缓冲液和核苷酸。循环条件是典型的: 1个循环45℃ 30分钟进行反转录; 然后1个循环94℃ 2分钟; 随后是40个循环的94℃ 30秒, 52℃ 30秒, 70℃ 1分钟和1个循环72℃ 2分钟, 随后冷却至5℃。将反应的扩增子连接到pGEM-T Easy中, 克隆至大肠杆菌中, 采用标准方法测序。

[0797] 除非另有说明, 否则均采用改良的末端转移酶方法分离cDNA的5' 末端, 其中使用2μg总RNA作为合成cDNA的模板。将10pmol的基因特异性反向引物加入到总RNA和10.8μL水中, 然后将混合物在65℃加热5分钟, 然后冰上冷却2分钟。然后加入以下成分: 4μL的Superscript III第一链cDNA缓冲液(Invitrogen), 1μL的0.1M二硫苏糖醇, 1μL RNaseOUT (Invitrogen) 和1μL的Superscript III反转录酶。混合物然后在55℃温育60分钟, 然后通

过在70℃进一步温育15分钟终止反应。在冰上短暂冷却后,然后将反应物以2单位的RNAseH 37℃处理20分钟。然后使用QIAQUICK PCR Purification Kit (QIAGEN, 目录#28106) 纯化cDNA。然后使用25μL的洗脱物进行A-加尾(A-tailed), 其中使用:10单位的TdT (NEB), 5μL的NEB Buffer#4, 5μL的2.5mM CoCl₂, 0.5μL的10mM dATP, 总体积为50μL。反应在37℃进行30分钟, 随后在70℃处理10分钟灭活酶。然后使用2.5单位的Taq DNA聚合酶 (NEB) 进行PCR反应, 反应混合物包括5μL的A-加尾cDNA, 10pmol的基因特异性反向引物, 30pmol的改良oligo-dT引物5'-ATTTAGGTGACACTATAG TTTTTTTTTTTTTTTTTT-3' (SEQ ID NO:41), 其中V代表A, G或C, 按照生产商的说明使用缓冲液和核苷酸。循环条件是典型的:1个循环94℃ 2分钟;5个循环的94℃ 20秒, 54℃ 1分钟, 72℃ 1分钟;30个循环的94℃ 20秒, 60℃ 30秒, 72℃ 1分钟;1个循环72℃ 5分钟;4℃放置。凝胶电泳后, 如果在预期大小的范围没有看到清晰的产物条带, 则切取胶上该区域并从胶中纯化DNA产物。取1μL的1:20稀释的洗脱物的样品用作第二轮PCR的模板。将反应的扩增子连接到pGEM-T Easy中并测序。

[0798] 酵母培养并加料前体脂肪酸

[0799] 热休克方法将质粒引入酵母, 并在含2%棉子糖作为单一碳源的酵母基本培养基(YMM)板上选择转化体。在含2%棉子糖作为单一碳源的液体YMM中建立克隆接种物培养物。从中取实验培养物接种在YMM+1%NP-40, 至初始OD600为~0.3。在30℃震荡生长培养物(~60rpm)至OD600为大约1.0。此时加入半乳糖至终浓度为2%, 加入前体脂肪酸至终浓度为0.5mM。培养物在20℃继续震荡培养48小时, 然后离心收集。以1%NP-40, 0.5%NP-40和水洗涤细胞沉淀, 以去除细胞表面的任何未掺入的脂肪酸。

[0800] 以瞬时表达系统在植物细胞中表达基因

[0801] 基本上按Voinnet等(2003)所述以瞬时表达系统在植物细胞中表达基因。含有待从强组成型启动子例如35S启动子表达的编码区的质粒被引入根癌土壤杆菌菌株AGL1。用于表达p19病毒沉默抑制物的嵌合基因35S:p19被另行引入AGL1。重组细胞生长在28℃的补充有50mg/mL卡那霉素和50mg/mL利福平的LB培养液中直至静止生长期。然后室温5000g离心15分钟沉淀细菌, 然后重悬于含10mM MES pH 5.7, 10mM MgCl₂和100μM乙酰丁香酮的渗入缓冲液中, 使得OD600=1.0。细胞在28℃震荡培养3小时, 然后与等体积的含有35S:p19和感兴趣的测试嵌合基因的土壤杆菌培养物混合, 随后渗入叶组织。渗入后, 植物通常再生长5天, 然后取叶盘用于GC分析脂肪酸。

[0802] 如果给叶组织补充外源性脂肪酸, 则通过将合适的脂肪酸在2M氢氧化铵溶液中60℃加热20分钟预处理脂肪酸, 随后同样在60℃将溶液蒸发。所得的盐重悬于0.1M磷酸盐缓冲液(pH 7.2)至终浓度为0.5μg/mL。在土壤杆菌渗入后4天将脂肪酸盐注入叶中, 在补料后不同时间点取叶盘用于分析脂肪酸组成, 例如加入外源性脂肪酸后2-48小时。所用对照为不加入外源性脂肪酸, 或用于渗入的土壤杆菌菌株不含有感兴趣的基因。

[0803] 气相色谱(GC)分析脂肪酸

[0804] 脂肪酸制备

[0805] 对于含有大量水分的样品, 包括所有本氏烟草叶样品和其他非种子组织, 先采用Bligh和Dyer(1959)所述的方法提取总脂质, 然后进行甲基化。将离心的酵母沉淀、拟南芥种子、本氏烟草的总脂质或其他总脂质样品与MeOH-CHCl₃-HCl(10:1:1, v/v/v)在配有Teflon-lined螺旋帽的玻璃试管中在90-100℃加热2小时, 进行转酯作用形成脂肪酸甲基

酯 (FAME)。FAME被抽提至己烷-二氯甲烷 (4:1, v/v) 中, 通过GC和GC-MS进行分析。

[0806] 毛细管气液相色谱 (GC)

[0807] 气相色谱 (GC) 分析FAME, 使用了Agilent Technologies 6890N GC (Palo Alto, California, USA), 其配备有Equity™-1熔融石英毛细管柱 (15m x 0.1mm i.d., 0.1μm薄膜厚度)、FID、分流/无分流进样器和Agilent Technologies 7683Series自动采样进样器。氦用作载体气体。以无分流模式进样, 炉温120℃。进样后炉温升至270℃, 速度为10℃·min⁻¹, 最后以5℃·min⁻¹升至310℃。以Agilent Technologies ChemStation软件 (Rev B.03.01 (317), \Palo Alto, California, USA) 定量峰。

[0808] 气相色谱-质谱 (GC-MS)

[0809] 在装有柱上进样器的Finnigan GCQ Plus GC-MS离子收集器上在4℃进行GC-MS。使用AS2000自动采样器将样品注入连接于HP-5Ultra 2键合相柱 (50m x 0.32mm i.d. x 0.17μm薄膜厚度) 的保留隘口。初始温度45℃保持1分钟, 随后以30℃·min⁻¹程控升温至140℃, 然后以3℃·min⁻¹至310℃, 并保持12分钟。氦用作载体气体。质谱仪操作条件为: 电子轰击能量70eV; 发射电流250μamp, 转运线310℃; 源温度240℃; 扫描速率0.8scans·s⁻¹, 质量范围40-650Dalton。获得质谱并以Xcalibur™软件处理。

[0810] 培养酵母并补料前体脂肪酸

[0811] 热休克方法将质粒引入酵母, 并在含2% 棉子糖作为单一碳源的酵母基本培养基 (YMM) 板上选择转化体。在含2% 棉子糖作为单一碳源的液体YMM中建立克隆接种物培养物。从中取实验培养物接种在YMM+1% NP-40, 至初始OD₆₀₀为~0.3。在30℃震荡生长培养物 (~60rpm) 至OD₆₀₀为大约1.0。此时加入半乳糖至终浓度为2%, 加入前体脂肪酸至终浓度为0.5mM。培养物在20℃继续震荡培养48小时, 然后离心收集。以1% NP-40、0.5% NP-40和水洗涤细胞沉淀, 以去除细胞表面的任何未掺入的脂肪酸。

[0812] 实施例2: 从微藻类分离并表征编码Δ6-延长酶的cDNA

[0813] 分离微胞藻CS-0170 Δ6-延长酶基因片段

[0814] CSIRO活微藻保藏中心的微胞藻CS-0170株 (W02005/103253) 被鉴定为具有高天然水平的Δ5-和Δ6-延长作用的微藻株 (表4)。

[0815] 表4. CSIRO活微藻保藏中心微胞藻CS-0170和塔胞藻CS-0140的脂肪酸转换

	类型	青绿藻纲 (PRASINOPHYCEAE)	青绿藻纲
	物种	细小微胞藻	<i>P. cordata</i>
[0816]	株	CS0170	CS0140
	生长期	对数	对数
	16:1 (n-7)	0.7	0.8

	18:1 (n-9)	0.3	0.2
	18:1 (n-7)	5.5	14.8
	16:2 (n-7)	0.2	0.0
	18:2 (n-6)	0.1	0.7
	18:3 (n-6)	0.0	0.0
	20:4 (n-6)	0.0	0.0
	16:3 (n-3)	0.0	0.0
[0817]	16:4 (n-3)	20.4	14.3
	18:3 (n-3)	1.4	4.6
	18:4 (n-3)	20.7	25.6
	18:5 (n-3)	16.7	3.4
	20:3 (n-3)	0.1	1.2
	20:4 (n-3)	0.0	0.0
	20:5 (n-3)	0.3	0.4
	22:5 (n-3)	0.3	4.1
	22:6 (n-3)	8.5	4.5

[0818] 为了鉴定保守序列,采用ClustalW算法比对来自GenBank登录号AAV67800、ABC18314、CAD58540、CAL55414、AAV67797、XP_001416454、AAW70157、AAV67799、ABC18313、AAY15135的延长酶氨基酸序列。在众多具有不同程度相同性的同源区中选择出共有氨基酸序列区段KXXXXXDT (SEQ ID NO:31) 和MYXY (SEQ ID NO:32) (其中每一X独立地是任意氨基酸),分别相应于AAY15135的氨基酸位置144-151和204-208。基于这两个区段的序列合成了简并引物5'-AAGWWCIKSGARYISYTCGACAC-3' (SEQ ID NO:42) 和5'-AIIMIRTARTASGTGTACAT-3' (SEQ ID NO:43),其中I=肌苷,W=A或T,R=A或G,Y=C或T,K=G或T,M=A或C,S=C或G。使用Superscript III One-Step RT-PCR系统 (Invitrogen) 进行RT-PCR扩增,体积为50μL,使用了20pmol的各种引物,MgSO₄至终浓度为2.5mM,200ng微胞藻CS-0170总RNA,按照说明使用缓冲液和核苷酸。循环条件为:初始48℃ 30分钟进行反转录,然后1个循环94℃ 2分钟,随后5个循环的94℃ 30秒,40℃ 30秒,70℃ 30秒;,然后40个循环的94℃ 30秒,45℃ 30秒,70℃ 30秒,然后72℃ 2分钟。产生了209bp的扩增子,将其连接至pGEM-T Easy并测序。

[0819] 分离编码微胞藻CS-0170 Δ6-延长酶的全长cDNA

[0820] 设计引物用于通过5'-和3'-RACE延伸所述209bp片段。如实施例1所述分离基因的3'末端,使用了基因特异性正向引物5'-GAACAACGACTGCATCGACGC-3' (SEQ ID NO:44) 和200ng微胞藻CS-0170总RNA。产生了454bp的扩增子,将其连接至pGEM-T Easy并测序。使用GeneRacer Kit (Invitrogen, 目录#L1500-01) 分离基因的5'末端,其中55℃反转录温育1小时以产生5'-适应的cDNA,按照所附说明操作。以GeneRacer 5'引物5'-CGACTGGAGCACGAGGACACTGA-3' (SEQ ID NO:45) 和基因特异性反向引物5'-TTGCGCAGCACCATAAAGACGGT-3' (SEQ ID NO:46) 进行PCR扩增,使用了PFU Ultra II Fusion DNA聚合酶,体积为50μL,使用10pmol的各种引物,1μl的GeneRacer cDNA模板,按照生产商的说明使用缓冲液和核苷酸 (Stratagene, 目录#600670)。循环条件为:1个循环94℃ 2分

钟;35个循环的94℃ 20秒,55℃ 30秒,72℃ 30秒;然后72℃ 2分钟,然后冷却至4℃。1:10稀释该产物并用1μl作为模板进行第二轮PCR,使用了GeneRacer 5'巢式引物5'-GGACACTGACATGGACTGAAGGAGTA-3'(SEQ ID NO:47)和基因特异性反向引物5'-TTGCGCAGCACCATAAAGACGGT-3'(SEQ ID NO:46),采用了与第一轮扩增相同的PCR条件。产生了522bp的扩增子,将其连接至pGEM-T Easy并测序。

[0821] 将这三种扩增子的核苷酸序列组装成一个序列,其为推定的全长序列。然后从微胞藻CS-0170株基因组DNA扩增具有短5'UTR区的全长编码区,使用了正向引物5'-CAGGCGACGCGCGCCAGAGTCC-3'(SEQ ID NO:48),反向引物5'-TTATTAGTTACTTGGCCTTTACCTTC-3'(SEQ ID NO:49)和PFU Ultra II Fusion DNA聚合酶(Stratagene)。产生了860bp的扩增子,将其连接至pGEM-T Easy并测序。该基因可读框的序列为SEQ ID NO:1。

[0822] 该基因编码的全长氨基酸序列为SEQ ID NO:2。BLAST分析蛋白质序列显示该分离的cDNA编码Δ5-延长酶或Δ6-延长酶。这两种延长酶在氨基酸水平具有相似性,但从氨基酸序列本身不确定编码何种活性。将其作为查询序列使用BLASTP搜索Genbank蛋白质序列数据库时,微胞藻CS-0170延长酶与其他延长酶之间的最高相同性程度为与登录号CAL55414具有65%相同性,后者是*Ostreococcus tauri*多不饱和脂肪酸延长酶2的序列。保守性GNS1/SUR4家族结构域(NCBI保守结构域pfam01151)位于该序列的氨基酸49至274,其通常表明该蛋白质参与长链脂肪酸延长系统。

[0823] 图3给出了基于对与微胞藻CS-0170延长酶具有相似性的序列(包括那些用于设计原始简并引物的序列)的多重比对的序列关系树。

[0824] 在酵母中功能性表征微胞藻CS-0170 Δ6-延长酶

[0825] 该克隆的完整蛋白质编码区位于pGEM-T Easy的SalI/SphI片段内,将该编码区插入pYES2的XhoI/SphI位点,产生载体pYES2+MicElo1,用于引入酵母进行功能性表征。用pYES2+MicElo1转化酵母INVSC1株的细胞,在不含尿嘧啶的培养基上选择转化体。含有pYES2+MicElo1的酵母细胞在培养物中生长,用半乳糖诱导GAL启动子以表达MicElo1基因。向培养基中加入ALA、SDA或EPA(0.5mM)并在30℃继续培养48小时后,分析总细胞脂质中的脂肪酸。当在培养基中加入ALA时,检测到酵母转化体的细胞内脂质中的ETrA占总脂肪酸的0.2%,这代表较低但仍可检测到的0.4%的转换效率。类似地,当在培养基中加入SDA时,检测到酵母转化体的细胞内脂质中的ETA占总脂肪酸的0.2%,这代表0.4%的转换效率,表明存在低水平的Δ6-延长酶活性。然而,当在培养基中加入EPA时,没有检测到酵母转化体的细胞内脂质中存在DPA,表明酵母细胞缺乏Δ5-延长酶活性(表5)。

[0826] 表5.以表达分离自微胞藻CS-0170和塔胞藻CS-0140的延长酶的遗传学构建体转化的酵母细胞中的脂肪酸转换

克隆	脂肪酸前体/ 占总 FA 的%	形成的脂肪酸/ 占总 FA 的%	转换率
pYES2+Mic-Elo1	ALA, 18:3 ω 3 / 52.2%	ETrA, 20:3 ω 3 / 0.2%	0.4%
pYES2+Mic-Elo1	SDA, 18:4 ω 3 / 54.3%	ETA, 20:4 ω 3 / 0.2%	0.4%
pYES2+Mic-Elo1	EPA, 20:5 ω 3 / 2.0%	DPA, 22:5 ω 3 / 0%	0%
[0827] pYES2+Pyrco-Elo1	ALA, 18:3 ω 3 / 51.4%	ETrA, 20:3 ω 3 / 5.3%	9.3%
pYES2+Pyrco-Elo1	SDA, 18:4 ω 3 / 17.9%	ETA, 20:4 ω 3 / 34.1%	65.6%
pYES2+Pyrco-Elo1	EPA, 20:5 ω 3 / 2.1%	DPA, 22:5 ω 3 / 痕量	-
pYES2+Pyrco-Elo2	ALA, 18:3 ω 3 / 56.4%	ETrA, 20:3 ω 3 / 0.3%	0.5%
pYES2+Pyrco-Elo2	SDA, 18:4 ω 3 / 51.7%	ETA, 20:4 ω 3 / 0.7%	1.3%
pYES2+Pyrco-Elo2	EPA, 20:5 ω 3 / 0.6%	DPA, 22:5 ω 3 / 1.8%	75.0%

[0828] 分离和表征塔胞藻CS-0140 Δ 6-延长酶

[0829] 分离塔胞藻CS-0140 Δ 6-延长酶基因片段

[0830] 通过对GenBank登录号AB094747、CAI58897、CAJ30869、CAL23339和AAV67797的延长酶氨基酸序列进行比对,我们鉴定到了共有氨基酸序列区段KIYEFVDT (SEQ ID NO:33) 和VHVCMYT (SEQ ID NO:34),它们分别对应于AAV67797的氨基酸位置143-150和199-205。基于这两个区段的序列合成了简并引物5'-AARATMTAYGAGTTYGTIGATAC-3' (SEQ ID NO:50) 和5'-TAIGTGACATGCACACRTGWACCC-3' (SEQ ID NO:51) (缩写如上)。使用Superscript III One-Step RT-PCR系统以100ng的塔胞藻CS-0140总RNA进行RT-PCR扩增。产生了191bp的扩增子,将其连接至pGEM-T Easy并测序。

[0831] 分离全长塔胞藻CS-0140 Δ 6-延长酶基因

[0832] 设计引物用于通过5'-和3'-RACE延伸所述191bp片段。如实施例1所述分离基因的3'末端,使用了基因特异性正向引物5'-TTCGTGGATACGTTTCATCATGC-3' (SEQ ID NO:52)。产生了945bp的扩增子,将其连接至pGEM-T Easy并测序。使用GeneRacer Kit从1 μ g塔胞藻CS-0140总RNA分离基因的5'末端,其中55℃反转录温育1小时以产生5'-适应的cDNA,按照所附说明操作。以GeneRacer 5'引物和基因特异性反向引物5'-AGTTGAGCGCCGCCGAGAAGTAC-3' (SEQ ID NO:53) 进行PCR扩增,使用了PFU Ultra II Fusion DNA聚合酶。1:10稀释该产物并用1 μ l作为模板进行第二轮PCR,使用了GeneRacer 5'巢式引物5'-GGACACTGACATGGACTGAAGGAGTA-3' (SEQ ID NO:47) 和基因特异性反向引物5'-ACCTGGTTGACGTTGCCCTTCA-3' (SEQ ID NO:54),采用了与第一轮扩增相同的PCR条件。产生了743bp的扩增子,将其连接至pGEM-T Easy并测序。然后将这三部分序列组装成推定的全长序列。

[0833] 然后通过RT-PCR从总RNA扩增出具有短5'UTR区的全长编码区。使用了正向引物5'-GCTATGGAGTTCGCTCAGCCT-3' (SEQ ID NO:55) 和反向引物5'-TTACTACTGCTTCTTGCTGGCCAGCT-3' (SEQ ID NO:56) 以及100ng的塔胞藻CS-0140总RNA。产生了900bp的扩增子,将其连接至pGEM-T Easy并测序。该扩增子的可读框的核苷酸序列为SEQ ID NO:3,所编码的蛋白质的氨基酸序列为SEQ ID NO:4。

[0834] BLAST分析显示该全长氨基酸序列 (SEQ ID NO:4) 与其他 Δ 5-延长酶和 Δ 6-延长酶具有相似性。塔胞藻CS-0140延长酶与其他蛋白质之间的最高相同性程度为与登录号AAV67797具有54%相同性 (BLASTX),后者是*Ostreococcus tauri*多不饱和脂肪酸延长酶1

的序列。图4给出了基于对与塔胞藻CS-0140延长酶具有相似性的序列(包括那些用于设计原始简并引物的序列)的多重比对的序列关系树。保守性GNS1/SUR4家族结构域(NCBI保守结构域pfam01151)位于该序列的氨基酸52至297,其通常表明该蛋白质参与长链脂肪酸延长系统。

[0835] 在酵母中功能性表征塔胞藻CS-0140 Δ 6-延长酶

[0836] 该克隆的完整蛋白质编码区位于pGEM-T Easy的EcoRI片段内,将该编码区插入pYES2的EcoRI位点,产生载体pYES2+Pyrco-Elo1,用于引入酵母进行功能性表征。用pYES2+Pyrco-Elo1转化酵母INVSC1株的细胞,在不含尿嘧啶的培养基上选择转化体。含有pYES2+Pyrco-Elo1的酵母细胞在培养物中生长,用半乳糖诱导GAL启动子以表达Pyrco-Elo1cDNA。向培养基中加入脂肪酸至终浓度为0.5mM并在30℃继续培养48小时后,然后分析总细胞脂质中的脂肪酸。当在培养基中加入ALA时,检测到酵母转化体的细胞内脂质中的ETrA占总脂肪酸的5.3%,这代表9.3%的转换效率(Δ 9-延长酶活性)。当在培养基中加入SDA时,检测到酵母转化体的细胞内脂质中的ETA占34.1%,这代表65.6%的转换效率,这是高水平的 Δ 6-延长酶活性。然而,当在培养基中加入EPA时,没有检测到酵母转化体的细胞内脂质中存在DPA(表5),表明该cDNA编码的 Δ 6-延长酶活性在酵母细胞中具有一些 Δ 9-延长酶活性,但没有 Δ 5-延长酶活性。

[0837] 有关这两种 Δ 6-延长酶基因的以上结果表明,来自塔胞藻的基因编码的酶的活性大大高于来自微胞藻的基因的活性。这一结果是出乎预料的。已经排出了从微胞藻基因组DNA扩增的编码区含有突变或该编码区是不完整的此类可能性。

[0838] 实施例3:从微藻类分离并表征编码 Δ 5-延长酶的cDNA

[0839] 分离塔胞藻CS-0140 Δ 5-延长酶基因片段

[0840] 已经鉴定了CSIRO活微藻保藏中心的塔胞藻CS-0140株是具有高天然水平的 Δ 5-和 Δ 6-延长作用的微藻株(表4)。

[0841] 对GenBank登录号AAV67798和AB098084的延长酶氨基酸序列进行了比对。从众多相配的序列中我们选出了共有氨基酸序列区段YLELLDT(SEQ ID NO:35)和MYSYY(SEQ ID NO:36),它们分别对应于AAV67798的氨基酸位置136-142和198-202。基于这两个区段的序列合成了简并引物5'-ARTAYYTSGARYTRYTGAYAC-3'(SEQ ID NO:57)和5'-CATKARRTARTASGAGTACAT-3'(SEQ ID NO:58)(缩写如上)。按照实施例1所述使用Superscript III One-Step RT-PCR系统进行RT-PCR扩增。然后将0.5 μ l的作为模板,使用Taq DNA聚合酶(NEB)和相同的引物进行第二轮PCR。产生了200bp的扩增子,将其连接至pGEM-T Easy并测序。

[0842] 分离全长塔胞藻CS-0140 Δ 5-延长酶基因

[0843] 设计引物用于通过5'-和3'-RACE延伸所述200bp片段。如实施例1所述分离基因的3'末端,使用了基因特异性正向引物5'-CATCATACCCTGTTGATCTGGTC-3'(SEQ ID NO:59)和oligo-dT反向引物。产生了408bp的扩增子,将其连接至pGEM-T Easy(Promega)并测序。使用GeneRacer Kit从1 μ g塔胞藻CS-0140总RNA分离基因的5'末端,其中55℃反转录温育1小时以产生5'-适应的cDNA,按照所附说明操作。按照生产商所述,使用PFU Ultra II Fusion DNA聚合酶,以基因特异性反向引物5'-CCAGATCAACAGGGTATGATGGT-3'(SEQ ID NO:60)进行PCR扩增。1:10稀释该产物并用1 μ l作为模板进行第二轮PCR,使用了GeneRacer 5'巢式引物

5'-GGACACTGACATGGACTGAAGGAGTA-3' (SEQ ID NO:47) 和基因特异性反向引物5'-CGAAAGCTGGTCAAACCTTCTTGCGCAT-3' (SEQ ID NO:61)。产生了514bp的扩增子,将其连接至pGEM-T Easy (Promega)并测序。然后将这三部分序列组装成全长序列。

[0844] 然后通过RT-PCR从总RNA扩增出具有短5'UTR区的全长编码区。使用正向引物5'-AACATGGCGTCTATTGCGATTCCGGCT-3' (SEQ ID NO:62) 和反向引物5'-TTATTACTGCTTCTTGGCACCTTGCT-3' (SEQ ID NO:63),按照实施例1所述进行RT-PCR扩增。产生了810bp的扩增子,将其连接至pGEM-T Easy并测序。该插入物的可读框的核苷酸序列为SEQ ID NO:5,该cDNA所编码的推定的氨基酸序列为SEQ ID NO:6。

[0845] BLAST分析显示该全长氨基酸序列与其他 $\Delta 5$ -延长酶和 $\Delta 6$ -延长酶具有同源性。BLASTP分析发现,塔胞藻CS-0140延长酶与Genbank数据库中其他蛋白质之间的最高相同性程度为与登录号ABR67690具有46%相同性,后者对应于绿色巴夫藻C20延长酶。图5给出了基于对与塔胞藻CS-0140延长酶具有相似性的序列(包括那些用于设计原始简并引物的序列)的多重比对的序列关系树。

[0846] 在酵母中功能性表征塔胞藻CS-0140 $\Delta 5$ -延长酶

[0847] 该克隆的完整蛋白质编码区位于pGEM-T Easy中的cDNA的EcoRI片段内,将该编码区插入pYES2的EcoRI位点,产生载体pYES2+Pyrco-Elo2,用于引入酵母进行功能性表征。用pYES2+Pyrco-Elo2转化酵母INVSC1株的细胞,在不含尿嘧啶的培养基上选择转化体。含有pYES2+Pyrco-Elo2的酵母细胞在培养物中生长,用半乳糖诱导表达cDNA。向培养基中加入脂肪酸并在30℃继续培养48小时,然后分析总细胞脂质中的脂肪酸。当在培养基中加入ALA时,检测到酵母转化体的细胞内脂质中的ETrA占总脂肪酸的0.3%,这代表0.5%的转换效率($\Delta 9$ -延长酶活性)。当在培养基中加入SDA时,检测到酵母转化体的细胞内脂质中的ETA占0.7%,这代表1.3%的转换效率($\Delta 6$ -延长酶活性)。当在培养基中加入EPA时,检测到酵母转化体的细胞内脂质中的DPA占1.8%,这代表出乎预料的高达75%的转换效率,表明酵母细胞中存在很强的 $\Delta 5$ -延长酶活性(表6)。

[0848] 本发明人认为,在重组细胞内如此高效地将EPA转换为DPA还未曾报道。预期该酶在植物内(in planta)会有类似的高转换效率。保守性GNS1/SUR4家族结构域(NCBI保守结构域pfam01151)存在于该序列的氨基酸50至267处,这通常表明该蛋白质参与长链脂肪酸延长系统。

[0849] 表6.以表达分离自微胞藻CS-0170和塔胞藻CS-0140的延长酶的遗传学构建体转化的酵母细胞中脂肪酸的转换

[0850]	克隆	脂肪酸前体/ 占总 FA 的%	形成的脂肪酸/ 占总 FA 的%	转换率
	pYES2+ Mic-Elo1	ALA, 18:3 ω 3 / 52.2%	ETrA, 20:3 ω 3 / 0.2%	0.4%
	pYES2+ Mic-Elo1	SDA, 18:4 ω 3 / 54.3%	ETA, 20:4 ω 3 / 0.2%	0.4%
	pYES2+ Mic-Elo1	EPA, 20:5 ω 3 / 2.0%	DPA, 22:5 ω 3 / 0%	0%
	pYES2+ Pyrco-Elo1	ALA, 18:3 ω 3 / 51.4%	ETrA, 20:3 ω 3 / 5.3%	9.3%
	pYES2+ Pyrco-Elo1	SDA, 18:4 ω 3 / 17.9%	ETA, 20:4 ω 3 / 34.1%	65.6%
	pYES2+ Pyrco-Elo1	EPA, 20:5 ω 3 / 2.1%	DPA, 22:5 ω 3 / 痕量	-
	pYES2+ Pyrco-Elo2	ALA, 18:3 ω 3 / 56.4%	ETrA, 20:3 ω 3 / 0.3%	0.5%
	pYES2+ Pyrco-Elo2	SDA, 18:4 ω 3 / 51.7%	ETA, 20:4 ω 3 / 0.7%	1.3%
	pYES2+ Pyrco-Elo2	EPA, 20:5 ω 3 / 0.6%	DPA, 22:5 ω 3 / 1.8%	75.0%

[0851] 实施例4:从微藻类分离并表征编码 Δ 6-去饱和酶的基因

[0852] 合成全长微胞藻CCMP1545 Δ 6-去饱和酶基因

[0853] 采用BLASTP程序分析美国能源部联合基因组研究所(US Department of Energy Joint Genome Institute) (<http://www.jgi.doe.gov/>) 所产生的微胞藻CCMP1545过滤蛋白质模型基因组序列,使用了*Ostreococcus tauri* Δ 6-去饱和酶氨基酸序列(Genbank登录号AAW70159)作为查询序列。这一分析发现微胞藻CCMP1545中存在与AAW70159具有同源性的推定的蛋白质。使用微胞藻CCMP1545推定的蛋白质序列设计并合成了最适合于在双子叶植物例如欧洲油菜中表达的密码子优化的核苷酸序列。该蛋白质编码区的核苷酸序列见SEQ ID NO:7。质粒构建体命名为pGA4。其氨基酸序列见SEQ ID NO:8。

[0854] 使用微胞藻CCMP1545去饱和酶氨基酸序列SEQ ID NO:8作为查询序列对Genbank数据库中的其他蛋白质进行BLASTP分析发现,该蛋白质与 Δ 6-去饱和酶具有同源性。最高相同性程度为与登录号AAW70159的氨基酸序列的全长具有66%相同性,后者对应于*Ostreococcus tauri* Δ 6-去饱和酶的序列。图6给出了基于对与微胞藻CCMP1545去饱和酶具有相似性的序列的多重比对的序列关系树。这种前端去饱和酶在氨基酸54至104含有细胞色素b5结构域(NCBI保守结构域pfam00173)并在氨基酸172至428含有 Δ 6-FADS-样保守结构域(NCBI保守结构域cd03506)。代表前端去饱和酶的三个组氨酸盒分别存在于该序列的第190-195、227-232和401-405位。同时含有这两种结构域的蛋白质通常是合成高度不饱和脂肪酸所需的前端去饱和酶。有趣的是,这种去饱和酶与AAW70159紧密成簇,而后者是目

前已知的唯一经生化方法确认的植物样酰基-CoA去饱和酶。

[0855] 在酵母细胞中功能性表征微胞藻CCMP1545 $\Delta 6$ -去饱和酶

[0856] 微胞藻去饱和酶的完整编码区位于质粒pGA4的KpnI-SacI片段内,将该编码区插入酵母载体pYES2的KpnI-SacI位点,产生pYES2+Micd6D,用于引入酵母进行功能性表征。用pYES2+Micd6D转化酵母INVSC1株的细胞,在不含尿嘧啶的培养基上选择转化体。含有pYES2+Micd6D的酵母细胞在培养物中生长,用半乳糖诱导。向培养基中加入0.5mM的LA、ALA、ETrA、DGLA或ETA并在30℃继续培养48小时,然后分析总细胞脂质中的脂肪酸。当在培养基中加入LA时,检测到酵母转化体的细胞内脂质中的GLA占总脂肪酸的3.9%,这代表11.4%的 $\Delta 6$ -去饱和作用转换效率。当在培养基中加入ALA时,检测到酵母转化体的细胞内脂质中的SDA占总脂肪酸的13.9%,这代表39.0%的 $\Delta 6$ -去饱和作用转换效率。也就是说,对 $\omega 3$ 脂肪酸底物的转换效率是对相应的 $\omega 6$ 脂肪酸底物的转换效率的3.5-倍。当在培养基中加入ETrA时,检测到酵母转化体的细胞内脂质中的ETA占总脂肪酸的0.21%,这代表8.0%的 $\Delta 8$ -去饱和作用转换效率。然而,当在培养基中加入DGLA或ETA时,分别没有检测到ARA或EPA。这表明缺乏任何 $\Delta 5$ -去饱和作用活性(表7)。

[0857] 表7.以表达分离自微胞藻CCMP1545、Ostreococcus lucimarinus和塔胞藻CS-0140的去饱和酶的遗传学构建体转化的酵母细胞中的脂肪酸转换

克隆	脂肪酸前体 / 占总 FA 的%	形成的脂肪酸 / 占总 FA 的%	转换率
pYES2+Mic-d6D	LA, 18:2 $\omega 6$ / 30.3%	GLA, 18:3 $\omega 6$ / 3.9%	11.4%
pYES2+Mic-d6D	ALA, 18:3 $\omega 3$ / 21.7%	SDA, 18:4 $\omega 3$ / 13.9%	39.0%
pYES2+Mic-d6D	ETrA, 20:3 $\omega 3$ / 2.4%	ETA, 20:4 $\omega 3$ / 0.21%	8.0%
pYES2+Mic-d6D	DGLA, 20:3 $\omega 6$ / 2.6%	ARA, 20:4 $\omega 6$ / 0%	-
pYES2+Mic-d6D	ETA, 20:4 $\omega 3$ / 6.2%	EPA, 20:5 $\omega 3$ / 0%	-
pYES2+Ostlu-d6D	LA, 18:2 $\omega 6$ / 29.5%	GLA, 18:3 $\omega 6$ / 2.1%	6.6%
pYES2+Ostlu-d6D	ALA, 18:3 $\omega 3$ / 21.8%	SDA, 18:4 $\omega 3$ / 13.8%	38.8%
pYES2+Ostlu-d6D	ETrA, 20:3 $\omega 3$ / 2.2%	ETA, 20:4 $\omega 3$ / 0%	-
pYES2+Ostlu-d6D	GLA, 18:3 $\omega 6$ / 29.2%	18:4 $\omega 6$ / 0%	-
pYES2+Ostlu-d6D	SDA, 18:4 $\omega 3$ / 41.7%	18:5 $\omega 3$ / 0%	-
pYES2+Ostlu-d6D	DGLA, 20:3 $\omega 6$ / 2.3%	ARA, 20:4 $\omega 6$ / 0%	-
pYES2+Ostlu-d6D	ETA, 20:4 $\omega 3$ / 4.9%	EPA, 20:5 $\omega 3$ / 0%	-
pYES2+Pyrco-d5D	LA, 18:2 $\omega 6$ / 35.1%	GLA, 18:3 $\omega 6$ / 0%	-
pYES2+Pyrco-d5D	ALA, 18:3 $\omega 3$ / 40.9%	SDA, 18:4 $\omega 3$ / 0%	-
pYES2+Pyrco-d5D	DGLA, 20:3 $\omega 6$ / 2.9%	ARA, 20:4 $\omega 6$ / 0.12%	4.0%
pYES2+Pyrco-d5D	ETA, 20:4 $\omega 3$ / 7.2%	EPA, 20:5 $\omega 3$ / 0.26%	3.5%

[0859] 在植物细胞中功能性表征微胞藻CCMP1545 $\Delta 6$ -去饱和酶

[0860] 使用增强型本氏烟草瞬时表达系统按照实施例1所述在植物内证实了微胞藻CCMP1545 $\Delta 6$ -去饱和酶(Mic1545-d6D)和车前叶蓝蓟 $\Delta 6$ -去饱和酶(Echp1-d6D; Zhou等, 2006)(用作阳性对样品)的酶活性。通过将含有35S启动子的PstI片段插入至载体pORE04的SfoI位点(经T4DNA聚合酶处理以提供平端)(Coutu等, 2007)而制备命名为35S-pORE04的载体。通过将位于SwaI片段内的pGA4的完整编码区插入至35S-pORE04的SmaI-EcoRV位点而

制备遗传学构建体35S:Mic1545-d6D,产生pJP2064。

[0861] 将这些嵌合载体引入根癌土壤杆菌菌株AGL1,来自这些培养物的细胞渗入温室中的本氏烟草植物的叶组织。渗入后,植物再生长5天,然后取叶盘进行GC分析,结果显示两种基因在本氏烟草中均发挥 $\Delta 6$ -去饱和酶的功能。

[0862] 以车前叶蓝蓟 $\Delta 6$ -去饱和酶转化的叶组织含有GLA(0.4%)和SDA(1.2%),分别代表3.8%和4.4%的转换效率。以微胞藻CCMP1545 $\Delta 6$ -去饱和酶转化的叶组织含有SDA(2.2%),代表转换效率为6.9%,但没有可检测到的GLA。叶组织缺乏GLA可能是由于在植物内极度偏向 $\omega 3$ 底物ALA而非 $\omega 6$ 底物LA,或是部分由于存在天然的本氏烟草 $\omega 3$ 去饱和酶活性,其可能将 $\Delta 6$ -去饱和作用所产生的部分GLA转换为SDA。以往用具有 $\omega 3$ 底物偏向性的酰基-PC $\Delta 6$ -去饱和酶进行的实验已经表明这是可能的(Sayanova等,2006),不过该研究中没有对其发生的程度进行定量。

[0863] 微胞藻 $\Delta 6$ -去饱和酶的 $\omega 3$ -底物偏向性

[0864] 分离自微胞藻的 $\Delta 6$ -去饱和酶在植物内与在酵母内一样对 $\omega 3$ 底物具有令人惊讶的强偏向性。在酵母中表达的该酶对 $\omega 3$ -去饱和的脂肪酸底物的活性是对相应的 $\omega 6$ -去饱和的脂肪酸底物的活性的3.5倍。鉴于所报道的*O. tauri*的酶缺乏偏向性(Domergue等,2005),因此所观察到的对 $\omega 3$ 底物的偏向性完全是令人惊讶且出乎预料的。对*O. tauri* $\Delta 6$ -去饱和酶在酵母或植物种子中表达的研究表明其对于LA和ALA具有相似的活性。

[0865] 因此可以预期,与使用对 $\omega 3$ -去饱和的底物不具有偏向性的去饱和酶相比,在植物内使用该基因或其他对 $\omega 3$ -去饱和的脂肪酸底物具有如此高特异性的基因连同作为重组VLC-PUFA途径的一部分的其他脂肪酸去饱和酶和延长酶,能够提高EPA、DPA和DHA的水平。推测导致这种升高的原因是降低了LA转换为GLA以及随后的 $\omega 6$ PUFA DGLA和ARA,真菌或酵母 $\Delta 17$ -去饱和酶在植物内不能高效地将它们转换为其 $\omega 3$ 对应物。尽管已经分离到了对 $\omega 3$ 脂肪酸底物具有偏向性的 $\Delta 6$ 去饱和酶(Sayanova等,2003),但其对磷脂连接的酰基链具有活性。相反,推测从微胞藻获得的所述去饱和酶对酰基辅酶A底物具有活性。

[0866] 微胞藻CCMP1545 $\Delta 6$ -去饱和酶的 $\Delta 6/\Delta 8$ 双重功能

[0867] 有趣的是注意到微胞藻CCMP1545 $\Delta 6$ -去饱和酶表现出显著水平的 $\Delta 8$ -去饱和酶活性且因此具有显著的双重活性,这不同于不具有可检测到的 $\Delta 8$ -去饱和酶活性的*Ostreococcus lucimarinus*酶(下文)。推测这种双重去饱和酶活性可用于构建植物中的双重 $\Delta 6/\Delta 8$ -去饱和酶途径,或用于当构建此类途径所使用的延长酶同时具有 $\Delta 9$ -延长酶和 $\Delta 6$ -延长酶活性的情况中。使用这种基因可能有助于通过将ETrA转换为ETA,其随后被 $\Delta 5$ -去饱和为EPA,从而降低ETrA的积聚。

[0868] 合成全长*Ostreococcus lucimarinus* $\Delta 6$ -去饱和酶基因

[0869] 使用*Ostreococcus tauri* $\Delta 6$ -去饱和酶核苷酸序列(登录号AY746357)作为查询序列,通过BLASTX分析非冗余蛋白序列的GenBank数据库。根据这一分析鉴定到了*Ostreococcus lucimarinus*的一种基因,其编码登录号为XP_001421073的氨基酸序列的一种部分长度的蛋白质。然后检查了编码XP_001421073的区域两侧的基因组DNA序列,以鉴定推定的翻译起始和终止密码子,以便确定全长蛋白质编码区,其核苷酸序列为SEQ ID NO:9。随后将该编码区翻译为蛋白质序列,其为SEQ ID NO:10。使用这一氨基酸序列来设计并合成最适合于在欧洲油菜和其他双子叶植物中表达的密码子优化的核苷酸序列,其核苷酸

序列为SEQ ID NO:11。

[0870] 使用*Ostreococcus lucimarinus*去饱和酶氨基酸序列作为查询序列对Genbank数据库中的其他蛋白质进行BLASTP分析,发现SEQ ID NO:10与 $\Delta 6$ -去饱和酶具有同源性。最高相同性程度为与登录号AAW70159的氨基酸序列的全长具有76%相同性,后者对应于*Ostreococcus tauri* $\Delta 6$ -去饱和酶的序列。图7给出了基于对与*Ostreococcus lucimarinus*去饱和酶具有相似性的序列的多重比对的序列关系树。这种前端去饱和酶在氨基酸55至108含有细胞色素b5结构域(NCBI保守结构域pfam00173)并在氨基酸198至444含有 $\Delta 6$ -FADS-样保守结构域(NCBI保守结构域cd03506)。代表前端去饱和酶的三个组氨酸盒分别存在于该序列的第207-212、244-249和417-421位。同时含有这两种结构域的蛋白质通常是合成高度不饱和脂肪酸所需的前端去饱和酶。有趣的是,这种去饱和酶与AAW70159紧密成簇,而后者是目前已知的唯一经生化方法确认的植物样酰基-CoA去饱和酶。

[0871] 在酵母中功能性表征*Ostreococcus lucimarinus* $\Delta 6$ -去饱和酶

[0872] *Ostreococcus*基因(SEQ ID NO:11)的完整编码区位于pGEM-T Easy的NotI片段内,将该编码区插入pYES2的NotI位点,产生嵌合载体pYES2+OstIud6D,用于引入酵母进行功能性表征。用pYES2+OstIud6D转化酵母INVSC1株的细胞,在不含尿嘧啶的培养基上选择转化体。含有pYES2+OstIud6D的酵母细胞在培养物中生长,然后用半乳糖诱导。向培养基中加入LA、ALA、SDA或EPA至每种的终浓度为0.5mM并在30°C继续培养48小时,然后分析细胞脂质中的脂肪酸。当在培养基中加入LA时,检测到酵母转化体的细胞内脂质中的GLA占总脂肪酸的2.1%,这代表 $\Delta 6$ -去饱和作用转换效率为6.6%。当在培养基中加入ALA时,检测到酵母转化体的细胞内脂质中的SDA占总脂肪酸的13.8%,这代表 $\Delta 6$ -去饱和作用转换效率为38.8%。然而,当在培养基中加入ETra、DGLA或ETA时,分别没有检测到ETA、ARA或EPA,表明没有任何 $\Delta 5$ -或 $\Delta 8$ -去饱和活性(表7),也表明与具有相同长度和不饱和模式的 $\omega 6$ 脂肪酸相比,其对 $\omega 3$ 脂肪酸底物具有偏向性。

[0873] 去饱和酶的酰基-CoA底物特异性

[0874] 与其他 $\Delta 6$ -去饱和酶相比,本实施例中所描述的去饱和酶与先前分离自*Ostreococcus tauri*的 $\Delta 6$ -去饱和酶具有更加密切的关联(图9)。在将这些基因与所述去饱和酶家族的其他成员放在一起的系统发生树中更可以看出这种相似性(图10)。已经报道了*Ostreococcus tauri* $\Delta 6$ -去饱和酶对酰基辅酶A底物具有活性(Domergue等,2005)。基于这些发现,推测上述基因编码的 $\Delta 6$ -去饱和酶可能对酰基辅酶A底物而非酰基-PC底物具有活性。有趣的是,盐生巴夫藻 $\Delta 5$ -去饱和酶也与*O. tauri* $\Delta 6$ -去饱和酶和所述 $\Delta 8$ -去饱和酶成簇,形成一个独立的分支。

[0875] 为了确定细小微胞藻(微胞藻CCMP1545) $\Delta 6$ -去饱和酶是否能够以酰基-CoA脂肪酸作为底物且由此产生 $\Delta 6$ -去饱和的酰基-CoA脂肪酸,仅以编码所述去饱和酶的基因构建体转化酿酒酵母,并将所转化的细胞系一式三份的培养物生长在存在250 μ M外源性18:3 $\Delta 9,12,15$ 的条件下。从这些培养物提取总脂质并通过薄层色谱(TLC)分级为中性脂质(NL)、磷脂酰胆碱(PC)、磷脂酰肌醇(PI)、磷脂酰丝氨酸(PS)和磷脂酰乙醇胺(PE)等类型,随后从每一类产生FAME,并通过GC进行分析。结果显示于表8。

[0876] 表8. 以克隆至pYES2的细小微胞藻 $\Delta 6$ -去饱和酶转化的酿酒酵母中总脂质和分级的中性脂质(NL)、磷脂酰胆碱(PC)、磷脂酰肌醇(PI)、磷脂酰丝氨酸(PS)和磷脂酰乙醇胺

(PE)的脂肪酸组成(占总脂肪酸的百分比)

	脂肪酸	总量	NL	PC	PI	PS	PE
	16:0	24.3±1.5	24.5±4.3	36.8±3.4	45.6±2.4	45.8±3.9	37.0±3.9
	16:1 ^{Δ3t}	14.7±0.4	12.1±0.9	15.7±0.5	8.1±0.7	19.4±1.0	21.5±3.5
	18:0	8.1±0.7	11.5±1.4	11.7±0.7	17.4±3.1	0.9±1.3	3.6±3.8
[0877]	18:1 ^{Δ9}	11.0±1.3	10.0±2.0	7.0±1.2	14.6±0.9	26.3±1.8	18.6±2.5
	18:3 ^{Δ9,12,15}	12.1±0.7	12.2±2.4	7.1±0.8	5.6±0.1	2.3±2.0	6.9±2.3
	18:4 ^{Δ6,9,12,15}	29.4±0.9	29.4±1.9	21.0±0.2	7.7±0.7	4.1±0.3	9.9±1.9
	其他	0.3	0.3	0.8	1.1	1.2	2.5
	总计	100	100	100	100	100	100

[0878] 在总脂质中,71%的18:3^{Δ9,12,15}已经被Δ6-去饱和为18:4^{Δ6,9,12,15}。与总脂质提取物相比,在PC级份中没有检测到产物富集。实际上,PC级份中18:4^{Δ6,9,12,15}的百分比显著低于总脂质中的百分比(21.0%vs.29.4%),表明所述去饱和酶以酰基-CoA硫酯的形式产生18:4^{Δ6,9,12,15}(Domergue等,2003)。

[0879] 通过转化将编码细小微胞藻Δ6-去饱和酶的基因引入拟南芥植物。通过将位于SwaI片段内的编码细小微胞藻Δ6-去饱和酶的完整编码区插入Linin-pWVEC8的SmaI位点产生遗传学构建体Linin:Micpu-d6D,产生linP-mic1545-d6D-linT。该构建体的启动子是来自亚麻的种子特异性linin启动子。将该构建体转化至拟南芥Columbia生态型,通过GC分析转化植物的T2种子内的脂肪酸组成(图11)。

[0880] 在酵母和本氏烟草中的生化研究均提示来自细小微胞藻的所述Δ6-去饱和酶是酰基-CoA去饱和酶。其他研究采用对确保延长步骤的动力学分析作为间接方法来确定去饱和酶产生酰基-CoA产物的能力:所述Δ6-去饱和产物(SDA)在随后Δ6-延长步骤(出现于酰基-CoA代谢库)中的可利用度受到所述Δ6-去饱和酶的底物特异性的影响(Domergue等,2003,2005;Hoffmann等,2008)。使用来自*O. tauri*和细小微胞藻的所述Δ6-去饱和酶获得了类似的Δ6-延长速度,这不同于使用车前叶蓝蓟酰基-PCΔ6-去饱和酶时所观察到的显著低水平的延长作用(图12a)。当分析酵母脂质类型中所述Δ6-去饱和酶产物SDA的分布时观察到了进一步的证据(表8)。与总脂质成分相比,没有观察到PC成分中的富集,不过,可预期的是如果是使用酰基-PC去饱和酶产生SDA,应该出现这种富集(Domergue等,2005)。我们的研究中所观察到的这种相对低水平的Δ6-去饱和作用(图12)是预料之中的,因为本氏烟草叶中的大部分底物LA和ALA位于质体中,不能用于去饱和作用。不过,由于在制备FAME过程中也分离到这些脂肪酸,它们的存在有效降低了所计算的总体转换效率。因此可以预期这些基因的种子特异性转换效率会高得多。

[0881] 酰基-CoAΔ6-去饱和酶和酰基-PCΔ6-去饱和酶之间的比较

[0882] 在植物细胞中进一步比较了微胞藻CCMP1545Δ6-去饱和酶、车前叶蓝蓟Δ6-去饱和酶和*Ostreococcus tauri*Δ6-去饱和酶(Domergue等,2005)。将实施例4所述的遗传学构建体35S:Mic1545-d6D和35S:Echp1-d6D与遗传学构建体35S:Ostta-d6D进行对比,后者是通过将位于SwaI片段内的*Ostreococcus tauri*Δ6-去饱和酶的完整编码区插入至35S-pORE04的SmaI-EcoRV位点而制备的,这产生pJP3065。

[0883] 车前叶蓝蓟与*O. tauri*和细小微胞藻EPA途径直接比较显示,酰基-CoA去饱和酶途径产生远远更高水平的EPA,这是因为两者均有更高效的 $\Delta 6$ -去饱和作用和更高效的随后的 $\Delta 6$ -延长作用(图12a)。车前叶蓝蓟 $\Delta 6$ -去饱和酶催化转换14%的 $\omega 3$ 底物($18:3^{\Delta 9,12,15}$ 转换为 $18:4^{\Delta 6,9,12,15}$)和30%的 $\omega 6$ 底物($18:2^{\Delta 9,12}$ 转换为 $18:3^{\Delta 6,9,12}$)。使用*O. tauri* $\Delta 6$ -去饱和酶导致24%的 $\omega 3$ 转换和40%的 $\omega 6$ 转换,而使用细小微胞藻 $\Delta 6$ -去饱和酶导致27%的 $\omega 3$ 转换和15%的 $\omega 6$ 转换。这些转换导致车前叶蓝蓟途径产生1.3%的 $20:4^{\Delta 5,8,11,14}$ 和3.4%的 $20:5^{\Delta 5,8,11,14,17}$,*O. tauri*途径产生1.2%的 $20:4^{\Delta 5,8,11,14}$ 和9.6%的 $20:5^{\Delta 5,8,11,14,17}$,而细小微胞藻途径产生0.6%的 $20:4^{\Delta 5,8,11,14}$ 和10.7%的 $20:5^{\Delta 5,8,11,14,17}$ 。

[0884] 与使用车前叶蓝蓟去饱和酶相比,当使用*O. tauri*或细小微胞藻 $\Delta 6$ -去饱和酶产生底物 $18:4^{\Delta 6,9,12,15}$ 时, $\Delta 6$ -延长作用远远更高(图12a)。除了细小微胞藻 $\Delta 6$ -去饱和酶所显示出的 $\omega 3$ 底物特异性之外,*P. cordata* $\Delta 6$ -延长酶(见实施例2)被证明具有更高的特异性,并且其转换 $\omega 3$ 底物 $18:4^{\Delta 6,9,12,15}$ 的速度远高于转换 $18:3^{\Delta 6,9,12}$ 的速度(对于细小微胞藻EPA途径分别为89%和21%)。

[0885] 使用双重 $\Delta 6$ -去饱和酶途径

[0886] 进行对比来研究通过使用含有两种 $\Delta 6$ -去饱和酶的途径来提高 $\Delta 6$ -去饱和作用的可能性。首先,车前叶蓝蓟酰基-PC去饱和酶和细小微胞藻酰基-CoA去饱和酶的组合没有使得转换效率显著高于仅含有细小微胞藻去饱和酶的途径(图12b)。当车前叶蓝蓟与*O. tauri* $\Delta 6$ -去饱和酶组合时得到了类似的结果。与使用*O. tauri*或细小微胞藻途径之一相比,*O. tauri*与细小微胞藻去饱和酶组合的双重酰基-CoA $\Delta 6$ -去饱和酶途径也没有导致 $\omega 3$ 转换效率升高(图12c)。

[0887] 也测试了在产生EPA的途径中使用双重 $\Delta 6$ -去饱和酶的效果。首先测试的是在分开的实验中将来自车前叶蓝蓟的酰基-PC去饱和酶与这两种酰基-CoA去饱和酶组合。假设加入脂质连接的去饱和酶可能会提高酰基-PC底物LA或ALA分别向GLA或SDA的转换。类似地,我们还测试了使用两种酰基-CoA去饱和酶是否有可能提高EPA的积聚。在本氏烟草中进行的瞬时测定中这些假设无一得到证实。

[0888] 实施例5:分离和表征编码来自微藻类的 $\Delta 5$ -去饱和酶的基因

[0889] 分离塔胞藻CS-0140 $\Delta 5$ -去饱和酶基因片段

[0890] 对来自GenBank登录号ABL96295、ABP49078、XP_001421073、AAM09687、AAT85661、AAW70159和AAX14505的去饱和酶氨基酸序列进行比对鉴定了共有氨基酸序列区段WKNMHKHHHA (SEQ ID NO:37) 和HHLFPSMP (SEQ ID NO:38),分别对应于ABL96295的氨基酸位置197-206和368-375。使用CODEHOP程序(Rose等,1998)基于这两个区段的序列设计了简并引物5'-GGTGGAAGAACAAGCACAAACrdncaycaygc-3' (SEQ ID NO:64) 和5'-GGGCATCGTGGGwanarrtgrtg-3' (SEQ ID NO:65)。使用Taq DNA聚合酶(NEB)进行降落PCR扩增,体积为20 μ L,使用了10pmol的各种引物,50ng的塔胞藻CS-0140基因组DNA,按照所附手册的说明使用缓冲液和核苷酸。循环条件为:1个循环94 $^{\circ}$ C 3分钟;20个循环的94 $^{\circ}$ C 1分钟,70 $^{\circ}$ C 2分钟(每一循环-1 $^{\circ}$ C),72 $^{\circ}$ C 1分钟;20个循环的94 $^{\circ}$ C 1分钟,55 $^{\circ}$ C 1分钟,72 $^{\circ}$ C 1分钟;1个循环72 $^{\circ}$ C 5分钟;4 $^{\circ}$ C 保持。产生了551bp的扩增子,将其连接至pGEM-T Easy (Promega)并测序。

[0891] 分离全长塔胞藻CS-0140 $\Delta 5$ -去饱和酶基因

[0892] 如实施例1所述设计引物用于通过5'-和3'-RACE延伸所述551bp片段。如实施例1所述分离编码所述 Δ 5-去饱和酶的基因的cDNA的3'末端,使用了基因特异性正向引物5'-AGCGAGTACCTGCATTGGGT-3'(SEQ ID NO:66)和修饰的oligo-dT反向引物。产生了477bp的扩增子,将其连接至pGEM-T Easy并测序。如实施例1所述通过改良的末端转移酶方法分离所述基因的5'末端。基因特异性反向引物为5'-ATAGTGCTTGGTGC GCAAGCTGTGCCT-3'(SEQ ID NO:67)。两轮PCR扩增之后产生了317bp的扩增子,将其连接至pGEM-T Easy (Promega)并测序。将这三部分序列组装成推定的全长基因序列。

[0893] 然后从基因组DNA扩增具有短5'UTR区的全长蛋白质编码区,PCR扩增使用了正向引物5'-CACCATGGGAAAGGGAGGCAATGCT-3'(SEQ ID NO:68)和反向引物5'-TTACTAGTGCGCCTTGGAGTGAGAT-3'(SEQ ID NO:69)和PFU Ultra II Fusion DNA聚合酶(Stratagene),体积为20 μ L,使用了4pmol的各种引物和50ng的塔胞藻CS-0140基因组DNA,按照所附PFU Ultra II Fusion手册的说明使用缓冲液成分。产生了代表全长cDNA的1336bp的扩增子,将其连接至pGEM-T Easy并测序。该cDNA的可读框的核苷酸序列为SEQ ID NO:12。

[0894] BLAST分析蛋白质序列显示该基因编码的全长氨基酸序列(SEQ ID NO:13),与 Δ 5-去饱和酶或 Δ 6-去饱和酶具有相似性。这两种去饱和酶在氨基酸水平具有相似性,但从氨基酸序列本身不确定编码何种活性。按照下文所述进行了酶活性分析,发现所编码的蛋白质具有 Δ 5-去饱和酶活性。经BLASTP确定,塔胞藻CS-0140去饱和酶与Genbank数据库中其他去饱和酶之间的最高相同性程度为与登录号EDQ92231具有52%相同性,后者是来自领鞭毛虫MX1的具有未确定酶活性的蛋白质的氨基酸序列。图8给出了基于对与塔胞藻CS-0140去饱和酶具有相似性的序列(包括那些用于设计原始简并引物的序列)的多重比对的序列关系树。这种前端去饱和酶在氨基酸16至67含有细胞色素b5结构域(NCBI保守结构域pfam00173)并在氨基酸159至411含有 Δ 6-FADS-样保守结构域(NCBI保守结构域cd03506)。代表前端去饱和酶的三个组氨酸盒分别存在于该序列的第175-180、212-217和384-388位氨基酸。含有这些结构域的蛋白质通常是合成多不饱和脂肪酸所需的前端去饱和酶。

[0895] 在酵母中功能性表征塔胞藻CS-0140 Δ 5-去饱和酶

[0896] 该克隆的完整编码区位于pGEM-T Easy的NotI片段内,将该编码区插入pYES2(Invitrogen)的NotI位点,产生pYES2+Pyrco-des2,用于引入酵母进行功能性表征。用pYES2+Pyrco-des2转化酵母INVSC1株的细胞,在不含尿嘧啶的培养基上选择转化体。含有pYES2+Pyrco-des2的酵母细胞在培养物中生长,然后用半乳糖诱导。向培养基中加入0.5mM LA、ALA、DGLA或ETA并在30 $^{\circ}$ C继续培养48小时,然后分析细胞脂质中的脂肪酸。当在培养基中加入DGLA时,检测到酵母转化体的细胞内脂质中的ARA占总脂肪酸的0.12%,这代表 Δ 5-去饱和作用转换效率为4.0%。当在培养基中加入ETA时,检测到酵母转化体的细胞内脂质中的EPA占总脂肪酸的0.26%,这代表 Δ 6-去饱和作用转换效率为3.5%。然而,当在培养基中加入LA或ALA时,酵母转化体分别没有产生GAL或SDA。这表明该蛋白质在酵母细胞中没有任何 Δ 6-去饱和活性(表7)。

[0897] 在植物细胞中表达P.cordata Δ 5-去饱和酶

[0898] 使用如实施例1所述的增强型本氏烟草瞬时表达系统在植物内证明以下的酶的酶活性:微胞藻CCMP1545 Δ 6-去饱和酶(SEQ ID NO:8,其编码序列为SEQ ID NO:7)、塔胞藻

CS-0140 Δ 6-延长酶 (SEQ ID NO:4, 其编码序列为SEQ ID NO:3) 和塔胞藻CS-0140 Δ 5-去饱和酶 (SEQ ID NO:13, 其编码序列为SEQ ID NO:12) 以及拟南芥DGAT1 (SEQ ID NO:74, 其编码序列为SEQ ID NO:75)。

[0899] 将位于EcoRI片段内的塔胞藻CS-0140 Δ 5-去饱和酶的完整编码区插入35S-pORE04 (实施例4, 见上) 的EcoRI位点, 产生35S:Pyrco-d5D, 由此制备编码置于组成型35S启动子控制下的 Δ 5-去饱和酶的遗传学构建体35S:Pyrco-d5D。将嵌合载体35S:Mic1545-d6D (实施例10)、35S:Pyrco-d6E (实施例10) 和35S:Pyrco-d5D分别引入根癌土壤杆菌菌株AGL1, 混合来自这些细胞的培养物的转基因细胞, 使得混合物渗入温室中的本氏烟草植物的叶组织。渗入后, 植物再生长5天, 然后取叶盘进行GC分析, 结果显示这些基因在本氏烟草中均发挥产生EPA的功能。以这些基因转化的叶组织含有SDA (1.0%)、ETA (0.1%)、EPA (10.0%)。叶组织还含有痕量水平的GLA、ETA和ARA。计算出的所述 Δ 5-去饱和酶转换效率为98.8%。

[0900] 本实验证实微藻 Δ 5-去饱和酶能够在植物细胞中以至少90%或至少95%的效率将ETA转换为EPA。

[0901] 实施例6: 分离和表征编码来自微藻类的 ω 3-去饱和酶的基因

[0902] 分离微胞藻CS-0170 ω 3-去饱和酶基因片段

[0903] 为了确定微藻类例如微胞藻是否具有编码 ω 3去饱和酶的基因, 且或许会鉴定到这样的基因, 搜索微胞藻RCC299株基因组序列中与FAD3显示出同源性的基因。但该搜索未能鉴定到任何候选基因。因此, 本发明人考虑微胞藻中是否有其他类型的去饱和酶能够充当 ω 3去饱和酶。这一假设得到了如下发现 (实施例4) 的支持, 即该株中的 Δ 6去饱和酶是属于前端、酰基-CoA依赖型的。然而, 经检查, 微胞藻RCC299基因组似乎含有至少30个推定的脂肪酸去饱和酶的基因, 而且不知道其中有哪些 (如果有的话) 可能编码 ω 3去饱和酶。

[0904] 在一个实验中, 对来自Genbank登录号BAD91495、ABL63813、BAD11952和AAR20444的去饱和酶氨基酸序列进行比对鉴定到了共有氨基酸序列区段WCIGHDCG (SEQ ID NO:39) 和TFLQHHEDEM (SEQ ID NO:40), 分部对应于BAD91495的氨基酸位置106-113和296-305。使用CODEHOP程序基于这两个区段的序列设计了简并引物5'-TGTGGTGCATCGGCCAYGANKSNGG-3' (SEQ ID NO:70) 和5'-TGTCCTCGTCGTTGTGCTGNARRWANGT-3' (SEQ ID NO:71)。使用TaqDNA聚合酶 (NEB) 进行降落PCR扩增, 体积为20 μ L, 使用了10pmol的各种引物, 50ng的微胞藻CS-0170基因组DNA, 按照所附手册的说明使用缓冲液和核苷酸。循环条件为: 1个循环94 $^{\circ}$ C 3分钟; 20个循环的94 $^{\circ}$ C 1分钟, 70 $^{\circ}$ C 2分钟 (每一循环-1 $^{\circ}$ C), 72 $^{\circ}$ C 1分钟; 35个循环的94 $^{\circ}$ C 1分钟, 56 $^{\circ}$ C 1分钟, 72 $^{\circ}$ C 1分钟; 1个循环72 $^{\circ}$ C 5分钟; 4 $^{\circ}$ C保持。产生了528bp的扩增子, 将其连接至pGEM-T Easy (Promega) 并测序。该扩增子的核苷酸序列为SEQ ID NO:14, 所编码的部分蛋白质序列为SEQ ID NO:15。

[0905] 合成全长微胞藻RCC299 ω 3-去饱和酶基因

[0906] 将简并PCR所产生的528bp片段与完整的微胞藻RCC299过滤蛋白质模型基因组序列 (由美国能源部联合基因组研究所 (<http://www.jgi.doe.gov/>) 产生) 进行比较。BLAST分析发现在微胞藻RCC299的13号染色体上的一个区域与SEQ ID NO:14之间存在具有高同源性的区域。基于这两个序列近乎相同, 因此微胞藻CS-0170和RCC299具有密切的关联 (微胞藻RCC299的核苷酸序列为SEQ ID NO:16)。使用微胞藻RCC299推定的蛋白质序列 (SEQ ID

N0:17) 设计并合成了最适合于在欧洲油菜或其他双子叶植物中表达的密码子优化的核苷酸序列(SEQ ID N0:18)。在酵母中测试了该基因从SEQ ID N0:18的164位核苷酸开始的一个较短形式,但没有检测到 ω 3去饱和酶活性。

[0907] BLAST分析表明全长氨基酸序列(SEQ ID N0:17)与FAT-1、FAT-2和 ω 3去饱和酶具有同源性。但不可能基于序列本身推断编码何种活性。经BLASTP确定,微胞藻CS-0170去饱和酶与Genbank数据库中其他蛋白质之间的最高相同性程度为与XP_001899085.1具有35%相同性,后者是脂肪酸去饱和酶家族中的*Brugia malayi*蛋白。该前端去饱和酶含有 Δ 12-FADS-样保守结构域(NCBI保守结构域cd03507)。含有这两种结构域的蛋白质通常是合成脂肪酸所需的前端去饱和酶,包括 ω 3去饱和酶家族。

[0908] 在植物中功能性表征微胞藻RCC299 ω 3-去饱和酶

[0909] 使用上述增强型本氏烟草瞬时表达系统在植物内测试分离自微胞藻RCC299的全长基因编码的推定的 ω 3-去饱和酶(Mic299-w3D,见上)和致病疫霉 Δ 17-去饱和酶(Phyind17D,GenBank登录号CAM55882)(用作阳性对样品)的酶功能。

[0910] 将位于EcoRI片段内的SEQ ID N0:18所示的完整蛋白质编码区克隆入载体35S-pORE04(实施例4)的EcoRI位点,产生命名为pJP2073的遗传学构建体,由此制备了35S:Mic299-w3D构建体。将位于EcoRI片段内的致病疫霉 Δ 17去饱和酶的完整编码区克隆入载体35S-pORE04的EcoRI位点,产生pJP2074,由此制备了35S:Phyind17D构建体。类似地,将位于EcoRI片段内的拟南芥DGAT1(AF051849)的完整编码区克隆入载体35S-pORE04的EcoRI位点,产生pJP2078,由此制备了35S:Arath-DGAT1构建体。

[0911] 根癌土壤杆菌菌株AGL1在28℃生长于LB培养液中,其中补充了50mg/mL卡那霉素和50mg/mL利福平,直至达到静止生长期。室温下5000g离心15分钟沉淀细菌,然后以OD600=1.0垂悬于渗入缓冲液,该渗入缓冲液含有10mM MES pH 5.7,10mM MgCl₂和100 μ M乙酰丁香酮。细胞随后在28℃震荡培养3小时,然后将等体积的含有35S:p19、35S:Arath-DGAT1以及35S:Phyind17D或35S:Mic299-w3D的土壤杆菌细胞培养物混合,随后渗入到叶组织中。制备花生四烯酸盐并如上所述补料给转化的叶组织,在补料底物后5小时和24小时取叶盘进行分析。以35S:Phyind17D构建体渗入的叶斑点,或以35S:Mic299-w3D构建体分开地渗入的叶斑点均证实分别以37%和50%的效率将ARA(20:4 ω 6)转换为EPA(20:5 ω 3)(图13),这表明该蛋白质具有 Δ 17-去饱和酶活性。

[0912] 讨论:表征了第一种具有 Δ 17-去饱和酶活性的微藻 ω 3-去饱和酶

[0913] 本研究所所述的微胞藻RCC299 ω 3去饱和酶是第一种微藻 Δ 17去饱和酶,即植物样 Δ 17去饱和酶,对C20或更长的脂肪酸底物具有活性。并不知道陆生植物具有前端去饱和酶类型的 ω 3去饱和酶,而是FAD3类型的。因此,发现微藻株(其比真菌更接近植物)具有前端去饱和酶类型的 ω 3去饱和酶是令人惊讶的。

[0914] 基于与其他去饱和酶的同源性,认为有可能的情况是,在上述实验中用作对照基因的真菌致病疫霉去饱和酶对酰基-PC底物具有活性,而微胞藻RCC299去饱和酶对酰基辅酶A底物具有活性。已知其他真菌去饱和酶对酰基-PC底物具有活性。关于微胞藻基因的这一结论与所观察到的其与来自相同株的 Δ 6-去饱和酶基因具有相似性这一点是一致的(实施例4)。可通过底物补料研究进一步测定这种底物优先性,在底物补料研究中,补料给转化组织的底物例如ARA可立即被酰基-CoA库利用,但却仅在被天然植物(例如,本氏烟草)酰基

转移酶转换后才能被酰基-PC库利用。

[0915] 微胞藻RCC299 ω 3去饱和酶基因对于构建用于在植物中产生EPA和下游脂肪酸DPA和DHA以及其他 ω 3 VLC-PUFA的重组途径是非常有用的,这特别是因为其能够将 ω 6底物例如ARA转换为 ω 3产物。当其与同样在酰基-CoA库中起作用的延长酶例如 Δ 5-延长酶组合时,对酰基辅酶A底物的活性增强了这种用途。此外,微胞藻株的脂肪酸谱表明微胞藻酶还能够将 ω 6 C18脂肪酸例如GLA或LA分别转换为它们的 ω 3对应物例如SDA或ALA。可在酵母细胞或在植物中通过以上就底物ARA所述的底物补料而证实GLA转换为SDA,而LA转换为ALA则在酵母细胞中得到更好的证实,这是因为植物中存在内源性 Δ 15去饱和酶。

[0916] 鉴定其他 ω 3-去饱和酶

[0917] 采用BLASTP程序分析美国能源部联合基因组研究所 (<http://www.jgi.doe.gov/>) 所产生的微胞藻CCMP1545过滤蛋白质模型基因组序列,使用了SEQ ID NO:17作为查询序列。这一分析发现微胞藻CCMP1545 (EuGene.0000150179) 中存在与SEQ ID NO:17具有同源性的基因。其可读框序列为SEQ ID NO:19,蛋白质序列为SEQ ID NO:20。

[0918] BLAST分析表明全长氨基酸序列SEQ ID NO:20与FAT-1、FAT-2和 ω 3去饱和酶具有同源性。微胞藻CCMP1545去饱和酶与Genbank数据库 (BLASTP) 中其他蛋白质之间的最高相同性程度为在全长上与SEQ ID NO:17具有59%相同性,后者是来自领鞭毛虫MX1的具有未确定酶活性的蛋白质的氨基酸序列。

[0919] 该前端去饱和酶含有 Δ 12-FADS-样保守结构域 (NCBI保守结构域cd03507)。含有这两种结构域的蛋白质通常是合成脂肪酸所需的前端去饱和酶,包括 ω 3去饱和酶家族。我们推测该蛋白质在植物内将同样作为具有 Δ 17-去饱和酶活性的 ω 3去饱和酶而起作用。

[0920] 实施例7:分离和表征编码来自微藻类的 Δ 9-延长酶的其他基因

[0921] 分离并表征Emiliana huxleyi CCMP1516 Δ 9-延长酶

[0922] 使用了GenBank登录号AF390174的氨基酸序列作为查询序列,采用BLASTP程序分析美国能源部联合基因组研究所 (<http://www.jgi.doe.gov/>) 所产生的Emiliana huxleyi CCMP1516过滤蛋白质模型基因组序列。这一分析发现Emiliana huxleyi CCMP1516中存在与AF390174具有同源性的推定的基因。其蛋白质序列为SEQ ID NO:28,编码核苷酸序列为SEQ ID NO:27。BLAST分析表明全长氨基酸序列与PUFA延长酶具有同源性。Emiliana huxleyi CCMP1516延长酶与其他蛋白质之间的最高相同性程度 (BLASTP) 为与AF390174具有80%相同性。该序列中存在保守性GNS1/SUR4家族结构域 (NCBI保守结构域pfam01151),这通常表明该蛋白质参与长链脂肪酸延长系统。

[0923] 使用Emiliana huxleyi CCMP1516推定的蛋白质序列设计并合成了最适合于在双子叶植物例如欧洲油菜中表达的密码子优化的核苷酸序列 (SEQ ID NO:29)。该质粒构建体命名为0835668_Emihu-d9E_pMA。

[0924] 分离并表征P. pinguis和盐生巴夫藻 Δ 9-延长酶

[0925] 为了鉴定P. pinguis和盐生巴夫藻 Δ 9-延长酶中可能存在的保守区域,对来自E. huxleyi Δ 9-延长酶、PLL00000665 (使用E. huxleyi延长酶序列作为查询序列,通过BLAST分析鉴定的来自TBestDB的路氏巴夫藻EST序列) 和Genbank登录号AAL37626 (球等鞭金藻 Δ 9-延长酶) 的推定的延长酶氨基酸序列进行比对。由此发现了共有氨基酸序列区段VDTRKGAYR (SEQ ID NO:76) 和FIHTIMYTY (SEQ ID NO:77),它们分别对应于Emihu-d9E的氨

基 酸 位 置 40-48 和 170-178。基 于 这 两 个 区 段 的 序 列 设 计 了 简 并 引 物 5'-TG GT GG AC AC A AG GA AG GG NG CN TA YM G-3' (SEQ ID NO:78) 和 5'-GTAGGTGTACATGATGGTRTGDATRAA-3' (SEQ ID NO:79), 并使用来自 *P. pinguis* 的 RNA 和来自盐生巴夫藻 (Zhou 等, 2007) 的 cDNA 文库进行 RT-PCR 和 PCR 扩增, 其中使用 Superscript IIITM Platinum[®] One-Step RT-PCR 系统或 Taq DNA 聚合酶 (NEB, Ipswich, MA, USA)。

[0926] 通过 RT-PCR 从 *P. pinguis* 中产生了 641bp 的扩增子, 将其连接至 pGEM-T Easy[®] 并测序。设计引物用于通过 5'-和 3'-RACE 延伸所述 641bp 片段。RT-PCR 分离基因的 3' 末端, 使用了基因特异性正向引物 5'-GTCCTTGCTCCAGGGCTCCACCA-3' (SEQ ID NO:80) 和 oligo-dT-SP6 反向引物 5'-ATTTAGGTGACACTATAGTTTTTTTTTTTTTTTTTTT-3' (SEQ ID NO:81)。产物以 1:10 进行稀释, 取 1.0 μ l 用作模板, 以 Taq DNA 聚合酶 (NEB) 进行第二轮 PCR, 使用了基因特异性正向引物 5'-TTCCAGAACGAGGGCATCTACGT-3' (SEQ ID NO:82) 和相同的反向引物。产生了 1079bp 的扩增子, 将其连接至 pGEM-T[®] Easy 并测序。从 1.0 μ g 的 *P. pinguis* cDNA 分离该基因的 5' 末端, 所述 cDNA 是使用基因特异性反向引物 5'-TTGGGTGATCTGCATGAGCGTGATG-3' (SEQ ID NO:83) 和末端转移酶 A-加尾产生的。然后将该 cDNA 用作模板, 使用 oligo-dT-SP6 引物和基因特异性引物 5'-CGAATACTTGAAGAGCTTGTTGGAGA-3' (SEQ ID NO:84) 进行 PCR 反应。产物以 1:10 进行稀释, 取 1.0 μ l 用作模板, 使用 oligo-dT-SP6 引物和基因特异性引物 5'-GGGCTACGAGCTGGCAGATGAAGCA-3' (SEQ ID NO:85) 进行第二轮 PCR。产生了 323bp 的扩增子, 将其连接至 pGEM-T[®] Easy 并测序。以这三部分序列组装成全长序列。然后从总 RNA RT-PCR 扩增具有短 5' UTR 区的全长编码区, RT-PCR 扩增使用了正向引物 5'-GAAAAAATGGTTGCGCCACCCATCA-3' (SEQ ID NO:86) 和反向引物 5'-TCACTACTTCTTCTTCTTGCCCGCGGC-3' (SEQ ID NO:87)。产生了 828bp 的扩增子, Pavpi-Elo1, 将其连接至 pGEM-T[®] Easy 并测序 (SEQ ID NO:93)。 *P. pinguis* Δ 9-延长酶的推定的氨基酸序列为 SEQ ID NO:94。

[0927] 类似地, 使用简并引物通过 PCR 从盐生巴夫藻产生了 425bp 的扩增子, 将其连接至 pGEM-T Easy[®] 并测序。设计引物用于通过 5'-和 3'-RACE 延伸所述 425bp 片段, 通过 RT-PCR 分离基因的 3' 末端, 使用了基因特异性正向引物 5'-TTCCGGTACTCAGCGGTGGCG-3' (SEQ ID NO:88) 和 oligo-dT-SP6 反向引物。产生了 776bp 的扩增子, 将其连接至 pGEM-T[®] Easy 并测序。通过 PCR 从盐生巴夫藻 cDNA 文库分离该基因的 5' 末端, 使用了 M13R 引物 5'-CAGGAAACAGCTATGAC-3' (SEQ ID NO:89) 和基因特异性反向引物 5'-ACGTAGATGCCCTCGTTCTG-3' (SEQ ID NO:90), 按照生产商的说明使用 PfuUltra II[®] Fusion DNA 聚合酶。产生了 710bp 的扩增子, 将其连接至 pGEM-T[®] Easy 并测序。从这三部分序列组装出全长序列。通过 RT-PCR 从总 RNA 扩增出具有具有短 5' UTR 区的全长编码区, 使用了正向引物 5'-CACCGAATGGCGACTGAAGGGATGCC-3' (SEQ ID NO:91) 和反向引物 5'-CTACTCGGTTTTTCATGCGGTTGCTGGA-3' (SEQ ID NO:92)。产生了 846bp 的扩增子, Pavsa-Elo3, 将其连接至 pGEM-T[®] Easy 并测序 (SEQ ID NO:95)。盐生巴夫藻 Δ 9-延长酶推定的氨基酸序列为 SEQ ID NO:96。

[0928] 在植物细胞中功能性表征 Δ 9-延长酶

[0929] 将位于EcoRI片段中的分别来自质粒0835668_Emihu-d9E_pMA、pGEMT+Pavpi-d9E和pGEMT+Pavsa-d9E的Emiliana延长酶(Emihu-d9E)、Pavlova pinguis延长酶(Pavpi-d9E)和盐生巴夫藻延长酶(Pavsa-d9E)的完整编码区插入35S-pORE04的EcoRI位点,产生了35S:Emihu-d9E(命名为pJP3027)、35S:Pavpi-d9E(命名为pJP3103)、35S:Pavsa-d9E(命名为pJP3081)和35S:Isoga-d9E(命名为pJP2062)。如实施例1所述,使用增强型本氏烟草瞬时表达系统在植物内证实了Emihu-d9E、Pavpi-d9E和Pavsa-d9E连同作为阳性对照样品的Isoga-d9E(Qi等,2002)的酶活性。

[0930] 将这些嵌合载体引入根癌土壤杆菌菌株AGL1,使来自这些培养物的细胞渗入温室中的本氏烟草植物的叶组织。渗入后,植物再生长5天,然后取叶盘进行GC分析,结果显示,产物脂肪酸的存在表明这两种基因在例如本氏烟草的植物细胞中均发挥 $\Delta 9$ -延长酶的功能。

[0931] 以Emiliana huxleyi CCMP1516 $\Delta 9$ -延长酶转化的叶组织含有20:2 ^{$\Delta 11,14$} (6.6%)和20:3 ^{$\Delta 11,14,17$} (6.4%),这代表对LA和ALA的转换效率分别为39.9%和12.4%。以Pavlova pinguis $\Delta 9$ -延长酶转化的叶组织含有20:2 ^{$\Delta 11,14$} (10.1%)和20:3 ^{$\Delta 11,14,17$} (6.6%),这代表转换效率分别为56.0%和13.3%。以盐生巴夫藻 $\Delta 9$ -延长酶转化的叶组织含有20:2 ^{$\Delta 11,14$} (7.7%)和20:3 ^{$\Delta 11,14,17$} (4.6%),这代表转换效率分别为45.0%和9.2%。以球等鞭金藻 $\Delta 9$ -延长酶转化的叶组织含有20:2 ^{$\Delta 11,14$} (9.2%)和20:3 ^{$\Delta 11,14,17$} (7.5%),这代表转换效率分别为48.9%和15.4%(表9)。

[0932] 表9.以 $\Delta 9$ -延长酶瞬时转化的本氏烟草叶组织的脂肪酸组成(占总脂肪酸的百分比)。给出了以一式三份进行的不同渗入之间的标准差。

	脂肪酸	对照	Emihu- $\Delta 9E$	Pavsa- $\Delta 9E$	Pavpi- $\Delta 9E$	Isoga- $\Delta 9E$
	普通 FA					
	16:0	15.7±0.6	14.6±0.1	15.2±0.6	14.5±0.8	14.2±0
[0933]	16:1 ^{$\Delta 3t$}	1.5±0	1.4±0	1.3±0.1	1.3±0	1.3±0.1
	16:3 ^{$\Delta 9,12,15$}	6.8±0.7	6.5±0.8	6.1±0.8	6.1±1.4	7.4±0.5
	18:0	3.0±0.1	3.4±0.2	4.2±0.2	3.4±0.5	3.1±0.2
	18:1 ^{$\Delta 9$}	2.2±0	2.9±0.2	3.2±0.2	3.6±0.7	3.2±0.1
	18:2 ^{$\Delta 9,12$}	11.8±0.4	9.9±0.1	9.4±0.2	7.9±0.5	9.7±0.4

	18:3 ^{Δ9,12,15}	56.0±1.4	45.3±2.2	46.0±1.4	43.4±1.8	41.5±1.3
	其他次要成分	3.0±0	3.0±0	2.3±0	3.1±0	2.9±0
	总计	100	87.0	87.7	83.3	83.3
	新的ω6 PUFA					
	20:2 ^{Δ8,11}	—	6.6±1.0	7.7±0.7	10.1±0.7	9.2±0.8
	20:3 ^{Δ8,11,14}	—	—	—	—	—
	20:4 ^{Δ5,8,11,14}	—	—	—	—	—
	22:4 ^{Δ7,10,13,16}	—	—	—	—	—
	22:5 ^{Δ4,7,10,13,16}	—	—	—	—	—
[0934]	总计	0	6.6	7.7	10.1	9.2
	新的ω3 PUFA					
	20:3 ^{Δ11,14,17}	—	6.4±1.5	4.6±0.7	6.6±0.4	7.5±0.2
	20:4 ^{Δ8,11,14,17}	—	—	—	—	—
	20:5 ^{Δ5,8,11,14,17}	—	—	—	—	—
	22:5 ^{Δ7,10,13,16,19}	—	—	—	—	—
	22:6 ^{Δ4,7,10,13,16,19}	—	—	—	—	—
	总计	0	6.4	4.6	6.6	7.5
	总的新 FA	0	13.0	12.3	16.7	16.7
	总 FA	100	100	100	100	100

[0935] 叶组织中这种对ω3底物ALA的明显的高偏向性是预料之中的,因为本氏烟草叶中的大部分底物ALA位于质体内,因此无法被质体外延长作用所利用,也因为在直接甲基化过程中质体和细胞质的ALA从叶中被分离,这人为降低了ω3转换率。*E. huxleyi*和球等鞭金藻Δ9-延长酶在本氏烟草中展现出相同的底物偏向性,ω3至ω6的转换率为0.31。观察到的ω6库中最高效的转换涉及盐生巴夫藻Δ9-延长酶,其将56.0%的底物发生了转换。相反,13.3%的ω3底物发生了转换,两者的比率为0.24。*P. pinguis*的酶对ω6底物展现出最高的偏向性,转换率为0.20,因为45.0%的ω6发生转换,而仅有9.2%的ω3发生转换。

[0936] 实施例8. 构建包括Δ9延长酶的生物合成途径以产生ARA

[0937] 构建转基因Δ-9延长酶途径

[0938] 从二元载体pJP101acq构建了含有球等鞭金藻Δ9-延长酶(氨基酸序列GenBank登录号AF390174-可读框如SEQ ID NO:21所示,氨基酸序列如SEQ ID NO:22所示)、盐生巴夫藻Δ8-去饱和酶(登录号ABL96296-可读框如SEQ ID NO:23所示,氨基酸序列如SEQ ID NO:24所示)、和盐生巴夫藻Δ5-去饱和酶(登录号ABL96295-可读框如SEQ ID NO:25所示,氨基酸序列如SEQ ID NO:26所示)的二元载体。此载体没有基因插入物的设计见图14。

[0939] 首先,将含有球等鞭金藻 $\Delta 9$ -延长酶的pBluescript克隆的SmaI-EcoRV片段连接至pJP101acq的SmaI位点以产生pJP105。将含有盐生巴夫藻 $\Delta 5$ -去饱和酶的pBluescript克隆的XhoI片段连接至pJP105的XhoI位点以产生pJP106。将含有盐生巴夫藻 $\Delta 8$ -去饱和酶的pBluescript克隆的NotI片段连接至pJP106的NotI位点以产生pJP107,其在图15中图示。

[0940] 关于该设计有若干要点值得关注。首先,所述3个基因中的2个在T-DNA上背驰转录,即彼此远离。这么做是为了防止从受关注的任一基因的转录潜在地干扰另一个基因的表达,由此使两者的表达最大化。其次,在该基因构建体中的第三个基因(在此情况下编码所述 $\Delta 8$ -去饱和酶)通过插入间隔区而与导向相同方向的第二个基因(编码所述 $\Delta 9$ -延长酶)分隔开。据认为上游基因的终止密码子与下游基因的启动密码子之间至少1.0kb的距离会降低前者的转录干扰后者或潜在地引起基因沉默的风险。第三,该三个基因每个的5' -UTR被修饰为包括TMV前导序列(已知其用于提供有效率的翻译)。任何其它已知赋予高翻译效率的5' UTR序列也可代替所述TMV序列而被使用。

[0941] 通过电穿孔将pJP107引入土壤杆菌菌株AGLI内并且经转化的菌株被用于将所述基因构建体引入至拟南芥内,所述拟南芥生态型MC49,其是fad3/fae1突变体具有高水平的LA作为所述 $\Delta 8$ -去饱和酶的潜在初始脂肪酸底物。使用花浸渍法(floral dipping method,参见Clough and Bent,1998)来进行植物转化和分析。将来自经处理的植物(T0植物)的种子(T1种子)铺板于潮霉素(20mg/L)选择性培养基上,选择转化的植物并将其转移至土壤以建立24株经确认的T1转基因植物。预期这些T1植物的大部分对于所引入的基因构建体都是杂合的。在成熟时收集来自该24株转基因植物的T2种子并分析其脂肪酸组成。这些T2株系包括所述基因构建体是纯合的株系以及所述基因构建体是杂合的株系。从所述T2种子中含有最高ARA水平的6个建立T2植物,在含潮霉素(20mg/mL)的MS培养基上选择,以确定所述转基因的存在。例如,所述T2种子栽培自称作FW-10的T1植物,含有5.8%的ARA并且在潮霉素培养基上显示出抗性后代对易感后代3:1的分离比,表明FW-10在单一的基因座上含有所述基因构建体。分析了来自FW-10的T3种子组的脂肪酸谱,数据在表10中给出。

[0942] 表10.转化了所述基因构建体pJP107(含有球等鞭金藻 $\Delta 9$ -延长酶基因、盐生巴夫藻 $\Delta 8$ -去饱和酶基因、和盐生巴夫藻 $\Delta 5$ -去饱和酶基因)的拟南芥种子的脂肪酸组成。

[0943]

	FW10-23	对照 MC49
样品	P1235	P1254
14:0	0.0	0.1
16:1 ω7	0.6	0.4
16:0	9.5	8.4
18:2 ω6	30.9	50.9
18:3 ω3	0.0	1.0
18:1 ω9	21.4	30.9
18:1 ω7	3.3	3.5
18:0	4.3	3.4
20:4 ω6	21.0	0.0
20:5 ω3	1.3	0.0
20:3 ω6	1.1	0.0
20:4 ω3*	0.2	0.0
20:2 ω6	2.6	0.0
20:3 ω3	0.2	0.0
20:1 ω9/ ω1	1.6	0.2
20:0	0.7	0.5
22:4 ω6	0.8	0.0
22:5 ω3	0.0	0.0
22:0	0.0	0.2
24:1 ω11/13	0.2	0.1
24:0	0.2	0.2
总和	100	100
总和ω6 PUFA	56	51

[0944]

%转换		
$\Delta 9E$		
LA $\rightarrow$ 20:2 $\omega 6$	45	0
ALA $\rightarrow$ 20:3 $\omega 3$	100	0
$\Delta 8D$		
20:2 $\omega 6 \rightarrow$ 20:3 $\omega 6$	90	0
20:3 $\omega 3 \rightarrow$ 20:4 $\omega 3$	88	0
$\Delta 5D$		
20:3 $\omega 6 \rightarrow$ 20:4 $\omega 6$	95	0
20:4 $\omega 3 \rightarrow$ 20:5 $\omega 3$	90	0

[0945] 如表10中所总结的,未经转化的拟南芥(生态型MC49)的种子含有显著量的前体 $\omega 6$ 底物LA,但却不含有任何ARA或预期在所述 $\Delta 9$ 延长酶途经中出现的中间体脂肪酸。相反,来自经转化的植物FW10-23(含有pJP107构建体)的种子含有显著量的20:2 $n-6$ 、20:3 $n-6$ 和20:4 $n-6$ (ARA),包括21%的ARA,用LA开始的三个酶步骤的产物。此外,种子油中低水平的ALA(在对照MC49中1.0%)被非常有效地转换成EPA,其在经转化的株系FW10-23中以1.3%的水平存在。

[0946] 讨论:转换效率及生化含义

[0947] 可以通过测试FW-10-23中底物脂肪酸转换成产物脂肪酸(包括随后的衍生物)的转换百分比来评估所述pJP107构建体所编码的单个酶步骤的相对效率。在 $\omega 6$ 库中,球等鞭金藻 $\Delta 9$ 延长酶显示出45%的LA转换成EDA以及随后的不饱和脂肪酸。在相同的种子里,盐生巴夫藻 $\Delta 8$ -去饱和酶和 $\Delta 5$ -去饱和酶分别显示了 $\omega 6$ 脂肪酸至其相关产物的90%和95%的转换效率。作为比较,在 $\omega 3$ 库中,球等鞭金藻 $\Delta 9$ 延长酶显示了基本上100%的ALA至延长的产物的转换,而盐生巴夫藻 $\Delta 8$ -去饱和酶和 $\Delta 5$ -去饱和酶分别显示了88%和90%的转换效率。这些酶步骤导致了1.3%EPA的合成,尽管拟南芥MC49的背景仅含有低水平的ALA。在最戏剧化的结果中,注意到在种子油中未检测到ALA,表明基本上100%的ALA由所述 $\Delta 9$ 延长酶转换成ALA的延长的产物。

[0948] 有趣的是,注意到FW-10-23中发现的不寻常中间体脂肪酸的水平很低(在 $\omega 3$ 库中<0.4%)并且与各种海鲜的食物链中已发现的那些相当(表11)。尽管未经转化的MC49种子油仅含有低水平的ALA,并且这可能造成了例如中间体脂肪酸ETrA的低观测水平;但是据预测,当以具有较高ALA水平的遗传背景来组建相同的途径时,所得的种子油依旧具有相对较低水平的ETrA(<3%)。存在如此低水平的这些中间体很可能是由于经 $\Delta 9$ 延长的中间体非常有效的去饱和作用。

[0949] 表11.转化了所述基因构建体pJP107(含有球等鞭金藻 $\Delta 9$ -延长酶基因、盐生巴夫藻 $\Delta 8$ -去饱和酶基因、和盐生巴夫藻 $\Delta 5$ -去饱和酶基因)的拟南芥种子中的脂肪酸,与在海鲜样品的范围中发现的中间体脂肪酸的比较。

[0950]

	P1235	海鲜	
样品		平均	范围 – 最大
14:0	0.0	1.7	31.1
16:1 ω7	0.6	2.9	8.2
16:0	9.5	18.7	53.6
18:2 ω6	30.9	1.9	14.6
18:3 ω3*	0.0	0	0
18:1 ω9	21.4	13.9	59.5
18:1 ω7	3.3	3	7.9
18:0	4.3	8.5	14.7
20:4 ω6	21.0	6.7	19.1
20:5 ω3	1.3	7.1	22.2
20:3 ω6	1.1	0.3	1.5
20:4 ω3	0.2	0.5	2.8
20:2 ω6	2.6	0.4	1.8
20:3 ω3*	0.2	0	0
20:1 ω9/ ω11	1.6	2.2	21.1
20:0	0.7	0.4	4.2
22:4 ω6	0.8	1	4.4
22:5 ω3	0.0	2.4	14.9
22:0	0.0	0.2	0.7
24:1 ω9 11/13*	0.2	0	0
24:0	0.2	0.2	1.6
总和ω6 PUFA	56	10	41

[0951]

% 转换			
$\Delta 9E$			
LA-->20:2 $\omega 6$	45	82	
ALA-->20:3 $\omega 3$	100	100	
$\Delta 8D$			
20:2 $\omega 6$ -->20:3 $\omega 6$	90	95	
20:3 $\omega 3$ -->20:4 $\omega 3$	88	100	
$\Delta 5D$			
20:3 $\omega 6$ -->20:4 $\omega 6$	95	96	
20:4 $\omega 3$ -->20:5 $\omega 3$	90	95	

[0952] 值得注意的是盐生巴夫藻 $\Delta 8$ -去饱和酶在将ETrA转换成ETA方面比起其它经报道的 $\Delta 8$ -去饱和酶要有效地多,特别是在与所述 $\Delta 9$ 延长酶和 $\Delta 5$ 去饱和酶共表达时。例如,据报道,当小眼虫 $\Delta 8$ -去饱和酶与小眼虫或球等鞭金藻 $\Delta 9$ -延长酶在大豆胚中共表达时, $\omega 3$ 和 $\omega 6$ 底物的转换效率分别是64%和73%。上述实验中所观测的每一步的效率以及ALA转换为EPA的总转换效率也比Qi等(2004)所报道的在拟南芥叶中的高得多,其中他们仅观测到3.0%的EPA以及实质水平的不期望的中间体包括ETrA(4.6%)。

[0953] 已知延长酶仅接近酰基-CoA库中的底物。随后的 $\Delta 8$ -去饱和酶和 $\Delta 5$ -去饱和酶步骤在经转化的种子中被观测到以极高的效率发挥作用,尽管所述经 $\Delta 9$ -延长的产物毫无疑问地在酰基-CoA库中产生,这一事实强烈表明两种盐生巴夫藻去饱和酶均能够高效地接近酰基辅酶A底物。

[0954] 用所述 $\Delta 9$ -延长酶途径生物合成高的ARA和EPA水平

[0955] 从单个步骤效率的这些数据和观测,预计到可以用修改的 $\Delta 9$ -延长酶途径在转基因植物例如拟南芥、芸苔、大豆、亚麻籽或棉花中产生高水平的ARA和EPA以及随后的DPA和DHA。进一步预计到通过再加入三种酶功能的任一种可达到甚至更高的水平,所述三种酶功能即:酰基-CoA $\Delta 12$ -去饱和酶功能以提高酰基-CoA库中可用的底物LA的数量用于 $\Delta 9$ -延长作用,第二是加入 $\Delta 15$ -去饱和酶以提高ALA的水平用于直接转换成EPA,以及第三是 $\Delta 17$ -去饱和酶其可将ARA转换为EPA例如实施例6中所述的。更优选地,加入酰基-CoA $\Delta 12$ -去饱和酶和所述 $\Delta 15$ -去饱和酶或所述 $\Delta 17$ -去饱和酶中的任意一种会提供最高的水平。因此,使用能够接近酰基-CoA库中底物的酶预期会导致更有效地转换成EPA、DHA和DHA。

[0956] 考虑到拟南芥MC49北京含有fad3突变,其导致低水平的ALA聚集(1-3%),则所观测的1.3%EPA的合成是值得注意的并且出乎意料的。我们预计,当此(或类似的) $\Delta 9$ -延长酶途径($\Delta 9$ -延长酶、 $\Delta 8$ -去饱和酶、和 $\Delta 5$ -去饱和酶)被转化至含有高水平ALA的植物时,会得到高水平的EPA。例如,我们预计将此途径转化至过表达紫苏(*Perilla frutescens*) $\Delta 15$ -去饱和酶或其它 $\Delta 15$ -去饱和酶基因的拟南芥株系,会导致EPA的水平为种子油中总脂肪酸的至少25%。

[0957] 实施例9:PUFA途径基因在植物细胞中的表达

[0958] 植物稳定转化的一种替代是在叶中转基因的瞬时表达,例如最早由Kapila等(1997)所介绍的。用此技术,允许叶细胞的细胞核(Zipfel等,2006)经由偏离轴的空气间隔的渗入而转化了在 T_{DNA} 界限内载有表达构建体的土壤杆菌培养物。转基因在叶中的表达显著地由病毒阻抑蛋白的共引入而增强,所述病毒阻抑蛋白如P19(Voinnet等,2003)和HC-Pro(Johansen and Carrington,2001;Kasschau等,2003),其抑制宿主细胞的转基因沉默装置并将转基因的表达扩展至更长的时间段。

[0959] 叶具有复杂的脂质代谢,其由质体半乳糖脂单半乳糖基二酰基甘油(MGDG)和二半乳糖基二酰基甘油(DGDG)的巨大的库所控制。更罕见的脂肪酸的库存在于质体的室之外包括经卵磷脂(PC)、辅酶A(CoA)以及单-和二-酰基甘油酯(MAG、DAG;Ohlrogge和Browse,1995)所酯化的那些。此实施例中所用的LC-PUFA合成的酶位于内质网(ER;Napier,2007)其中他们可接触相对较少见的经PC和CoA酯化的叶脂质库。PC-CoA-连接的反应的代谢产物,例如在ER上有活性的那些,可通过二酰基甘油酯-0-酰基转移酶(DGAT;Bouvier-Nave等,2000)的过表达而在三酰基甘油酯(TAG)中聚积。与MAG或DAG相比,位于TAG上的脂肪酸更具代谢惰性并且有比较小的可能会重进入脂质生物合成途径或进入质体。重要的是,使用标准薄膜层析技术(TLC)可轻易地将TAG从位于叶质体中的丰度更高的脂质类别中分离出来。因此增强的TAG聚积和TAG/脂质类别纯化的组合可以有助于更完全地理解叶ER上的LC-PUFA酶反应。

[0960] 在此实施例中用编码去饱和酶和延长酶的基因来测试此系统的LC-PUFA生产。

[0961] 用于瞬时表达的质粒构建体

[0962] 通过将基因的编码区克隆至经修饰版本的Coutu et al(2007)所述的pORE04二元载体中来制备二元载体,所述pORE04二元载体中花椰菜花叶病病毒(CaMV)35S启动子已被克隆至SfoI位点以产生35S-pORE04。将球等鞭金藻 $\Delta 9$ -延长酶基因编码区(Genbank登记号AAL37626)(SEQ ID NO:21)从基因组DNA扩增并克隆至35S-pORE04内的EcoRI位点。将进过植物表达密码子优化的所述三种盐生巴夫藻去饱和酶(Genbank登记号ABL96296、ABL96295和AAY15136-每个都在WO 2005/103253中有描述)作为SwaI插入物而克隆至35S-pORE04的EcoRV-SmaI位点内。将未经优化的盐生巴夫藻 $\Delta 5$ -延长酶(Genbank登记号AAY15135)作为an XhoI-XbaI片段克隆至35S-pORE04内的XhoI-NheI位点。由CaMV 35S-推动的版本的P19病毒阻抑物由Dr Peter Waterhouse所提供。拟南芥的DGAT1基因编码区(Genbank登记号AAF19262)(SEQ ID NO:74)通过RT-PCR获得并作为BamHI-EcoRV片段被克隆至35S-pORE04中的相应位点。总RNA用RNeasy mini kit(QIAGEN)分离自致病疫霉以及用PlatinumSuperscript III One-Step(QIAGEN)进行RT-PCR。将所得的含有致病疫霉 $\Delta 17$ -去饱和酶蛋白编码区(WO 2005/012316)的扩增子克隆pGEMT-Easy(Promega)内并测序。继而将EcoRI片段克隆至35S-pORE04内。

[0963] 土壤杆菌渗入和本氏烟草生长条件

[0964] 载有每个二元载体的根癌土壤杆菌菌株AGL1在28℃于补加了适当抗生素的LB培养液中生长。将培养物离心并小心地将其重悬浮两体积的渗入缓冲液(5mM MES,5mM $MgSO_4$,pH 5.7,100 μ M乙酰丁香酮)以及进一步生长3小时。测量每个培养物的光密度并制备了培养物的最终组合,从而使每个土壤杆菌构建体等于 OD_{600nm} 0.2或如图16所另外标注。如

Voinnet等(2003)所述将细胞渗入一个月大的本氏烟草植物的叶的下面,所述本氏烟草植物已被安置在23℃的植物生长室内,且具有10:14的光:暗循环。用记号笔将渗入的区域画圈。在渗入之后,将该植物置于28℃一小时,之后将他们转移到24℃的植物生长室直至分析。除非另有指定,所有的本氏烟草农杆菌渗入均在存在分离的含P19病毒阻抑蛋白的二元构建体的情况下进行

[0965] 脂质分析

[0966] 本实施例中在用甲醇/HCl/二氯甲烷(DCM;体积10/1/1)溶液于80℃转甲基化2h以产生脂肪酸甲基酯(FAME)之后,通过GC和GC-MS分析叶组织的脂肪谱,或脂质类别样品。所述FAME在己烷:DCM(4:1,v/v)中提取并在用GC和GC-MS分析前于DCM中进行重建。

[0967] 对于脂质类别分析,用Bligh和Dyer(1959)所述的方法,从大约50mg鲜重的经渗入的叶组织中提取两次总脂质。中性脂质通过TLC在预包被了硅胶的平板(Silica gel 60, Merck)上用己烷/乙醚/乙酸(体积70/30/1)纯化,而极性脂质用二维TLC分馏,氯仿/甲醇/水(体积65/25/4)用于第一个方向,而氯仿/甲醇/ NH_4OH /乙基正丙胺(体积130/70/10/1by vol.)用于第二个方向(Khozin等,1997)。用碘蒸汽使得脂质的斑点可见,收集入小管并转甲基以产生FAME用于如上所述的用于GC分析。TAG被作为所存在的脂肪酸的总量的百分比而进行定量,其如上所述通过GC分析并根据对每个脂肪酸所注射的已知量的外部标准进行评估。

[0968] 用Agilent Technologies 6890N GC(Palo Alto,California,USA)来进行GC,其转配了非极性EquityTM-1熔融石英毛细管柱(15m x 0.1mm i.d.,0.1 μm 薄膜厚度)、FID、分流/无分流进样器和Agilent Technologies 7683Series自动取样器以及注射器并使用氦作为载体气。以无分流模式在120℃的炉温注射样品,在注射后将炉温以10℃·min⁻¹升至201℃并最终升至270℃且维持20分钟。用Agilent Technologies ChemStation软件(Rev B.03.01(317),Palo Alto,California,USA)对峰进行定量。峰应答类似于可靠的Nu-Check GLC standard-411(Nu-Check Prep Inc,MN,USA)的脂肪酸,其含有等比例的31种不同的脂肪酸甲基酯,范围从碘苯腈辛酸酯到DHA以及几种其它的LC-PUFA。峰之间峰应答的轻微差异通过用峰面积乘以每个峰的正态化因子来进行平衡。每种脂肪酸在总脂肪酸中的比例基于脂肪酸的单个和总的峰面积进行计算。

[0969] 进行GC-MS以确认所有形成的新脂肪酸的身份,并且其在FinniganGCQ Plus GC-MS离子收集器(安装了柱上进样器)上于at 45℃进行。用AS2000自动取样器将样品注射至附着于非极性HP-5Ultra 2键合相柱(50m x 0.32mm i.d.x 0.17 μm 薄膜厚度)的保留隘口上。45℃的初始温度保持1分钟,随后通过温度程序以30℃·min⁻¹升高至140℃继而以3℃·min⁻¹升至310℃并将其保持12分钟。氦被用作载体气。质谱仪操作条件是:电子轰击能量70eV;发射电流250 μamp ,转运线310℃;源温度240℃;扫描速率0.8次扫描·s⁻¹以及质量范围40-650道尔顿。质谱由XcaliburTM软件获取并处理。

[0970] 用瞬时表达的脂肪酸延长酶修改本氏烟草脂肪酸谱

[0971] 为了评估所需的土壤杆菌的浓度以产生功能性转基因酶的最大生产,表达编码球等鞭金藻 $\Delta 9$ -延长酶的基因(Ig $\Delta 9\text{elo}$;Qi等,2002),其已知作用于CoA-连接的亚油酸(LA)和已知在本氏烟草叶中丰裕的ALA底物。在将此基因转移至花椰菜花叶病毒(CaMV) 35S启动子下游的二元载体后,将此构建体农杆菌渗入至本氏烟草叶,其中存在P19病毒阻抑蛋白以

阻抑宿主介导的转基因沉默,并评估了 $\Delta 9$ -延长作用的水平。所述延长作用的产物LA和ALA、EDA和EtrA分别被检测为接近最大基因活性(得自具有 $OD_{600}=0.2$ 的土壤杆菌培养物)(图16)。然而有趣的是,注意到以浓度十分稀的所述培养物(低至 $OD_{600}=0.05$)进行农杆菌渗入也导致了可轻易检测水平的酶活性。

[0972] 瞬时DGAT表达对三酰基甘油聚积的作用

[0973] 接着研究了是否可将本氏烟草叶中的TAG库的大小增加提供更大的接收端(sink)以在其中捕获所引入的在ER上有活性的脂肪酸生物合成酶的产物。含有拟南芥DGAT1(AtDGAT1)基因(其催化Kennedy途径的TAG生物合成中的组后一步)的构建体被作为可能的方法进行测试以增加叶中的TAG库,因为叶天然仅产生低水平的TAG。将该构建体通过如山所述的农杆菌渗入而引入到本氏烟草叶中。为了测试TAG的存在,将经渗入的叶的大约 1cm^2 大小的区段浸入小的含1%水成Nile Blue(BDH, Poole, UK)的皮氏培养皿中,真空渗入3分钟,简单用水清洗,在1%乙酸中温育3分钟,并装载到水中用于保存。在570-670nm用488nm的激发于Leica SP2激光扫描共聚焦显微镜上(Leica Microsystems, Sydney, Australia)收集荧光发射。相同叶的未转化区域被用作对照。用ImageJ软件评估了每个测定中TAG聚积的相对的量。

[0974] 拟南芥DGAT1(AtDGAT1)的瞬时表达导致了显著更多的脂质体的产生,所述脂质体由Nile Blue染色并用共聚焦显微镜来观测。TAG的增加是通过分馏总脂质至TAG内来量定的,其中使用中性相的TLC与其他叶脂质类别分离,之后TAG的量被测量为TAG级分中总脂肪酸的量。P19或P19与AtDGAT1一起的瞬时表达分别导致TAG从 $46\mu\text{g}\cdot\text{g}^{-1}$ 鲜重增加至 $206\mu\text{g}\cdot\text{g}^{-1}$ 鲜重,现实出DGAT1基因的加入提高了叶组织中聚积的TAG的水平。因此,除非另外指明,则此基因被包括在随后的实验中。

[0975] 外源性脂肪酸底物对瞬时表达的基因的可用性

[0976] 接着测试了对于本氏烟草非天然的外源性脂肪酸底物是否能提供给叶子并变得可用于转基因介导的转换,从而允许单个的酶步骤在分离中进行测试。为了测试此问题,将编码致病疫霉 $\Delta 17$ -去饱和酶(Pi $\Delta 17\text{des}$)(其作用于非天然存在于本氏烟草中的底物ARA)的基因进行农杆菌渗入以产生EPA。用Pi $\Delta 17\text{des}$ 渗入4天后,以类似于用土壤杆菌培养物转化叶子的方式通过注射而像叶提供ARA铵盐。继而在从叶组织中提取总脂质之前,允许该叶代谢底物4小时。这些总脂质的GC和GC-M分析显示出37%的外源提供的ARA由 $\Delta 17$ -去饱和作用转换成EPA,其效果与基于酵母测定中所报道的相当(WO 2005/012316)。

[0977] 从分离的二元载体快速组装五步LC-PUFA途径

[0978] 已经确立本氏烟草系统在确定单一基因的功能以及增强的TAG聚积方面是有用的工具,研究了该系统可用于组装the extent to which the system整个LC-PUFA途径的程度。在此研究中,测试了编码5种LC-PUFA代谢酶的基因,其产生了平行的两个线性LC-PUFA途径,即 $\omega 6$ -pathway,将LA转换为 $\text{DPA}^{\omega 6}$,以及 $\omega 3$ -途径,将ALA转换成DHA(图1)。所使用的生物合成基因是球等鞭金藻 $\Delta 9$ -延长酶(Ig $\Delta 9\text{elo}$)、盐生巴夫藻 $\Delta 8$ -去饱和酶(Ps $\Delta 8\text{des}$)、盐生巴夫藻 $\Delta 5$ -去饱和酶(Ps $\Delta 5\text{des}$)、盐生巴夫藻 $\Delta 5$ -延长酶(Ps $\Delta 5\text{elo}$)和盐生巴夫藻 $\Delta 4$ -去饱和酶(Ps $\Delta 4\text{des}$; Qi等,2004; Robert等,2009; Zhou等,2007)。将每个基因分别克隆至如上所述的CaMV 35S启动子下游的植物二元表达载体内,并将这些构建体的混合物(每个以 $OD_{600\text{nm}}=0.2$ 的浓度存在)连同AtDGAT1和P19,农杆菌渗入至本氏烟草叶的离轴表面

内,得到总共7个单个的构建体以及总 $OD_{600nm}=1.4$ 。

[0979] 渗入后5天对叶盘取样并直接从新鲜组织产生脂肪酸甲基酯 (FAME) 并通过GC/MS来分析和鉴定(表12)。很清楚的是所有的途径酶都能够接受 $\omega 6$ 或 $\omega 3$ PUFA的任一作为底物and th并且他们对LA或ALA的顺序作用分别导致LC-PUFA、ARA和DHA的合成。鉴定出新产生的LC-PUFA的百分比是16.9%,包括9.8%的 $\omega 6$ LC-PUFA和7.1%的 $\omega 6$ LC-PUFA。在所有这新形成的LC-PUFA中,ARA、EPA和DHA被认为在营养上很重要,并分别构成叶组织中总脂肪酸的3.6%、2.6%和1.1%。对 $\omega 6$ 和 $\omega 3$ 途径的每一步计算酶转换效率并与先前的报道进行比较(图17)。 $\omega 6$ 和 $\omega 3$ 的5步途径的前三步在效果上与Qi等的报道类似(Qi等,2004),而途径最后两步的效果与Robert等(Robert等,2005)所受用的相同。瞬时表达的基因与稳定表达的基因的此种对比表明,两种引入途径的方法均产生类似的流通和效果。这些在总脂肪酸谱上计算的转换效率很可能是一种低估,特别是对于第一步, $\Delta 9$ -延长作用,只是由于质体中巨大的LA库和ALA库的稀释作用。此议题通如下所述的脂质的分馏而完成。

[0980] 表12.产生 $\omega 6$ -和 $\omega 3$ -LC-PUFA二者的本氏烟草叶斑点的脂肪酸谱。每个渗入含有土壤杆菌培养物的混合物,所述土壤杆菌培养物载有分离自番茄丛矮病毒和拟南芥二酰基甘油0-酰基转移酶(AtDGAT1)的P19病毒阻抑物异位表达构建体。LC-PUFA途径的渗入包括额外的5个基因,即,球等鞭金藻 $\Delta 9$ 延长酶(Ig $\Delta 9elo$)、盐生巴夫藻 $\Delta 8$ -去饱和酶(Ps $\Delta 8des$)、盐生巴夫藻 $\Delta 5$ -去饱和酶(Ps $\Delta 5des$)、盐生巴夫藻 $\Delta 5$ -延长酶(Ps $\Delta 5elo$)和盐生巴夫藻 $\Delta 4$ -去饱和酶(Ps $\Delta 4des$)。为了清楚,表中不包括饱和的和较少见的脂肪酸,但其却用于计算百分比。(-)表示无可检测的量的脂肪酸。数据产生自3次重复并显示出标准差。

[0981] 总FAME(%)

	脂肪酸	导致	5 LC-PUFA 基因
[0982]	16:0	15.9±0.2	20.1±0.9
	16:1 ^{Δ3t}	1.7±0.1	1.5±0.2

	16:3 ^{Δ9,12,15}	6.3±0.3	5.2±0.3
	18:0	3.6±0.3	3.7±0.2
	18:1 ^{Δ9}	2.8±0.1	3.1±0.7
	18:2 ^{Δ9,12} (LA)	18.6±0.1	8.0±0.7
	18:3 ^{Δ9,12,15} (ALA)	45.5±1.4	38.0±1.9
	20:0	1.3±0.4	0.8±0
	其它较少见得	4.1	2.6
	Total	100	83.1
New ω6 PUFA			
	20:2 ^{Δ11,14} (EDA)	0	1.4±0.2
	20:3 ^{Δ8,11,14} (DGLA)	0	0.3±0
	20:4 ^{Δ5,8,11,14} (AA)	0	3.6±0.4
[0983]	22:4 ^{Δ7,10,13,16} (DTA)	0	1.5±0.1
	22:5 ^{Δ4,7,10,13,16} (DPA ^{ω6})	0	3.0±0.4
	总的	0	9.8
新的ω3 PUFA			
	20:3 ^{Δ11,14,17} (ETrA)	0	2.3±0.1
	20:4 ^{Δ8,11,14,17} (ETA)	0	0.2±0
	20:5 ^{Δ5,8,11,14,17} (EPA)	0	2.6±0.3
	22:5 ^{Δ7,10,13,16,19} (DPA ^{ω3})	0	0.9±0.1
	22:6 ^{Δ4,7,10,13,16,19} (DHA)	0	1.1±0.1
	总的	0	7.1
	总的新脂肪酸	0	16.9
	总脂肪酸	100	100

[0984] LC-PUFA的脂质类别区分

[0985] 为了评估新合成的LC-PUFA在TAG磷脂和质体半乳糖脂之间的区分将瞬时表达LC-PUFA途径基因的本氏烟草叶的总脂质进行如上所述的脂质类别划并确定了其脂肪酸谱(表13)。本氏烟草叶脂质含有高等植物叶中典型的脂质类别和脂肪酸谱(Fraser等,2004; Moreau等,1998)。新合成的ω6和ω3LC-PUFA均主要限于通常在质体外发现的脂质,而质体脂质基本上没有这些脂肪酸。例如,TAG和磷脂(PC、PE和PA)-主要的质体外叶脂质-分别包含了高至20.4%和16.9%的新合成的ω6和ω3LC-PUFA。值得注意的是,表达全LC-PUFA途

径的叶, AtDGAT1和P19产生TAG富集了37%的LC-PUFA. 具有特别兴趣的是营养学上重要的脂肪酸的聚积ARA、EPA和DHA, 在叶TAG中分别存在7.2%、5.9%和3%。区分显示主要的质体脂质类别, MGDG、DGDG和PG, 只分别包含了1.1%和0.3%的新合成的 ω 6和 ω 3LC-PUFA。尽管这些质体脂质类别代表了叶中脂肪酸最巨大的库, 总而言之, 这些类别与Tag相比只包含少量的 ω 6和 ω 3LC-PUFA。有趣的是, SQDG脂质类别完全不存在于新合成的LC-PUFA。

[0986] 脂质类别区分也被用于计算LC-PUFA途径每一步的酶效果, 所述LC-PUFA途径与用于TAG中的脂肪酸的ER关联, 其无法接触质体脂质(图17)。从这些计算中去除质体脂质类别对ALA转换成EtrA的 Δ 9-延长作用步骤有最显著的效果, 将转换效率从16%提高至55%。此步骤的酶转换效率3-倍的增长是由于质体中ALA巨大的库其对于此ER结合的酶而言不可用(表13)。

[0987]

表 13. 本氏烟草的脂肪酸谱表达 P10、*AtDGAT1* 和 5 种 LC-PUFA 基因的的脂质类别如表 12 中所述。TAG 的数据产生自用一维 TLC 进行的脂质分类划分，而其它脂质类别在二维 TLC 上进行。一式三份进行是按并显示标准。

脂肪酸	质体外					质体				
	TAG	PC	PE	PA	MGDG	DGDG	SQDG	PG		
16:0	22.6±0.6	24.3±0.8	23.0±0.4	22.5±2.8	4.7±0.5	20.9±0.7	52.7±2.1	31.9±1.9		
16:1 ^{Δ3t}	0.2±0	0	0	0.8±0.7	0	0	0.5±0.5	21.6±1.5		
16:3 ^{Δ9,12,15}	0.4±0	0.1±0.1	0	0	17.1±1.0	1.4±0.2	1.2±0.1	0		
18:0	7.8±0.7	8.9±1.5	8.6±1.6	8.7±2.2	0.8±0.1	3.8±0.3	7.2±1.1	5.2±0.7		
18:1 ^{Δ9}	3.2±0.9	6.7±2.3	1.8±0.6	5.6±2.2	0.9±0.1	1.3±0.1	2.3±0.3	11.8±0.2		
18:2 ^{Δ9,12} (LA)	9.2±0.2	16.6±1.2	21.2±1.5	18.1±0.6	2.7±0.1	3.9±0.1	5.3±0.1	12.9±0.5		
18:3 ^{Δ9,12,15} (ALA)	13.8±1.6	12.7±1.6	13.6±1.7	14.4±1.2	70.5±1.9	66.8±0.9	30.7±2.2	15.1±1.8		
20:0	2.2±0.1	0.6±0	1.0±0.1	0.4±0.3	0	0	0	0		
其它较少数的	3.3	2.8	4.2	1.1	1.3	1.0	0.1	0.9		
总共	62.7	72.7	73.4	71.6	98	99.1	100	99.4		
新的ω6 PUFA										
20:2 ^{Δ11,14} (EDA)	4.2±0.3	3.2±0.2	1.9±0.1	2.8±0.1	0.2±0	0.1±0.1	0	0		
20:3 ^{Δ8,11,14} (DGLA)	0.6±0	0.8±0	0.7±0.1	0.5±0.4	0	0	0	0		
20:4 ^{Δ5,8,11,14} (AA)	7.2±0.7	5.3±0.3	8.4±0.3	7.3±0.2	0.5±0.1	0	0	0.5±0		
22:4 ^{Δ7,10,13,16} (DTA)	1.7±0.5	5.8±0.1	3.7±0.3	5.6±0.2	0.2±0	0	0	0.1±0.2		

[0988]	22:5 ^{Δ4,7,10,13,16} (DPA ^{ω6})	6.7±0.4	1.9±0.1	2.5±0.3	1.8±0.4	0.2±0	0.2±0	0	0
	总新 ω6 LC-PUFA	20.4	17.0	17.2	18.0	1.1	0.3	0	0.6
	新ω3 PUFA								
	20:3 ^{Δ11,14,17} (ETra)	6.7±0.6	3.3±0.2	2.2±0.1	3.0±0.2	0.5±0.1	0.6±0	0	0
	20:4 ^{Δ8,11,14,17} (ETA)	0.5±0	0.1±0.2	0.1±0.2	0.1±0.2	0	0	0	0
	20:5 ^{Δ5,8,11,14,17} (EPA)	5.9±0.4	2.3±0.3	3.4±0.2	3.1±0.3	0.3±0	0	0	0
	22:5 ^{Δ7,10,13,16,19} (DPA ^{ω3})	0.8±0	3.2±0.3	2.1±0.1	3.2±0.6	0.1±0	0	0	0
	22:6 ^{Δ4,7,10,13,16,19} (DHA)	3.0±0.2	1.4±0.1	1.6±0.2	1.0±0.8	0	0	0	0
	总的新 ω3 LC-PUFA	16.9	10.3	9.4	10.4	0.9	0.6	0	0
	总的新的 总脂肪酸	37.3 100	27.3 100	26.6 100	28.4 100	2.0 100	0.9 100	0 100	0.6 100

[0989] 讨论

[0990] 这些实验显示出,本氏烟草或其它植物叶中的一系列途径基因的瞬时表达模拟了经稳定转化植物中的表达,因此很适宜于且可预计用于产生LC-PUFA油料的种子中途径的表达。所述瞬时表达系统提供了可互换的表达平台,其给出了快速且可靠的整个途径的结果,其中在多步骤的重组途径中单个的组分很容易被更换。LC-PUFA生物合成的瞬时组装是有力的且可重复的。一式三份进行测定并给出紧凑的数据点,通常标准擦低于5%。

[0991] 所组装的LC-PUFA途径中的第一步,Δ 9延长作用,显示出对LA比对ALA更高比率的延长作用,与在稳定转化的植物中先前所观测的一样 (Fraser等,2004)。这种不同不能用酶

的脂肪酸底物偏好来解释,酵母中的表达显示出Ig Δ 9延长酶具有对LA和ALA同样的偏好(Qi等,2002)。更可能的是对LA所观测到的较高比率的延长作用反映出在叶的质体外酰基-CoA库(延长酶的作用位点)中存在的LA的量高于ALA的量(Domergue等,2003;Fraser等,2004)。

[0992] 脂质分类分析证明了所有新形成的LC-PUFA在PC库和TAG中发现有大致相同的比率。此比率在途径的倒数第二部(产生DTA和DPA的 Δ 5-延长作用)有轻微的差异,其在TAG上丰度低于PC。相反,最终的 Δ 4去饱和作用的产物,DPA^{ω6}和DHA,相较于PC而更倾向于在TAG上聚集。这些差异可能会揭示出膜编辑的酶的微妙的偏好或AtDGAT1对这些产物的微妙的偏好。值得注意的是 Δ 5-延长作用和DGAT活性均在CoA库上出现,这些酶对底物的竞争可以改变脂肪酸在这些PC和TAG库中的存在。

[0993] 从这个研究可以得到几个提示。首先,基于叶的瞬时测定显示出适宜于测定脂肪酸酶(单一的或者组合的)。这对于产生容易地与本氏烟草的内源脂肪酸谱区别的脂肪酸的酶而言特别合适。例如本研究中的LC-PUFA。对分离的酶的脂肪供给测定以及LC-PUFA快速组装成油的证明表明,瞬时叶测定很适合用于ER-先关的去饱和作用、延长作用和TAG组装。第二,尽管LC-PUFA油是植物转化技术现时的目标,但是叶细胞提供了相对于其它异源表达平台的一定范围的优势。叶细胞提供了其它表达宿主不能提供的广泛的代谢产物,而这些现在可以成为需要修改的重组途径的靶标。此外,植物更忠实地处理真核的转基因,包括RNA编辑、翻译后修饰和细胞器定位。

[0994] 最后,所述本氏烟草测定可能适用于cDNA文库筛选测定。此建议是基于在叶上单一的渗入区中检测到7个不同基因的接近最大的活性,这表明按照此种设置克隆至二元表达载体内的cDNA文库可以全身性的渗透到叶内。计算表明,至少7个不同的克隆,包括OD_{600nm} 0.2的P19,可在单一的斑点中表达。或者,可将形成不完整的或部分途径的基因加入到每个渗入并由此使得汇集的库的克隆能够常规地测试新的步骤或改善的流通。

[0995] 实施例10:在植物细胞中高效的DHA生物合成

[0996] 使用如实施例1中所述的增强的本氏烟草瞬时表达系统在植物体中证明了微胞藻CCMP1545 Δ 6-去饱和酶(SEQ ID NO:8由SEQ ID NO:7编码)、塔胞藻CS-0140 Δ 6-延长酶(SEQ ID NO:4由SEQ ID NO:3编码)、盐生巴夫藻 Δ 5-去饱和酶(SEQ ID NO:26由SEQ ID NO:25编码)、塔胞藻CS-0140 Δ 5-延长酶(SEQ ID NO:6由SEQ ID NO:5编码)和盐生巴夫藻 Δ 4-去饱和酶(SEQ ID NO:73由SEQ ID NO:72编码)连同拟南芥DGAT1(SEQ ID NO:74由SEQ ID NO:75编码)的酶活性。

[0997] 通过将包含在SwaI片段内的pGA4A的整个编码区插入到35S-pORE04(上述实施例4)内的SmaI-EcoRV位点,制备了在组成性35S启动子控制下的编码 Δ 6-去饱和酶的基因构建体35S:Mic1545-d6D,产生pJP2064。通过将包含在SwaI片段内的0804673_Pyrco-elol_pGA18的整个编码区插入到35S-pORE04内的SmaI-EcoRV位点,制备了编码 Δ 6-延长酶的基因构建体35S:Pyrco-d6E,产生pJP2060。通过将包含在SwaI片段内的0804674_Pavsa-d5D_pGA15的整个编码区插入到35S-pORE04内的SmaI-EcoRV位点,制备了编码 Δ 5-去饱和酶的基因构建体35S:Pavsa-d5D,产生pJP2067。通过将包含在SwaI片段内的0804675_Pyrco-elol2_pGA4的整个编码区插入到35S-pORE04内的SmaI-EcoRV位点,制备了编码 Δ 5-延长酶的基因构建体35S:Pyrco-d5E,产生pJP2061。通过将包含在SwaI片段内的0804676_Pavsa-

d4D_pGA15的整个编码区插入到35S-pORE04内的SmaI-EcoRV位点,制备了编码 $\Delta 4$ -去饱和酶的基因构建体35S:Pavsa-d4D,产生pJP2068。通过将包含在BamHI-EcoRV片段内的pXZP513E的整个编码区插入到35S-pORE04内的BamHI-EcoRV位点,制备了编码酶DGAT1的基因构建体35S:Arath-DGAT1,产生pJP2078。

[0998] 将这些嵌合载体个别地引入到根癌土壤杆菌菌株AGL1内,以及将来自这些的培养物的转基因细胞混合并将混合物在温室内渗入到本氏烟草植物的叶组织中。在渗入后将植物再培养5天,随后取叶盘用于GC分析,其表明这些基因在发挥功能以在本氏烟草中产生DHA。用这些基因转化的叶组织含有SDA(2.3%)、ETA(0.7%)、EPA(0.8%)、DPA(2.8%)和DHA(4.4%)(表14)。叶组织还含有痕量水平的GLA、ETA和ARA。转换效率如下:细胞中所产生的ALA的17.4%被转化成EPA(包括EPA随后被转换成DPA或DHA);15.5%的ALA被转换成DPA或DHA;细胞中所产生的ALA的9.6%被转换成DHA;而细胞中所产生的 $\Delta 6$ -不饱和的ALA的40%随后被转换为DHA。由于本实验中转基因的瞬时表达,在稳定转化的细胞中会预期有比上述更高的功效。

[0999] 当由TLC分馏从叶组织提取的总脂质以分开脂质类别时,以及由FAME来分析TAG和极性脂质级分的脂肪酸组成时,观测到在TAG中的DHA水平作为总脂肪酸百分比是7%,并且在极性脂质中DHA的水平是2.8%。极性脂质类别中较低的水平被认为是由于叶中叶绿体脂质的相关贡献,倾向于极性脂质,以及基因的瞬时表达,而不是转基因稳定的插入宿主细胞基因组。

[1000] 表14. 来自经去饱和酶和延长酶的组合转化的叶的脂质的脂肪酸组成。

[1001]

脂肪酸	脂肪酸	未转化的细胞	经转化的细胞
16:0	棕榈酸	17.1	20.4
16:1d7		0.1	0.4
16:1d9		0.2	0.2
C6:1d?		0.5	0.4
16:1d?		0.5	0.4
17:1d9		0.9	0.7
16:2		0.9	0.9
16:3		6.6	5.4
18:0	硬脂酸	2.1	3.6
18:1d7		0.0	0.0
18:1d9	油酸	0.8	2.6
18:1d11		0.0	0.0
18:1d12		0.3	0.6
18:1d13		0.2	0.2
18:2n6	LA	5.1	10.8
18:3n6	GLA	0.7	1.9
18:3n3	ALA	57.8	35.0

[1002]

20:0		0.4	0.7
20:1d5		0.4	0.3
18:4n3	SDA	0.4	2.3
20:1d8		0.0	0.0
20:1d11		0.0	0.0
20:2n6	EDA	0.1	0.2
20:3n6	DGLA	0.3	0.4
20:4n6	ARA	0.4	0.5
20:3n3	ETrA	0.2	0.2
22:0		0.2	0.2
20:4n3	ETA	0.0	0.7
22:1d9		0.0	0.9
20:5n3	EPA	0.8	0.8
22:2n6		0.1	0.2
22:4n6		0.1	0.1
22:3n3		1.1	1.2
24:0		0.4	0.4
22:5n6		0.4	0.2
24:1d9		0.8	0.1
22:5n3	DPA	0.2	2.8
22:6n3	DHA	0.1	4.4

[1003] 此实验证明了,分离的微胞藻CCMP1545 Δ 6-去饱和酶对 ω 3底物ALA与 ω 6底物LA相比具有实质的偏好。该实验也证明了适宜基因的表达可以导致实质百分比的LC-PUFA在叶中的聚积,最值得注意的是EPA、DPA和DHA。

[1004] 次实验也显示了本氏烟草瞬时测定系统用于快速测试各种脂肪酸生物合成途径以及用于选择最佳的基因组合的用途。此系统可用于快速地对具有同源功能的基因的相对活性进行比较,以及比较整个生物合成途径。

[1005] 讨论:在叶和种子组织中高效的DHA合成

[1006] 在此数据的基础上预计到,当种子特异性启动子被用于表达此基因组合或者类似的组时,在种子中汇产生相同水平EPA、DPA和DHA,或甚至更高的水平。所观测到的从ALA到EPA直到DPA和DHA的脂肪酸的高效转变被认为是由于高效的延长酶与酰基-CoA去饱和酶的组合,由此主要在酰基-CoA库中对脂肪酸起作用。此外,预计可以通过使用具有适当的组织特异性的启动子或者启动子的组合,来实现在转基因植物的叶和种子中或者在种子和叶之外的其它组织中产生EPA、DPA、DHA和其它LC-PUFA。融合的启动子将能够推动在该两种组织类型中酶的产生。所得的植物将或许可用于油料的提取(特别是从种子中),以及具有最少加工的原料。

[1007] 实施例11:植物细胞中更有效的DHA生物合成

[1008] 利用如实施例1和实施例10所述的增强型本氏烟草瞬时表达系统,除了拟南芥DGAT1 (SEQ ID NO:74,由SEQ ID NO:75编码)以外,植物内(in planta)证实了微胞藻CCMP1545 Δ 6-去饱和酶 (SEQ ID NO:8,由SEQ ID NO:7编码)、塔胞藻CS-0140 Δ 6-延长酶 (SEQ ID NO:4,由SEQ ID NO:3编码)、盐生巴夫藻 Δ 5-去饱和酶 (SEQ ID NO:26,由SEQ ID NO:25编码)、塔胞藻CS-0140 Δ 5-延长酶 (SEQ ID NO:6,由SEQ ID NO:5编码)和盐生巴夫藻 Δ 4-去饱和酶 (SEQ ID NO:73,由SEQ ID NO:72编码)的酶活性。该实验利用更年幼和健康的本氏烟草植物进行了优化。

[1009] 将实施例10所述的这些嵌合载体分别引入根癌土壤杆菌菌株AGL1,并且在温室中

将来自这些培养物的转基因细胞混合,并将混合物渗入本氏烟草植物的叶组织。在渗入后使植物再生长5天,然后取叶盘进行GC分析,这表明这些基因在本氏烟草中发挥功能以产生DHA(表15和16)。用这些基因转化的叶组织含有SDA(2.0%)、ETA(0.4%)、EPA(0.7%)、DPA(4.3%)和DHA(4.4%)。叶组织还有痕量水平的GLA、ETA和ARA。转换效率为:23.4%的细胞中产生的ALA转换为EPA(包括随后转换为DPA或DHA的EPA);21.6%的ALA转换为DPA或DHA;10.9%的细胞中产生的ALA转换为DHA;而37.2%的细胞中产生的D6-去饱和的ALA随后转换为DHA。由于该实验中转基因的瞬时表达,在稳定转化的细胞中预期比以上效率更高的效率。

[1010] 当通过TLC分级从叶组织提取的总脂质以分离类型,并且通过FAME分析TAG和极性脂质组分的脂肪酸组成时,观察到TAG中的DHA水平是总脂肪酸的15.9%,并且在极性脂质中DHA水平为4.4%。极性脂质类型中较低的水平据认为是由于叶子中叶绿体的相对贡献、有利于极性脂质以及基因的瞬时表达而不是转基因稳定插入宿主细胞基因组。

[1011] 表15.来自用去饱和酶和延长酶的组合转化的叶子的脂质的脂肪酸组成

[1012]

	对照	Pavsa-D5E DHA 途径		Pyrco-D5E DHA 途径	
脂肪酸	总脂质	总脂质	TAG	总脂质	TAG
一般 FA					
16:0	15.9±0.2	17.0±0.1	20.2	16.6±0.1	21.6
16:1 ^{Δ3t}	1.7±0.1	1.5±0.2	0.3	1.5±0.1	0.3
16:3 ^{Δ9,12,15}	6.3±0.3	5.2±0.1	0.4	5.6±0.1	0.6
18:0	3.6±0.3	3.5±0.1	5.9	3.3±0.1	6.5
18:1 ^{Δ9}	2.8±0.1	3.4±0.1	5.1	2.8±0.2	5.6
18:2 ^{Δ9,12}	18.7±0.1	14.1±0.4	15.2	13.0±0.1	17.3
18:3 ^{Δ9,12,15}	45.6±1.4	39.1±0.4	6.9	40.2±0.5	6.7
20:0	1.3±0.4	0.7±0	1.8	0.6±0	2.0
其他少量	4.1	2.5	6.3	2.3	6.2
总计	100	87.0	62.1	85.9	66.8
新 w6 PUFA					
18:3 ^{Δ6,9,12}	-	1.6±0.1	3.3	2.1±0.2	4.3
20:3 ^{Δ8,11,14}	-	-	-	-	-
20:4 ^{Δ5,8,11,14}	-	0.3±0.1	0.5	0.2±0	-
总计	0	1.9	3.8	2.3	4.3
New ω3 PUFA					
18:4 ^{Δ6,9,12,15}	-	1.5±0.1 (22% D6D)	3.8	2.0±0 (23% D6D)	6.4
20:4 ^{Δ8,11,14,17}	-	0.5±0 (86% D6E)	1.5	0.4±0 (83% D6E)	1.6
20:5 ^{Δ5,8,11,14,17}	-	4.1±0.2 (95% D5D)	14.2	0.7±0 (96% D5D)	1.5
22:5 ^{Δ7,10,13,16,19}	-	2.4±0.1 (55% D5E)	1.6	4.3±0 (93% D5E)	3.5
22:6 ^{Δ4,7,10,13,16,19}	-	2.6±0.1 (52% D4D)	13.0	4.4±0.1 (51% D4D)	15.9
总计	0	11.1	34.1	11.8	28.9
总新 FA	0	13.0	37.9	14.0	33.2
总 FA	100	100	100	100	100

[1013] 表16. 来自用去饱和酶和延长酶的组合转化的叶子的转换效率

[1014]

脂肪酸	总计		酶	转换效率			
LA	13						
GLA	2.1	15.0%	d6D				
DGLA	0	8.7%	d6E				
AA	0.2	100.0%	d5D				
ALA	40.2						
SDA	2	22.7%	d6D	18.1%	d6D 转换为 EPA+DPA+DHA		
ETA	0.4	83.1%	d6E	16.7%	d6D 转换为 DPA+DHA		
EPA	0.7	95.9%	d5D	8.5%	ALA 转换为 DHA		
DPA	4.3	92.6%	d5E	46.8%	EPA 转换为 DHA		
DHA	4.4	50.6%	d4D				
	TAG						
LA	17.3						
GLA	4.3	19.9%	d6D				
DGLA	0	0.0%	d6E				
AA	0	0.0%	d5D				
ALA	6.7						
SDA	6.4	81.2%	d6D	58.7%	d6D 转换为 EPA+DPA+DHA		
ETA	1.6	77.9%	d6E	54.5%	d6D 转换为 DPA+DHA		
EPA	1.5	92.9%	d5D	44.7%	ALA 转换为 DHA		
DPA	3.5	92.8%	d5E	76.1%	EPA 转换为 DHA		
DHA	15.9	82.0%	d4D				

[1015] 讨论:叶子和种子组织中更有效的DHA合成

[1016] 由于叶子和种子组织之间质体外脂质合成机制的重要转换,该结果可能为种子TAG的类似进展做好准备(Ohlrogge and Browse,2004;Bates等,2007)。我们推测几种因素可能是产量如此大量增加的原因:1. ω 3-特异性酰基-CoA Δ 6-去饱和酶的使用增加了 ω 3途径的流向(flux down)并减少了与平行 ω 6底物对随后的代谢步骤的竞争;2. 高效 Δ 5-延长酶明显增加了可用于 Δ 4-去饱和作用而转换为DHA的DPA的量;3. 通过使用独立转录单元和使用病毒阻抑蛋白(P19)减少了基因沉默。

[1017] 当试图改造积累 ω 3LC-PUFA EPA和DHA的陆生植物时,所述 Δ 6-去饱和酶所表现出的强 ω 3优选是明显期望的,因为并不要求额外的 Δ 17-去饱和酶活性来将AA(20:4^{D5},8,11,14)转换为EPA,因而简化了代谢工程和调控挑战。

[1018] 除了上述优化步骤以外,使用高效P.cordata Δ 5-延长酶导致了非常类似于金枪鱼油的TAG(油)组分的脂肪酸谱(profile),所述金枪鱼油是一种高DHA和低中间体含量的鱼油(图18)。此外,本氏烟草中P.cordata Δ 5-延长酶所表现出的活性是迄今为止我们所知的最有效的 Δ 5-延长作用,并且该基因的使用有效地克服了在转基因DHA生产的其他尝试中所经历的大 Δ 5-延长作用瓶颈。

[1019] 最后,尽管明显要求使用最优基因,但是我们认为这些转基因引入的方法(即,作

为独立表达盒和在基因沉默阻抑物的存在下) 在本研究中所实现的高水平DHA中起到关键作用。因此, LC-PUFA生产的代谢工程迄今为止依赖于随机插入宿主基因组的大多基因构建体, 尽管许多组通过该方法获得了良好的结果, 但是还有迹象表明难以获得表现出所有转基因相等表达的事件 (WO 2004/017467)。此外, 沉默效果可能降低产生的效率 (Matzke等, 2001)。诸如涉及在独立组装单元和改造的小染色体上从头着丝粒形成的人工染色体的替代性转化方法最后可能是成功的稳定LC-PUFA代谢工程所要求的 (Yu等, 2007)。

[1020] 实施例12: 利用单个生物的基因的完全ALA至DHA途径的转基因组装

[1021] 利用编码来自盐生巴夫藻的酶的基因重建完全ALA至DHA途径并在本氏烟草组装, 该途径由所述 $\Delta 9$ - 延长酶、 $\Delta 8$ - 去饱和酶、 $\Delta 5$ - 去饱和酶、 $\Delta 5$ - 延长酶和 $\Delta 4$ - 去饱和酶组成。土壤杆菌渗入后5天的总叶组织的GC分析证实产生了0.7%DHA (表17)。这是首次报导的由单个生物的基因组成的ALA至DHA转基因途径。

[1022] 表17. 瞬时转化了单基因CaMV 35S二元构建体的本氏烟草叶组织的脂肪酸组成 (总脂肪酸的百分比)。DHA途径由盐生巴夫藻 $\Delta 9$ - 延长酶、 $\Delta 8$ - 去饱和酶、 $\Delta 5$ - 去饱和酶、 $\Delta 5$ - 延长酶和 $\Delta 4$ - 去饱和酶组成。示出了一式三份进行的单独渗入之间的标准偏差。

	脂肪酸	对照	Pavpi- $\Delta 9$ E DHA
	一般 FA		
[1023]	16:0	15.7±0.6	15.9±0.2
	16:1 ^{$\Delta 3t$}	1.5±0	1.4±0.1
	16:3 ^{$\Delta 9,12,15$}	6.8±0.7	5.9±0.3

	18:0	3.0±0.1	3.7±0.2
	18:1 ^{Δ9}	2.2±0	2.7±0.3
	18:2 ^{Δ9,12}	11.8±0.4	8.6±0.6
	18:3 ^{Δ9,12,15}	56.0±1.4	49.0±1.4
	其他少量	3.0±0	2.8±0
	总计	100	90.0
新 ω6 PUFA			
	20:2 ^{Δ8,11}	—	1.7±0.2
	20:3 ^{Δ8,11,14}	—	0.5±0
	20:4 ^{Δ5,8,11,14}	—	2.4±0.1
	22:4 ^{Δ7,10,13,16}	—	1.2±0
[1024]	22:5 ^{Δ4,7,10,13,16}	—	—
	总计	0	5.8
新 ω3 PUFA			
	20:3 ^{Δ11,14,17}	—	1.5±0.2
	20:4 ^{Δ8,11,14,17}	—	0.2±0
	20:5 ^{Δ5,8,11,14,17}	—	1.2±0.1
	22:5 ^{Δ7,10,13,16,19}	—	0.6±0
	22:6 ^{Δ4,7,10,13,16,19}	—	0.7±0.1
	总计	0	4.2
	总新 FA	0	10.0
	总 FA	100	100

[1025] 实施例13:植物的发育中的种子中VSP的特异性表达

[1026] 将35S启动子控制之下的5个病毒阻抑蛋白(VSP)即P19、V2、P38、PePo和RPV-P0的蛋白编码区初始插入二元载体pART27(Gleave,1992)以用于植物组织中的组成型表达。这些蛋白如下表征为VSP。P19是来自番茄丛矮病毒(TBSV)的阻抑蛋白,其结合21个核苷酸的长siRNA然后指导Argonaute指导的同源RNA切割(Voinnet等,2003)。V2是来自番茄黄化曲叶病毒(TYLRV)的阻抑蛋白,其结合植物蛋白SGS3(Glick等,2008),这是一种据认为是从ssRNA底物产生双链RNA中间所需的蛋白(Beclin等,2002)。P38是来自芜菁皱病毒(TCV)的阻抑蛋白,其抑制对产生siRNA关键的依赖于RNA的聚合酶活性(RdRP)并结合切酶蛋白DCL4(Ding and Voinnet,2007)。诸如PePo和RPV-P0的P0蛋白来自马铃薯卷叶病毒,其靶向Argonaut以增强降解(Baumberger等,2007;Bortolamiol等,2007)。为了确立这些蛋白作为

沉默阻抑物增加转基因表达的功能,将土壤杆菌中5个35S驱动的VSP构建体与35S驱动的GFP构建体共渗入本氏烟草叶子中。在所有的情况下,VSP的存在增加并延长了GFP的表达,这赋予了GFP基因活性的增加水平,特别是在用土壤杆菌柱接种4天后,从而证实在这种测定模式中所述蛋白作为沉默阻抑物的功能。

[1027] 将5个VSP编码区各自插入基于pART27骨架载体的第二二元载体pXRZ393,从而VSP的表达在种子特异性FP1启动子的控制之下(Ellerstrom等,1996),这提供了VSP在双子叶植物的发育中的种子的子叶中的表达。根据实施例1所述的方法,将构建体用于产生转化的拟南芥植物。对于每个编码VSP的嵌合基因获得至少20个转化的植物。如通过在选择培养基和土壤中生长、正常生长为产生成活的种子的可繁殖成体植物所证实的,所述植物是成活的并且大部分表型上是正常的。唯一在幼苗中看起来被改变的形态学表型是在P19、PePo和RPV-P0萌发后从一些种子出现的子叶。与野生型植株不同,用种子特异性编码P19的构建体转化的植株具有扁平的棒状子叶并且没有向下卷曲。用编码PePo的构建体转化的植株产生了“芭蕾舞短裙”表型,其中子叶以向内或凹面卷曲向上。这些植物上的真叶正常发育。表达RPV-P0的植株是茂密的,并且在整个营养生长中趋向于保留茂密的生长习性。V2和P38的植株没有明显的显著表型。

[1028] 除了子叶发育以外,用P19和PoPe转化的植株正常生长,并且在随后的生长和发育中与对照植物没有区别。在该实验中,表达VSP的启动子FP1的特征为拟南芥种子发育期间发育的子叶中有限但强的表达(Ellerstrom等,1996)。在新生子叶表型的基础上,我们认为发育中的种子中P19、PePo或RPV-P0的FP1-驱动的表达可能与正常子叶发育所需的小RNA生物发生重叠。子叶发育中这些VSP相关的变化并不影响转基因植物的总体发育,并且植物正常的随后生长和发育表明除了发育中的种子以外的组织中来自FP1启动子的任何VSP渗漏表达是微量且不明显的。这与植物组织中许多VSP的组成型表达是非常有害的之前研究形成对比(Mallory等,2002;Chapman等,2004;Chen等,2004;Dunoyer等,2004;Zhang等,2006;Lewsey等,2007;Meng等,2008)。

[1029] 缺少VSP V2和P38的表型表明这些VSP以不影响发育的小RNA生物发生或识别的方式靶向小RNA代谢。利用本氏烟草中的GFP测定检查了每种VSP的功能。

[1030] 实施例14:发现和评价用VSP构建体转化的T1种子的种子特异性可视标记物的开发

[1031] 实施例13的数据表明种子以及随后的子代植物可以耐受来自种子特异性启动子的VSP表达而没有明显的有害效果。本发明人还考虑了如上文所述产生的转基因植物是否仅表达低水平的VSP,从而允许种子存活,有效地针对潜在地表达致死量的VSP并因此在所用条件下无法恢复的转化种子进行选择。

[1032] 为了更准确地评价T1种子中转基因的表达,开发了用于转基因种子中的转化和表达的可视标记物。对于其他双子叶植物种子,拟南芥种子含有被母本种皮所围绕的父本胚和胚乳。在拟南芥的土壤杆菌介导的转化中,父本组织被T-DNA转化,而母本组织保持未转化。直到产生下一代的T2种子才获得转化的母本组织。为了提供有用的筛选系统,将编码绿色荧光蛋白(GFP)的基因进行修饰以促进细胞外蛋白的强分泌。这样的GFP定位据证实允许转化的T1种子的可视恢复(Nishizawa等,2003),这是通过检测转化的T1胚和胚乳中产生的GFP通过薄但是未转化的种皮的分泌而实现的(Fuji等,2007)。

[1033] 编码分泌GFP的嵌合基因如下构建。所述基因含有两个内含子,一个在5' 非翻译区(5' UTR),第二个在GFP的蛋白编码区。包含这些内含子以增强嵌合基因的表达(Chung等,2006),并保证种子中检测到的任何GFP信号只能来自植物细胞中基因的表达而不是土壤杆菌细胞的渗漏表达。根据Nishizawa等(2003)的报导,将编码定位于胞质的GFP蛋白的内含子中断形式的人源化GFP(Clontech)(Brosnan等,2007)修饰以促进通过内质网(ER)向质外体的分泌。用于质外体分泌的修饰的GFP构建体包括在GFP的N端作为框内(in-frame)翻译融合而加入的伴大豆球蛋白(conglycinin)分泌肽和在C端加入的4个甘氨酸以促进从ER的分泌。对嵌合GFP基因的这些加入通过5' 区和用于C端甘氨酸增加的PCR介导的序列修饰的基因合成来进行。将该分泌GFP的编码区插入FP1启动子的下游,其在5' UTR中具有来自过氧化氢酶的内含子。将完全FP1-分泌GFP序列插入pORE系列二元载体中的nos3' 加A信号的上游。将含有编码分泌GFP的嵌合基因的构建体命名为pCW141。

[1034] 为了证实GFP蛋白的表达和分泌,将编码分泌GFP序列但不具有FP1启动子或pCW141的5' UTR的基因区亚克隆至pCaMV35S-OCS3' 表达盒以产生pCW228,并通过土壤杆菌介导的转化引入本氏烟草叶子。基因的表达和GFP蛋白的分泌利用共焦显微术(Leica)来证实。

[1035] 在构建体是正确的基础上,通过实施例1所述的土壤杆菌介导的方法将pCW141引入拟南芥Col-0生态型的植物。将拟南芥植物的种子用pCW141,FP1-驱动分泌GFP构建体浸泡,收集并利用配备了荧光检测的解剖镜(Leica MFZIII)筛选GFP阳性种子。发出绿色荧光的种子是容易鉴定的,即使在群体中大部分T1种子是未转化的。筛选这些GFP阳性种子,并在选择培养基上生长以证实T-DNA上连接至嵌合基因构建体的卡那霉素抗性选择标记物基因的存在。汇集其他20个阳性种子,并通过利用抗GFP抗体的蛋白印迹来证实GFP蛋白的表达。

[1036] 利用分泌的GFP标记物选择T1种子中VSP表达

[1037] 将各自在发育的子叶特异性表达的FP1启动子控制之下的编码VSP:P19、P38、V2和P38的4个嵌合基因的每一个分别插入上文所述的GFP选择载体pCW141以产生pCW161、pCW162、pCW163和pCW164。这些二元载体各自具有用于VSP表达和分泌GFP的连接嵌合基因,从而允许通过GFP表型来鉴定、选择和分析用构建体转化的转基因种子而无需在选择培养基上生长。预期在大部分转化子中,编码VSP的基因会被整合从而与编码GFP的基因连接。

[1038] 收集获得自拟南芥植物的种子,并如上文所述筛选GFP荧光,所述拟南芥植物根据实施例1中的方法用含有VSP-GFP构建体组合的土壤杆菌接种。通过手工收集发出绿色荧光的种子,并在选择培养基上生长以确定它们是否也转化了选择标记物基因。在所有的情况下,GFP阳性种子在选择培养基上生长,并表现出与上文所述用VSP基因转化但没有GFP基因的植物相同的子叶表型。所观察到的GFP样子种子无论如何在选择培养基上生长。这样的种子可以是预期的,如果某些转化事件导致发育中的种子中的细胞,其中VSP表达水平过高并引起致死。转化的群体中不存在此类种子表明,当从FP1启动子表达时,种子中VSP表达是耐受的。不存在此类有害效果与当许多VSP组成型表达时的有害效果的报导形成对比(Mallory等,2002;Chapman等,2004;Chen等,2004;Dunoyer等,2004;Zhang等,2006;Lewsey等,2007;Meng等,2008)。

[1039] 因此,使用诸如GFP的可视检测的标记物证实了鉴定、选择和分析连接的基因的有

力且有效的方式,所述连接的基因例如编码VSP、去饱和酶、延长酶或其他脂肪酸修饰酶的基因。

[1040] 胚后表达VSP的种子中GFP表达的定量

[1041] 利用荧光显微术和数字成像分析,在从FP1启动子表达VSP的T1和T2种子中定量GFP表达。这些分析清楚地表明,GFP表达不受发育的子叶-特异性启动子控制之下的编码VSP的基因的共引入的影响。分析了GFP-VSP构建体在后续代和独立转化事件中的性能的扩展研究。预计VSP的存在会导致种子的连续代中更稳定且更高的平均表达水平。

[1042] 实施例15:种子中VSP与LC-PUFA合成基因的共表达

[1043] 为了确立VSP能够保护或增强种子中的转基因性能,将一些表达载体设计为考虑对宿主介导的抑制(沉默)比对仅具有单基因的载体易感,以增加VSP的相对效率。利用各个基因的相同配置,构建了各自含有种子中DHA合成的5个去饱和酶或延长酶基因的一系列载体。认为使得这些构建体对沉默(降低的表达)更易感的一个因素是使用相同的启动子(FP1)来驱动每个基因。使用FP1启动子,因为其相对较小,并且减小了总载体大小和每个编码区之间的间隔。此外,每个基因盒具有相同的朝向,我们认为这会增强沉默的可能性。LC-PUFA途径的三个基因具有编码区,其对于最优植物表达是密码子优化(A-B-C)的;而两个(E-D)是获得自微藻类的天然序列。相同组的5个基因之前在叶子中表达以产生组装完整的LC-PUFA生物合成途径(实施例11)。在该系列的第一载体中包含编码VSP P19的另一基因,在第二载体中包含编码V2的基因,而该系列的第三载体则不具有VSP。

[1044] 如下构建了这些载体pJP3057(图19)、pJP3059(图20)和(图21)并平行转化。首先在克隆载体中包含的FP1启动子和nos终止子之间克隆去饱和酶或延长酶基因或病毒阻抑物基因。然后顺序并以相同的朝向将启动子-基因-终止子盒克隆至二元载体骨架。pJP3057含有整个DHA途径,而pJP3059和pJP3060的差异分别是加入了FP1-P19-NOS或FP1-V2-NOS盒。构建步骤如下。首先,将含有细小微胞藻 $\Delta 6$ -去饱和酶的AscI-PacI片段克隆至pJP2015的AscI-PacI位点,然后将来自该载体的含有整个启动子-基因-终止子盒的SwaI片段克隆至pORE02的EcoRV位点以产生pJP3050。随后,将含有P.cordata $\Delta 6$ -延长酶的AscI-PacI片段克隆至pJP2015TMV(pJP2015略微修饰的形式,其中TMV前导序列存在于启动子的下游和基因的上游)的AscI-PacI位点,然后将来自该载体的含有整个启动子-基因-终止子盒的SwaI片段克隆至pJP3050的T4DNA聚合酶处理的SacI位点以产生pJP3051。随后,将含有盐生巴夫藻 $\Delta 5$ -去饱和酶的AscI-PacI片段克隆至pJP2015TMV的AscI-PacI位点,然后将来自该载体的含有整个启动子-基因-终止子盒的SwaI片段克隆至pJP3051的SmaI位点以产生pJP3052。随后,将含有盐生巴夫藻 $\Delta 5$ -去饱和酶的AscI-PacI片段克隆至pJP2015TMV的AscI-PacI位点,然后将来自该载体的含有整个启动子-基因-终止子盒的SwaI片段克隆至pORE02的SmaI位点以产生pJP3054。随后,将含有P.cordata $\Delta 5$ -延长酶的AscI-PacI片段克隆至pJP2015TMV的AscI-PacI位点,然后将来自该载体的含有整个启动子-基因-终止子盒的SwaI片段克隆至pJP3054的StuI位点以产生pJP3055。随后,将含有盐生巴夫藻 $\Delta 4$ -去饱和酶的AscI-PacI片段克隆至pJP2015TMV的AscI-PacI位点,然后将来自该载体的含有整个启动子-基因-终止子盒的SwaI片段克隆至pJP3056的SfoI位点以产生pJP3056。随后将pJP3056的PmeI-NotI片段克隆至pJP3051的PmeI-NotI位点以产生pJP3057,其是含有从ALA产生DHA的5个基因的二元载体。

[1045] 随后将含有编码P19病毒阻抑物的嵌合基因的AscI-PacI片段克隆至pJP2015TMV的AscI-PacI位点,然后将来自该载体的含有整个启动子-基因-终止子盒的SwaI片段克隆至pJP3057的ZraI位点以产生pJP3059。类似地,将含有编码V2病毒阻抑物的嵌合基因的AscI-PacI片段克隆至pJP2015TMV的AscI-PacI位点,然后将来自该载体的含有整个启动子-基因-终止子盒的SwaI片段克隆至pJP3057的ZraI位点以产生pJP3060。

[1046] 将所有三个构建体在拟南芥中转化(生态型Columbia)。用每个构建体转化拟南芥植物(Col-0生态型)并用pJP3057转化油菜。收集T1种子,在含有除草剂的培养基上分析,并分析所得T2种子的一般形态学变化和LC-PUFA合成。

[1047] 使用三个构建体pJP3057、pJP3059和pJP3060产生的转化的植物(拟南芥,生态型Columbia)自受精并收集T1种子。将这些种子播种在含有卡那霉素的培养基上以确定T1植物的杂合性/纯合性,并且分析从每个T1植物所得的T2种子的一般形态学变化和LC-PUFA合成(表18)。

[1048] 用pJP3057转化的植物的代表性T2种子在种子油中的总脂肪酸中含有SDA(0.4%)、ETA(0.6%)、EPA(0.2%)、DPA(0.3%)和DHA(2.4%)。T2植物的种子油还含有GLA(1.4%)以及痕量的EtrA和ARA。种子中的转换效率如下:细胞中产生的18.4%的ALA为 $\Delta 6$ -去饱和的;细胞中产生的89.7%的SDA为 $\Delta 6$ -延长的;细胞中82.9%的ETA为 $\Delta 5$ -去饱和的;细胞中93.1%的EPA为 $\Delta 5$ -延长的;细胞中88.9%的DPA为 $\Delta 4$ -去饱和以产生DHA(表18)。

[1049] 用pJP3059转化的植物的代表性T2种子含有SDA(0.7%)、ETA(0.3%)、EPA(0.2%)、DPA(0.9%)和DHA(1.3%)。种子油还含有GLA(0.8%)以及痕量的EtrA和ARA。转换效率如下:细胞中产生的15.7%的ALA为 $\Delta 6$ -去饱和的;细胞中产生的79.4%的SDA为 $\Delta 6$ -延长的;细胞中88.9%的ETA为 $\Delta 5$ -去饱和的;细胞中91.7%的EPA为 $\Delta 5$ -延长的;细胞中59.1%的DPA为 $\Delta 4$ -去饱和以产生DHA(表18)。

[1050] 表18. 用pJP3057、pJP3059、pJP3060转化的T2拟南芥种子的代表性脂肪酸分布

	样品	Columbia	pJP3057	pJP3059	pJP3060
[1051]	16:0	7.7	7.6	8.3	7.4
	16:1d9	0.3	0.3	0.3	0.3
	18:0	3.1	3.7	3.8	3.4
	20:0	2.1	1.8	1.8	1.9
	22:0	0.3	0.3	0.3	0.3
	24:0	0.2	0.2	0.2	0.2

[1052]	18:1d9	12.9	12.8	12.2	13.6
	18:1d11	1.5	1.8	1.9	1.6
	20:1d11	18.3	16.3	14.7	16.0
	20:1 d13	1.7	1.5	2.0	1.9
	22:1d13	1.6	1.2	1.2	1.3
	24:1d15	0.2	0.2	0.2	0.2
	其他	2.5	2.3	2.6	2.5
	18:2n6	27.8	27.2	28.2	27.9
	18:3n6	0.0	1.4	0.8	0.4
	20:3n6	0.0	0.0	0.0	0.0
	20:4n6	0.0	0.0	0.0	0.0
	18:3n3	19.7	17.3	18.2	18.2
	18:4n3	0.0	0.4	0.7	0.6
	20:4n3	0.0	0.6	0.3	0.7
	20:5n3	0.0	0.2	0.2	0.3
	22:5n3	0.0	0.3	0.9	0.2
	22:6n3	0.0	2.4	1.3	1.0
		100.0	100.0	100.0	100.0

[1053] 用pJP3060转化的代表性T2种子含有SDA (0.6%)、ETA (0.7%)、EPA (0.3%)、DPA (0.2%)和DHA (1.0%)。种子油还含有痕量的GLA、EtrA和ARA。转换效率如下：细胞中产生的13.3%的ALA为 $\Delta 6$ -去饱和的；细胞中产生的78.6%的SDA为 $\Delta 6$ -延长的；细胞中68.2%的ETA为 $\Delta 5$ -去饱和的；细胞中80.0%的EPA为 $\Delta 5$ -延长的；细胞中83.3%的DPA为 $\Delta 4$ -去饱和以产生DHA (表18)。

[1054] 结果

[1055] 构建体pJP3057中除 $\Delta 6$ -去饱和酶之外的所有基因表现出高活性/高转换效率。这表明 $\Delta 6$ -、 $\Delta 5$ -和 $\Delta 4$ -去饱和酶可能作用于酰基辅酶A底物，因为天然底物ALA由酰基-PC去饱和酶产生，导致较低的 $\Delta 6$ -去饱和作用，并且转基因去饱和酶底物ETA和DPA由已知在酰基-CoA库中作用的延长酶产生。此外， $\Delta 6$ -和 $\Delta 5$ -延长酶步骤的高效率(>80%效率)表明前面紧邻的那个步骤的去饱和酶(分别为 $\Delta 6$ -和 $\Delta 5$ -去饱和酶)作用于酰基辅酶A底物。可以合理地预期当达到纯合性时，在转基因植物的后代中这些基因的活性会增加，并且作为结果LC-PUFA产物的水平会增加。

[1056] 在构建体pJP3057和pJP3059中沉默阻抑物的存在增加了种子油中新脂肪酸的总水平和途径的终产物DHA的水平。

[1057] 讨论

[1058] 对于实施例13-15，引入的外源核酸可以被植物作为外源DNA或RNA检测，从而由于宿主介导的转基因抑制机制而导致降低的表达。这些抑制机制可以通过小RNA群体的生物

发生而靶向转基因,并且这些小RNA指导抑制装置来限制转基因的表达(Matzke等,2001)。转基因可以通过各种方式来限制,包括染色体插入位点处的DNA直接修饰,例如通过甲基化,或者通过RNA(Hamilton and Baulcombe,1999;Voinnet等,2003)或蛋白(Brodersen等,2008)水平的转录后沉默。引发这样的抑制机制的外源DNA或RNA的特征尚未完全了解(Lindbo等,1993;Lechtenberg等,2003)。但是这样的宿主介导的转基因表达抑制更可能具有这样的特性,其要求高表达、多转基因以及具有互相类似或与宿主基因组类似的区域的转基因(Schubert等,2004)。而且,转基因性能可以随着随后的每代而逐渐降低,这很可能是由于转基因的启动子和编码区的甲基化(Hagan等,2003)。

[1059] 在本文中,我们证实拟南芥中从胚发生后种子特异性启动子表达的病毒阻抑蛋白(VSP)在发育上是耐受的。各种VSP与可定量特性GFP的共表达表明,重组特性在表达VSP的种子中也是耐受的(实施例14)。由于已知VSP阻断构成转基因抑制装置的小RNA代谢,所以我们认为种子中具有重组特性的VSP的共表达会保证,这些特性会以高且不降低的水平在许多代中保持。

[1060] 由于植物耐受胚发生后诸如P19和PePo的VSP表达,所以这表明内源性发育信号,至少那些利用小RNA的信号在植物发育的这一晚期阶段是较低的或较不重要的。本研究所选的4种VSP很可能在小RNA生物发生的不同部分发挥作用,并因此发挥不同程度的功能。通过使用共引入的VSP来降低多转基因盒中的沉默效果,可以预见转基因表达策略的许多变化。首先,可以重复使用具有较少要求的相同表达盒以避免调控序列或编码区之间的序列重复。因此,这一特性可以允许利用相同的启动子-加A信号来构建大多基因表达载体。可选地,单基因的多个拷贝可以用于增加表达水平,并具有降低的可能性或程度的沉默效果的发生或者在植物的代中增加表达稳定性。

[1061] 实施例16:利用种子特异性启动子在植物的叶细胞中瞬时表达基因

[1062] 如下利用增强的本氏烟草瞬时表达系统,在植物的叶组织中证实了微胞藻CCMP1545 Δ 6-去饱和酶(SEQ ID NO:8,由SEQ ID NO:7编码)、塔胞藻CS-0140 Δ 6-延长酶(SEQ ID NO:4,由SEQ ID NO:3编码)、盐生巴夫藻 Δ 5-去饱和酶(SEQ ID NO:26,由SEQ ID NO:25编码)、塔胞藻CS-0140 Δ 5-延长酶(SEQ ID NO:6,由SEQ ID NO:5编码)和盐生巴夫藻 Δ 4-去饱和酶(SEQ ID NO:73,由SEQ ID NO:72编码)基因编码的蛋白的酶活性,每个基因均在种子特异性启动子的控制之下。

[1063] 将实施例15所述含有5个DHA生物合成基因的嵌合载体pJP3057引入根癌土壤杆菌菌株AGL1,每个基因均在种子特异性截短的napin启动子FP1的控制之下。通过将密码子优化的拟南芥多叶子叶2(Arath-LEC2)基因克隆至35S:pORE04的EcoRI位点来产生命名为35S:LEC2的嵌合载体。将35S:LEC2构建体单独引入根癌土壤杆菌菌株AGL1。将来自含有pJP3057或35S:LEC2的AGL1单独培养物的转基因细胞混合并且将混合物渗入本氏烟草植物的叶组织。渗入之后使植物再生长4天,然后取叶盘用于叶子脂质中的总脂肪酸和分离的脂质组分的GC分析。这表明这些基因在本氏烟草中发挥功能以产生DHA(表19)。用这些基因转化的叶组织含有SDA(1.2%)、ETA(2.0%)、EPA(0.6%)、DPA(1.7%)和DHA(2.5%)。所述叶组织还含有GLA(2.4%)和痕量的其他长链 ω 6脂肪酸。

[1064] 将嵌合载体pJP3115和pJP3116(实施例17)引入根癌土壤杆菌菌株AGL1。将来自4个含有pJP3115、pJP3116、35S:P19和35S:LEC2中任一个的AGL1单独培养物的转基因细胞混

合并并且将混合物渗入本氏烟草植物的叶组织。渗入之后使植物再生长4天,然后取叶盘用于GC分析,显示这些基因在本氏烟草中发挥功能以产生DHA(表19)。用这些基因转化的叶组织含有SDA(5.6%)、ETA(1.4%)、EPA(0.2%)、DPA(1.7%)和DHA(2.4%)。所述叶组织还含有痕量的长链 ω 6脂肪酸。

[1065] 该实验证实双构建体pJP3115和pJP3116组合发挥功能以和含有所有8个基因的单构建体一样有效地产生DHA。

[1066] 表19. 用各种构建体瞬时转化的本氏烟草叶组织的脂肪酸组成(总脂肪酸的百分比)。误差表示一式三份进行的单独渗入之间的标准偏差。

[1067]	一般 FA	仅 P19	35S:LEC2	pJP3057	pJP3057 + 35S:LEC2
	16:0	13.2 ± 0.5	12.8 ± 1.0	13.3 ± 0.1	13.2 ± 0.6
	16:1 ^{Δ3t}	1.5 ± 0.1	1.4 ± 0.2	1.3 ± 0.1	1.1 ± 0.0
	16:3 ^{Δ9,12,15}	7.6 ± 0.3	8.2 ± 0.4	7.1 ± 0.3	7.5 ± 0.4
	18:0	1.7 ± 0.2	2.0 ± 0.4	1.8 ± 0.1	2.4 ± 0.3
	18:1 ^{Δ9}	0.9 ± 0.1	0.9 ± 0.1	1.1 ± 0.1	1.5 ± 0.2
	18:2 ^{Δ9,12}	12.6 ± 1.1	12.8 ± 0.6	13.8 ± 0.1	12.7 ± 0.4
	18:3 ^{Δ9,12,15}	58.3 ± 1.9	56.3 ± 2.6	56.3 ± 0.7	44.8 ± 2.1
	20:0	0.3 ± 0.0	0.4 ± 0.1	0.3 ± 0.0	0.5 ± 0.1
	其他少量	3.8	4.6	3.9	4.8
	总计	99.9	99.6	98.9	88.5
	新Δ6 PUFA				
	18:3 ^{Δ6,9,12}	—	0.1 ± 0.0	—	2.4 ± 0.1
	20:3 ^{Δ8,11,14}	0.1 ± 0.1	0.3 ± 0.1	0.2 ± 0.1	0.2 ± 0.1
	20:4 ^{Δ5,8,11,14}	—	—	—	—
	22:4 ^{Δ7,10,13,16}	—	—	—	0.6 ± 0.1
	22:5 ^{Δ4,7,10,13,16}	—	—	—	0.3 ± 0.0
	总计	0.1	0.4	0.2	3.5
	新Δ3 PUFA				
	18:4 ^{Δ6,9,12,15}	—	—	0.9 ± 0.1	1.2 ± 0.1
	20:4 ^{Δ8,11,14,17}	—	—	—	2.0 ± 0.1
	20:5 ^{Δ5,8,11,14,17}	—	—	—	0.6 ± 0.0
	22:5 ^{Δ7,10,13,16,19}	—	—	—	1.7 ± 0.1
	22:6^{Δ4,7,10,13,16,19}	—	—	—	2.5 ± 0.2
	总计	—	—	0.9	8.0
	总新 FA	0.1	0.4	1.1	11.5
	总 FA	100.0	100.0	100.0	100.0

[1068] 讨论:种子特异性构建体的快速失败和验证

[1069] 利用转录因子,在这种情况下为LEC2,联合每个在诸如种子特异性启动子的组织特异性启动子的控制之下的一套基因的实验表明,这样的构建体可以在组织特异性启动子不会正常表达的诸如叶的异源系统中测试,并且预测在种子中表达。在叶细胞中瞬时表达

种子特异性启动子的能力允许构建体设计的快速验证。确定种子特异性启动子的效果的实验,尤其是在多基因构建体的背景下,先前依赖于稳定转化至油料种子模式植物或农作物,之后产生后代系,然后表型分析可以确定构建体在植物种子中的效果。在本氏烟草中获得的脂肪酸的水平与用如实施例15所述的该相同构建体稳定转化的拟南芥中所见的水平相似的事实增加了对该测定的适用性的信心。

[1070] 实施例17:用于DHA生物合成的双构建体

[1071] 如下构建载体pJP3115(图22)。首先,将载体pJP101acq(图14)的SbfI-ApaI片段克隆至pORE03的PstI-ApaI位点以产生pJP3011。然后,将含有密码子优化的细小微胞藻 $\Delta 6$ -去饱和酶(SEQ ID NO:125)的SwaI片段克隆至pJP3011中T4DNA聚合酶处理的XhoI位点以产生pJP3108。然后将含有密码子优化的盐生巴夫藻 $\Delta 5$ -去饱和酶(SEQ ID NO:127)的SwaI片段克隆至pJP3108中T4DNA聚合酶处理的NotI位点以产生pJP3109。将含有密码子优化的P.cordata $\Delta 6$ -延长酶(SEQ ID NO:126)的SwaI片段克隆至pJP3109中的SmaI位点以产生pJP3110。然后将pJP3110的BsiWI-AsiSI片段克隆至pORE04的BsiWI-AsiSI位点来将构建体从抗BASTA构建体转换为抗卡那霉素构建体,产生pJP3111。将含有截短的napin启动子FP1和Crepis palestina $\Delta 12$ -去饱和酶的NcoI(T4DNA聚合酶处理)-SbfI片段克隆至pJP3111中的EcoRV-PstI位点以产生pJP3115。

[1072] 如下构建载体pJP3116(图23)。首先将含有密码子优化的P.cordata $\Delta 5$ -延长酶(SEQ ID NO:128)的SwaI片段克隆至pJP3011中T4DNA聚合酶处理的XhoI位点以产生pJP3112。将含有密码子优化的盐生巴夫藻 $\Delta 4$ -去饱和酶(SEQ ID NO:129)的SwaI片段克隆至pJP3112中的SmaI位点以产生pJP3113。然后将含有Perilla frutescens $\Delta 15$ -去饱和酶的NotI片段克隆至pJP3113中的NotI位点以产生pJP3114。然后将含有潮霉素抗性盒(由花椰菜花叶病病毒35S启动子随后从二元载体pWVEC8获得的CAT-1内含子断裂潮霉素B磷酸转移酶基因和NOS终止子组成)的XbaI-MluI片段克隆至pJP3114的AvrII-MluI位点来将构建体从抗BASTA构建体转变为抗潮霉素构建体以产生pJP3116。

[1073] 将嵌合载体pJP3115和pJP3116分别引入根癌土壤杆菌菌株AGL1,并且将来自这些培养物的转基因细胞与用35S:P19转化的AGL1混合,然后在温室中将混合物渗入本氏烟草植物的叶组织。渗入之后使植物再生长5天,然后取叶盘用于GC分析,显示这些基因在本氏烟草中发挥功能以产生DHA(表20)。用这些基因转化的叶组织含有SDA(5.6%)、ETA(1.4%)、EPA(0.2%)、DPA(1.7%)和DHA(2.4%)。所述叶组织还含有痕量的GLA、ETA和ARA。转换效率如下:细胞中98.9%的油酸为 $\Delta 12$ -去饱和的(与对照样品没有显著不同);细胞中95.4%的LA为 $\Delta 15$ -去饱和的;细胞中产生的18.1%的ALA为 $\Delta 6$ -去饱和的;细胞中产生的50.4%的SDA为 $\Delta 6$ -延长的;细胞中75.4%的ETA为 $\Delta 5$ -去饱和的;细胞中95.4%的EPA为 $\Delta 5$ -延长的;细胞中58.5%的DPA为 $\Delta 4$ -去饱和以产生DHA。

[1074] 两个构建体均用于转化油菜。收集T1种子并分析一般形态学变化和LC-PUFA合成水平。

[1075] 表20.用各种构建体瞬时转化的本氏烟草叶组织的脂肪酸组成(总脂肪酸百分比)。误差表示一式三份进行的单独渗入之间的标准偏差。

	一般 FA	仅 P19	35S:LEC2	pJP3115 + pJP3116 + 35S:LEC2
[1076]	16:0	13.2 ± 0.5	12.8 ± 1.0	16.1 ± 0.1
	16:1 ^{Δ3t}	1.5 ± 0.1	1.4 ± 0.2	1.3 ± 0.1
	16:3 ^{Δ9,12,15}	7.6 ± 0.3	8.2 ± 0.4	6.8 ± 0.1
	18:0	1.7 ± 0.2	2.0 ± 0.4	3.4 ± 0.0
	18:1 ^{Δ9}	0.9 ± 0.1	0.9 ± 0.1	0.7 ± 0.0
	18:2 ^{Δ9,12}	12.6 ± 1.1	12.8 ± 0.6	2.6 ± 0.1
	18:3 ^{Δ9,12,15}	58.3 ± 1.9	56.3 ± 2.6	51.0 ± 0.1
	20:0	0.3 ± 0.0	0.4 ± 0.1	0.6 ± 0.0
	其他少量	3.8	4.6	5.7
	总计	99.9	99.6	88.3
	新Δ6 PUFA			
	18:3 ^{Δ6,9,12}	—	0.1 ± 0.0	0.2 ± 0.0
	20:3 ^{Δ8,11,14}	0.1 ± 0.1	0.3 ± 0.1	0.1 ± 0.1
	20:4 ^{Δ5,8,11,14}	—	—	—
	22:4 ^{Δ7,10,13,16}	—	—	—
	22:5 ^{Δ4,7,10,13,16}	—	—	0.1 ± 0.0
	总计	0.1	0.4	0.4
	新Δ3 PUFA			
	18:4 ^{Δ6,9,12,15}	—	—	5.6 ± 0.1
[1077]	20:4 ^{Δ8,11,14,17}	—	—	1.4 ± 0.1
	20:5 ^{Δ5,8,11,14,17}	—	—	0.2 ± 0.0
	22:5 ^{Δ7,10,13,16,19}	—	—	1.7 ± 0.0
	22:6^{Δ4,7,10,13,16,19}	—	—	2.4 ± 0.1
	总计	—	—	11.3
	总新 FA	0.1	0.4	11.7
	总 FA	100.0	100.0	100.0

[1078] 讨论

[1079] pJP3115和pJP3116设计为提供产生DHA的所有基因的组合,即两个重组载体互补以组成途径。pJP3115编码的Δ12-去饱和酶所产生的脂肪酸被用作pJP3116编码的Δ15-去饱和酶的底物,pJP3116还含有随后的Δ6-去饱和酶、Δ6-延长酶和Δ5-去饱和酶的基因。然后pJP3115编码的Δ5-延长酶作用于Δ5-去饱和酶的产物EPA,其产物被也由pJP3115编码的Δ4-去饱和酶转换为DHA。在两个构建体之间划分转基因的原理,其用于单独稳定转化植物与随后的优良株的杂交以组成整个途径,避免了与在单个构建体中含有众多转基因相关的一些问题,例如因为增加的大小而减少的转化效率和减少的基因表达。这些构建体的稳定转化的组合,通过超级转化或通过两个转基因系的杂交,会导致含有DHA合成所需的全套基因的转基因植物。

[1080] 还注意到构建体pJP3115含有4个以颠倒形式表达的基因,即2个基因以一个方向而2个基因以另一个方向,这样基因对以发散方式转录(远离对方)。当与构建体pJP107(实

施例8)中用于表达3个基因的颠倒设计比较时,得出的结论是在该形式中第4个基因的加入不会阻碍基因的表达。

[1081] 有趣的是注意到与上述其他实验相比相对较低的 $\Delta 6$ -延长作用效率(50.4%),其可能是因为编码用于前三个去饱和作用步骤的酶的基因全部从FP1启动子表达而编码 $\Delta 6$ -延长酶的基因由拟南芥FAE1启动子驱动这一事实。这被认为引起启动子时机的不同,FAE1启动子在FP1启动子之后活化。与 $\Delta 6$ -延长酶由FP1启动子驱动的前述实验相比,这导致SDA的较高积累,然后在 $\Delta 6$ -延长酶可以对SDA作用之前从 $\Delta 6$ -延长酶到达的代谢库去除SDA。

[1082] 实施例18:编码微藻DGAT2的基因的分离和表征

[1083] 全长细小微胞藻DGAT2基因的合成

[1084] 利用来自*Ostreococcus lucimarinus*的推定的氨基酸序列,Genbank登录号XP_00141576,作为查询序列用BLASTP程序分析美国能源部联合基因组研究所(<http://www.jgi.doe.gov/>)产生的微胞藻CCMP1545过滤蛋白质模型基因组序列。该分析显示在微胞藻CCMP1545中与XP_00141576具有同源性的推定的蛋白的存在。将微胞藻CCMP1545推定的蛋白序列用于设计和合成最适合在诸如欧洲油菜的双子叶植物中表达的密码子优化的核苷酸序列。蛋白编码区的核苷酸序列在SEQ ID NO:107中给出。质粒构建体命名为0928814_Mic1545-DGAT2_pMA。氨基酸序列如SEQ ID NO:108所示。利用微胞藻CCMP1545去饱和酶氨基酸序列SEQ ID NO:108在Genbank数据库中查询其他蛋白的BLASTP分析显示蛋白与DGAT具有同源性。最高程度的同一性为与登录号XP_002503155的氨基酸序列,微胞藻RCC299推定的蛋白的序列的全长具有53%的同一性。该基因在氨基酸74-334含有二酰基甘油酰基转移酶基序(NCBI保守结构域pfam03982)。

[1085] 通过将EcoRI片段内含有的0928814_Mic1545-DGAT2_pMA的整个编码区在EcoRI位点插入35S-pORE04(以上实施例4)来产生在组成型35S启动子控制之下的编码DGAT2的遗传构建体35S:Mic1545-DGAT2,产生pJP3128。将该嵌合载体引入根癌土壤杆菌菌株AGL1并将来自这些培养物的转基因细胞与35S:P19AGL1混合,然后在温室中将混合物渗入本氏烟草植物的叶组织。渗入之后使植物再生长4天,然后取叶盘用于脂质分析,显示编码DGAT2的基因发挥功能以增加转化的叶细胞中的总TAG,偏好多不饱和脂肪酸(表21)。具体地,在转化的细胞中TAG中多不饱和脂肪酸的水平增加了至少3倍。

[1086] 本领域人员理解可以对具体实施方式中所给出的发明进行多种变化和/或改动而不脱离所宽泛描述的本发明的实质或范围。以上实施方式在各方面均是示例性的,而不是限制性的。

[1087] 本申请要求2008年11月18日提交的US 61/199,669和2009年7月9日提交的US 61/270,710的优先权,在此通过引用将其内容并入本申请。

[1088] 本文中所讨论和/或援引的所有文献均以其全部内容并入本申请。

[1089] 本说明书中对文献、法案、材料、装置、物品等等的任何讨论的目的仅在于提供本发明的背景。不应理解为承认它们的任何或全部如同在本申请各项权利要求的优先权日之前已经存在一样构成了现有技术基础或者是本发明所属技术领域的公知常识。

[1090]

表 21. 在本氏烟草叶中表达细小微胞藻 DGAT2 的影响。P19 为对照，DGAT2 也含有 P19。当 DGAT2 表达时叶组织中 TAG 的总量增加 2 倍并且多不饱和脂肪酸受惠。

TAG 量, ug		TAG 谱		PL 量, ug		PL 谱	
P19	DGAT2	P19	DGAT2	P19	DGAT2	P19	DGAT2
C16:0	0.73	0.86	32.7	C16:0	26.33	14.2	14.5
C16:1d7	0.00	0.00	0.0	C16:1d7	0.00	0.0	0.0
16:1d13t	0.00	0.00	0.0	16:1d13t	3.40	1.8	2.1
16:2w6	0.00	0.00	0.0	16:2w6	1.60	0.9	0.8
16:2w4	0.00	0.05	1.7	16:2w4	1.15	0.6	0.7
16:3w3	0.00	0.00	0.0	16:3w3	12.65	6.8	6.4
C18:0	0.27	0.32	12.0	C18:0	4.50	2.4	2.5
C18:1d9	0.00	0.21	8.0	C18:1d9	1.89	1.0	1.3
C18:1d11	0.00	0.00	0.0	C18:1d11	0.85	0.5	0.6
C18:2n6	0.14	0.68	25.9	C18:2n6	24.36	13.2	11.7
C18:3n6	0.00	0.00	0.0	C18:3n6	0.00	0.0	0.0
C18:3n3	0.16	0.52	19.6	C18:3n3	106.38	57.6	58.4
C20:0	0.00	0.00	0.0	C20:0	0.59	0.3	0.4
C18:4n3	0.00	0.00	0.0	C18:4n3	0.00	0.0	0.0
C20:3n3	0.00	0.00	0.0	C20:3n3	0.34	0.2	0.0
C22:0	0.00	0.00	0.0	C22:0	0.37	0.2	0.3
C20:4n3	0.00	0.00	0.0	C20:4n3	0.00	0.0	0.0
C20:5n3	0.00	0.00	0.0	C20:5n3	0.00	0.0	0.0
C22:3n3	0.00	0.00	0.0	C22:3n3	0.00	0.0	0.0
C24:0	0.00	0.00	0.0	C24:0	0.43	0.2	0.3
C22:5n6	0.00	0.00	0.0	C22:5n6	0.00	0.0	0.0
C22:5n3	0.00	0.00	0.0	C22:5n3	0.00	0.0	0.0
C22:6n3	0.00	0.00	0.0	C22:6n3	0.00	0.0	0.0
TFA	26.2	52.7	100.0	TFA	3696.9	100.0	100.0
(ug/g lf FW)		(ug/g lf FW)		2814.0			

[1091] 参考文献

[1092] Abbadi et al. (2004) Plant Cell 16:2734-2748.

- [1093] Abbott et al. (1998) *Science* 282:2012-2018.
- [1094] Abdullah et al. (1986) *Biotech.* 4:1087.
- [1095] Agaba et al. (2004) *Marine Biotechnol.* (NY) 6:251-261.
- [1096] Al-Mariri et al. (2002) *Infect. Immun.* 70:1915-1923.
- [1097] Alvarez et al. (2000) *Theor Appl Genet* 100:319-327.
- [1098] Bates et al. (2007) *J. Biol. Chem.* Vol. 282:31206-31216.
- [1099] Baumberger et al. (2007) *Curr. Biol.* 17:1609-1614.
- [1100] Baumlein et al. (1991) *Mol. Gen. Genet.* 225:459-467.
- [1101] Baumlein et al. (1992) *Plant J.* 2:233-239.
- [1102] Beaudoin et al. (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97:6421-6426.
- [1103] Beclin et al. (2002) *Curr. Biol.* 12:684-688.
- [1104] Berberich et al. (1998) *Plant Mol. Biol.* 36:297-306.
- [1105] Bligh and Dyer (1959) *Canadian J. Biochem.* 37:911-917.
- [1106] Bortolamiol et al. (2007) *Curr. Biol.* 17:1615-1621.
- [1107] Bouvier-Nave et al. (2000) *Euro. J. Biochem.* 267:85-96.
- [1108] Brodersen et al. (2008) *Science* 320:1185-1190.
- [1109] Brothaerts et al. (2005) *Nature* 433:629-633.
- [1110] Brosnan et al. (2007) *Proc. Natl. Acad. Sci. U.S.A.* 104:14741-14746.
- [1111] Broun et al. (1998) *Plant J.* 13:201-210.
- [1112] Capecchi (1980) *Cell* 22:479-488.
- [1113] Chapman et al. (2004) *Gen. Dev.* 18:1179-1186.
- [1114] Chen et al. (2004) *The Plant Cell* 16:1302-1313.
- [1115] Cheng et al. (1996) *Plant Cell Rep.* 15:653-657.
- [1116] Chikwamba et al. (2003) *Proc. Natl. Acad. Sci. U.S.A.* 100:11127-11132.
- [1117] Cho et al. (1999a) *J. Biol. Chem.* 274:471-477.
- [1118] Cho et al. (1999b) *J. Biol. Chem.* 274:37335-37339.
- [1119] Chung et al. (2006) *BMC Genomics* 7:120.
- [1120] Clapp (1993) *Clin. Perinatol.* 20:155-168.
- [1121] Clough and Bent (1998) *Plant J.* 16:735-43.
- [1122] Courvalin et al (1995) *Life Sci.* 318:1209-1212.
- [1123] Coutu et al. (2007) *Transgenic Res.* 16:771-781.
- [1124] Curiel et al. (1992) *Hum. Gen. Ther.* 3:147-154.
- [1125] Darji et al. (1997) *Cell* 91:765-775.
- [1126] Denic and Weissman (2007) *Cell* 130:663-677.
- [1127] Dietrich et al. (1998) *Nature Biotech.* 18:181-185.
- [1128] Dietrich et al. (2001) *Vaccine* 19:2506.
- [1129] Ding and Voinnet (2007) *Cell* 130:413-426.
- [1130] Domergue et al. (2002) *Eur. J. Biochem.* 269:4105-4113.
- [1131] Domergue et al. (2003) *J. Biol. Chem.* 278:35115-35126.

- [1132] Domergue et al. (2005) *Biochem. J.* 1 389:483-490.
- [1133] Dunoyer et al. (2004) *The Plant Cell* 16:1235-1250.
- [1134] Eglitis et al. (1988) *Biotechniques* 6:608-614.
- [1135] Ellerstrom et al. (1996) *Plant Mol. Biol.* 32:1019-1027.
- [1136] Fennelly et al. (1999) *J. Immunol.* 162:1603-1610.
- [1137] Fraser et al. (2004) *Plant Physiol.* 135:859-866.
- [1138] Fuji et al. (2007) *Plant Cell* 19:597-609.
- [1139] Fujimura et al. (1985) *Plant Tissue Culture Lett.* 2:74.
- [1140] Garcia-Maroto et al. (2002) *Lipids* 37:417-426.
- [1141] Girke et al. (1998) *Plant J.* 15:39-48.
- [1142] Gleave (1992) *Plant Mol. Biol.* 20:1203-1207.
- [1143] Glevin et al (2003) *Microbiol. Mol. Biol. Rev.* 67:16-37.
- [1144] Glick et al. (2008) *Proc. Natl. Acad. Sci. U.S.A.* 105:157-161.
- [1145] Graham et al. (1973) *Virology* 54:536-539.
- [1146] Grant et al. (1995) *Plant Cell Rep.* 15:254-258.
- [1147] Guillard and Rythers (1962) *Can. J. Microbiol.* 8:229-239.
- [1148] Grillot-Courvalin et al. (1998) *Nature Biotech.* 16:862-866.
- [1149] Grillot-Courvalin (1999) *Curr. Opin. Biotech.* 10:477-481.
- [1150] Hagan et al. (2003) *Plant. Biotech. J.* 1:479-490.
- [1151] Hamilton and Baulcombe (1999) *Science* 286:950-952.
- [1152] Harayama (1998) *Trends Biotechnol.* 16:76-82.
- [1153] Hastings et al. (2001) *Proc. Natl. Acad. Sci. U.S.A.* 98:14304-14309.
- [1154] Hense et al. (2001) *Cell Microbiol.* 3:599-609.
- [1155] Hinchee et al. (1988) *Biotechnology* 6:915-922.
- [1156] Hoffmann et al. (2008) *J Biol. Chem.* 283:22352-22362.
- [1157] Hong et al. (2002a) *Lipids* 37:863-868.
- [1158] Horiguchi et al. (1998) *Plant Cell Physiol.* 39:540-544.
- [1159] Horvath et al. (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97:1914-1919.
- [1160] Huang et al. (1999) *Lipids* 34:649-659.
- [1161] Huang et al. (2004) *Biochimie* 86 (11) :793-8.
- [1162] Inagaki et al. (2002) *Biosci. Biotechnol. Biochem.* 66:613-621.
- [1163] Johansen and Carrington (2001) *Plant Physiol.* 126:930-938.
- [1164] Kajikawa et al. (2004) *Plant Mol. Biol.* 54:335-52.
- [1165] Kajikawa et al. (2006) *FEBS Lett* 580:149-154.
- [1166] Kapila et al. (1997) *Plant Sci.* 122:101-108.
- [1167] Kasschau et al. (2003) *Devel. Cell* 4:205-217.
- [1168] Khozin et al. (1997) *Plant Physiol.* 114:223-230.
- [1169] Knutzon et al. (1998) *J. Biol. Chem.* 273:29360-6.
- [1170] Koziel et al. (1996) *Plant Mol. Biol.* 32:393-405.

- [1171] Kunik et al. (2001) *Proc.Natl.Acad.Sci.U.S.A.* 98:1871-1876.
- [1172] Lacroix et al. *Proc.Natl.Acad.Sci.U.S.A.* 105:15429-15434.
- [1173] Lazo et al. (1991) *Biotechnol.* 9:693-697.
- [1174] Lechtenberg et al. (2003) *Plant J.* 507-517.
- [1175] Leonard et al. (2000) *Biochem.J.* 347:719-724.
- [1176] Leonard et al. (2000b) *Biochem.J.* 350:765-770.
- [1177] Leonard et al. (2002) *Lipids* 37:733-740.
- [1178] Lewsey et al. (2007) *Plant J.* 50:240-252.
- [1179] Lindbo et al. (1993) *Plant Cell* 5:1749-1759.
- [1180] Lo et al. (2003) *Genome Res.* 13:455-466.
- [1181] Lu et al. (1993) *J.Exp.Med.* 178:2089-2096.
- [1182] Mallory et al (2002) *Nat.Biotech.* 20:622-625.
- [1183] Marillonnet et al. (2005) *Nature Biotechnology* 23:718-723.
- [1184] Matzke et al. (2001) *Science* 293:1080-1083.
- [1185] Meng et al. (2008) *J.Gen.Virol.* 89:2349-2358.
- [1186] Meyer et al. (2003) *Biochem.* 42:9779-9788.
- [1187] Meyer et al. (2004) *Lipid Res* 45:1899-1909.
- [1188] Michaelson et al. (1998a) *J.Biol.Chem.* 273:19055-19059.
- [1189] Michaelson et al. (1998b) *FEBS Lett.* 439:215-218.
- [1190] Moreau et al. (1998) *Progress Lip.Res.* 37:371-391.
- [1191] Napier (2007) *Ann.Rev.Plant.Biol.* 58:295-319.
- [1192] Napier et al. (1998) *Biochem.J.* 330:611-614.
- [1193] Needleman and Wunsch (1970) *J.Mol.Biol.* 48:443-453.
- [1194] Niedz et al (1995) *Plant Cell Reports* 14:403.
- [1195] Nishizawa et al. (2003) *Plant J.* 34:647-659.
- [1196] Ohlrogge and Browse (1995) *Plant Cell* 7:957-970.
- [1197] Ow et al. (1986) *Science* 234:856-859.
- [1198] Parker-Barnes et al. (2000) *Proc.Natl.Acad.Sci.USA* 97:8284-8289.
- [1199] Pereira et al. (2004a) *Biochem.J.* 378:665-671.
- [1200] Pereira et al. (2004b) *Biochem.J.* 384:357-366.
- [1201] Perrin et al. (2000) *Mol Breed* 6:345-352.
- [1202] Potenza et al. (2004) *In Vitro Cell Dev Biol-Plant* 40:1-22.
- [1203] Prasher et al (1985) *Biochem.Biophys.Res.Comm.* 127:31-36.
- [1204] Qi et al. (2002) *FEBS Lett.* 510:159-165.
- [1205] Qi et al. (2004) *Nat.Biotech.* 22:739-745.
- [1206] Qiu et al. (2001) *J.Biol.Chem.* 276:31561-31566.
- [1207] Reddy and Thomas (1996) *Nat.Biotech.* 14:639-642.
- [1208] Reddy et al. (1993) *Plant Mol.Biol.* 22:293-300.
- [1209] Robert et al. (2005) *Func.Plant Biol.* 32:473-479.

- [1210] Robert et al. (2009) *Marine Biotech* 11:410-418.
- [1211] Rose et al. (1998) *Nucleic Acids Res.* 26:1628-1635.
- [1212] Saha et al. (2006) *Plant Physiol.* 141:1533-1543.
- [1213] Saito et al. (2000) *Eur. J. Biochem.* 267:1813-1818.
- [1214] Sakuradani et al. (1999) *Gene* 238:445-453.
- [1215] Sato et al. (2004) *Crop Sci.* 44:646-652.
- [1216] Sayanova et al. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94:4211-4216.
- [1217] Sayanova et al. (2003) *FEBS Lett.* 542:100-104.
- [1218] Sayanova et al. (2006) *Planta* 224:1269-1277.
- [1219] Sayanova et al. (2007) *Plant Physiol* 144:455-467.
- [1220] Schaffner et al (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77:2163-2167.
- [1221] Schubert et al. (2004) *Plant Cell* 16:2561-2572.
- [1222] Singh et al. (2005) *Curr. Opin. in Plant Biol.* 8:197-203.
- [1223] Sizemore et al (1995) *Science* 270:299-302.
- [1224] Shianu et al (2001) *Vaccine* 19:3947-3956.
- [1225] Sperling et al. (2000) *Eur. J. Biochem.* 267:3801-3811.
- [1226] Sperling et al. (2001) *Arch. Biochem. Biophys.* 388:293-8.
- [1227] Sprecher et al. (1995) *J. Lipid Res.* 36:2471-2477.
- [1228] Spsychalla et al. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94:1142-1147.
- [1229] Stalker et al (1998) *J. Biol. Chem.* 263:6310-6314.
- [1230] Thillet et al (1988) *J. Biol. Chem* 263:12500-12508.
- [1231] Tonon et al. (2003) *FEBS Lett.* 553:440-444.
- [1232] Toriyama et al. (1986) *Theor. Appl. Genet.* 205:34.
- [1233] Trautwein (2001) *European J. Lipid Sci. and Tech.* 103:45-55.
- [1234] Tvrdik (2000) *J. Cell Biol.* 149:707-718.
- [1235] Tzfira & Citovsky (2006) *Curr. Opin. Biotech.* 17:147-154.
- [1236] Voinnet et al., (2003) *Plant J.* 33:949-956.
- [1237] Wagner et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:6099-6103.
- [1238] Wallis and Browse (1999) *Arch. Biochem. Biophys.* 365:307-316.
- [1239] Watts and Browse (1999b) *Arch. Biochem. Biophys.* 362:175-182.
- [1240] Weiss et al (2003) *Int. J. Med. Microbiol.* 293:95-106.
- [1241] Whitney et al. (2003) *Planta* 217:983-992.
- [1242] Wu et al. (2005) *Nat. Biotech.* 23:1013-1017.
- [1243] Yang et al. (2003) *Planta* 216:597-603.
- [1244] Yu et al. (2007) *Proc. Natl. Acad. Sci U.S.A.* 104:8924-8929.
- [1245] Zank et al. (2002) *Plant J.* 31:255-268.
- [1246] Zank et al. (2005) WO 2005/012316
- [1247] Zhang et al. (2004) *FEBS Lett.* 556:81-85.
- [1248] Zhang et al. (2006) 20:3255-3268.

- [1249] Zhang et al. (2007) Yeast 25:21-27.
- [1250] Zhou et al. (2006) Plant Sci.170:665-673.
- [1251] Zhou et al. (2007) Phytochem.68:785-796.
- [1252] Zipfel et al. (2006) Cell 125:749-760.

- [0001] 序列表
- [0002] <110> 联邦科学技术研究组织
- [0003] <120> 产生omega-3脂肪酸的酶和方法
- [0004] <130> 508730
- [0005] <150> US 61/199,669
- [0006] <151> 2008-11-18
- [0007] <150> US 61/270,710
- [0008] <151> 2009-07-09
- [0009] <160> 129
- [0010] <170> PatentIn version 3.5
- [0011] <210> 1
- [0012] <211> 825
- [0013] <212> DNA
- [0014] <213> Micromonas CS-0170
- [0015] <400> 1
- [0016] atggctttcc ccgccgtcgc gaccgctgtc ttcggtatct atgcgtacgt gatgtactgg 60
- [0017] tactccgtcc ccggcgccgg cttcctcgcc gagggcaagc aggacgtcgt ggcggtggatc 120
- [0018] ggcgatctgt ccgtggcgt ccccgccacc gccaccgtcc tgtacctctc catgtgctac 180
- [0019] ttcgggccga agatcatggc caagaggag gcgttcgagc cgaaggggtt catgctgggtg 240
- [0020] tacaacgcgt accagaccgt gttcaacatc atcacgtcg ccatcttcat cgctgagctc 300
- [0021] aaccggctgg gcgtgaagc gtggggcggc cgctgtcgt ggtccgatcc caacagcttc 360
- [0022] tacatctgcc tcgccatctg gctccactac aacaacaagt acctggagct cctcgacacc 420
- [0023] gtcttcatgg tgctgcgcaa gaagaacaac cagctcagct tcctccacat ctaccaccac 480
- [0024] tgctcctca tctgggcgtg gtggatgggtg tgcttcgtca tcaagaacaa cgactgcatc 540
- [0025] gacgcgtact ttggcgctg catgaacgt ggaattcacg tcatcatgta ctctattac 600
- [0026] ctcatggcgg cgctgaagat caagtcccc tggaagcagt acatcaccat ggcgagatg 660
- [0027] ctccagttcg ccatcgtctt tgcgcacagc tgctacgtca tctacgacgg gcaactgccc 720
- [0028] gcgattctgc cgtggtcgca gatgttcgtg atgaccaaca tgctgggtgct gttcgggcag 780
- [0029] ttctacgtgc agacgtacac gaagaagaag gtaaaggcca agtaa 825
- [0030] <210> 2
- [0031] <211> 274
- [0032] <212> PRT
- [0033] <213> Micromonas CS-0170
- [0034] <400> 2
- [0035] Met Ala Phe Pro Ala Val Ala Thr Ala Val Phe Gly Ile Tyr Ala Tyr
- [0036] 1 5 10 15
- [0037] Val Met Tyr Trp Tyr Ser Val Pro Gly Ala Gly Phe Leu Ala Glu Gly
- [0038] 20 25 30
- [0039] Lys Gln Asp Val Val Ala Trp Ile Gly Asp Leu Ser Val Ala Leu Pro
- [0040] 35 40 45
- [0041] Ala Thr Ala Thr Val Leu Tyr Leu Ser Met Cys Tyr Phe Gly Pro Lys

[0042]	50	55	60
[0043]	Ile Met Ala Lys Arg Glu Ala Phe Glu Pro Lys Gly Phe Met Leu Val		
[0044]	65	70	75
[0045]	Tyr Asn Ala Tyr Gln Thr Val Phe Asn Ile Ile Thr Val Ala Ile Phe		
[0046]	85	90	95
[0047]	Ile Ala Glu Leu Asn Arg Leu Gly Val Lys Ala Trp Gly Gly Arg Leu		
[0048]	100	105	110
[0049]	Ser Trp Ser Asp Pro Asn Ser Phe Tyr Ile Cys Leu Ala Ile Trp Leu		
[0050]	115	120	125
[0051]	His Tyr Asn Asn Lys Tyr Leu Glu Leu Leu Asp Thr Val Phe Met Val		
[0052]	130	135	140
[0053]	Leu Arg Lys Lys Asn Asn Gln Leu Ser Phe Leu His Ile Tyr His His		
[0054]	145	150	155
[0055]	Cys Leu Leu Ile Trp Ala Trp Trp Met Val Cys Phe Val Ile Lys Asn		
[0056]	165	170	175
[0057]	Asn Asp Cys Ile Asp Ala Tyr Phe Gly Ala Cys Met Asn Ala Gly Ile		
[0058]	180	185	190
[0059]	His Val Ile Met Tyr Ser Tyr Tyr Leu Met Ala Ala Leu Lys Ile Lys		
[0060]	195	200	205
[0061]	Cys Pro Trp Lys Gln Tyr Ile Thr Met Ala Gln Met Leu Gln Phe Ala		
[0062]	210	215	220
[0063]	Ile Val Phe Ala His Ser Cys Tyr Val Ile Tyr Asp Gly His Cys Pro		
[0064]	225	230	235
[0065]	Ala Ile Leu Pro Trp Ser Gln Met Phe Val Met Thr Asn Met Leu Val		
[0066]	245	250	255
[0067]	Leu Phe Gly Gln Phe Tyr Val Gln Thr Tyr Thr Lys Lys Lys Val Lys		
[0068]	260	265	270
[0069]	Ala Lys		
[0070]	<210> 3		
[0071]	<211> 894		
[0072]	<212> DNA		
[0073]	<213> Pyramimonas CS-0140		
[0074]	<400> 3		
[0075]	atggagttcg ctcagcctct tgttgctatg gcacaggagc agtatgccgc aattgacgcg 60		
[0076]	gtggtagccc ctgcaatttt ctcagctacc gacagcatcg gttgggggtct taagcccatt 120		
[0077]	agcagcgcga caaaggatct tcctctcggt gagagtccga cgccgctcat actgagcctg 180		
[0078]	ttggcctatt ttgcgatcgt cggctctggg ctggtgtacc gcaaagtatt ccctcgcaca 240		
[0079]	gtaaaggggc aagaccctt cctgctgaag gcgctcatgc ttgcgcacaa cgtgttcttc 300		
[0080]	attggcctca gtctatacat gtgcttgaag cttgtctacg aggcttacgt caacaagtac 360		
[0081]	tccttctggg gaaacgccta caacccgcga cagaccgaga tggcgaaggt catctggatt 420		
[0082]	ttctacgtct ccaagatcta tgagttcatg gacacgttca tcatgctctt gaagggaac 480		
[0083]	gtcaaccagg tctcttttct gcatgtgtac catcatggct ccatctctgg tatctggtgg 540		

[0084] atgatcacct acgctgcccc tggcgggtgac gcgtacttct cggcggcgct caactcgtgg 600
 [0085] gtgcacgtgt gcatgtacac gtactacttc atggcggcgg tgctgccc aa ggacgagaag 660
 [0086] accaagcgca agtacctctg gtggggccgc tacctgaccc agatgcagat gttccagttc 720
 [0087] ttcatgaacc tgctccaggc ggtctacctc ctctactcct ctagccccta ccccaagttc 780
 [0088] atgcccagc tgctgggtgt gtacatggtc acgctgctga tgctcttcgg caacttctac 840
 [0089] tacatgaagc accacgcgag caagaagcag aagctggcca gcaagaagca gtag 894
 [0090] <210> 4
 [0091] <211> 297
 [0092] <212> PRT
 [0093] <213> Pyramimonas CS-0140
 [0094] <400> 4
 [0095] Met Glu Phe Ala Gln Pro Leu Val Ala Met Ala Gln Glu Gln Tyr Ala
 [0096] 1 5 10 15
 [0097] Ala Ile Asp Ala Val Val Ala Pro Ala Ile Phe Ser Ala Thr Asp Ser
 [0098] 20 25 30
 [0099] Ile Gly Trp Gly Leu Lys Pro Ile Ser Ser Ala Thr Lys Asp Leu Pro
 [0100] 35 40 45
 [0101] Leu Val Glu Ser Pro Thr Pro Leu Ile Leu Ser Leu Leu Ala Tyr Phe
 [0102] 50 55 60
 [0103] Ala Ile Val Gly Ser Gly Leu Val Tyr Arg Lys Val Phe Pro Arg Thr
 [0104] 65 70 75 80
 [0105] Val Lys Gly Gln Asp Pro Phe Leu Leu Lys Ala Leu Met Leu Ala His
 [0106] 85 90 95
 [0107] Asn Val Phe Leu Ile Gly Leu Ser Leu Tyr Met Cys Leu Lys Leu Val
 [0108] 100 105 110
 [0109] Tyr Glu Ala Tyr Val Asn Lys Tyr Ser Phe Trp Gly Asn Ala Tyr Asn
 [0110] 115 120 125
 [0111] Pro Ala Gln Thr Glu Met Ala Lys Val Ile Trp Ile Phe Tyr Val Ser
 [0112] 130 135 140
 [0113] Lys Ile Tyr Glu Phe Met Asp Thr Phe Ile Met Leu Leu Lys Gly Asn
 [0114] 145 150 155 160
 [0115] Val Asn Gln Val Ser Phe Leu His Val Tyr His His Gly Ser Ile Ser
 [0116] 165 170 175
 [0117] Gly Ile Trp Trp Met Ile Thr Tyr Ala Ala Pro Gly Gly Asp Ala Tyr
 [0118] 180 185 190
 [0119] Phe Ser Ala Ala Leu Asn Ser Trp Val His Val Cys Met Tyr Thr Tyr
 [0120] 195 200 205
 [0121] Tyr Phe Met Ala Ala Val Leu Pro Lys Asp Glu Lys Thr Lys Arg Lys
 [0122] 210 215 220
 [0123] Tyr Leu Trp Trp Gly Arg Tyr Leu Thr Gln Met Gln Met Phe Gln Phe
 [0124] 225 230 235 240
 [0125] Phe Met Asn Leu Leu Gln Ala Val Tyr Leu Leu Tyr Ser Ser Ser Pro

[0126]	245	250	255
[0127]	Tyr Pro Lys Phe Ile Ala Gln Leu Leu Val Val Tyr Met Val Thr Leu		
[0128]	260	265	270
[0129]	Leu Met Leu Phe Gly Asn Phe Tyr Tyr Met Lys His His Ala Ser Lys		
[0130]	275	280	285
[0131]	Lys Gln Lys Leu Ala Ser Lys Lys Gln		
[0132]	290	295	
[0133]	<210> 5		
[0134]	<211> 804		
[0135]	<212> DNA		
[0136]	<213> Pyramimonas CS-0140		
[0137]	<400> 5		
[0138]	atggcgtcta ttgcgattcc ggctgcgctg gcagggactc ttggttatgt gacgtacaat	60	
[0139]	gtcgcaaacc cagatattcc tgcattccgag aagggtgcctg cttactttat gcaggctcgag	120	
[0140]	tattgggggc caacgattgg gaccatcggg tatcttctgt tcatctactt tggtaaacgg	180	
[0141]	attatgcaaa acaggagcca gccgtttggc ctgaagaacg ctatgctggt gtacaacttc	240	
[0142]	tatcagactt tcttcaactc gtactgcata tacctttttg tcacgtcgca ccgcgctcag	300	
[0143]	gggctgaaag tttggggaaa catccccgat atgactgcca acagctgggg gatctcacag	360	
[0144]	gtgatctggc tgcactacaa caacaagtac gttgagctgc tggacacgtt cttcatggtc	420	
[0145]	atgcgcaaga agtttgacca gctttcgctc ctgcacattt accatcatac cctgttgatc	480	
[0146]	tggtcttggt tcgtgtgat gaaattggag cccgttgggg actgctactt tggctctagc	540	
[0147]	gtcaacacgt ttgtgcacgt cattatgtac tcgtactatg gccttgccgc gctcgggggtg	600	
[0148]	aattgcttct ggaagaagta cattacgcag attcagatgc tgcagttctg tatctgcgct	660	
[0149]	tcgcaactga tttataccgc ctatgtgcag aacaccgcgt tctggttggc ttacttgcag	720	
[0150]	ctgtgggtga tggatgaacat gttcgtgttg ttccccaact tctatcgcaa gcgctacaag	780	
[0151]	agcaagggtg ccaagaagca gtaa	804	
[0152]	<210> 6		
[0153]	<211> 267		
[0154]	<212> PRT		
[0155]	<213> Pyramimonas CS-0140		
[0156]	<400> 6		
[0157]	Met Ala Ser Ile Ala Ile Pro Ala Ala Leu Ala Gly Thr Leu Gly Tyr		
[0158]	1 5 10 15		
[0159]	Val Thr Tyr Asn Val Ala Asn Pro Asp Ile Pro Ala Ser Glu Lys Val		
[0160]	20 25 30		
[0161]	Pro Ala Tyr Phe Met Gln Val Glu Tyr Trp Gly Pro Thr Ile Gly Thr		
[0162]	35 40 45		
[0163]	Ile Gly Tyr Leu Leu Phe Ile Tyr Phe Gly Lys Arg Ile Met Gln Asn		
[0164]	50 55 60		
[0165]	Arg Ser Gln Pro Phe Gly Leu Lys Asn Ala Met Leu Val Tyr Asn Phe		
[0166]	65 70 75 80		
[0167]	Tyr Gln Thr Phe Phe Asn Ser Tyr Cys Ile Tyr Leu Phe Val Thr Ser		

[0168]	85	90	95
[0169]	His Arg Ala Gln Gly Leu Lys Val Trp Gly Asn Ile Pro Asp Met Thr		
[0170]	100	105	110
[0171]	Ala Asn Ser Trp Gly Ile Ser Gln Val Ile Trp Leu His Tyr Asn Asn		
[0172]	115	120	125
[0173]	Lys Tyr Val Glu Leu Leu Asp Thr Phe Phe Met Val Met Arg Lys Lys		
[0174]	130	135	140
[0175]	Phe Asp Gln Leu Ser Phe Leu His Ile Tyr His His Thr Leu Leu Ile		
[0176]	145	150	155
[0177]	Trp Ser Trp Phe Val Val Met Lys Leu Glu Pro Val Gly Asp Cys Tyr		
[0178]	165	170	175
[0179]	Phe Gly Ser Ser Val Asn Thr Phe Val His Val Ile Met Tyr Ser Tyr		
[0180]	180	185	190
[0181]	Tyr Gly Leu Ala Ala Leu Gly Val Asn Cys Phe Trp Lys Lys Tyr Ile		
[0182]	195	200	205
[0183]	Thr Gln Ile Gln Met Leu Gln Phe Cys Ile Cys Ala Ser His Ser Ile		
[0184]	210	215	220
[0185]	Tyr Thr Ala Tyr Val Gln Asn Thr Ala Phe Trp Leu Pro Tyr Leu Gln		
[0186]	225	230	235
[0187]	Leu Trp Val Met Val Asn Met Phe Val Leu Phe Ala Asn Phe Tyr Arg		
[0188]	245	250	255
[0189]	Lys Arg Tyr Lys Ser Lys Gly Ala Lys Lys Gln		
[0190]	260	265	
[0191]	<210> 7		
[0192]	<211> 1392		
[0193]	<212> DNA		
[0194]	<213> Micromonas CCMP1545		
[0195]	<400> 7		
[0196]	atgtgtcctc ctaagaccga tggaagatct tctcctagat ctctctctcac caggtctaag	60	
[0197]	tcatctgctg aggtcttga tgctaaggat gcttctaccg ctctgttga tcttaagacc	120	
[0198]	cttgagcctc atgaacttgc tgctaccttc gagactagat gggttagggt tgaggatgtt	180	
[0199]	gagtacgacg tgaccaactt caaacatcct ggtggaagcg tgatcttcta catgcttgct	240	
[0200]	aacctgggtg ctgatgctac tgaggctttc aaagaatttc acatgcgtag cctcaaggct	300	
[0201]	tggaagatgc ttagagcttt gccttctaga cctgctgaga tcaagagatc tgagtctgag	360	
[0202]	gatgctccta tgcttgagga tttcgctagg tggagagctg aacttgagag ggacggattc	420	
[0203]	ttcaagcctt ctatcaccca tgttgcttac cgtcttttgg agcttcttgc tactttcgtc	480	
[0204]	cttggaaaccg ctcttatgta cgctggatac cctatcattg ctagegttgt gtacgggtgt	540	
[0205]	ttcttcggag ctagatgtgg atgggttcaa catgagggtg gacacaactc tcttaccgga	600	
[0206]	tctgtgtacg tggataagag acttcaggct atgacttgcg gattcggact ttctaccage	660	
[0207]	ggagagatgt ggaaccagat gcataacaag caccatgcta cccctcagaa agttagacac	720	
[0208]	gacatggatc ttgataccac tcctgctgtg gctttcttca acaccgctgt ggaggataat	780	
[0209]	agacctaggg gattctctag agcttgggct agacttcaag cttggacctt cgttctgtt	840	

[0210]	acttctggac ttctcgttca ggctttcttg atctacgttc tccatcctag acaggtgctc	900
[0211]	aggaagaaga actacgagga agcttcttgg atgctcgttt ctcacgttgt tagaaccgct	960
[0212]	gttatcaagc ttgctaccgg atactcttgg cctgttgctt actggtggtt cactttcgga	1020
[0213]	aactggatcg cttacatgta cctcttcgct cactttctta cttctcacac tcacctcct	1080
[0214]	gttgttccat ctgacaagca ccttagctgg gttaactacg ctgttgatca caccgttgac	1140
[0215]	atcgatcctt ctctggata cgtaaactgg cttatgggat accttaactg ccaggttatc	1200
[0216]	caccatctct tcctgatat gcctcaattc agacagcctg aggtgtcaag aagattcgctc	1260
[0217]	cctttcgcta agaagtgggg actcaactac aaggtgctct cttactacgg tgcttggaag	1320
[0218]	gctactttca gcaacctcga caaagttgga cagcactact acgttaacgg aaaggctgag	1380
[0219]	aaggctcact ga	1392
[0220]	<210>	8
[0221]	<211>	463
[0222]	<212>	PRT
[0223]	<213>	Micromonas CCMP1545
[0224]	<400>	8
[0225]	Met Cys Pro Pro Lys Thr Asp Gly Arg Ser Ser Pro Arg Ser Pro Leu	
[0226]	1 5 10 15	
[0227]	Thr Arg Ser Lys Ser Ser Ala Glu Ala Leu Asp Ala Lys Asp Ala Ser	
[0228]	20 25 30	
[0229]	Thr Ala Pro Val Asp Leu Lys Thr Leu Glu Pro His Glu Leu Ala Ala	
[0230]	35 40 45	
[0231]	Thr Phe Glu Thr Arg Trp Val Arg Val Glu Asp Val Glu Tyr Asp Val	
[0232]	50 55 60	
[0233]	Thr Asn Phe Lys His Pro Gly Gly Ser Val Ile Phe Tyr Met Leu Ala	
[0234]	65 70 75 80	
[0235]	Asn Thr Gly Ala Asp Ala Thr Glu Ala Phe Lys Glu Phe His Met Arg	
[0236]	85 90 95	
[0237]	Ser Leu Lys Ala Trp Lys Met Leu Arg Ala Leu Pro Ser Arg Pro Ala	
[0238]	100 105 110	
[0239]	Glu Ile Lys Arg Ser Glu Ser Glu Asp Ala Pro Met Leu Glu Asp Phe	
[0240]	115 120 125	
[0241]	Ala Arg Trp Arg Ala Glu Leu Glu Arg Asp Gly Phe Phe Lys Pro Ser	
[0242]	130 135 140	
[0243]	Ile Thr His Val Ala Tyr Arg Leu Leu Glu Leu Leu Ala Thr Phe Ala	
[0244]	145 150 155 160	
[0245]	Leu Gly Thr Ala Leu Met Tyr Ala Gly Tyr Pro Ile Ile Ala Ser Val	
[0246]	165 170 175	
[0247]	Val Tyr Gly Ala Phe Phe Gly Ala Arg Cys Gly Trp Val Gln His Glu	
[0248]	180 185 190	
[0249]	Gly Gly His Asn Ser Leu Thr Gly Ser Val Tyr Val Asp Lys Arg Leu	
[0250]	195 200 205	
[0251]	Gln Ala Met Thr Cys Gly Phe Gly Leu Ser Thr Ser Gly Glu Met Trp	

[0252]	210	215	220
[0253]	Asn Gln Met His Asn Lys His His Ala Thr Pro Gln Lys Val Arg His		
[0254]	225	230	235 240
[0255]	Asp Met Asp Leu Asp Thr Thr Pro Ala Val Ala Phe Phe Asn Thr Ala		
[0256]	245	250	255
[0257]	Val Glu Asp Asn Arg Pro Arg Gly Phe Ser Arg Ala Trp Ala Arg Leu		
[0258]	260	265	270
[0259]	Gln Ala Trp Thr Phe Val Pro Val Thr Ser Gly Leu Leu Val Gln Ala		
[0260]	275	280	285
[0261]	Phe Trp Ile Tyr Val Leu His Pro Arg Gln Val Leu Arg Lys Lys Asn		
[0262]	290	295	300
[0263]	Tyr Glu Glu Ala Ser Trp Met Leu Val Ser His Val Val Arg Thr Ala		
[0264]	305	310	315 320
[0265]	Val Ile Lys Leu Ala Thr Gly Tyr Ser Trp Pro Val Ala Tyr Trp Trp		
[0266]	325	330	335
[0267]	Phe Thr Phe Gly Asn Trp Ile Ala Tyr Met Tyr Leu Phe Ala His Phe		
[0268]	340	345	350
[0269]	Ser Thr Ser His Thr His Leu Pro Val Val Pro Ser Asp Lys His Leu		
[0270]	355	360	365
[0271]	Ser Trp Val Asn Tyr Ala Val Asp His Thr Val Asp Ile Asp Pro Ser		
[0272]	370	375	380
[0273]	Arg Gly Tyr Val Asn Trp Leu Met Gly Tyr Leu Asn Cys Gln Val Ile		
[0274]	385	390	395 400
[0275]	His His Leu Phe Pro Asp Met Pro Gln Phe Arg Gln Pro Glu Val Ser		
[0276]	405	410	415
[0277]	Arg Arg Phe Val Pro Phe Ala Lys Lys Trp Gly Leu Asn Tyr Lys Val		
[0278]	420	425	430
[0279]	Leu Ser Tyr Tyr Gly Ala Trp Lys Ala Thr Phe Ser Asn Leu Asp Lys		
[0280]	435	440	445
[0281]	Val Gly Gln His Tyr Tyr Val Asn Gly Lys Ala Glu Lys Ala His		
[0282]	450	455	460
[0283]	<210> 9		
[0284]	<211> 1449		
[0285]	<212> DNA		
[0286]	<213> <i>Ostreococcus lucimarinus</i>		
[0287]	<400> 9		
[0288]	atgtgcgtcg aaacgaccga aggcacatcg cgaacgatgg cgaacgaacg cacgagctcg	60	
[0289]	tcgtcgtcgc tgagcgaagg cggaacccg acggtgacgg tcgggatggg aagcgaagac	120	
[0290]	gcggggaaga agactcgaaa cgcgagcgtc acggcgtgga cgaaagagtt ggagccgcac	180	
[0291]	gcgatcgcga agacgttcga acggcggtac gtagcatcg aaggcgtgga atacgatgtg	240	
[0292]	acggatttta agcatcccgg aggatcgggtt atttattaca tgctgtcgaa cacgggagcg	300	
[0293]	gacgcgacgg aggtttttaa agagtttcat tatcggtcga aaaaggcgcg caaggcgttg	360	

[0294]	gcggcggttgc cgcataagcc agtggacgcg gcgacgcggg aaccgatcga agatgaggcg	420
[0295]	atgctgaagg atttcgcgca gtggcgcaag gaattggagc gtgagggatt ttttaagccc	480
[0296]	tcgccggcgc acgtggcgta tcgattcgcc gagctcgcg c gatgttcgc gctcggcacg	540
[0297]	gcgttgatgc acgcgcgttg gcacgtcgt tccgtgatcg tgtactcgtg tttcttcggc	600
[0298]	gcgcgatgcg gttgggtgca gcacgagggt gggcacaatt cgttgactgg aaacatttgg	660
[0299]	tgggacaagc gaatccaagc cttcgccgcg gggttcggct tggcgtcgag tggcgacatg	720
[0300]	tggacaaca tgcacaaca gcatcacgcg acgccccaaa aggtgcgaca cgatatggat	780
[0301]	ctcgacacca ctcccacggt ggcgttcttc aactccgcgg ttgaagaaaa tcgcccgcg	840
[0302]	ggattcagta agttgtggtt gcgccttcaa gcgtggacct tcgtgcccgt gacgtccggt	900
[0303]	atggttttgt tcttctggat gttcgtcttg caccgcgta acgcgctgcg acgcaaaagc	960
[0304]	ttcgaagaag cggttggat gttttccgcg cagtcattc gcacggcggt tatcaaagcc	1020
[0305]	gtcaccggt actcctggat cgctcgtac ggcttggtcg cggcgacgat gtgggcgagc	1080
[0306]	ggatgttact tgttcgcgca cttttccacg tctcacacgc acttgatgt cgtgccgagc	1140
[0307]	gataaacacc tctcgtgggt gcgatacgcc gtcgatcaca cgatcgacat caatccgaac	1200
[0308]	aacagcgtcg tcaactggtt gatgggttac ttgaactgcc aagtcacca tcacctgttc	1260
[0309]	ccgatatgc ctacgttcg ccaaccgaa gtctcccgcc gattcgtccc gtttgcaag	1320
[0310]	aagtggaact taaactacaa ggtcttgacg tattatgggg cctggaagc gacgttcggc	1380
[0311]	aactgaacg acgtcgggaa gcactattac gtgcacggat ctcagcgcgt caaatcaaag	1440
[0312]	tcggcgtga	1449
[0313]	<210>	10
[0314]	<211>	482
[0315]	<212>	PRT
[0316]	<213>	Ostreococcus lucimarinus
[0317]	<400>	10
[0318]	Met Cys Val Glu Thr Thr Glu Gly Thr Ser Arg Thr Met Ala Asn Glu	
[0319]	1 5 10 15	
[0320]	Arg Thr Ser Ser Ser Ser Leu Ser Glu Gly Gly Thr Pro Thr Val	
[0321]	20 25 30	
[0322]	Thr Val Gly Met Gly Ser Glu Asp Ala Gly Lys Lys Thr Arg Asn Ala	
[0323]	35 40 45	
[0324]	Ser Val Thr Ala Trp Thr Lys Glu Leu Glu Pro His Ala Ile Ala Lys	
[0325]	50 55 60	
[0326]	Thr Phe Glu Arg Arg Tyr Val Thr Ile Glu Gly Val Glu Tyr Asp Val	
[0327]	65 70 75 80	
[0328]	Thr Asp Phe Lys His Pro Gly Gly Ser Val Ile Tyr Tyr Met Leu Ser	
[0329]	85 90 95	
[0330]	Asn Thr Gly Ala Asp Ala Thr Glu Ala Phe Lys Glu Phe His Tyr Arg	
[0331]	100 105 110	
[0332]	Ser Lys Lys Ala Arg Lys Ala Leu Ala Ala Leu Pro His Lys Pro Val	
[0333]	115 120 125	
[0334]	Asp Ala Ala Thr Arg Glu Pro Ile Glu Asp Glu Ala Met Leu Lys Asp	
[0335]	130 135 140	

[0336]	Phe	Ala	Gln	Trp	Arg	Lys	Glu	Leu	Glu	Arg	Glu	Gly	Phe	Phe	Lys	Pro
[0337]	145					150					155					160
[0338]	Ser	Pro	Ala	His	Val	Ala	Tyr	Arg	Phe	Ala	Glu	Leu	Ala	Ala	Met	Phe
[0339]					165					170						175
[0340]	Ala	Leu	Gly	Thr	Ala	Leu	Met	His	Ala	Arg	Trp	His	Val	Ala	Ser	Val
[0341]				180					185					190		
[0342]	Ile	Val	Tyr	Ser	Cys	Phe	Phe	Gly	Ala	Arg	Cys	Gly	Trp	Val	Gln	His
[0343]			195					200					205			
[0344]	Glu	Gly	Gly	His	Asn	Ser	Leu	Thr	Gly	Asn	Ile	Trp	Trp	Asp	Lys	Arg
[0345]	210					215						220				
[0346]	Ile	Gln	Ala	Phe	Ala	Ala	Gly	Phe	Gly	Leu	Ala	Ser	Ser	Gly	Asp	Met
[0347]	225					230						235				240
[0348]	Trp	Asn	Asn	Met	His	Asn	Lys	His	His	Ala	Thr	Pro	Gln	Lys	Val	Arg
[0349]				245						250						255
[0350]	His	Asp	Met	Asp	Leu	Asp	Thr	Thr	Pro	Thr	Val	Ala	Phe	Phe	Asn	Ser
[0351]				260						265						270
[0352]	Ala	Val	Glu	Glu	Asn	Arg	Pro	Arg	Gly	Phe	Ser	Lys	Leu	Trp	Leu	Arg
[0353]			275					280					285			
[0354]	Leu	Gln	Ala	Trp	Thr	Phe	Val	Pro	Val	Thr	Ser	Gly	Met	Val	Leu	Phe
[0355]	290						295					300				
[0356]	Phe	Trp	Met	Phe	Val	Leu	His	Pro	Arg	Asn	Ala	Leu	Arg	Arg	Lys	Ser
[0357]	305					310					315					320
[0358]	Phe	Glu	Glu	Ala	Ala	Trp	Met	Phe	Ser	Ala	His	Val	Ile	Arg	Thr	Ala
[0359]				325						330					335	
[0360]	Val	Ile	Lys	Ala	Val	Thr	Gly	Tyr	Ser	Trp	Ile	Ala	Ser	Tyr	Gly	Leu
[0361]			340						345					350		
[0362]	Phe	Ala	Ala	Thr	Met	Trp	Ala	Ser	Gly	Cys	Tyr	Leu	Phe	Ala	His	Phe
[0363]			355						360				365			
[0364]	Ser	Thr	Ser	His	Thr	His	Leu	Asp	Val	Val	Pro	Ser	Asp	Lys	His	Leu
[0365]	370						375					380				
[0366]	Ser	Trp	Val	Arg	Tyr	Ala	Val	Asp	His	Thr	Ile	Asp	Ile	Asn	Pro	Asn
[0367]	385					390					395					400
[0368]	Asn	Ser	Val	Val	Asn	Trp	Leu	Met	Gly	Tyr	Leu	Asn	Cys	Gln	Val	Ile
[0369]				405						410					415	
[0370]	His	His	Leu	Phe	Pro	Asp	Met	Pro	Gln	Phe	Arg	Gln	Pro	Glu	Val	Ser
[0371]				420						425					430	
[0372]	Arg	Arg	Phe	Val	Pro	Phe	Ala	Lys	Lys	Trp	Asn	Leu	Asn	Tyr	Lys	Val
[0373]			435						440					445		
[0374]	Leu	Thr	Tyr	Tyr	Gly	Ala	Trp	Lys	Ala	Thr	Phe	Gly	Asn	Leu	Asn	Asp
[0375]	450						455					460				
[0376]	Val	Gly	Lys	His	Tyr	Tyr	Val	His	Gly	Ser	Gln	Arg	Val	Lys	Ser	Lys
[0377]	465						470				475					480

[0378] Ser Ala
 [0379] <210> 11
 [0380] <211> 1449
 [0381] <212> DNA
 [0382] <213> Artificial Sequence
 [0383] <220>
 [0384] <223> Codon-optimized nucleotide sequence for production in plants of
 [0385] SEQ ID NO:10
 [0386] <400> 11
 [0387] atgtgtgttg agactactga gggaacctct agaactatgg ctaacgagag gacctcttct 60
 [0388] tcttcttcac tctctgaggg tggaactcct actgttactg tgggaatggg atctgaggat 120
 [0389] gctggaaaga aaaccagaaa cgcttctgtt actgcttgga ccaaagagct tgagcctcac 180
 [0390] gctatcgcta agaccttga gagaagatac gttaccatcg aggggtgttg gtacgatgtg 240
 [0391] accgatttca aacaccctgg tggatctgtg atctactaca tgctctctaa cactgggtgct 300
 [0392] gatgctactg aggctttcaa agagttccac taccgttcta agaaggctag aaaggctctt 360
 [0393] gctgctcttc ctcaacaagg tgttgatgct gctactagag agcctattga ggacgagget 420
 [0394] atgcttaagg atttcgctca gtggagaaaa gagttggaga gagagggtt cttcaagcct 480
 [0395] tctcctgctc atgttgctta ccgtttcgtc gaactcgtg ctatgttcgc tcttgggaacc 540
 [0396] gctcttatgc atgctagatg gcacgttgct agcgttatcg tgtactcctg tttcttcgga 600
 [0397] gctagatgtg gatgggttca acatgagggt ggacacaact ctcttaccgg aaacatctgg 660
 [0398] tgggataaga gaatccaagc tttcgtgctc ggattcggac ttgcttcttc tggtgacatg 720
 [0399] tggaacaaca tgcacaacaa gcacatgct actcctcaga aagtgagaca cgatatggat 780
 [0400] cttgatacca cccctaccgt tgctttcttc aactctgctg tggaggaaaa cagacctagg 840
 [0401] ggattctcta agctttggct cagacttcaa gcttggacct tcgttctgt tacctctgga 900
 [0402] atgggtgctc tcttctggat gttcgttctc catcctagaa acgctctccg tcgtaagtct 960
 [0403] ttcgaagagg ctgcttgat gttctctgct cacgttatca gaaccgctg tatcaaggct 1020
 [0404] gttaccgat actcttgat cgctagctac ggacttttcg ctgctactat gtgggcttct 1080
 [0405] ggatgctacc ttttcgctca cttctctact tctcacaccc acctcgatgt tgttccatct 1140
 [0406] gataagcacc ttagctgggt taggtacgct gttgatcaca ccatcgacat caaccctaac 1200
 [0407] aactctgttg tgaactggt tatgggatac cttaactgcc aggttatcca ccatctcttc 1260
 [0408] cctgatatgc ctcaattcag acagcctgag gtgtcaagaa gattcgtccc tttcgctaag 1320
 [0409] aagtgaacc tcaactacaa ggtgctcact tactacgggt cttggaaggc tactttcgga 1380
 [0410] aacctcaacg atgttggaag gcactactac gttcacggat ctgagagagt gaagagcaag 1440
 [0411] agcgcttga 1449
 [0412] <210> 12
 [0413] <211> 1329
 [0414] <212> DNA
 [0415] <213> Pyramimonas CS-0140
 [0416] <400> 12
 [0417] atgggaaagg gaggcaatgc tagcgtcct actgcgaaga aggaggtgtt gatcgagggg 60
 [0418] aagttttacg atgtcaccga cttcaggcac cccggtggtt cgatcatcaa gtttctctcg 120
 [0419] ggttctggtg ctgacgccac cgcttctac cgcgagttcc acgttaggtc agcgaaggca 180

[0420]	gacaagttct tgaagacgct gccctcccg	gaagccactc cccaggagct gaagcaggcg	240
[0421]	gttgagttct ccaagctcaa cccgccctcc	gcggagagtg cctctgctcc cctgaccgac	300
[0422]	cttgccaagg tggaagcgt gaacaaggac	ttcgaggctt tccgtgagca gctcattcag	360
[0423]	gagggttct ttaagcccaa tatcccgcat	gtggtcaagc gcatcacgga agtcgtggcg	420
[0424]	atgatggccg tagcctcctg gatgatggtg	cagaccaacg ctcttgttgt gaccctcgga	480
[0425]	gttctgatcc gcggcattgc acagggccgg	tgcggttggc ttatgcacga gggcgccac	540
[0426]	tatagtctta ctgggaagat ctccattgat	aggcgtctgc aggagtcaat ttacggattc	600
[0427]	ggctgtggaa tgtccggcgc ctggtggcgc	aaccagcaca acaagcacca cgcaacccca	660
[0428]	cagaagctgc agcatgacgt cgacctggag	acccttcctc tgatggcttt caacaacgct	720
[0429]	gttaccgata gacgcaaggt gaagcctggt	agtctccagg ctctgtggct caagtaccag	780
[0430]	gccttcctct tcttccccgt gacctcctt	ctggtcggcc tcggttggac caccgtcctc	840
[0431]	caccccaggc acagcttgcg caccaagcac	tatttcgagc tgctctgcat ggctgctcgt	900
[0432]	tacgcgagtt tcgctgctct tttcgctccc	aagtacggac ttgcaggagc tgccgggctc	960
[0433]	tacctcgcca cttcgctgt cgggtgcaac	tatattttca tcaacttctc ggtctctcac	1020
[0434]	actcacctgc ccgtgagcgg tgcgagcgag	tacctgcatt gggctcgtga ttcggccatc	1080
[0435]	cacaccacta acatcaaate cagcatgctg	tgcgattggt ggatgtcatt cctcaacttc	1140
[0436]	cagatcgagc atcacctgtt ccttcaatg	ccccagttcc gccacaagat tatctccccg	1200
[0437]	cgtgtaaagg cttgtttga gaagcacggt	cttgtgtatg atgtgcgccc ctattggggg	1260
[0438]	gccatggctg acaccttcaa gaacttgaat	gacgttggca ctcacgcatc tcaactcaag	1320
[0439]	gcgcactag		1329
[0440]	<210>	13	
[0441]	<211>	442	
[0442]	<212>	PRT	
[0443]	<213>	Pyramimonas CS-0140	
[0444]	<400>	13	
[0445]	Met Gly Lys Gly Gly Asn Ala Ser Ala Pro Thr Ala Lys Lys Glu Val		
[0446]	1 5 10 15		
[0447]	Leu Ile Glu Gly Lys Phe Tyr Asp Val Thr Asp Phe Arg His Pro Gly		
[0448]	20 25 30		
[0449]	Gly Ser Ile Ile Lys Phe Leu Ser Gly Ser Gly Ala Asp Ala Thr Ala		
[0450]	35 40 45		
[0451]	Ser Tyr Arg Glu Phe His Val Arg Ser Ala Lys Ala Asp Lys Phe Leu		
[0452]	50 55 60		
[0453]	Lys Thr Leu Pro Ser Arg Glu Ala Thr Pro Gln Glu Leu Lys Gln Ala		
[0454]	65 70 75 80		
[0455]	Val Glu Phe Ser Lys Leu Asn Pro Pro Ser Ala Glu Ser Ala Ser Ala		
[0456]	85 90 95		
[0457]	Pro Leu Thr Asp Leu Ala Lys Val Glu Ala Leu Asn Lys Asp Phe Glu		
[0458]	100 105 110		
[0459]	Ala Phe Arg Glu Gln Leu Ile Gln Glu Gly Phe Phe Lys Pro Asn Ile		
[0460]	115 120 125		
[0461]	Pro His Val Val Lys Arg Ile Thr Glu Val Val Ala Met Met Ala Val		

[0462]	130	135	140
[0463]	Ala Ser Trp Met Met Val Gln Thr Asn Ala Leu Val Val Thr Leu Gly		
[0464]	145	150	155
[0465]	Val Leu Ile Arg Gly Ile Ala Gln Gly Arg Cys Gly Trp Leu Met His		
[0466]	165	170	175
[0467]	Glu Gly Gly His Tyr Ser Leu Thr Gly Lys Ile Ser Ile Asp Arg Arg		
[0468]	180	185	190
[0469]	Leu Gln Glu Ser Ile Tyr Gly Phe Gly Cys Gly Met Ser Gly Ala Trp		
[0470]	195	200	205
[0471]	Trp Arg Asn Gln His Asn Lys His His Ala Thr Pro Gln Lys Leu Gln		
[0472]	210	215	220
[0473]	His Asp Val Asp Leu Glu Thr Leu Pro Leu Met Ala Phe Asn Asn Ala		
[0474]	225	230	235
[0475]	Val Thr Asp Arg Arg Lys Val Lys Pro Gly Ser Leu Gln Ala Leu Trp		
[0476]	245	250	255
[0477]	Leu Lys Tyr Gln Ala Phe Leu Phe Phe Pro Val Thr Ser Leu Leu Val		
[0478]	260	265	270
[0479]	Gly Leu Gly Trp Thr Thr Val Leu His Pro Arg His Ser Leu Arg Thr		
[0480]	275	280	285
[0481]	Lys His Tyr Phe Glu Leu Leu Cys Met Ala Ala Arg Tyr Ala Ser Phe		
[0482]	290	295	300
[0483]	Ala Ala Leu Phe Ala Pro Lys Tyr Gly Leu Ala Gly Ala Ala Gly Leu		
[0484]	305	310	315
[0485]	Tyr Leu Ala Thr Phe Ala Val Gly Cys Asn Tyr Ile Phe Ile Asn Phe		
[0486]	325	330	335
[0487]	Ser Val Ser His Thr His Leu Pro Val Ser Gly Ala Ser Glu Tyr Leu		
[0488]	340	345	350
[0489]	His Trp Val Val Tyr Ser Ala Ile His Thr Thr Asn Ile Lys Ser Ser		
[0490]	355	360	365
[0491]	Met Leu Cys Asp Trp Trp Met Ser Phe Leu Asn Phe Gln Ile Glu His		
[0492]	370	375	380
[0493]	His Leu Phe Pro Ser Met Pro Gln Phe Arg His Lys Ile Ile Ser Pro		
[0494]	385	390	395
[0495]	Arg Val Lys Ala Leu Phe Glu Lys His Gly Leu Val Tyr Asp Val Arg		
[0496]	405	410	415
[0497]	Pro Tyr Trp Gly Ala Met Ala Asp Thr Phe Lys Asn Leu Asn Asp Val		
[0498]	420	425	430
[0499]	Gly Thr His Ala Ser His Ser Lys Ala His		
[0500]	435	440	
[0501]	<210> 14		
[0502]	<211> 528		
[0503]	<212> DNA		

[0504] <213> Micromonas CS-0170
 [0505] <400> 14
 [0506] tgtggtgcat cggccatgat tgggggcacg gatccttctc caaaaacgcc ctggtcaacg 60
 [0507] gcgtcgtcgg acacctcacg cactcgccgc tgctcgtgcc cttttaccgc tgggcgtact 120
 [0508] cgcacaagca gcaccacagg ttccacaacc acgaggtacg ggacatgtcg caccctgtga 180
 [0509] tgtccaagga ggagtacgcg gacgtgaacc ccgtcgtgag ggcgctggcg ctggacggtt 240
 [0510] ggtggggaac cttctcgcgc ttcccccggg acctctctct ggaaccgcag tgggccggga 300
 [0511] cggacgggtg ccacttcaac ccgaactcca ggctcttcga cagggcgccc aaggatgagc 360
 [0512] ggggtgaagtg cgcggtttcc accgtggcgt gcgcggcggt tctcggcgcg tccttcgtcg 420
 [0513] cgtgcgactc gaaccccggt cactggttcg ctacgtactt ggcgccgtac ttgtgcttct 480
 [0514] cgtggtggct cttaccgtg acatacttac agcacaacga cgaggaca 528
 [0515] <210> 15
 [0516] <211> 175
 [0517] <212> PRT
 [0518] <213> Micromonas CS-0170
 [0519] <400> 15
 [0520] Trp Cys Ile Gly His Asp Trp Gly His Gly Ser Phe Ser Lys Asn Ala
 [0521] 1 5 10 15
 [0522] Leu Val Asn Gly Val Val Gly His Leu Thr His Ser Pro Leu Leu Val
 [0523] 20 25 30
 [0524] Pro Phe Tyr Pro Trp Ala Tyr Ser His Lys Gln His His Arg Phe His
 [0525] 35 40 45
 [0526] Asn His Glu Val Arg Asp Met Ser His Pro Trp Met Ser Lys Glu Glu
 [0527] 50 55 60
 [0528] Tyr Ala Asp Val Asn Pro Val Val Arg Ala Leu Ala Leu Asp Gly Trp
 [0529] 65 70 75 80
 [0530] Trp Gly Thr Phe Leu Gly Phe Pro Gly Tyr Leu Leu Leu Glu Pro Gln
 [0531] 85 90 95
 [0532] Trp Ala Gly Thr Asp Gly Cys His Phe Asn Pro Asn Ser Arg Leu Phe
 [0533] 100 105 110
 [0534] Asp Arg Ala Pro Lys Asp Glu Arg Val Lys Cys Ala Val Ser Thr Val
 [0535] 115 120 125
 [0536] Ala Cys Ala Ala Phe Leu Gly Ala Ser Phe Val Ala Cys Asp Ser Asn
 [0537] 130 135 140
 [0538] Pro Val His Trp Phe Ala Gln Tyr Leu Ala Pro Tyr Leu Cys Phe Ser
 [0539] 145 150 155 160
 [0540] Trp Trp Leu Phe Thr Val Thr Tyr Leu Gln His Asn Asp Glu Asp
 [0541] 165 170 175
 [0542] <210> 16
 [0543] <211> 1461
 [0544] <212> DNA
 [0545] <213> Micromonas RCC299

[0546]	<400> 16	
[0547]	atgaccgccg tcggcggcgg gcgtcgacac tcccgcgccg cgcgtcccga gcggactcga	60
[0548]	gaccgtcgga agcgtctgtc ggattccaca gactcggatt ccgtcgtctc ccggacaccc	120
[0549]	ggcgacgtac ctcccgccgc gctcgttcac gcgatccccg cgatgatgac ctcatccacc	180
[0550]	tgctctcggg ccacgtcgcc ccggatggaa tcgacacgcc gcggctccct ccgcgtctgc	240
[0551]	gacgccgcgg cgcccgcacg caccggcgca tccaagacgg gtcggtccgt cccgcgtcaa	300
[0552]	ctcccccgcg ccatttcgcg gccgacgacg aacctcggaa ccgcggacgc gccaaactccc	360
[0553]	gccgccggtc ccgtgccgtc cctccgagag ctgcgcgcac ccatcccga ggagtgttc	420
[0554]	gagcccgacc tgggggagtc gctcaagtac gccgcgtacg atctcgccgc gctcgccgcg	480
[0555]	tgcttcggcg tcatctcccc gcacgtggtg gaccaccgt ggctgttacc gctgtacgcg	540
[0556]	ccctcacgg gcaccgtcat gtggatgaac ttgctcgtcg gccacgactg cgggcacggc	600
[0557]	tccttctcca aatctccgt cgtcaacggc gtcgtcggcc acctcacga ctgcgcgtg	660
[0558]	ctcgtgccct tctaccgtg ggcgtactcg cacaagcagc accacaggt tcacaaccac	720
[0559]	caggtaaagg acatgtcgca cccgtggatg actgcagagg agtacgccga ggtgaacccc	780
[0560]	atcgtccggg ctctggcgct tgacgggtgg tggggaacct tcctcggctt tcccggtac	840
[0561]	ctctctctgg aaccgcgtg gcccgggacg gacgggtgcc acttcaaccc acaatccagg	900
[0562]	ctcttcgaca gggcgcccaa ggacgagcgg gtgaagtgcg cggtttccac cgtggcgtgc	960
[0563]	gcggcgtttc tcggcgcgtc ctctcgtcgc tgcgactcga acccggcgca ctggttcgct	1020
[0564]	cagtacttgg cgccgtactt gtgcttctcg tgggtgctct tcaccgtgac ctacctccag	1080
[0565]	caccacgact acgacacaac gacgtacgaa gagggccagt gggagtacgt gctggggggt	1140
[0566]	ttggagacga tcgaccgga gtttgggcac ggcgtggacg agctgacga tcacatcacc	1200
[0567]	gactgccacg tcgcgcacca catgttctcg gacatgcccc actaccgcct gcccgcgcg	1260
[0568]	accgccgggg tgcgatcggg gctggagccc cgtgggtgtg acaagcgacg ggacacgcgg	1320
[0569]	gatttcgtga cgaagtggt cgagttgcac ggggacgtcg ggcactgcgt ggagagcgag	1380
[0570]	gcgggggtga ggccgagggc gacgcgcgac gactgcgcgg cggcgttcaa cggcgaggag	1440
[0571]	tggcgcgagg ttcgcaagta g	1461
[0572]	<210> 17	
[0573]	<211> 486	
[0574]	<212> PRT	
[0575]	<213> Micromonas RCC299	
[0576]	<400> 17	
[0577]	Met Thr Ala Val Gly Gly Gly Arg Arg His Ser Arg Ala Ala Arg Pro	
[0578]	1 5 10 15	
[0579]	Glu Arg Thr Arg Asp Arg Arg Lys Arg Ser Ser Asp Ser Thr Glu Ser	
[0580]	20 25 30	
[0581]	Asp Ser Val Val Ser Arg Thr Pro Gly Asp Val Pro Pro Ala Ala Leu	
[0582]	35 40 45	
[0583]	Val His Ala Met Pro Ala Met Met Thr Ser Ser Thr Cys Leu Arg Ala	
[0584]	50 55 60	
[0585]	Thr Ser Pro Arg Met Glu Ser Thr Arg Arg Gly Ser Leu Arg Val Val	
[0586]	65 70 75 80	
[0587]	Asp Ala Ala Ala Pro Ala Arg Thr Gly Ala Ser Lys Thr Gly Arg Ser	

[0588]	85	90	95
[0589]	Val Pro Arg Gln Leu Pro Arg Ala Ile Ser Ala Pro Thr Thr Asn Leu		
[0590]	100	105	110
[0591]	Gly Thr Ala Asp Ala Pro Thr Pro Ala Ala Gly Pro Val Pro Ser Leu		
[0592]	115	120	125
[0593]	Arg Glu Leu Arg Ala Ser Ile Pro Lys Glu Cys Phe Glu Pro Asp Leu		
[0594]	130	135	140
[0595]	Gly Glu Ser Leu Lys Tyr Ala Ala Tyr Asp Leu Ala Ala Leu Ala Ala		
[0596]	145	150	155
[0597]	Cys Phe Gly Val Ile Ser Pro His Val Val Asp His Pro Trp Leu Leu		
[0598]	165	170	175
[0599]	Pro Leu Tyr Ala Pro Leu Thr Gly Thr Val Met Trp Met Asn Phe Val		
[0600]	180	185	190
[0601]	Val Gly His Asp Cys Gly His Gly Ser Phe Ser Lys Ser Ser Val Val		
[0602]	195	200	205
[0603]	Asn Gly Val Val Gly His Leu Thr His Ser Pro Leu Leu Val Pro Phe		
[0604]	210	215	220
[0605]	Tyr Pro Trp Ala Tyr Ser His Lys Gln His His Arg Phe His Asn His		
[0606]	225	230	235
[0607]	Gln Val Lys Asp Met Ser His Pro Trp Met Thr Ala Glu Glu Tyr Ala		
[0608]	245	250	255
[0609]	Glu Val Asn Pro Ile Val Arg Ala Leu Ala Leu Asp Gly Trp Trp Gly		
[0610]	260	265	270
[0611]	Thr Phe Leu Gly Phe Pro Gly Tyr Leu Leu Leu Glu Pro Arg Trp Ala		
[0612]	275	280	285
[0613]	Gly Thr Asp Gly Cys His Phe Asn Pro Gln Ser Arg Leu Phe Asp Arg		
[0614]	290	295	300
[0615]	Ala Pro Lys Asp Glu Arg Val Lys Cys Ala Val Ser Thr Val Ala Cys		
[0616]	305	310	315
[0617]	Ala Ala Phe Leu Gly Ala Ser Phe Val Ala Cys Asp Ser Asn Pro Ala		
[0618]	325	330	335
[0619]	His Trp Phe Ala Gln Tyr Leu Ala Pro Tyr Leu Cys Phe Ser Trp Trp		
[0620]	340	345	350
[0621]	Leu Phe Thr Val Thr Tyr Leu Gln His His Asp Tyr Asp Thr Thr Thr		
[0622]	355	360	365
[0623]	Tyr Glu Glu Gly Gln Trp Glu Tyr Val Leu Gly Gly Leu Glu Thr Ile		
[0624]	370	375	380
[0625]	Asp Arg Glu Phe Gly His Gly Val Asp Glu Leu Thr His His Ile Thr		
[0626]	385	390	395
[0627]	Asp Cys His Val Ala His His Met Phe Ser Asp Met Pro His Tyr Arg		
[0628]	405	410	415
[0629]	Leu Pro Ala Ala Thr Ala Gly Val Arg Ser Val Leu Glu Pro Arg Gly		

[0630]	420	425	430
[0631]	Leu Tyr Lys Arg Arg Asp Thr Arg Asp Phe Val Thr Lys Val Phe Glu		
[0632]	435	440	445
[0633]	Leu His Gly Asp Val Gly His Cys Val Glu Ser Glu Ala Gly Val Arg		
[0634]	450	455	460
[0635]	Pro Arg Ala Thr Arg Asp Glu Cys Ala Ala Ala Phe Asn Gly Glu Glu		
[0636]	465	470	475
[0637]	Trp Arg Glu Val Arg Lys		480
[0638]	485		
[0639]	<210> 18		
[0640]	<211> 1461		
[0641]	<212> DNA		
[0642]	<213> Artificial Sequence		
[0643]	<220>		
[0644]	<223> Codon-optimized nucleotide sequence for production in plants of		
[0645]	SEQ ID NO:17		
[0646]	<400> 18		
[0647]	atgactgctg ttggaggtgg aagaaggcat tctagagctg ctagacctga gagaaccagg	60	
[0648]	gatcgtagaa agagatcttc tgactctacc gagtctgatt ctgtggtttc taggaccctc	120	
[0649]	ggtgatgttc ctctgctgc tcttggtcat gctatgcctg ctatgatgac ctcttctacc	180	
[0650]	tgccttagag ctactctcc taggatggaa tctaccagaa ggggatctct cagagttgtt	240	
[0651]	gacgctgctg ctccagctag aactggtgct tctaagaccg gaagatctgt tcctagacag	300	
[0652]	ctccctagag ctatctctgc tcctactact aaccttggaa ccgctgatgc tcctactcct	360	
[0653]	gctgctggac ctgttccttc tcttagagag cttagggcta gcatccctaa agagtgttc	420	
[0654]	gagcctgac ttggagagtc tctcaagtac gctgcttacg atcttgctgc tctgctgct	480	
[0655]	tgtttcggag ttatctctcc tcacgttgtt gatcatcctt ggcttctccc actttacgct	540	
[0656]	cctctaccg gaactgtgat gtggatgaac ttcgttgtt gacatgattg cggacacgga	600	
[0657]	tctttctcta agtctagcgt ggtaaaccgt gttgttgac acctacca ctctcctct	660	
[0658]	ctcgttctt tctacccttg ggcttactct cataagcagc accacagatt ccataaccac	720	
[0659]	caggtgaagg atatgtctca cccttggatg actgctgagg aatacgtga ggtaaccct	780	
[0660]	attgtgaggg ctcttgctct tgatggatgg tggggaactt tccttggatt ccctggatac	840	
[0661]	cttctcttg agcctagatg ggctggaact gatggatgac acttcaacc tcagtctaga	900	
[0662]	cttttcgaca gggtccctaa ggatgaacgt gttaagtgcg ctgtttctac tgttgcttgc	960	
[0663]	gctgctttcc ttggagcttc tttcgttgct tgcgattcta acctgctca ttggttcgct	1020	
[0664]	cagtacctg ctcttacct ttgcttctct tgggtgctct tctactgtac ttacctcag	1080	
[0665]	caccacgatt acgataccac tacctacgag gaaggacagt gggaatacgt tcttgagga	1140	
[0666]	ctcgagacta tcgatagaga atttggacac ggtgttgatg agcttactca ccacatcact	1200	
[0667]	gattgtcacg tggtcacca catgttctct gacatgcctc attacagact tccagctgct	1260	
[0668]	actgctggtg ttagatctgt tctcgacct agaggactct acaagagaag agacaccagg	1320	
[0669]	gatttcgtta ccaaggtgtt cgagcttcat ggtgatgtt gacactgcgt tgagtctgaa	1380	
[0670]	gctggtgtta ggcctagagc tactagagat gagtgcgctg ctgcttttaa tggatgaag	1440	
[0671]	tggaggaag ttaggaagtg a	1461	

[0672]	<210> 19
[0673]	<211> 1389
[0674]	<212> DNA
[0675]	<213> Micromonas CCMP1545
[0676]	<400> 19
[0677]	atgacggcgg cgacgacgct ctgcgcgcgc tcctccgtcg tcggcgctcg cgcaggcgcc 60
[0678]	acgtcgagge gtccggtcga tcgcgcgcgc tcgcgcgcgc cgacgtctc ctctatttcg 120
[0679]	cgcgcgcgc gcgcgcgcgc gaccgtccgc gcggcgctcca gccgcacgic gtctctccgc 180
[0680]	acgacgacga cgacgacgac gacgcgctcc gctctggcca tggagacccc gtccacgacg 240
[0681]	acgacggacg cggacgacga cgacgacaag accggcgctc tcccgtcgct ccgagagctc 300
[0682]	cgagccgcga tccccgcgga gtgcttcgic ccgtgcctga gggagtcgat gaagtacgcc 360
[0683]	gcgatcgacc tcggtctcct cgcgcgcgc tcggtctgt ggtctccgct cgtggtggat 420
[0684]	aaccgtgga tgctccgct gtacgcccc gtcaccggga cgatcatgtg gatgtgcttc 480
[0685]	gtcgtcgggc acgactgcgc gcacggcagt ttcaggtgcg tcagttctac aaacgtcgag 540
[0686]	aacaagtgga tcaacggcgt cgtcggtcac gcgacgcact cgcgcgctgt cgtgccgttc 600
[0687]	taccgtggg cgtactcgca caagcagcac caccggttc ataaccacga ggagaaggac 660
[0688]	atgtcgcacc cgtggatgag cgcggagcgc tacaaggaca cgaacgcgat cgttcggtgc 720
[0689]	gcggcgggcg gtcggtttc tgggttttg ttcctcgcgc tcgatcacc atggggcgcg 780
[0690]	tttctcggt tccccggta tctctctc gagccaaagt ggagcagcac ggacggctcg 840
[0691]	cacttcaacc cggccggtc tctgttcgac cgcgccccca aggacgaacg tctcaagtgc 900
[0692]	gcggtctcga cggtcgcgtg cgcggcgctc ctctcgcga cgttcgccgc gtgcgacggc 960
[0693]	cccgcgcagt gggcgacgca gtacctcgc ccgtatctgt gcttctcgtg gtggctcttc 1020
[0694]	acgtgacct acctccagca ccacgaccac gacacgaaga cgtacaagga aggcgagtgg 1080
[0695]	gagtacgtct tggggggatt agagacgac gatcgcgagt tcgggtacgg cgtggacgag 1140
[0696]	gcgacgcac acatcacgga ctgccacgic gcgcaccaca tgttttcgga catgccgcac 1200
[0697]	tacaacttgg agaaggccac cgcgggggtg cggggcgctc tcgagcccag gggactgtat 1260
[0698]	aagaagaggg acacgaggga tttcgcgacg aagattttg ggttgacaaa agacgtcggg 1320
[0699]	cactgcgtcg agccggaccg gccgcgggcg acgaaggagg agctcgtgaa ggcgttgggg 1380
[0700]	cgggagtga 1389
[0701]	<210> 20
[0702]	<211> 462
[0703]	<212> PRT
[0704]	<213> Micromonas CCMP1545
[0705]	<400> 20
[0706]	Met Thr Ala Ala Thr Thr Leu Ser Arg Ala Ser Ser Val Val Gly Val
[0707]	1 5 10 15
[0708]	Gly Ala Gly Ala Thr Ser Arg Arg Pro Val Asp Arg Ala Val Ser Ala
[0709]	20 25 30
[0710]	Val Ala Thr Ser Ser Ser Tyr Ser Arg Arg Ala Arg Arg Val Ala Thr
[0711]	35 40 45
[0712]	Val Arg Ala Ala Ser Ser Arg Thr Ser Ser Ser Ala Thr Thr Thr Thr
[0713]	50 55 60

[0714]	Thr Thr Thr Thr Arg Ser Ala Leu Ala Met Glu Thr Pro Ser Thr Thr		
[0715]	65	70	75 80
[0716]	Thr Thr Asp Ala Asp Asp Asp Asp Asp Lys Thr Gly Val Val Pro Ser		
[0717]		85 90	95
[0718]	Leu Arg Glu Leu Arg Ala Ala Ile Pro Ala Glu Cys Phe Val Pro Cys		
[0719]		100 105	110
[0720]	Leu Arg Glu Ser Met Lys Tyr Ala Ala Ile Asp Leu Gly Leu Leu Ala		
[0721]		115 120	125
[0722]	Ala Cys Phe Gly Leu Trp Ser Pro Leu Val Val Asp Asn Pro Trp Met		
[0723]		130 135	140
[0724]	Leu Pro Leu Tyr Ala Pro Val Thr Gly Thr Ile Met Trp Met Cys Phe		
[0725]		145 150	155 160
[0726]	Val Val Gly His Asp Cys Gly His Gly Ser Phe Arg Cys Val Ser Ser		
[0727]		165 170	175
[0728]	Thr Asn Val Glu Asn Lys Trp Ile Asn Gly Val Val Gly His Ala Thr		
[0729]		180 185	190
[0730]	His Ser Pro Leu Leu Val Pro Phe Tyr Pro Trp Ala Tyr Ser His Lys		
[0731]		195 200	205
[0732]	Gln His His Arg Phe His Asn His Glu Glu Lys Asp Met Ser His Pro		
[0733]		210 215	220
[0734]	Trp Met Ser Ala Glu Arg Tyr Lys Asp Thr Asn Ala Ile Val Arg Cys		
[0735]		225 230	235 240
[0736]	Ala Ala Gly Gly Ser Val Ser Gly Phe Trp Phe Leu Ala Leu Asp His		
[0737]		245 250	255
[0738]	Pro Trp Gly Ala Phe Leu Gly Phe Pro Gly Tyr Leu Leu Leu Glu Pro		
[0739]		260 265	270
[0740]	Lys Trp Ser Ser Thr Asp Gly Ser His Phe Asn Pro Ala Gly Arg Leu		
[0741]		275 280	285
[0742]	Phe Asp Arg Ala Pro Lys Asp Glu Arg Leu Lys Cys Ala Val Ser Thr		
[0743]		290 295	300
[0744]	Val Ala Cys Ala Ala Phe Leu Leu Ala Thr Phe Ala Ala Cys Asp Gly		
[0745]		305 310	315 320
[0746]	Pro Ala Gln Trp Ala Thr Gln Tyr Leu Ala Pro Tyr Leu Cys Phe Ser		
[0747]		325 330	335
[0748]	Trp Trp Leu Phe Thr Val Thr Tyr Leu Gln His His Asp His Asp Thr		
[0749]		340 345	350
[0750]	Lys Thr Tyr Lys Glu Gly Glu Trp Glu Tyr Val Leu Gly Gly Leu Glu		
[0751]		355 360	365
[0752]	Thr Ile Asp Arg Glu Phe Gly Tyr Gly Val Asp Glu Ala Thr His His		
[0753]		370 375	380
[0754]	Ile Thr Asp Cys His Val Ala His His Met Phe Ser Asp Met Pro His		
[0755]		385 390	395 400

[0756]	Tyr Asn Leu Glu Lys Ala Thr Ala Gly Val Arg Gly Val Leu Glu Pro	
[0757]	405	410 415
[0758]	Arg Gly Leu Tyr Lys Lys Arg Asp Thr Arg Asp Phe Ala Thr Lys Ile	
[0759]	420	425 430
[0760]	Phe Gly Leu His Lys Asp Val Gly His Cys Val Glu Pro Asp Arg Pro	
[0761]	435	440 445
[0762]	Arg Ala Thr Lys Glu Glu Leu Val Lys Ala Leu Gly Arg Glu	
[0763]	450	455 460
[0764]	<210> 21	
[0765]	<211> 792	
[0766]	<212> DNA	
[0767]	<213> Isochrysis galbana	
[0768]	<400> 21	
[0769]	atggccctcg caaacgacgc gggagagcgc atctgggcgg ctgtgaccga cccggaaatc	60
[0770]	ctcattggca ctttctcgta cttgctactc aaaccgctgc tccgcaattc cgggctgggtg	120
[0771]	gatgagaaga agggcgcata caggacgtcc atgatctggt acaacgttct gctggcgctc	180
[0772]	ttctctgcgc tgagcttcta cgtgacggcg accgccctcg gctgggacta tggtaggggc	240
[0773]	gcgtggctgc gcaggcaaac cggcgacaca ccgcagccgc tcttccagtg cccgtccccg	300
[0774]	gtttgggact cgaagctctt cacatggacc gccaaaggcat tctattactc caagtacgtg	360
[0775]	gagtacctcg acacggcctg gctgggtgctc aagggaaga gggctctcctt tctccaggcc	420
[0776]	ttccaccact ttggcgcgcc gtgggatgtg tacctcgga ttcggctgca caacgagggc	480
[0777]	gtatggatct tcatgttttt caactcgttc attcacacca tcatgtacac ctactacggc	540
[0778]	ctcaccgccg ccgggtataa gttcaaggcc aagccgctca tcaccgcgat gcagatctgc	600
[0779]	cagttcgtgg gcggcttctt gttggtctgg gactacatca acgtccctg cttcaactcg	660
[0780]	gacaaaggga agttgttcag ctgggcttcc aactatgcat acgtcggtc ggtcttcttg	720
[0781]	ctctctgcc acttttctta ccaggacaac ttggcaacga agaaatcggc caaggcgggc	780
[0782]	aagcagctct ag	792
[0783]	<210> 22	
[0784]	<211> 263	
[0785]	<212> PRT	
[0786]	<213> Isochrysis galbana	
[0787]	<400> 22	
[0788]	Met Ala Leu Ala Asn Asp Ala Gly Glu Arg Ile Trp Ala Ala Val Thr	
[0789]	1	5 10 15
[0790]	Asp Pro Glu Ile Leu Ile Gly Thr Phe Ser Tyr Leu Leu Leu Lys Pro	
[0791]	20	25 30
[0792]	Leu Leu Arg Asn Ser Gly Leu Val Asp Glu Lys Lys Gly Ala Tyr Arg	
[0793]	35	40 45
[0794]	Thr Ser Met Ile Trp Tyr Asn Val Leu Leu Ala Leu Phe Ser Ala Leu	
[0795]	50	55 60
[0796]	Ser Phe Tyr Val Thr Ala Thr Ala Leu Gly Trp Asp Tyr Gly Thr Gly	
[0797]	65	70 75 80

[0798]	Ala Trp Leu Arg Arg Gln Thr Gly Asp Thr Pro Gln Pro Leu Phe Gln	
[0799]	85	90 95
[0800]	Cys Pro Ser Pro Val Trp Asp Ser Lys Leu Phe Thr Trp Thr Ala Lys	
[0801]	100	105 110
[0802]	Ala Phe Tyr Tyr Ser Lys Tyr Val Glu Tyr Leu Asp Thr Ala Trp Leu	
[0803]	115	120 125
[0804]	Val Leu Lys Gly Lys Arg Val Ser Phe Leu Gln Ala Phe His His Phe	
[0805]	130	135 140
[0806]	Gly Ala Pro Trp Asp Val Tyr Leu Gly Ile Arg Leu His Asn Glu Gly	
[0807]	145	150 155 160
[0808]	Val Trp Ile Phe Met Phe Phe Asn Ser Phe Ile His Thr Ile Met Tyr	
[0809]	165	170 175
[0810]	Thr Tyr Tyr Gly Leu Thr Ala Ala Gly Tyr Lys Phe Lys Ala Lys Pro	
[0811]	180	185 190
[0812]	Leu Ile Thr Ala Met Gln Ile Cys Gln Phe Val Gly Gly Phe Leu Leu	
[0813]	195	200 205
[0814]	Val Trp Asp Tyr Ile Asn Val Pro Cys Phe Asn Ser Asp Lys Gly Lys	
[0815]	210	215 220
[0816]	Leu Phe Ser Trp Ala Phe Asn Tyr Ala Tyr Val Gly Ser Val Phe Leu	
[0817]	225	230 235 240
[0818]	Leu Phe Cys His Phe Phe Tyr Gln Asp Asn Leu Ala Thr Lys Lys Ser	
[0819]	245	250 255
[0820]	Ala Lys Ala Gly Lys Gln Leu	
[0821]	260	
[0822]	<210> 23	
[0823]	<211> 1284	
[0824]	<212> DNA	
[0825]	<213> Pavlova salina	
[0826]	<400> 23	
[0827]	atgggacgcg gcggagacag cagtgggcag gcgcatccgg cggcggagct ggcgggtcccg	60
[0828]	agcgaccgcg cggaggtgag caacgctgac agcaaagcgc tgcacatcgt gctgtatggc	120
[0829]	aagcgcgtgg atgtgaccaa gttccaacgc acgcacccgg gtggtagcaa ggtcttccgg	180
[0830]	atcttccagg accgcgatgc gacggagcag ttcgagtcct accactcgaa gcgcgcgatc	240
[0831]	aagatgatgg agggcatgct caagaagtct gaggatgctc ccgccgacac gcccttgccc	300
[0832]	tcccagtcac cgatggggaa ggacttcaag gcgatgatcg agcggcacgt tgcagcgggt	360
[0833]	tactacgatc catgccgct cgatgagctg ttcaagctca gcctcgtgct cctcccgacc	420
[0834]	tttgcgggca tgtacatgct caaggcgggc gtcggctccc cgctctgcgg cgccctcatg	480
[0835]	gtgagctttg gctggtacct cgatggctgg ctgcgcacg actatctgca ccaactccgtc	540
[0836]	ttcaaggggt ccgtcgcacg caccgtcggg tggaacaacg cggcgggcta cttcctcggc	600
[0837]	ttcgtgcagg ggtatgcggt cgagtggctg cgcgcgcggc ataacacgca ccacgtgtgc	660
[0838]	accaatgagg acggctcgga ccccgacatc aaaacggcgc cgctgctcat atacgtgcgc	720
[0839]	aacaagccga gcatcgccaa gcgcctgaac gccttcacg gctaccagca gtactactat	780

[0840]	gtgccggtga tggcaatcct cgacctgtac tggcggctcg agtcgatcgc ctacgtcgcg	840
[0841]	atgcgcctgc cgaagatgct gccgcaggcc ctgcactcg tcgcgacta cgccatcgtc	900
[0842]	gcgtgggtct ttgcgggcaa ctaccactg ctcccgtcg tgacggttct gcgcgggttt	960
[0843]	ggcactggga tcaccgtttt cgcgacgcac tacggtgagg acattctcga cgcggaccag	1020
[0844]	gtgcgtcaca tgacgtcgt cgagcagacg gcactcacct cgcgcaacat ctggggcggc	1080
[0845]	tggctcgtga acgtgctcac cggcttcac tctactgcaga cggagcacca cctgttccc	1140
[0846]	atgatgcaa ccggcaacct catgactatc cagcccagg tgcgcgctt cttcaagaag	1200
[0847]	cacggacttg agtaccgga gggcaacctc attgagtgcg tgcggcagaa catccgtgcg	1260
[0848]	cttgcatcag agcacctgct ttga	1284
[0849]	<210>	24
[0850]	<211>	427
[0851]	<212>	PRT
[0852]	<213>	Pavlova salina
[0853]	<400>	24
[0854]	Met Gly Arg Gly Gly Asp Ser Ser Gly Gln Ala His Pro Ala Ala Glu	
[0855]	1 5 10 15	
[0856]	Leu Ala Val Pro Ser Asp Arg Ala Glu Val Ser Asn Ala Asp Ser Lys	
[0857]	20 25 30	
[0858]	Ala Leu His Ile Val Leu Tyr Gly Lys Arg Val Asp Val Thr Lys Phe	
[0859]	35 40 45	
[0860]	Gln Arg Thr His Pro Gly Gly Ser Lys Val Phe Arg Ile Phe Gln Asp	
[0861]	50 55 60	
[0862]	Arg Asp Ala Thr Glu Gln Phe Glu Ser Tyr His Ser Lys Arg Ala Ile	
[0863]	65 70 75 80	
[0864]	Lys Met Met Glu Gly Met Leu Lys Lys Ser Glu Asp Ala Pro Ala Asp	
[0865]	85 90 95	
[0866]	Thr Pro Leu Pro Ser Gln Ser Pro Met Gly Lys Asp Phe Lys Ala Met	
[0867]	100 105 110	
[0868]	Ile Glu Arg His Val Ala Ala Gly Tyr Tyr Asp Pro Cys Pro Leu Asp	
[0869]	115 120 125	
[0870]	Glu Leu Phe Lys Leu Ser Leu Val Leu Leu Pro Thr Phe Ala Gly Met	
[0871]	130 135 140	
[0872]	Tyr Met Leu Lys Ala Gly Val Gly Ser Pro Leu Cys Gly Ala Leu Met	
[0873]	145 150 155 160	
[0874]	Val Ser Phe Gly Trp Tyr Leu Asp Gly Trp Leu Ala His Asp Tyr Leu	
[0875]	165 170 175	
[0876]	His His Ser Val Phe Lys Gly Ser Val Ala Arg Thr Val Gly Trp Asn	
[0877]	180 185 190	
[0878]	Asn Ala Ala Gly Tyr Phe Leu Gly Phe Val Gln Gly Tyr Ala Val Glu	
[0879]	195 200 205	
[0880]	Trp Trp Arg Ala Arg His Asn Thr His His Val Cys Thr Asn Glu Asp	
[0881]	210 215 220	

[0882]	Gly Ser Asp Pro Asp Ile Lys Thr Ala Pro Leu Leu Ile Tyr Val Arg		
[0883]	225	230	235 240
[0884]	Asn Lys Pro Ser Ile Ala Lys Arg Leu Asn Ala Phe Gln Arg Tyr Gln		
[0885]	245	250	255
[0886]	Gln Tyr Tyr Tyr Val Pro Val Met Ala Ile Leu Asp Leu Tyr Trp Arg		
[0887]	260	265	270
[0888]	Leu Glu Ser Ile Ala Tyr Val Ala Met Arg Leu Pro Lys Met Leu Pro		
[0889]	275	280	285
[0890]	Gln Ala Leu Ala Leu Val Ala His Tyr Ala Ile Val Ala Trp Val Phe		
[0891]	290	295	300
[0892]	Ala Gly Asn Tyr His Leu Leu Pro Leu Val Thr Val Leu Arg Gly Phe		
[0893]	305	310	315 320
[0894]	Gly Thr Gly Ile Thr Val Phe Ala Thr His Tyr Gly Glu Asp Ile Leu		
[0895]	325	330	335
[0896]	Asp Ala Asp Gln Val Arg His Met Thr Leu Val Glu Gln Thr Ala Leu		
[0897]	340	345	350
[0898]	Thr Ser Arg Asn Ile Ser Gly Gly Trp Leu Val Asn Val Leu Thr Gly		
[0899]	355	360	365
[0900]	Phe Ile Ser Leu Gln Thr Glu His His Leu Phe Pro Met Met Pro Thr		
[0901]	370	375	380
[0902]	Gly Asn Leu Met Thr Ile Gln Pro Glu Val Arg Ala Phe Phe Lys Lys		
[0903]	385	390	395 400
[0904]	His Gly Leu Glu Tyr Arg Glu Gly Asn Leu Ile Glu Cys Val Arg Gln		
[0905]	405	410	415
[0906]	Asn Ile Arg Ala Leu Ala Phe Glu His Leu Leu		
[0907]	420	425	
[0908]	<210> 25		
[0909]	<211> 1278		
[0910]	<212> DNA		
[0911]	<213> Pavlova salina		
[0912]	<400> 25		
[0913]	atgccgccgc gcgatagcta ctgtacgcc gccccgccgt cggcccagct gcacgaggtc	60	
[0914]	gataccccgc aggagcatga taagaaggag ctgtcatcg gtgaccgcgc gtacgacgtg	120	
[0915]	accaactttg tgaagcgcca cccgggtggc aagatcatcg cataccaggt tggcacagat	180	
[0916]	gcgacggacg cgtacaagca gttccatgtg cggcttgcca aggcggacaa gatgctcaag	240	
[0917]	tcgtgcctt cgcgcccggt gcacaaggc tactcgcccc gccgcgtga cctcattgcc	300	
[0918]	gacttccagg agttcaccaa gcagctggag gcggagggca tgtttgagcc gtcgctgccg	360	
[0919]	cacgtggcat accgcctggc ggaggtgatc gcgatgcacg tggccggcgc cgcgctcadc	420	
[0920]	tggcacgggt acaccttgc gggcattgcc atgctcggcg ttgtgcaggg ccgctgcggc	480	
[0921]	tggctcatgc acgagggcgg ccactactcg ctacaggca acattgcttt tgaccgtgcc	540	
[0922]	atccaagtgc cgtgctacgg ctttggtgc ggcattgctg gcgctggtg gcgcaaccag	600	
[0923]	cacaacaagc accacgcgac gccgcagaag ttgcagcacg acgtcgacct cgacaccctc	660	

[0924]	ccgctcgtcg ccttcacga gcgtagacc gccagggtga agagccccgc gatgaaggcg	720
[0925]	tggcttagta tgcaggcgaa gctcttcgcg ccagtaccca cgctgctggt cgcgctgggc	780
[0926]	tggcagctgt acctgcaccc gcgccatatg ctgcgcacca agcactacga cgagctcgcg	840
[0927]	atgctcggca ttcgctacgg ccttgctcgc tacctcgcg cgaactacgg cgcggggtac	900
[0928]	gtgctcgcgt gctacctgct gtacgtgcag ctcggcgcca tgtacatctt ctgcaacttt	960
[0929]	gccgtgtcgc acacacacct gccggttgtc gaggcctaag agcacgcaac gtgggtggag	1020
[0930]	tacgccgcca accacacgac caactgctcg ccctcgtggt ggtgcgactg gtggatgtcg	1080
[0931]	tacctcaact accagatcga gcaccacctc taccgtcca tgccgcagtt ccgccacccg	1140
[0932]	aagattgcgc cgcggtgaa gcagctcttc gagaagcacg gcctgcacta cgacgtgcgt	1200
[0933]	ggctacttcg aggccatggc ggacacgttt gccaaccttg acaacgtcgc gcacgcgccg	1260
[0934]	gagaagaaga tgcagtga	1278
[0935]	<210>	26
[0936]	<211>	425
[0937]	<212>	PRT
[0938]	<213>	Pavlova salina
[0939]	<400>	26
[0940]	Met Pro Pro Arg Asp Ser Tyr Ser Tyr Ala Ala Pro Pro Ser Ala Gln	
[0941]	1 5 10 15	
[0942]	Leu His Glu Val Asp Thr Pro Gln Glu His Asp Lys Lys Glu Leu Val	
[0943]	20 25 30	
[0944]	Ile Gly Asp Arg Ala Tyr Asp Val Thr Asn Phe Val Lys Arg His Pro	
[0945]	35 40 45	
[0946]	Gly Gly Lys Ile Ile Ala Tyr Gln Val Gly Thr Asp Ala Thr Asp Ala	
[0947]	50 55 60	
[0948]	Tyr Lys Gln Phe His Val Arg Ser Ala Lys Ala Asp Lys Met Leu Lys	
[0949]	65 70 75 80	
[0950]	Ser Leu Pro Ser Arg Pro Val His Lys Gly Tyr Ser Pro Arg Arg Ala	
[0951]	85 90 95	
[0952]	Asp Leu Ile Ala Asp Phe Gln Glu Phe Thr Lys Gln Leu Glu Ala Glu	
[0953]	100 105 110	
[0954]	Gly Met Phe Glu Pro Ser Leu Pro His Val Ala Tyr Arg Leu Ala Glu	
[0955]	115 120 125	
[0956]	Val Ile Ala Met His Val Ala Gly Ala Ala Leu Ile Trp His Gly Tyr	
[0957]	130 135 140	
[0958]	Thr Phe Ala Gly Ile Ala Met Leu Gly Val Val Gln Gly Arg Cys Gly	
[0959]	145 150 155 160	
[0960]	Trp Leu Met His Glu Gly Gly His Tyr Ser Leu Thr Gly Asn Ile Ala	
[0961]	165 170 175	
[0962]	Phe Asp Arg Ala Ile Gln Val Ala Cys Tyr Gly Leu Gly Cys Gly Met	
[0963]	180 185 190	
[0964]	Ser Gly Ala Trp Trp Arg Asn Gln His Asn Lys His His Ala Thr Pro	
[0965]	195 200 205	

[0966]	Gln Lys Leu Gln His Asp Val Asp Leu Asp Thr Leu Pro Leu Val Ala		
[0967]	210	215	220
[0968]	Phe His Glu Arg Ile Ala Ala Lys Val Lys Ser Pro Ala Met Lys Ala		
[0969]	225	230	235 240
[0970]	Trp Leu Ser Met Gln Ala Lys Leu Phe Ala Pro Val Thr Thr Leu Leu		
[0971]	245	250	255
[0972]	Val Ala Leu Gly Trp Gln Leu Tyr Leu His Pro Arg His Met Leu Arg		
[0973]	260	265	270
[0974]	Thr Lys His Tyr Asp Glu Leu Ala Met Leu Gly Ile Arg Tyr Gly Leu		
[0975]	275	280	285
[0976]	Val Gly Tyr Leu Ala Ala Asn Tyr Gly Ala Gly Tyr Val Leu Ala Cys		
[0977]	290	295	300
[0978]	Tyr Leu Leu Tyr Val Gln Leu Gly Ala Met Tyr Ile Phe Cys Asn Phe		
[0979]	305	310	315 320
[0980]	Ala Val Ser His Thr His Leu Pro Val Val Glu Pro Asn Glu His Ala		
[0981]	325	330	335
[0982]	Thr Trp Val Glu Tyr Ala Ala Asn His Thr Thr Asn Cys Ser Pro Ser		
[0983]	340	345	350
[0984]	Trp Trp Cys Asp Trp Trp Met Ser Tyr Leu Asn Tyr Gln Ile Glu His		
[0985]	355	360	365
[0986]	His Leu Tyr Pro Ser Met Pro Gln Phe Arg His Pro Lys Ile Ala Pro		
[0987]	370	375	380
[0988]	Arg Val Lys Gln Leu Phe Glu Lys His Gly Leu His Tyr Asp Val Arg		
[0989]	385	390	395 400
[0990]	Gly Tyr Phe Glu Ala Met Ala Asp Thr Phe Ala Asn Leu Asp Asn Val		
[0991]	405	410	415
[0992]	Ala His Ala Pro Glu Lys Lys Met Gln		
[0993]	420	425	
[0994]	<210> 27		
[0995]	<211> 801		
[0996]	<212> DNA		
[0997]	<213> <i>Emiliana huxleyi</i> CCMP1516		
[0998]	<400> 27		
[0999]	atgctcgatc ggcctcgtc cgacgcggcc atctggtctg cgggtgtccga tccggaaatc 60		
[1000]	ctgatcggca ctttctccta cctgctgctc aagccgctgc tacgcaactc agggctcgtg 120		
[1001]	gacgagcggga aaggcgcta ccggacctcg atgatctggt acaacgtggt gctcgcgctc 180		
[1002]	ttctccgcga cgagcttcta cgtgactgcg accgcgctcg ggtgggacaa gggcaccggc 240		
[1003]	gagtggctcc gcagtctcac gggcgacagc ccgcagcagc tgtggcaatg cccgtcgagg 300		
[1004]	gtatgggact ccaagctgtt cctgtggacg gccaaaggcct tctactactc aaagtacgtg 360		
[1005]	gagtacctcg acacggcgtg gctcgtcctc aaggggaaga aggtctcctt cctgcagggc 420		
[1006]	ttccaccact ttggcgcgcc gtgggacgtg tacctgggca ttcggctgaa gaacgagggc 480		
[1007]	gtgtggatct tcatgttctt caactcgttc atccacacgg tcatgtacac gtactacggc 540		

[1008] ctcaccgccg cgggctacaa gatccgcggc aagccgatca tcaccgcgat gcaaataagc 600
 [1009] cagttcgtcg gcggctttgt cctagtgtgg gactacatca acgtgccgtg cttccacgcc 660
 [1010] gacgccgggc aggtcttcag ctgggtcttt aactatgctt acgtcggtc cgtctttctg 720
 [1011] ctgttctgcc acttcttcta catggacaac atcgcaagg ccaaggccaa gaaggccgtc 780
 [1012] gctaccgcga aggcgtgtg a 801
 [1013] <210> 28
 [1014] <211> 266
 [1015] <212> PRT
 [1016] <213> *Emiliana huxleyi* CCMP1516
 [1017] <400> 28
 [1018] Met Leu Asp Arg Ala Ser Ser Asp Ala Ala Ile Trp Ser Ala Val Ser
 [1019] 1 5 10 15
 [1020] Asp Pro Glu Ile Leu Ile Gly Thr Phe Ser Tyr Leu Leu Leu Lys Pro
 [1021] 20 25 30
 [1022] Leu Leu Arg Asn Ser Gly Leu Val Asp Glu Arg Lys Gly Ala Tyr Arg
 [1023] 35 40 45
 [1024] Thr Ser Met Ile Trp Tyr Asn Val Val Leu Ala Leu Phe Ser Ala Thr
 [1025] 50 55 60
 [1026] Ser Phe Tyr Val Thr Ala Thr Ala Leu Gly Trp Asp Lys Gly Thr Gly
 [1027] 65 70 75 80
 [1028] Glu Trp Leu Arg Ser Leu Thr Gly Asp Ser Pro Gln Gln Leu Trp Gln
 [1029] 85 90 95
 [1030] Cys Pro Ser Arg Val Trp Asp Ser Lys Leu Phe Leu Trp Thr Ala Lys
 [1031] 100 105 110
 [1032] Ala Phe Tyr Tyr Ser Lys Tyr Val Glu Tyr Leu Asp Thr Ala Trp Leu
 [1033] 115 120 125
 [1034] Val Leu Lys Gly Lys Lys Val Ser Phe Leu Gln Gly Phe His His Phe
 [1035] 130 135 140
 [1036] Gly Ala Pro Trp Asp Val Tyr Leu Gly Ile Arg Leu Lys Asn Glu Gly
 [1037] 145 150 155 160
 [1038] Val Trp Ile Phe Met Phe Phe Asn Ser Phe Ile His Thr Val Met Tyr
 [1039] 165 170 175
 [1040] Thr Tyr Tyr Gly Leu Thr Ala Ala Gly Tyr Lys Ile Arg Gly Lys Pro
 [1041] 180 185 190
 [1042] Ile Ile Thr Ala Met Gln Ile Ser Gln Phe Val Gly Gly Phe Val Leu
 [1043] 195 200 205
 [1044] Val Trp Asp Tyr Ile Asn Val Pro Cys Phe His Ala Asp Ala Gly Gln
 [1045] 210 215 220
 [1046] Val Phe Ser Trp Val Phe Asn Tyr Ala Tyr Val Gly Ser Val Phe Leu
 [1047] 225 230 235 240
 [1048] Leu Phe Cys His Phe Phe Tyr Met Asp Asn Ile Ala Lys Ala Lys Ala
 [1049] 245 250 255

[1050] Lys Lys Ala Val Ala Thr Arg Lys Ala Leu
 [1051] 260 265
 [1052] <210> 29
 [1053] <211> 801
 [1054] <212> DNA
 [1055] <213> Artificial Sequence
 [1056] <220>
 [1057] <223> Codon-optimized nucleotide sequence for production in plants of
 [1058] SEQ ID NO:28
 [1059] <400> 29
 [1060] atgcttgata gagcttcata tgatgctgct atttggagcg ctgtttctga tcctgagatc 60
 [1061] cttatcggaa cttctcttta ctttttgctt aagcctctcc tcagaaactc tggacttggt 120
 [1062] gatgagagaa agggagctta ccgtacttct atgatctggt acaacgttgt tcttgctctt 180
 [1063] ttctctgcta cctctttcta cgttactgct actgctcttg gatgggataa gggaactggt 240
 [1064] gattggctta gatctcttac tggtgattct cctcaacaac tttggcagtg cccttctaga 300
 [1065] gtttgggaca gcaaaactct cttgtggact gctaaagcct tctactactc caagtacgtt 360
 [1066] gagtaccttg atactgcttg gcttgttctc aagggaaga aggtttcatt cctccaggga 420
 [1067] ttccatcatt tcggtgctcc atgggatgtt taccttgga tcaggcttaa gaacgaggga 480
 [1068] gtttgatct tcatgttctt caacagcttc atccacactg ttatgtacac ttactacgga 540
 [1069] cttactgctg ctggatacaa gatcagagga aagcctatca tcaccgctat gcaaatctct 600
 [1070] caattcgttg gtggattcgt tcttggtggt gactacatca acgttccttg tttccatgct 660
 [1071] gatgctggac aagttttctc ttgggtgttc aactacgctt atgtgggac tgttttcctt 720
 [1072] cttttctgcc acttcttcta catggacaac attgctaagg ctaaggctaa aaaggctgtt 780
 [1073] gctaccagaa aggtctcttg a 801
 [1074] <210> 30
 [1075] <211> 456
 [1076] <212> PRT
 [1077] <213> *Ostreococcus tauri*
 [1078] <400> 30
 [1079] Met Cys Val Glu Thr Glu Asn Asn Asp Gly Ile Pro Thr Val Glu Ile
 [1080] 1 5 10 15
 [1081] Ala Phe Asp Gly Glu Arg Glu Arg Ala Glu Ala Asn Val Lys Leu Ser
 [1082] 20 25 30
 [1083] Ala Glu Lys Met Glu Pro Ala Ala Leu Ala Lys Thr Phe Ala Arg Arg
 [1084] 35 40 45
 [1085] Tyr Val Val Ile Glu Gly Val Glu Tyr Asp Val Thr Asp Phe Lys His
 [1086] 50 55 60
 [1087] Pro Gly Gly Thr Val Ile Phe Tyr Ala Leu Ser Asn Thr Gly Ala Asp
 [1088] 65 70 75 80
 [1089] Ala Thr Glu Ala Phe Lys Glu Phe His His Arg Ser Arg Lys Ala Arg
 [1090] 85 90 95
 [1091] Lys Ala Leu Ala Ala Leu Pro Ser Arg Pro Ala Lys Thr Ala Lys Val

[1092]	100	105	110
[1093]	Asp Asp Ala Glu Met Leu Gln Asp Phe Ala Lys Trp Arg Lys Glu Leu		
[1094]	115	120	125
[1095]	Glu Arg Asp Gly Phe Phe Lys Pro Ser Pro Ala His Val Ala Tyr Arg		
[1096]	130	135	140
[1097]	Phe Ala Glu Leu Ala Ala Met Tyr Ala Leu Gly Thr Tyr Leu Met Tyr		
[1098]	145	150	155
[1099]	Ala Arg Tyr Val Val Ser Ser Val Leu Val Tyr Ala Cys Phe Phe Gly		
[1100]	165	170	175
[1101]	Ala Arg Cys Gly Trp Val Gln His Glu Gly Gly His Ser Ser Leu Thr		
[1102]	180	185	190
[1103]	Gly Asn Ile Trp Trp Asp Lys Arg Ile Gln Ala Phe Thr Ala Gly Phe		
[1104]	195	200	205
[1105]	Gly Leu Ala Gly Ser Gly Asp Met Trp Asn Ser Met His Asn Lys His		
[1106]	210	215	220
[1107]	His Ala Thr Pro Gln Lys Val Arg His Asp Met Asp Leu Asp Thr Thr		
[1108]	225	230	235
[1109]	Pro Ala Val Ala Phe Phe Asn Thr Ala Val Glu Asp Asn Arg Pro Arg		
[1110]	245	250	255
[1111]	Gly Phe Ser Lys Tyr Trp Leu Arg Leu Gln Ala Trp Thr Phe Ile Pro		
[1112]	260	265	270
[1113]	Val Thr Ser Gly Leu Val Leu Leu Phe Trp Met Phe Phe Leu His Pro		
[1114]	275	280	285
[1115]	Ser Lys Ala Leu Lys Gly Gly Lys Tyr Glu Glu Leu Val Trp Met Leu		
[1116]	290	295	300
[1117]	Ala Ala His Val Ile Arg Thr Trp Thr Ile Lys Ala Val Thr Gly Phe		
[1118]	305	310	315
[1119]	Thr Ala Met Gln Ser Tyr Gly Leu Phe Leu Ala Thr Ser Trp Val Ser		
[1120]	325	330	335
[1121]	Gly Cys Tyr Leu Phe Ala His Phe Ser Thr Ser His Thr His Leu Asp		
[1122]	340	345	350
[1123]	Val Val Pro Ala Asp Glu His Leu Ser Trp Val Arg Tyr Ala Val Asp		
[1124]	355	360	365
[1125]	His Thr Ile Asp Ile Asp Pro Ser Gln Gly Trp Val Asn Trp Leu Met		
[1126]	370	375	380
[1127]	Gly Tyr Leu Asn Cys Gln Val Ile His His Leu Phe Pro Ser Met Pro		
[1128]	385	390	395
[1129]	Gln Phe Arg Gln Pro Glu Val Ser Arg Arg Phe Val Ala Phe Ala Lys		
[1130]	405	410	415
[1131]	Lys Trp Asn Leu Asn Tyr Lys Val Met Thr Tyr Ala Gly Ala Trp Lys		
[1132]	420	425	430
[1133]	Ala Thr Leu Gly Asn Leu Asp Asn Val Gly Lys His Tyr Tyr Val His		

[1134]	435	440	445
[1135]	Gly Gln His Ser Gly Lys Thr Ala		
[1136]	450	455	
[1137]	<210> 31		
[1138]	<211> 8		
[1139]	<212> PRT		
[1140]	<213> Artificial Sequence		
[1141]	<220>		
[1142]	<223> Elongase consensus domain 1		
[1143]	<220>		
[1144]	<221> MISC_FEATURE		
[1145]	<222> (2) .. (6)		
[1146]	<223> X = Any amino acid		
[1147]	<400> 31		
[1148]	Lys Xaa Xaa Xaa Xaa Xaa Asp Thr		
[1149]	1	5	
[1150]	<210> 32		
[1151]	<211> 5		
[1152]	<212> PRT		
[1153]	<213> Artificial Sequence		
[1154]	<220>		
[1155]	<223> Elongase consensus domain 2		
[1156]	<220>		
[1157]	<221> MISC_FEATURE		
[1158]	<222> (3) .. (3)		
[1159]	<223> X = any amino acid		
[1160]	<400> 32		
[1161]	Met Tyr Xaa Tyr Tyr		
[1162]	1	5	
[1163]	<210> 33		
[1164]	<211> 8		
[1165]	<212> PRT		
[1166]	<213> Artificial Sequence		
[1167]	<220>		
[1168]	<223> Elongase consensus domain 3		
[1169]	<400> 33		
[1170]	Lys Ile Tyr Glu Phe Val Asp Thr		
[1171]	1	5	
[1172]	<210> 34		
[1173]	<211> 7		
[1174]	<212> PRT		
[1175]	<213> Artificial Sequence		

[1176]	<220>	
[1177]	<223> Elongase consensus domain 4	
[1178]	<400> 34	
[1179]	Val His Val Cys Met Tyr Thr	
[1180]	1	5
[1181]	<210> 35	
[1182]	<211> 7	
[1183]	<212> PRT	
[1184]	<213> Artificial Sequence	
[1185]	<220>	
[1186]	<223> Elongase consensus domain 5	
[1187]	<400> 35	
[1188]	Tyr Leu Glu Leu Leu Asp Thr	
[1189]	1	5
[1190]	<210> 36	
[1191]	<211> 5	
[1192]	<212> PRT	
[1193]	<213> Artificial Sequence	
[1194]	<220>	
[1195]	<223> Elongase consensus domain 6	
[1196]	<400> 36	
[1197]	Met Tyr Ser Tyr Tyr	
[1198]	1	5
[1199]	<210> 37	
[1200]	<211> 10	
[1201]	<212> PRT	
[1202]	<213> Artificial Sequence	
[1203]	<220>	
[1204]	<223> Desaturase consensus domain 1	
[1205]	<400> 37	
[1206]	Trp Lys Asn Met His Asn Lys His His Ala	
[1207]	1	5 10
[1208]	<210> 38	
[1209]	<211> 8	
[1210]	<212> PRT	
[1211]	<213> Artificial Sequence	
[1212]	<220>	
[1213]	<223> Desaturase consensus domain 2	
[1214]	<400> 38	
[1215]	His His Leu Phe Pro Ser Met Pro	
[1216]	1	5
[1217]	<210> 39	

[1218] <211> 8
[1219] <212> PRT
[1220] <213> Artificial Sequence
[1221] <220>
[1222] <223> Desaturase consensus domain 3
[1223] <400> 39
[1224] Trp Cys Ile Gly His Asp Cys Gly
[1225] 1 5
[1226] <210> 40
[1227] <211> 10
[1228] <212> PRT
[1229] <213> Artificial Sequence
[1230] <220>
[1231] <223> Desaturase consensus domain 4
[1232] <400> 40
[1233] Thr Phe Leu Gln His His Asp Glu Asp Met
[1234] 1 5 10
[1235] <210> 41
[1236] <211> 37
[1237] <212> DNA
[1238] <213> Artificial Sequence
[1239] <220>
[1240] <223> Oligonucleotide primer
[1241] <400> 41
[1242] atttagtgta cactatagtt tttttttttt ttttttv 37
[1243] <210> 42
[1244] <211> 23
[1245] <212> DNA
[1246] <213> Artificial Sequence
[1247] <220>
[1248] <223> Oligonucleotide primer
[1249] <220>
[1250] <221> misc_feature
[1251] <222> (7) .. (7)
[1252] <223> n= inosine
[1253] <220>
[1254] <221> misc_feature
[1255] <222> (14) .. (14)
[1256] <223> n= inosine
[1257] <400> 42
[1258] aagwwcnksg arynsytcga cac 23
[1259] <210> 43

[1260]	<211> 20
[1261]	<212> DNA
[1262]	<213> Artificial Sequence
[1263]	<220>
[1264]	<223> Oligonucleotide primer
[1265]	<220>
[1266]	<221> misc_feature
[1267]	<223> n = inosine
[1268]	<220>
[1269]	<221> misc_feature
[1270]	<222> (2) .. (3)
[1271]	<223> n = inosine
[1272]	<220>
[1273]	<221> misc_feature
[1274]	<222> (5) .. (5)
[1275]	<223> n = inosine
[1276]	<400> 43
[1277]	annmnrart asgtgtacat 20
[1278]	<210> 44
[1279]	<211> 21
[1280]	<212> DNA
[1281]	<213> Artificial Sequence
[1282]	<220>
[1283]	<223> Oligonucleotide primer
[1284]	<400> 44
[1285]	gaacaacgac tgcacgacg c 21
[1286]	<210> 45
[1287]	<211> 23
[1288]	<212> DNA
[1289]	<213> Artificial Sequence
[1290]	<220>
[1291]	<223> Oligonucleotide primer
[1292]	<400> 45
[1293]	cgactggagc acgaggacac tga 23
[1294]	<210> 46
[1295]	<211> 23
[1296]	<212> DNA
[1297]	<213> Artificial Sequence
[1298]	<220>
[1299]	<223> Oligonucleotide primer
[1300]	<400> 46
[1301]	ttgcgcagca ccataaagac ggt 23

[1302] <210> 47
[1303] <211> 26
[1304] <212> DNA
[1305] <213> Artificial Sequence
[1306] <220>
[1307] <223> Oligonucleotide primer
[1308] <400> 47
[1309] ggacactgac atggactgaa ggagta 26
[1310] <210> 48
[1311] <211> 22
[1312] <212> DNA
[1313] <213> Artificial Sequence
[1314] <220>
[1315] <223> Oligonucleotide primer
[1316] <400> 48
[1317] caggcgacgc gcgccagagt cc 22
[1318] <210> 49
[1319] <211> 26
[1320] <212> DNA
[1321] <213> Artificial Sequence
[1322] <220>
[1323] <223> Oligonucleotide primer
[1324] <400> 49
[1325] ttattagtta cttggccttt accttc 26
[1326] <210> 50
[1327] <211> 23
[1328] <212> DNA
[1329] <213> Artificial Sequence
[1330] <220>
[1331] <223> Oligonucleotide primer
[1332] <220>
[1333] <221> misc_feature
[1334] <222> (18) .. (18)
[1335] <223> n = inosine
[1336] <400> 50
[1337] aaratmtayg agttygtnga tac 23
[1338] <210> 51
[1339] <211> 25
[1340] <212> DNA
[1341] <213> Artificial Sequence
[1342] <220>
[1343] <223> Oligonucleotide primer

[1344] <220>
[1345] <221> misc_feature
[1346] <222> (3) .. (3)
[1347] <223> n = inosine
[1348] <400> 51
[1349] tangtgtaca tgcacacrtg waccc 25
[1350] <210> 52
[1351] <211> 22
[1352] <212> DNA
[1353] <213> Artificial Sequence
[1354] <220>
[1355] <223> Oligonucleotide primer
[1356] <400> 52
[1357] ttcgtggata cgttcatcat gc 22
[1358] <210> 53
[1359] <211> 23
[1360] <212> DNA
[1361] <213> Artificial Sequence
[1362] <220>
[1363] <223> Oligonucleotide primer
[1364] <400> 53
[1365] agttgagcgc cgccgagaag tac 23
[1366] <210> 54
[1367] <211> 22
[1368] <212> DNA
[1369] <213> Artificial Sequence
[1370] <220>
[1371] <223> Oligonucleotide primer
[1372] <400> 54
[1373] acctggttga cgttgccctt ca 22
[1374] <210> 55
[1375] <211> 21
[1376] <212> DNA
[1377] <213> Artificial Sequence
[1378] <220>
[1379] <223> Oligonucleotide primer
[1380] <400> 55
[1381] gctatggagt tcgctcagcc t 21
[1382] <210> 56
[1383] <211> 26
[1384] <212> DNA
[1385] <213> Artificial Sequence

[1386] <220>
[1387] <223> Oligonucleotide primer
[1388] <400> 56
[1389] ttactactgc ttcttgctgg ccagct 26
[1390] <210> 57
[1391] <211> 22
[1392] <212> DNA
[1393] <213> Artificial Sequence
[1394] <220>
[1395] <223> Oligonucleotide primer
[1396] <400> 57
[1397] artayytsga rytrytggay ac 22
[1398] <210> 58
[1399] <211> 21
[1400] <212> DNA
[1401] <213> Artificial Sequence
[1402] <220>
[1403] <223> Oligonucleotide primer
[1404] <400> 58
[1405] catkarrrtar tasgagtaca t 21
[1406] <210> 59
[1407] <211> 22
[1408] <212> DNA
[1409] <213> Artificial Sequence
[1410] <220>
[1411] <223> Oligonucleotide primer
[1412] <400> 59
[1413] catcatccct gttgatctgg tc 22
[1414] <210> 60
[1415] <211> 23
[1416] <212> DNA
[1417] <213> Artificial Sequence
[1418] <220>
[1419] <223> Oligonucleotide primer
[1420] <400> 60
[1421] ccagatcaac agggatatgat ggt 23
[1422] <210> 61
[1423] <211> 27
[1424] <212> DNA
[1425] <213> Artificial Sequence
[1426] <220>
[1427] <223> Oligonucleotide primer

[1428] <400> 61
[1429] cgaaagctgg tcaaacttct tgcgcat 27
[1430] <210> 62
[1431] <211> 27
[1432] <212> DNA
[1433] <213> Artificial Sequence
[1434] <220>
[1435] <223> Oligonucleotide primer
[1436] <400> 62
[1437] aacatggcgt ctattgcgat tccggct 27
[1438] <210> 63
[1439] <211> 27
[1440] <212> DNA
[1441] <213> Artificial Sequence
[1442] <220>
[1443] <223> Oligonucleotide primer
[1444] <400> 63
[1445] ttattactgc ttcttggcac ccttgct 27
[1446] <210> 64
[1447] <211> 31
[1448] <212> DNA
[1449] <213> Artificial Sequence
[1450] <220>
[1451] <223> Oligonucleotide primer
[1452] <220>
[1453] <221> misc_feature
[1454] <222> (23) .. (23)
[1455] <223> n = inosine
[1456] <400> 64
[1457] ggtggaagaa caagcacaac rdncaycayg c 31
[1458] <210> 65
[1459] <211> 24
[1460] <212> DNA
[1461] <213> Artificial Sequence
[1462] <220>
[1463] <223> Oligonucleotide primer
[1464] <220>
[1465] <221> misc_feature
[1466] <222> (16) .. (16)
[1467] <223> n = inosine
[1468] <400> 65
[1469] gggcatcgtg gggwanarrt grtg 24

[1470]	<210> 66
[1471]	<211> 20
[1472]	<212> DNA
[1473]	<213> Artificial Sequence
[1474]	<220>
[1475]	<223> Oligonucleotide primer
[1476]	<400> 66
[1477]	agcgagtacc tgcattgggt 20
[1478]	<210> 67
[1479]	<211> 27
[1480]	<212> DNA
[1481]	<213> Artificial Sequence
[1482]	<220>
[1483]	<223> Oligonucleotide primer
[1484]	<400> 67
[1485]	atagtgttg gtgcgcaagc tgtgcct 27
[1486]	<210> 68
[1487]	<211> 25
[1488]	<212> DNA
[1489]	<213> Artificial Sequence
[1490]	<220>
[1491]	<223> Oligonucleotide primer
[1492]	<400> 68
[1493]	caccatggga aaggaggca atgct 25
[1494]	<210> 69
[1495]	<211> 25
[1496]	<212> DNA
[1497]	<213> Artificial Sequence
[1498]	<220>
[1499]	<223> Oligonucleotide primer
[1500]	<400> 69
[1501]	ttactagtgc gccttgagt gagat 25
[1502]	<210> 70
[1503]	<211> 25
[1504]	<212> DNA
[1505]	<213> Artificial Sequence
[1506]	<220>
[1507]	<223> Oligonucleotide primer
[1508]	<220>
[1509]	<221> misc_feature
[1510]	<222> (20) .. (20)
[1511]	<223> n = inosine

[1512] <220>
 [1513] <221> misc_feature
 [1514] <222> (23) .. (23)
 [1515] <223> n = inosine
 [1516] <400> 70
 [1517] tgtggtgcat cggccaygan ksngg 25
 [1518] <210> 71
 [1519] <211> 28
 [1520] <212> DNA
 [1521] <213> Artificial Sequence
 [1522] <220>
 [1523] <223> Oligonucleotide primer
 [1524] <220>
 [1525] <221> misc_feature
 [1526] <222> (20) .. (20)
 [1527] <223> n = inosine
 [1528] <220>
 [1529] <221> misc_feature
 [1530] <222> (26) .. (26)
 [1531] <223> n = inosine
 [1532] <400> 71
 [1533] tgtcctcgtc gttgtgctgn arrwangt 28
 [1534] <210> 72
 [1535] <211> 1344
 [1536] <212> DNA
 [1537] <213> Pavlova salina
 [1538] <400> 72
 [1539] atgcctccga gcgcggcgaa gcagatgggc gcgagcacgg gcgtgcatgc gggcgtcaca 60
 [1540] gattcgtegg cttcacgcg caaggatgtc gccacaggc cggacctcac gatcgtaggt 120
 [1541] gacagcgtgt acgatgcgaa ggcgttccgc tccgagcatc cgggtggcgc gcactttgtg 180
 [1542] tcgctgttcg gcgggcgcga tgccacggag gcgttcattg agtaccaccg gcgcgcctgg 240
 [1543] cccaagtcgc gcatgtcgcg cttccacgtc ggctctcttg catcgaccga ggagcccgtc 300
 [1544] gccgccgatg agggctacct ccagctgtgc gctcgcatcg ccaagatggt gccgtcggtc 360
 [1545] agcagcgggt tcgcgccggc gtcgtactgg gtgaaggccg ggctgatcct cggctccgcg 420
 [1546] atcgcgctcg aggcgtacat gctgtacgcg ggcaagcgcc tgctcccgtc gatcgtgctc 480
 [1547] gggtagctgt ttgcgtgat tggcctgaac atccagcacg atgccaacca cggcgcgctc 540
 [1548] tccaagtcgg ctcggtcaa cctggcgctc gggtttgtcc aggactggat cggcggggagc 600
 [1549] atgatcctct ggctgcagga gcacgttgc atgcaccact tgcacaccaa cgacgttgac 660
 [1550] aaggaccgg accagaaggc gcacggcgcc ctgcggctca agccgaccga cgcgtggagc 720
 [1551] ccgatgcact ggctgcagca cctctacctg ctgcctgggg agacgatgta cgccttcaag 780
 [1552] ctgctgtttc tcgacatcag cgagctggtg atgtggcggg gggagggcga gcccatcagc 840
 [1553] aagctggcgg ggtacctctt catgccctcg ctgctcctca agctcacctt ctgggcgcgc 900

[1554] tttgtcgcgc tgccgctgta cctcgcgccc agcgtgcaca cggcgggtgtg catcgcggcg 960
 [1555] acggtaatga cggggagctt ctacctcgcc ttctttcttct tcatctcgca caacttcgag 1020
 [1556] ggcgtggcga gcgtcggacc ggacggcagc atcaccagca tgacgcgcgg cgcatecttc 1080
 [1557] ctcaagcggc aggccgagac ctcgtccaac gtgggcggcc cgctgctcgc cacgctcaac 1140
 [1558] ggcgccctca actaccaaact cgagcaccac ctcttcccca gggtgcacca cggcttctac 1200
 [1559] cctcgctcg cgccgttggc caaggcggag ctcgaggcgc gcggcattga gtacaagcac 1260
 [1560] taccacacca tatggagcaa cctggcatcc acgctgaggc acatgtacgc gctcggccgc 1320
 [1561] aggccgcgca gcaaggcggg gtga 1344
 [1562] <210> 73
 [1563] <211> 447
 [1564] <212> PRT
 [1565] <213> Pavlova salina
 [1566] <400> 73
 [1567] Met Pro Pro Ser Ala Ala Lys Gln Met Gly Ala Ser Thr Gly Val His
 [1568] 1 5 10 15
 [1569] Ala Gly Val Thr Asp Ser Ser Ala Phe Thr Arg Lys Asp Val Ala Asp
 [1570] 20 25 30
 [1571] Arg Pro Asp Leu Thr Ile Val Gly Asp Ser Val Tyr Asp Ala Lys Ala
 [1572] 35 40 45
 [1573] Phe Arg Ser Glu His Pro Gly Gly Ala His Phe Val Ser Leu Phe Gly
 [1574] 50 55 60
 [1575] Gly Arg Asp Ala Thr Glu Ala Phe Met Glu Tyr His Arg Arg Ala Trp
 [1576] 65 70 75 80
 [1577] Pro Lys Ser Arg Met Ser Arg Phe His Val Gly Ser Leu Ala Ser Thr
 [1578] 85 90 95
 [1579] Glu Glu Pro Val Ala Ala Asp Glu Gly Tyr Leu Gln Leu Cys Ala Arg
 [1580] 100 105 110
 [1581] Ile Ala Lys Met Val Pro Ser Val Ser Ser Gly Phe Ala Pro Ala Ser
 [1582] 115 120 125
 [1583] Tyr Trp Val Lys Ala Gly Leu Ile Leu Gly Ser Ala Ile Ala Leu Glu
 [1584] 130 135 140
 [1585] Ala Tyr Met Leu Tyr Ala Gly Lys Arg Leu Leu Pro Ser Ile Val Leu
 [1586] 145 150 155 160
 [1587] Gly Trp Leu Phe Ala Leu Ile Gly Leu Asn Ile Gln His Asp Ala Asn
 [1588] 165 170 175
 [1589] His Gly Ala Leu Ser Lys Ser Ala Ser Val Asn Leu Ala Leu Gly Leu
 [1590] 180 185 190
 [1591] Cys Gln Asp Trp Ile Gly Gly Ser Met Ile Leu Trp Leu Gln Glu His
 [1592] 195 200 205
 [1593] Val Val Met His His Leu His Thr Asn Asp Val Asp Lys Asp Pro Asp
 [1594] 210 215 220
 [1595] Gln Lys Ala His Gly Ala Leu Arg Leu Lys Pro Thr Asp Ala Trp Ser

[1596]	225	230	235	240
[1597]	Pro Met His Trp Leu Gln His Leu Tyr Leu Leu Pro Gly Glu Thr Met			
[1598]		245	250	255
[1599]	Tyr Ala Phe Lys Leu Leu Phe Leu Asp Ile Ser Glu Leu Val Met Trp			
[1600]		260	265	270
[1601]	Arg Trp Glu Gly Glu Pro Ile Ser Lys Leu Ala Gly Tyr Leu Phe Met			
[1602]		275	280	285
[1603]	Pro Ser Leu Leu Leu Lys Leu Thr Phe Trp Ala Arg Phe Val Ala Leu			
[1604]		290	295	300
[1605]	Pro Leu Tyr Leu Ala Pro Ser Val His Thr Ala Val Cys Ile Ala Ala			
[1606]	305	310	315	320
[1607]	Thr Val Met Thr Gly Ser Phe Tyr Leu Ala Phe Phe Phe Phe Ile Ser			
[1608]		325	330	335
[1609]	His Asn Phe Glu Gly Val Ala Ser Val Gly Pro Asp Gly Ser Ile Thr			
[1610]		340	345	350
[1611]	Ser Met Thr Arg Gly Ala Ser Phe Leu Lys Arg Gln Ala Glu Thr Ser			
[1612]		355	360	365
[1613]	Ser Asn Val Gly Gly Pro Leu Leu Ala Thr Leu Asn Gly Gly Leu Asn			
[1614]		370	375	380
[1615]	Tyr Gln Ile Glu His His Leu Phe Pro Arg Val His His Gly Phe Tyr			
[1616]	385	390	395	400
[1617]	Pro Arg Leu Ala Pro Leu Val Lys Ala Glu Leu Glu Ala Arg Gly Ile			
[1618]		405	410	415
[1619]	Glu Tyr Lys His Tyr Pro Thr Ile Trp Ser Asn Leu Ala Ser Thr Leu			
[1620]		420	425	430
[1621]	Arg His Met Tyr Ala Leu Gly Arg Arg Pro Arg Ser Lys Ala Glu			
[1622]		435	440	445
[1623]	<210> 74			
[1624]	<211> 1563			
[1625]	<212> DNA			
[1626]	<213> Arabidopsis thaliana			
[1627]	<400> 74			
[1628]	atggcgattt tggattctgc tggcggtact acggtgacgg agaacggtgg cggagagttc	60		
[1629]	gtcgatcttg ataggcttcg tcgacggaaa tcgagatcgg attcttctaa cggacttctt	120		
[1630]	ctctctggtt ccgataataa ttctccttcg gatgatgttg gagctccgc cgacgttagg	180		
[1631]	gatcggttg attccgttgt taacgatgac gctcaggga cagccaattt ggccggagat	240		
[1632]	aataacggtg gtggcgataa taacggttgt ggaagaggcg gcggagaagg aagaggaaac	300		
[1633]	gccgatgcta cgtttacgta tcgaccgtcg gttccagctc atcggagggc gagagagagt	360		
[1634]	ccacttagct ccgacgcaat cttcaaacag agccatgccg gattattcaa cctctgtgta	420		
[1635]	gtagttctta ttgctgtaaa cagtagactc atcatcgaaa atcttatgaa gtatggttgg	480		
[1636]	ttgatcagaa cggatttctg gtttagttca agatcgctgc gagattggcc gcttttcagt	540		
[1637]	tgttgtatat ccctttcgat ctttcctttg gctgccttta cggttgagaa attggtactt	600		

[1638]	cagaaataca tatcagaacc tgttgtcatc tttcttcata ttattatcac catgacagag	660
[1639]	gttttgtatc cagtttacgt caccctaagg tgtgattctg cttttttatc aggtgtcact	720
[1640]	ttgatgctcc tcacttgcac tgtgtggcta aagttggttt cttatgctca tactagctat	780
[1641]	gacataagat ccctagccaa tgcagctgat aaggccaatc ctgaagtctc ctactacgtt	840
[1642]	agcttgaaga gcttggcata tttcatggtc gctccacat tgtgttatca gccaagttat	900
[1643]	ccacgttctg catgtatacg gaagggttg gtggctcgtc aatttgcaaa actggtcata	960
[1644]	ttcaccgat tcatgggatt tataatagaa caatatataa atcctattgt caggaactca	1020
[1645]	aagcatcctt tgaaaggcga tcttctatat gctattgaaa gagtgttgaa gctttcagtt	1080
[1646]	ccaaatttat atgtgtggct ctgcatgttc tactgcttct tccaccttg gttaacata	1140
[1647]	ttggcagagc ttctctgctt cggggatcgt gaattctaca aagattggtg gaatgcaaaa	1200
[1648]	agtgtgggag attactggag aatgtggaat atgcctgttc ataaatggat ggttcgacat	1260
[1649]	atatacttcc cgtgcttgcg cagcaagata ccaaagacac tcgccattat cattgctttc	1320
[1650]	ctagtctctg cagtctttca tgagctatgc atgcagttc cttgtcgtct cttcaagcta	1380
[1651]	tgggcttttc ttgggattat gtttcaggtg cctttggtct tcatcacaaa ctatctacag	1440
[1652]	gaaaggtttg gctcaacggt ggggaacatg atcttctggt tcatcttctg cattttcgga	1500
[1653]	caaccgatgt gtgtgcttct ttattaccac gacctgatga accgaaaagg atcgatgtca	1560
[1654]	tga	1563
[1655]	<210>	75
[1656]	<211>	520
[1657]	<212>	PRT
[1658]	<213>	Arabidopsis thaliana
[1659]	<400>	75
[1660]	Met Ala Ile Leu Asp Ser Ala Gly Val Thr Thr Val Thr Glu Asn Gly	
[1661]	1 5 10 15	
[1662]	Gly Gly Glu Phe Val Asp Leu Asp Arg Leu Arg Arg Arg Lys Ser Arg	
[1663]	20 25 30	
[1664]	Ser Asp Ser Ser Asn Gly Leu Leu Leu Ser Gly Ser Asp Asn Asn Ser	
[1665]	35 40 45	
[1666]	Pro Ser Asp Asp Val Gly Ala Pro Ala Asp Val Arg Asp Arg Ile Asp	
[1667]	50 55 60	
[1668]	Ser Val Val Asn Asp Asp Ala Gln Gly Thr Ala Asn Leu Ala Gly Asp	
[1669]	65 70 75 80	
[1670]	Asn Asn Gly Gly Gly Asp Asn Asn Gly Gly Gly Arg Gly Gly Gly Glu	
[1671]	85 90 95	
[1672]	Gly Arg Gly Asn Ala Asp Ala Thr Phe Thr Tyr Arg Pro Ser Val Pro	
[1673]	100 105 110	
[1674]	Ala His Arg Arg Ala Arg Glu Ser Pro Leu Ser Ser Asp Ala Ile Phe	
[1675]	115 120 125	
[1676]	Lys Gln Ser His Ala Gly Leu Phe Asn Leu Cys Val Val Val Leu Ile	
[1677]	130 135 140	
[1678]	Ala Val Asn Ser Arg Leu Ile Ile Glu Asn Leu Met Lys Tyr Gly Trp	
[1679]	145 150 155 160	

[1680]	Leu	Ile	Arg	Thr	Asp	Phe	Trp	Phe	Ser	Ser	Arg	Ser	Leu	Arg	Asp	Trp
[1681]						165				170					175	
[1682]	Pro	Leu	Phe	Met	Cys	Cys	Ile	Ser	Leu	Ser	Ile	Phe	Pro	Leu	Ala	Ala
[1683]						180				185					190	
[1684]	Phe	Thr	Val	Glu	Lys	Leu	Val	Leu	Gln	Lys	Tyr	Ile	Ser	Glu	Pro	Val
[1685]						195				200					205	
[1686]	Val	Ile	Phe	Leu	His	Ile	Ile	Ile	Thr	Met	Thr	Glu	Val	Leu	Tyr	Pro
[1687]						210				215					220	
[1688]	Val	Tyr	Val	Thr	Leu	Arg	Cys	Asp	Ser	Ala	Phe	Leu	Ser	Gly	Val	Thr
[1689]						225				230					235	
[1690]	Leu	Met	Leu	Leu	Thr	Cys	Ile	Val	Trp	Leu	Lys	Leu	Val	Ser	Tyr	Ala
[1691]						245				250					255	
[1692]	His	Thr	Ser	Tyr	Asp	Ile	Arg	Ser	Leu	Ala	Asn	Ala	Ala	Asp	Lys	Ala
[1693]						260				265					270	
[1694]	Asn	Pro	Glu	Val	Ser	Tyr	Tyr	Val	Ser	Leu	Lys	Ser	Leu	Ala	Tyr	Phe
[1695]						275				280					285	
[1696]	Met	Val	Ala	Pro	Thr	Leu	Cys	Tyr	Gln	Pro	Ser	Tyr	Pro	Arg	Ser	Ala
[1697]						290				295					300	
[1698]	Cys	Ile	Arg	Lys	Gly	Trp	Val	Ala	Arg	Gln	Phe	Ala	Lys	Leu	Val	Ile
[1699]						305				310					315	
[1700]	Phe	Thr	Gly	Phe	Met	Gly	Phe	Ile	Ile	Glu	Gln	Tyr	Ile	Asn	Pro	Ile
[1701]						325				330					335	
[1702]	Val	Arg	Asn	Ser	Lys	His	Pro	Leu	Lys	Gly	Asp	Leu	Leu	Tyr	Ala	Ile
[1703]						340				345					350	
[1704]	Glu	Arg	Val	Leu	Lys	Leu	Ser	Val	Pro	Asn	Leu	Tyr	Val	Trp	Leu	Cys
[1705]						355				360					365	
[1706]	Met	Phe	Tyr	Cys	Phe	Phe	His	Leu	Trp	Leu	Asn	Ile	Leu	Ala	Glu	Leu
[1707]						370				375					380	
[1708]	Leu	Cys	Phe	Gly	Asp	Arg	Glu	Phe	Tyr	Lys	Asp	Trp	Trp	Asn	Ala	Lys
[1709]						385				390					395	
[1710]	Ser	Val	Gly	Asp	Tyr	Trp	Arg	Met	Trp	Asn	Met	Pro	Val	His	Lys	Trp
[1711]						405				410					415	
[1712]	Met	Val	Arg	His	Ile	Tyr	Phe	Pro	Cys	Leu	Arg	Ser	Lys	Ile	Pro	Lys
[1713]						420				425					430	
[1714]	Thr	Leu	Ala	Ile	Ile	Ile	Ala	Phe	Leu	Val	Ser	Ala	Val	Phe	His	Glu
[1715]						435				440					445	
[1716]	Leu	Cys	Ile	Ala	Val	Pro	Cys	Arg	Leu	Phe	Lys	Leu	Trp	Ala	Phe	Leu
[1717]						450				455					460	
[1718]	Gly	Ile	Met	Phe	Gln	Val	Pro	Leu	Val	Phe	Ile	Thr	Asn	Tyr	Leu	Gln
[1719]						465				470					475	
[1720]	Glu	Arg	Phe	Gly	Ser	Thr	Val	Gly	Asn	Met	Ile	Phe	Trp	Phe	Ile	Phe
[1721]						485				490					495	

[1722] Cys Ile Phe Gly Gln Pro Met Cys Val Leu Leu Tyr Tyr His Asp Leu
 [1723] 500 505 510
 [1724] Met Asn Arg Lys Gly Ser Met Ser
 [1725] 515 520
 [1726] <210> 76
 [1727] <211> 9
 [1728] <212> PRT
 [1729] <213> Artificial Sequence
 [1730] <220>
 [1731] <223> Elongase consensus domain 7
 [1732] <400> 76
 [1733] Val Asp Thr Arg Lys Gly Ala Tyr Arg
 [1734] 1 5
 [1735] <210> 77
 [1736] <211> 9
 [1737] <212> PRT
 [1738] <213> Artificial Sequence
 [1739] <220>
 [1740] <223> Elongase consensus domain 8
 [1741] <400> 77
 [1742] Phe Ile His Thr Ile Met Tyr Thr Tyr
 [1743] 1 5
 [1744] <210> 78
 [1745] <211> 28
 [1746] <212> DNA
 [1747] <213> Artificial Sequence
 [1748] <220>
 [1749] <223> Oligonucleotide primer
 [1750] <220>
 [1751] <221> misc_feature
 [1752] <222> (20) .. (20)
 [1753] <223> n is a, c, g, or t
 [1754] <220>
 [1755] <221> misc_feature
 [1756] <222> (23) .. (23)
 [1757] <223> n is a, c, g, or t
 [1758] <400> 78
 [1759] tgggtggacac aaggaagggn gcntaymg 28
 [1760] <210> 79
 [1761] <211> 27
 [1762] <212> DNA
 [1763] <213> Artificial Sequence

[1764]	<220>
[1765]	<223> Oligonucleotide primer
[1766]	<400> 79
[1767]	gtaggtgtac atgatggtrt gdatraa 27
[1768]	<210> 80
[1769]	<211> 24
[1770]	<212> DNA
[1771]	<213> Artificial Sequence
[1772]	<220>
[1773]	<223> Oligonucleotide primer
[1774]	<400> 80
[1775]	gtccttgetc cagggttcc acca 24
[1776]	<210> 81
[1777]	<211> 36
[1778]	<212> DNA
[1779]	<213> Artificial Sequence
[1780]	<220>
[1781]	<223> Oligonucleotide primer
[1782]	<400> 81
[1783]	atttaggtga cactatagtt tttttttttt tttttt 36
[1784]	<210> 82
[1785]	<211> 23
[1786]	<212> DNA
[1787]	<213> Artificial Sequence
[1788]	<220>
[1789]	<223> Oligonucleotide primer
[1790]	<400> 82
[1791]	ttccagaacg agggcatcta cgt 23
[1792]	<210> 83
[1793]	<211> 25
[1794]	<212> DNA
[1795]	<213> Artificial Sequence
[1796]	<220>
[1797]	<223> Oligonucleotide primer
[1798]	<400> 83
[1799]	ttgggtgatc tgcattgagcg tgatg 25
[1800]	<210> 84
[1801]	<211> 26
[1802]	<212> DNA
[1803]	<213> Artificial Sequence
[1804]	<220>
[1805]	<223> Oligonucleotide primer

[1806]	<400> 84
[1807]	cgaatacttg aagagcttgt tggaga 26
[1808]	<210> 85
[1809]	<211> 25
[1810]	<212> DNA
[1811]	<213> Artificial Sequence
[1812]	<220>
[1813]	<223> Oligonucleotide primer
[1814]	<400> 85
[1815]	gggctacgag ctggcagatg aagca 25
[1816]	<210> 86
[1817]	<211> 25
[1818]	<212> DNA
[1819]	<213> Artificial Sequence
[1820]	<220>
[1821]	<223> Oligonucleotide primer
[1822]	<400> 86
[1823]	gaaaaaatgg ttgcgccacc catca 25
[1824]	<210> 87
[1825]	<211> 27
[1826]	<212> DNA
[1827]	<213> Artificial Sequence
[1828]	<220>
[1829]	<223> Oligonucleotide primer
[1830]	<400> 87
[1831]	tcactacttc ttcttcttgc ccgcggc 27
[1832]	<210> 88
[1833]	<211> 21
[1834]	<212> DNA
[1835]	<213> Artificial Sequence
[1836]	<220>
[1837]	<223> Oligonucleotide primer
[1838]	<400> 88
[1839]	ttccgtact cagcgtggc g 21
[1840]	<210> 89
[1841]	<211> 17
[1842]	<212> DNA
[1843]	<213> Artificial Sequence
[1844]	<220>
[1845]	<223> Oligonucleotide primer
[1846]	<400> 89
[1847]	caggaaacag ctatgac 17

[1848]	<210> 90
[1849]	<211> 20
[1850]	<212> DNA
[1851]	<213> Artificial Sequence
[1852]	<220>
[1853]	<223> Oligonucleotide primer
[1854]	<400> 90
[1855]	acgtagatgc cctcgttctg 20
[1856]	<210> 91
[1857]	<211> 26
[1858]	<212> DNA
[1859]	<213> Artificial Sequence
[1860]	<220>
[1861]	<223> Oligonucleotide primer
[1862]	<400> 91
[1863]	caccgaatgg cgactgaagg gatgcc 26
[1864]	<210> 92
[1865]	<211> 27
[1866]	<212> DNA
[1867]	<213> Artificial Sequence
[1868]	<220>
[1869]	<223> Oligonucleotide primer
[1870]	<400> 92
[1871]	ctactcgggtt ttcatgcggt tgctgga 27
[1872]	<210> 93
[1873]	<211> 819
[1874]	<212> DNA
[1875]	<213> Pavlova pinguis
[1876]	<400> 93
[1877]	atggttgcgc caccatcac gctcgagtgg ctgctttcgc cgaagctcaa ggatgcagtg 60
[1878]	ttcgggtgggg aggtgctcta cttctccatt gcctacctgt ttcttgccgc cattttgaag 120
[1879]	cgcaccccggt tgggtggacac gcggaagggc gcgtataaga gtggtatgat cgcgtacaac 180
[1880]	gtgatcatgt gcgtgttctc gctgggtgtgc ttcatctgcc agctcgcagc cctgggctat 240
[1881]	gacatgggct acttgcagtg ggtgcgtgac ctcacagggg acgagattgt cccctctac 300
[1882]	caggacgtgt ccccgctccc cgccttctcc aacaagctct tcaagtattc gtctattgcc 360
[1883]	ttccactact ccaagtatgt tgagtacatg gacaccgcat ggctggtgat gaagggaag 420
[1884]	cccgtgtcct tgctccaggg cttccaccac tttggcgccg cctgggacac ctactttggc 480
[1885]	atcaccttcc agaacgaggg catctacgtg ttctgtgtgc tcaacgcctt catecacag 540
[1886]	atcatgtacg catactacgc ggccactgcg gcgggtctca agttctcact gaagtctgc 600
[1887]	atcacgtca tgcagatcac ccaattcaac gtgggcttcg taatggtgta tcactacatc 660
[1888]	accttgaggt acttccgcaa ctcaccggag ctcgtcttct cctacctttt caactatgcg 720
[1889]	tacgtctgca cggttctcct cctcttcatg cagttcttct acatggacaa ctttggcaag 780

[1890] aagaaggccg ctgccgccgc gggcaagaag aagaagtag 819
 [1891] <210> 94
 [1892] <211> 272
 [1893] <212> PRT
 [1894] <213> Pavlova pinguis
 [1895] <400> 94
 [1896] Met Val Ala Pro Pro Ile Thr Leu Glu Trp Leu Leu Ser Pro Lys Leu
 [1897] 1 5 10 15
 [1898] Lys Asp Ala Val Phe Gly Gly Glu Val Leu Tyr Phe Ser Ile Ala Tyr
 [1899] 20 25 30
 [1900] Leu Phe Leu Ala Pro Ile Leu Lys Arg Thr Pro Leu Val Asp Thr Arg
 [1901] 35 40 45
 [1902] Lys Gly Ala Tyr Lys Ser Gly Met Ile Ala Tyr Asn Val Ile Met Cys
 [1903] 50 55 60
 [1904] Val Phe Ser Leu Val Cys Phe Ile Cys Gln Leu Ala Ala Leu Gly Tyr
 [1905] 65 70 75 80
 [1906] Asp Met Gly Tyr Leu Gln Trp Val Arg Asp Leu Thr Gly Asp Glu Ile
 [1907] 85 90 95
 [1908] Val Pro Leu Tyr Gln Asp Val Ser Pro Ser Pro Ala Phe Ser Asn Lys
 [1909] 100 105 110
 [1910] Leu Phe Lys Tyr Ser Ser Ile Ala Phe His Tyr Ser Lys Tyr Val Glu
 [1911] 115 120 125
 [1912] Tyr Met Asp Thr Ala Trp Leu Val Met Lys Gly Lys Pro Val Ser Leu
 [1913] 130 135 140
 [1914] Leu Gln Gly Phe His His Phe Gly Ala Ala Trp Asp Thr Tyr Phe Gly
 [1915] 145 150 155 160
 [1916] Ile Thr Phe Gln Asn Glu Gly Ile Tyr Val Phe Val Val Leu Asn Ala
 [1917] 165 170 175
 [1918] Phe Ile His Thr Ile Met Tyr Ala Tyr Tyr Ala Ala Thr Ala Ala Gly
 [1919] 180 185 190
 [1920] Leu Lys Phe Ser Leu Lys Phe Val Ile Thr Leu Met Gln Ile Thr Gln
 [1921] 195 200 205
 [1922] Phe Asn Val Gly Phe Val Met Val Tyr His Tyr Ile Thr Leu Glu Tyr
 [1923] 210 215 220
 [1924] Phe Arg Asn Ser Pro Glu Leu Val Phe Ser Tyr Leu Phe Asn Tyr Ala
 [1925] 225 230 235 240
 [1926] Tyr Val Cys Thr Val Leu Leu Leu Phe Met Gln Phe Phe Tyr Met Asp
 [1927] 245 250 255
 [1928] Asn Phe Gly Lys Lys Lys Ala Ala Ala Ala Ala Gly Lys Lys Lys Lys
 [1929] 260 265 270
 [1930] <210> 95
 [1931] <211> 840

[1932]	<212> DNA
[1933]	<213> Pavlova salina
[1934]	<400> 95
[1935]	atggcgactg aagggatgcc ggcgataacg ctggactggc tgctctcgcc cgggctgaag 60
[1936]	gatgccgtaa ttggcgggga ggtgctctac ttttcgcttg ggtatctgct gctcgagccc 120
[1937]	atcctcaagc gctcaccgtt tgtggacaag cgcaaggcg cataccgcaa cggcatgata 180
[1938]	gcgtacaaca tcctcatgtg cggttttctg ctggatatgt tcgtgtgcca gatggcggcg 240
[1939]	ctcggccttg atcgcggcca cctgcagttt gtccgcgacc tcacgggcga cagcgtggtg 300
[1940]	cagctctacc aggacgtgag cccatcccct gcattcgca acaagctctt ccggtactca 360
[1941]	gcggtggcgt tccactactc aaagtacgtg gagtacatgg acacagcgtg gcttgtgctg 420
[1942]	aagggaagc ccgtctcggt cctgcagggc ttccaccact tcggcgccgc gtgggacacc 480
[1943]	tactttggca tcacgtttca gaacgagggc acctacgtct ttgtgtgct caacgcattc 540
[1944]	atccacacaa tcatgtacac ctactacggc gcgacggcag cgggcatcaa aatctcgatg 600
[1945]	aagccgctga tcaccctcat gcagatcacg cagttcctgc tgggcttcgc gctcgtctac 660
[1946]	ccgtacattg acctcgcta cttccgtgcg tcgcccagc tcgtgtggag ctacctgttc 720
[1947]	aactatgcgt acgtactcat ggtgctcttc ctcttcacgc gcttcttcta ccacgacaac 780
[1948]	tttagcaagc acaagccaat ctgcgcgcatc gactccagca accgcatgaa aaccgagtag 840
[1949]	<210> 96
[1950]	<211> 279
[1951]	<212> PRT
[1952]	<213> Pavlova salina
[1953]	<400> 96
[1954]	Met Ala Thr Glu Gly Met Pro Ala Ile Thr Leu Asp Trp Leu Leu Ser
[1955]	1 5 10 15
[1956]	Pro Gly Leu Lys Asp Ala Val Ile Gly Gly Glu Val Leu Tyr Phe Ser
[1957]	20 25 30
[1958]	Leu Gly Tyr Leu Leu Leu Glu Pro Ile Leu Lys Arg Ser Pro Phe Val
[1959]	35 40 45
[1960]	Asp Lys Arg Lys Gly Ala Tyr Arg Asn Gly Met Ile Ala Tyr Asn Ile
[1961]	50 55 60
[1962]	Leu Met Cys Gly Phe Ser Leu Val Cys Phe Val Cys Gln Met Ala Ala
[1963]	65 70 75 80
[1964]	Leu Gly Leu Asp Arg Gly His Leu Gln Phe Val Arg Asp Leu Thr Gly
[1965]	85 90 95
[1966]	Asp Ser Val Val Gln Leu Tyr Gln Asp Val Ser Pro Ser Pro Ala Phe
[1967]	100 105 110
[1968]	Ala Asn Lys Leu Phe Arg Tyr Ser Ala Val Ala Phe His Tyr Ser Lys
[1969]	115 120 125
[1970]	Tyr Val Glu Tyr Met Asp Thr Ala Trp Leu Val Leu Lys Gly Lys Pro
[1971]	130 135 140
[1972]	Val Ser Phe Leu Gln Gly Phe His His Phe Gly Ala Ala Trp Asp Thr
[1973]	145 150 155 160

[1974]	Tyr Phe Gly Ile Thr Phe Gln Asn Glu Gly Thr Tyr Val Phe Val Leu
[1975]	165 170 175
[1976]	Leu Asn Ala Phe Ile His Thr Ile Met Tyr Thr Tyr Tyr Gly Ala Thr
[1977]	180 185 190
[1978]	Ala Ala Gly Ile Lys Ile Ser Met Lys Pro Leu Ile Thr Leu Met Gln
[1979]	195 200 205
[1980]	Ile Thr Gln Phe Leu Leu Gly Phe Ala Leu Val Tyr Pro Tyr Ile Asp
[1981]	210 215 220
[1982]	Leu Gly Tyr Phe Arg Ala Ser Pro Glu Leu Val Trp Ser Tyr Leu Phe
[1983]	225 230 235 240
[1984]	Asn Tyr Ala Tyr Val Leu Met Val Leu Phe Leu Phe Met Arg Phe Phe
[1985]	245 250 255
[1986]	Tyr His Asp Asn Phe Ser Lys His Lys Pro Ile Ser Arg Ile Asp Ser
[1987]	260 265 270
[1988]	Ser Asn Arg Met Lys Thr Glu
[1989]	275
[1990]	<210> 97
[1991]	<211> 172
[1992]	<212> PRT
[1993]	<213> Tomato bushy stunt virus
[1994]	<400> 97
[1995]	Met Glu Arg Ala Ile Gln Gly Asn Asp Ala Arg Glu Gln Ala Asn Ser
[1996]	1 5 10 15
[1997]	Glu Arg Trp Asp Gly Gly Ser Gly Gly Thr Thr Ser Pro Phe Lys Leu
[1998]	20 25 30
[1999]	Pro Asp Glu Ser Pro Ser Trp Thr Glu Trp Arg Leu His Asn Asp Glu
[2000]	35 40 45
[2001]	Thr Asn Ser Asn Gln Asp Asn Pro Leu Gly Phe Lys Glu Ser Trp Gly
[2002]	50 55 60
[2003]	Phe Gly Lys Val Val Phe Lys Arg Tyr Leu Arg Tyr Asp Arg Thr Glu
[2004]	65 70 75 80
[2005]	Ala Ser Leu His Arg Val Leu Gly Ser Trp Thr Gly Asp Ser Val Asn
[2006]	85 90 95
[2007]	Tyr Ala Ala Ser Arg Phe Phe Gly Phe Asp Gln Ile Gly Cys Thr Tyr
[2008]	100 105 110
[2009]	Ser Ile Arg Phe Arg Gly Val Ser Ile Thr Val Ser Gly Gly Ser Arg
[2010]	115 120 125
[2011]	Thr Leu Gln His Leu Cys Glu Met Ala Ile Arg Ser Lys Gln Glu Leu
[2012]	130 135 140
[2013]	Leu Gln Leu Ala Pro Ile Glu Val Glu Ser Asn Val Ser Arg Gly Cys
[2014]	145 150 155 160
[2015]	Pro Glu Gly Thr Glu Thr Phe Glu Lys Glu Ser Glu

[2016]		165	170
[2017]	<210>	98	
[2018]	<211>	116	
[2019]	<212>	PRT	
[2020]	<213>	Tomato yellow leaf curl virus	
[2021]	<400>	98	
[2022]	Met Trp Asp Pro Leu Leu Asn Glu Phe Pro Glu Ser Val His Gly Phe		
[2023]	1	5	10 15
[2024]	Arg Cys Met Leu Ala Ile Lys Tyr Leu Gln Ser Val Glu Glu Thr Tyr		
[2025]		20	25 30
[2026]	Glu Pro Asn Thr Leu Gly His Asp Leu Ile Arg Asp Leu Ile Ser Val		
[2027]		35	40 45
[2028]	Val Arg Ala Arg Asp Tyr Val Glu Ala Thr Arg Arg Tyr Asn His Phe		
[2029]		50	55 60
[2030]	His Ala Arg Leu Glu Gly Ser Pro Lys Ala Glu Leu Arg Gln Pro Ile		
[2031]	65	70	75 80
[2032]	Gln Gln Pro Cys Cys Cys Pro His Cys Pro Arg His Lys Gln Ala Thr		
[2033]		85	90 95
[2034]	Ile Met Asp Val Gln Ala His Val Pro Glu Ala Gln Asn Ile Gln Asn		
[2035]		100	105 110
[2036]	Val Ser Lys Pro		
[2037]		115	
[2038]	<210>	99	
[2039]	<211>	351	
[2040]	<212>	PRT	
[2041]	<213>	Turnip crinkle virus	
[2042]	<400>	99	
[2043]	Met Glu Asn Asp Pro Arg Val Arg Lys Phe Ala Ser Glu Gly Ala Gln		
[2044]	1	5	10 15
[2045]	Trp Ala Ile Lys Trp Gln Lys Lys Gly Trp Ser Ser Leu Thr Ser Arg		
[2046]		20	25 30
[2047]	Gln Lys Gln Thr Ala Arg Ala Ala Met Gly Ile Lys Leu Ser Pro Val		
[2048]		35	40 45
[2049]	Ala Gln Pro Val Gln Lys Val Thr Arg Leu Ser Ala Pro Val Ala Leu		
[2050]		50	55 60
[2051]	Ala Tyr Arg Glu Val Ser Thr Gln Pro Arg Val Ser Thr Ala Arg Asp		
[2052]	65	70	75 80
[2053]	Gly Ile Thr Arg Ser Gly Ser Glu Leu Ile Thr Thr Leu Lys Lys Asn		
[2054]		85	90 95
[2055]	Thr Asp Thr Glu Pro Lys Tyr Thr Thr Ala Val Leu Asn Pro Ser Glu		
[2056]		100	105 110
[2057]	Pro Gly Thr Phe Asn Gln Leu Ile Lys Glu Ala Ala Gln Tyr Glu Lys		

[2058]	115	120	125
[2059]	Tyr Arg Phe Thr Ser Leu Arg Phe Arg Tyr Ser Pro Met Ser Pro Ser		
[2060]	130	135	140
[2061]	Thr Thr Gly Gly Lys Val Ala Leu Ala Phe Asp Arg Asp Ala Ala Lys		
[2062]	145	150	155
[2063]	Pro Pro Pro Asn Asp Leu Ala Ser Leu Tyr Asn Ile Glu Gly Cys Val		
[2064]	165	170	175
[2065]	Ser Ser Val Pro Trp Thr Gly Phe Ile Leu Thr Val Pro Thr Asp Ser		
[2066]	180	185	190
[2067]	Thr Asp Arg Phe Val Ala Asp Gly Ile Ser Asp Pro Lys Leu Val Asn		
[2068]	195	200	205
[2069]	Phe Gly Lys Leu Ile Met Ala Thr Tyr Gly Gln Gly Ala Asn Asp Ala		
[2070]	210	215	220
[2071]	Ala Gln Leu Gly Glu Val Arg Val Glu Tyr Thr Val Gln Leu Lys Asn		
[2072]	225	230	235
[2073]	Arg Thr Gly Ser Thr Ser Asp Ala Gln Ile Gly Asp Phe Ala Gly Val		
[2074]	245	250	255
[2075]	Lys Asp Gly Pro Arg Leu Val Ser Trp Ser Lys Thr Lys Gly Thr Ala		
[2076]	260	265	270
[2077]	Gly Trp Glu His Asp Cys His Phe Leu Gly Thr Gly Asn Phe Ser Leu		
[2078]	275	280	285
[2079]	Thr Leu Phe Tyr Glu Lys Ala Pro Val Ser Gly Leu Glu Asn Ala Asp		
[2080]	290	295	300
[2081]	Ala Ser Asp Phe Ser Val Leu Gly Glu Ala Ala Ala Gly Ser Val Gln		
[2082]	305	310	315
[2083]	Trp Ala Gly Val Lys Val Ala Glu Arg Gly Gln Ser Val Lys Met Val		
[2084]	325	330	335
[2085]	Thr Thr Glu Glu Gln Pro Arg Gly Lys Trp Gln Ala Leu Arg Ile		
[2086]	340	345	350
[2087]	<210> 100		
[2088]	<211> 303		
[2089]	<212> PRT		
[2090]	<213> Pea mosaic virus		
[2091]	<400> 100		
[2092]	Met His Gly Ile Glu Gln Pro Gln Leu Pro Leu Asp Tyr Val His Arg		
[2093]	1	5	10
[2094]	Cys Ala Ser Thr Ser Phe Leu Leu Ala Ser Leu Asp Gly Leu Leu Ser		
[2095]	20	25	30
[2096]	Glu Ala Arg Glu Leu Ser Gly Pro Leu Ala Leu Ile Thr Ser Ser Tyr		
[2097]	35	40	45
[2098]	Tyr Leu Leu Val Ser Ile Ala Leu Cys Trp Ala Ile Pro Gly Ser Phe		
[2099]	50	55	60

[2100]	Trp Tyr Arg Pro Gly Cys Trp Leu Gln Pro Val Ser Gly Arg Asn Leu
[2101]	65 70 75 80
[2102]	Ile Phe Cys Gly Pro Thr Glu Ala Leu Gln Arg Phe Arg Leu Tyr Ala
[2103]	85 90 95
[2104]	Ala Arg Leu Gly Leu Val Leu Ser Glu Asn Cys Pro Arg His Gly Gln
[2105]	100 105 110
[2106]	Ser Ala Ala Ile Thr Leu Gln Ser Tyr Trp Ala Leu Pro Asn Asn Ile
[2107]	115 120 125
[2108]	Trp Met Asp Met Ala Gln Leu Asp Leu Leu Thr Phe Ser Met Pro Ile
[2109]	130 135 140
[2110]	Ala Asn Thr Phe Ala Tyr Leu Ala Asp Cys Glu Ala Arg Phe Pro Pro
[2111]	145 150 155 160
[2112]	Ile Val Glu Gly Val Gly Ser Ala Tyr Tyr Val Pro Thr Leu Leu Gly
[2113]	165 170 175
[2114]	Leu Thr His Gln Asp Pro Arg Leu Tyr Leu Ala Leu Arg Arg Arg Asn
[2115]	180 185 190
[2116]	Leu Asp Leu Ser Gly Glu Pro His Arg Val Arg Pro Gly Val Leu Glu
[2117]	195 200 205
[2118]	Ser Met Ala Leu Leu Cys Ser Ser Val Arg Ser Thr Ser Arg Ser Arg
[2119]	210 215 220
[2120]	Gln Ile Pro Pro Leu Tyr Gly Ser Val Leu His His Val Leu Gly Leu
[2121]	225 230 235 240
[2122]	Ala Glu Arg Asp Cys Ile Leu Phe Asp Thr Asp Ser Asn Tyr Ser Ser
[2123]	245 250 255
[2124]	Tyr Thr His Arg Val Leu Glu Gln Asp Arg Asn Arg Ala Asp Gln Ser
[2125]	260 265 270
[2126]	Leu Phe Ser Ile Asp Leu Glu Tyr Val His Asp Leu Glu Leu Ile Ala
[2127]	275 280 285
[2128]	Leu Gly Tyr Ser Asp Glu Asp Asp Glu Asp Leu Asp Asn Phe Phe
[2129]	290 295 300
[2130]	<210> 101
[2131]	<211> 256
[2132]	<212> PRT
[2133]	<213> Cereal yellow dwarf virus
[2134]	<400> 101
[2135]	Met Phe Ile Ala Gln Pro Cys Gly Arg Val Leu Val Phe Asp Val Ala
[2136]	1 5 10 15
[2137]	Ser Arg Thr Pro Ser Phe Phe Thr Arg Tyr Ser Val Glu Leu Ser Leu
[2138]	20 25 30
[2139]	Arg Val Leu Asp Pro Phe Phe Thr Arg Ala Val Thr Asp Phe Arg Tyr
[2140]	35 40 45
[2141]	Thr Gln Asn Glu Ile Asp Leu Phe Cys Val Ser Leu Gly Phe Leu Leu

[2142]	50	55	60
[2143]	Pro Ile Leu Leu Thr Gly Glu Ser Tyr Ser Trp Arg Gly His Leu Asn		
[2144]	65	70	75
[2145]	Leu Pro Leu Ser Tyr Thr Glu Leu Leu Val Arg Trp Gly Leu Ala Val		
[2146]	85	90	95
[2147]	Gly Tyr Phe Pro Thr Phe Ser Thr Asp Gly Asp Ile Arg Gln Asn Pro		
[2148]	100	105	110
[2149]	Glu Leu Arg Ile Asp Leu Ser Thr Met Ser Thr Arg Ser Phe Tyr Glu		
[2150]	115	120	125
[2151]	Gln Phe Leu Leu Arg Tyr Asn Thr Ser Gly Leu Ala Lys Ala Ile Val		
[2152]	130	135	140
[2153]	Gly Gln Gln Glu Cys Phe Gln Ser Gly Met Glu Ser Phe Lys Arg Phe		
[2154]	145	150	155
[2155]	Leu His Tyr Arg Leu Thr Cys Phe Glu Ser Cys Leu Pro Arg Pro Arg		
[2156]	165	170	175
[2157]	Trp Glu Ser Pro Leu Ala Pro Gly Pro Tyr Leu Asp Arg Ala Phe Glu		
[2158]	180	185	190
[2159]	Ala Thr Leu Leu Gly Arg Met Val Gly His Asn Gln Leu Leu Phe Thr		
[2160]	195	200	205
[2161]	Gly Leu Ser Ser Asp Ile Thr Arg Tyr Tyr Asn Glu Leu Val Val Glu		
[2162]	210	215	220
[2163]	Gly Val Pro Val Ala Phe Trp Asp Ala Ala Gly Ile Thr Leu His His		
[2164]	225	230	235
[2165]	Ala Gly Glu Glu Tyr Phe Ser Asn Ser Tyr Ile Gln Lys Ile Leu Gln		
[2166]	245	250	255
[2167]	<210> 102		
[2168]	<211> 519		
[2169]	<212> DNA		
[2170]	<213> Tomato bushy stunt virus		
[2171]	<400> 102		
[2172]	atggaacgag ctatacaagg aaacgacgct agggacaacg ctaacagtga acgttgggat	60	
[2173]	ggaggatcag gaggtaccac ttctcccttc aaacttcctg acgaaagtcc gagttggact	120	
[2174]	gagtggcggc tacataacga tgagacgaat tcgaatcaag ataatccctt tggtttcaag	180	
[2175]	gaaagctggg gtttcgggaa agttgtatit aagagatatc tcagatacga caggacggaa	240	
[2176]	gcttcactgc acagagtcct tggatcttgg acgggagatt cggttaacta tgcagcatct	300	
[2177]	cgatttttcg gtttcgacca gatcgatgt acctatagta ttcggtttcg aggagttagt	360	
[2178]	atcaccgttt ctggagggtc gcgaactcct cagcatctct gtgagatggc aattcgggtc	420	
[2179]	aagcaagaac tgctacagct tgccccaatc gaagtggaaa gtaatgtatc aagaggatgc	480	
[2180]	cctgaaggta ctgagacctt cgaaaaagaa agcgagtaa	519	
[2181]	<210> 103		
[2182]	<211> 351		
[2183]	<212> DNA		

[2184]	<213> Tomato yellow leaf curl virus
[2185]	<400> 103
[2186]	atgtgggatac cacttctaaa tgaatttcct gaatctgttc acggatttcg ttgtatgtta 60
[2187]	gctattaaat atttgacgtc cgttgaggaa acttacgagc ccaatacatt gggccacgat 120
[2188]	ttaattaggg atcttatatc tgttgtaagg gcccgtagact atgtcgaagc gaccaggcga 180
[2189]	tataatcatt tccacgccc cctcgaaggt tcgccgaagg ctgaacttcg acagcccata 240
[2190]	cagcagccgt gctgctgtcc ccattgtcca aggcacaaac aagcagcagat catggacgta 300
[2191]	caggcccatg taccggaagc ccagaatata cagaatgtat cgaagccctg a 351
[2192]	<210> 104
[2193]	<211> 1056
[2194]	<212> DNA
[2195]	<213> Turnip crinkle virus
[2196]	<400> 104
[2197]	atggaaaatg atcctagagt ccgaaagtgc gcatccgagg gcgccaatg ggcgataaag 60
[2198]	tggcagaaga agggctgtgc atccctaacc agcagacaga aacagaccgc ccgcgcagcg 120
[2199]	atggggatca agctctcccc tgtggcgcaa cctgtgcaga aagtgactcg actgagtgt 180
[2200]	ccggtgtgtc tcgcctaccg cgaggtttcc acccagcctc gggtttctac tgccagggac 240
[2201]	ggcataacca gaagcgggtc tgaactgac acaaccctga agaagaacac tgacactgaa 300
[2202]	cctaagtaca ccacagctgt gcttaaccca agcgaaccgc gaacattcaa ccaactcatc 360
[2203]	aaggaggcgg ccagtatga aaaataccga ttcacgtcac tcagatttag gtactctccc 420
[2204]	atgagccctt caaccaccg gggcaaggtg gctctggcat tcgaccgaga cgctgccaaa 480
[2205]	cctccgccc aagacctgc ttccctctac aacatagagg gttgtgtatc tagcgtgccc 540
[2206]	tggacagggt ttattttgac cgtcccaaca gattctactg accgctttgt ggcggatggt 600
[2207]	atcagcgatc caaagcttgt caatttcggc aagctcatca tggccaccta tggccaagga 660
[2208]	gccaatgatg ccgccaact cggatgaagt cgagtcgagt acaccgtgca gctcaagaac 720
[2209]	agaactggct caaccagcga cgcccagatt ggggacttcg cgggtgttaa ggacggaccc 780
[2210]	aggttgggtc cgtgggtcaa gaccaaggga acagctgggt gggagcacga ttgtcatttt 840
[2211]	ctcgaaccg gaaactctc gttgacactg ttctacgaga aggcgcccgt ttcggggcta 900
[2212]	gaaaacgcag acgctctga cttctcggtc ctgggagaag ccgcagcagg tagtgtccaa 960
[2213]	tgggctggag tgaaggtagc agaaaggga caaagcgtga aaatgggtcac aactgaggag 1020
[2214]	cagccaaggg gaaaatggca agcactcagg atttag 1056
[2215]	<210> 105
[2216]	<211> 912
[2217]	<212> DNA
[2218]	<213> Pea mosaic virus
[2219]	<400> 105
[2220]	atgcacggaa ttgagcagcc tcaactaccg ctagattacg ttcaccgttg cgcatacaac 60
[2221]	tccttcttgc tcgcatcact agatggcctc ctttctgaag cccgtgaact ctcagggcct 120
[2222]	ctggctctca ttacttctag ctattactta cttgtttcta ttgccctctg ctgggcaatc 180
[2223]	cctggatcct tctggtatag gcctggctgc tggttgcagc cagtctcagg gcggaatctc 240
[2224]	atcttttgcg gccctaccga ggccttcaa cgattccgtc tgtacgtgc cagacttggg 300
[2225]	ttggtcctgt cagagaactg cccaagacac ggccaatcag cagcaatcac ccttcaatca 360

[2226]	tactgggcac ttcctaaca catctggatg gacatggccc aattggactt gctcaccttc	420
[2227]	tcaatgccaa ttgctaatac atttgcctac ttggcagatt gtgaagcaag atttcctcct	480
[2228]	attgttgaag gagtgggatc tgcttactat gtgccaacgc tgctcgact tactaccaaa	540
[2229]	gaccccaggc tttatcttgc gcttcgcagg agaaacctg atcttagtgg cgaacctcat	600
[2230]	agagtctgct ctggtgtcct ggagtctatg gctttgctct gttctagtgt acgtagcaca	660
[2231]	agccgttcca ggcaaattcc tcctttatat ggcagcgttt tgcaccacgt tttgggcctg	720
[2232]	gccgagagag actgcatcct ctttgatacg gatagtaact actcctctta cactcatcgg	780
[2233]	gttcttgaac aagaccgga tcgggctgat cagtcattgt ttagcattga cttggaatat	840
[2234]	gttcatgacc tggagcttat tgccctgggt tactctgatg aagatgatga agatcttgat	900
[2235]	aacttcttct ag	912
[2236]	<210>	106
[2237]	<211>	771
[2238]	<212>	DNA
[2239]	<213>	Cereal yellow dwarf virus
[2240]	<400>	106
[2241]	atgttcatcg cccaaccttg cgggcgagtt cttgtgttcg acgtcgctc cgcacgcca	60
[2242]	tcgttcttca ctcttatag tgttgaactc tcgtccgtg ttctagacc attcttcacg	120
[2243]	cgagcagtaa cagatttccg atacaccaa aatgaaatcg atttattttg tgtgtctctt	180
[2244]	ggctttctgt tgccaattct cctcacagga gaatcttact cttggcgcgg tcacctaac	240
[2245]	ctccccctt cttacaccga attacttgt cgatgggggc tcgcagtgg gtacttccct	300
[2246]	accttctcca ctgatgtga cattcgacag aaccagaaac tccgcatcga cctgtccacc	360
[2247]	atgtcaacc gctctttcta cgagcagttc ctactcagat ataacacgag tgggttggca	420
[2248]	aaagctatcg tcggacagca agagtgttt caaagcggca tggagtctt taaaagattc	480
[2249]	ctacactacc gcctcacgtg ctttgaaagc tgccttcac gacctcgtt ggaaagtct	540
[2250]	ttggctcctg gtcttatct ggacagggt tttgaggcaa ctcttctcgg ccgtatggtc	600
[2251]	ggtcataacc aactactctt taccggttg tcttctgata tctactagta ttataacgag	660
[2252]	ttggttgtgg aaggcgtgcc ggtggcttt tgggacgtg ccggcattac tttgcatcac	720
[2253]	gctggtgaag aatatcttc gaattcttac attcaaaaga ttcttcaatg a	771
[2254]	<210>	107
[2255]	<211>	1005
[2256]	<212>	DNA
[2257]	<213>	Micromonas CCMP1545
[2258]	<400>	107
[2259]	atgaccgtg ctccagcttc tgctagagat cctgctcttg ctactggatc tggatctctt	60
[2260]	tacaaccacc tcatgatgtt cgcttggtt ggagttgtc atgctgtgtg tgctcttgtt	120
[2261]	gctgttgett gctacttct tcctcatcct gttgctactt gtgctatcgc tctcgttget	180
[2262]	cttgcgtctc ttactcctgt tactactcct catcctgctt ggggacttgc tatcgctaga	240
[2263]	gctatcacta aggetgctgt gagatactc cctcttacta tggaatggga ggatgagaga	300
[2264]	gcttaccttg acgtgctgc taagggtgtt cctgctgtta tcggacttga gcctcattct	360
[2265]	gttctccctt tgtctatcgt tgctttcgga aactacttct tcttcaccga atctacacct	420
[2266]	gagtgcgtga gaaactctag agctttggct actggaacca tcttcgttat ccctgttctt	480
[2267]	aagcaccttt ggtcttggct tggaatggat gctatctcta ggcgtgctat gaagactctt	540

[2268]	ctcgaatgatg gaagatctgt tctcatcatc cctgggtggag ttgctgagtg ccttcaaagt 600
[2269]	aggcctgggtg ttgagactat ctacctcaag aagagggttcg gattcggttaa gcttgctatc 660
[2270]	caaaccgggtg cttctcttgt tcttgccttc actttcggac agaccagatc ttactcttac 720
[2271]	tggcgtcttg gacctcctct tgttccatcat gttgttgctg aggttttcgc tagagcttgt 780
[2272]	agagttgctc ctatgggttt ctggggaaag tggggatctc ctatccctaa catgggttcct 840
[2273]	atgcacactg ttgttggaag gcctatccct gttaagaagc agtctgagcc ttctaacgag 900
[2274]	tacgttcaag agaagctcaa cgagttcggt gctgctatgg aatctcttta cgctagacac 960
[2275]	aagggaagc acggatacgc tgagtctact ctcgttggtc ttga 1005
[2276]	<210> 108
[2277]	<211> 334
[2278]	<212> PRT
[2279]	<213> Micromonas CCMP1545
[2280]	<400> 108
[2281]	Met Thr Ala Ala Pro Ala Ser Ala Arg Asp Pro Ala Leu Ala Thr Gly
[2282]	1 5 10 15
[2283]	Ser Gly Ser Leu Tyr Asn His Leu Met Met Phe Ala Trp Leu Gly Val
[2284]	20 25 30
[2285]	Val His Ala Val Val Ala Leu Val Ala Val Ala Cys Tyr Phe Leu Pro
[2286]	35 40 45
[2287]	His Pro Val Ala Thr Cys Ala Ile Ala Leu Val Ala Leu Ala Ala Leu
[2288]	50 55 60
[2289]	Thr Pro Val Thr Thr Pro His Pro Ala Trp Gly Leu Ala Ile Ala Arg
[2290]	65 70 75 80
[2291]	Ala Ile Thr Lys Ala Ala Val Arg Tyr Phe Pro Leu Thr Met Glu Trp
[2292]	85 90 95
[2293]	Glu Asp Glu Arg Ala Tyr Leu Asp Ala Ala Ala Lys Gly Val Pro Ala
[2294]	100 105 110
[2295]	Val Ile Gly Leu Glu Pro His Ser Val Leu Pro Leu Ser Ile Val Ala
[2296]	115 120 125
[2297]	Phe Gly Asn Tyr Phe Phe Phe Thr Glu Ser Thr Pro Glu Cys Val Arg
[2298]	130 135 140
[2299]	Asn Ser Arg Ala Leu Ala Thr Gly Thr Ile Phe Val Ile Pro Val Leu
[2300]	145 150 155 160
[2301]	Lys His Leu Trp Ser Trp Leu Gly Met Asp Ala Ile Ser Arg Arg Ala
[2302]	165 170 175
[2303]	Met Lys Thr Leu Leu Asp Asp Gly Arg Ser Val Leu Ile Ile Pro Gly
[2304]	180 185 190
[2305]	Gly Val Ala Glu Cys Leu Gln Met Arg Pro Gly Val Glu Thr Ile Tyr
[2306]	195 200 205
[2307]	Leu Lys Lys Arg Phe Gly Phe Val Lys Leu Ala Ile Gln Thr Gly Ala
[2308]	210 215 220
[2309]	Ser Leu Val Pro Ala Phe Thr Phe Gly Gln Thr Arg Ser Tyr Ser Tyr

[2310]	225	230	235	240
[2311]	Trp Arg Leu Gly Pro Pro Leu Val Pro His Val Val Ala Glu Val Phe			
[2312]		245	250	255
[2313]	Ala Arg Ala Cys Arg Val Ala Pro Met Val Phe Trp Gly Lys Trp Gly			
[2314]		260	265	270
[2315]	Ser Pro Ile Pro Asn Met Val Pro Met His Thr Val Val Gly Lys Pro			
[2316]		275	280	285
[2317]	Ile Pro Val Lys Lys Gln Ser Glu Pro Ser Asn Glu Tyr Val Gln Glu			
[2318]		290	295	300
[2319]	Lys Leu Asn Glu Phe Val Ala Ala Met Glu Ser Leu Tyr Ala Arg His			
[2320]	305	310	315	320
[2321]	Lys Gly Lys His Gly Tyr Ala Glu Ser Thr Leu Val Val Leu			
[2322]		325	330	
[2323]	<210> 109			
[2324]	<211> 17			
[2325]	<212> DNA			
[2326]	<213> Artificial Sequence			
[2327]	<220>			
[2328]	<223> Transfer nucleic acid border sequence			
[2329]	<220>			
[2330]	<221> misc_feature			
[2331]	<222> (2) .. (2)			
[2332]	<223> n is a, c, g, or t			
[2333]	<220>			
[2334]	<221> misc_feature			
[2335]	<222> (6) .. (6)			
[2336]	<223> n is a, c, g, or t			
[2337]	<220>			
[2338]	<221> misc_feature			
[2339]	<222> (10) .. (15)			
[2340]	<223> n is a, c, g, or t			
[2341]	<400> 109			
[2342]	angatntatn nnnnngt 17			
[2343]	<210> 110			
[2344]	<211> 25			
[2345]	<212> DNA			
[2346]	<213> Artificial Sequence			
[2347]	<220>			
[2348]	<223> Transfer nucleic acid border sequence			
[2349]	<400> 110			
[2350]	tgacaggata tattggcggg taaac 25			
[2351]	<210> 111			

[2352]	<211> 25
[2353]	<212> DNA
[2354]	<213> Artificial Sequence
[2355]	<220>
[2356]	<223> Transfer nucleic acid border sequence
[2357]	<400> 111
[2358]	tggcaggata tattgtggtg taaac 25
[2359]	<210> 112
[2360]	<211> 25
[2361]	<212> DNA
[2362]	<213> Artificial Sequence
[2363]	<220>
[2364]	<223> Transfer nucleic acid border sequence
[2365]	<400> 112
[2366]	tggcaggata tataccgttg taatt 25
[2367]	<210> 113
[2368]	<211> 25
[2369]	<212> DNA
[2370]	<213> Artificial Sequence
[2371]	<220>
[2372]	<223> Transfer nucleic acid border sequence
[2373]	<400> 113
[2374]	cggcaggata tattcaattg taatt 25
[2375]	<210> 114
[2376]	<211> 25
[2377]	<212> DNA
[2378]	<213> Artificial Sequence
[2379]	<220>
[2380]	<223> Transfer nucleic acid border sequence
[2381]	<400> 114
[2382]	tggtaggata tataccgttg taatt 25
[2383]	<210> 115
[2384]	<211> 25
[2385]	<212> DNA
[2386]	<213> Artificial Sequence
[2387]	<220>
[2388]	<223> Transfer nucleic acid border sequence
[2389]	<400> 115
[2390]	tggcaggata tatggtactg taatt 25
[2391]	<210> 116
[2392]	<211> 25
[2393]	<212> DNA

[2394]	<213> Artificial Sequence
[2395]	<220>
[2396]	<223> Transfer nucleic acid border sequence
[2397]	<220>
[2398]	<221> misc_feature
[2399]	<222> (1) .. (1)
[2400]	<223> y = c or t
[2401]	<220>
[2402]	<221> misc_feature
[2403]	<222> (3) .. (3)
[2404]	<223> r = a or g
[2405]	<220>
[2406]	<221> misc_feature
[2407]	<222> (4) .. (4)
[2408]	<223> y = c or t
[2409]	<220>
[2410]	<221> misc_feature
[2411]	<222> (14) .. (14)
[2412]	<223> w = a or t
[2413]	<220>
[2414]	<221> misc_feature
[2415]	<222> (15) .. (15)
[2416]	<223> s = c or g
[2417]	<220>
[2418]	<221> misc_feature
[2419]	<222> (16) .. (16)
[2420]	<223> n is a, c, g, or t
[2421]	<220>
[2422]	<221> misc_feature
[2423]	<222> (17) .. (17)
[2424]	<223> v = a, c or g
[2425]	<220>
[2426]	<221> misc_feature
[2427]	<222> (18) .. (18)
[2428]	<223> b = c, g or t
[2429]	<220>
[2430]	<221> misc_feature
[2431]	<222> (19) .. (19)
[2432]	<223> k = g or t
[2433]	<220>
[2434]	<221> misc_feature
[2435]	<222> (24) .. (24)

- [2436] <223> w = a or t
[2437] <220>
[2438] <221> misc_feature
[2439] <222> (25) .. (25)
[2440] <223> y = c or t
[2441] <400> 116
[2442] ygryaggata tatwsnvbkg taawy 25
[2443] <210> 117
[2444] <211> 25
[2445] <212> DNA
[2446] <213> Artificial Sequence
[2447] <220>
[2448] <223> Transfer nucleic acid border sequence
[2449] <400> 117
[2450] cggcaggata tatcctgatg taaat 25
[2451] <210> 118
[2452] <211> 25
[2453] <212> DNA
[2454] <213> Artificial Sequence
[2455] <220>
[2456] <223> Transfer nucleic acid border sequence
[2457] <400> 118
[2458] tggcaggagt tattcgaggg taaac 25
[2459] <210> 119
[2460] <211> 25
[2461] <212> DNA
[2462] <213> Artificial Sequence
[2463] <220>
[2464] <223> Transfer nucleic acid border sequence
[2465] <400> 119
[2466] tgacaggata tatcgtgatg tcaac 25
[2467] <210> 120
[2468] <211> 25
[2469] <212> DNA
[2470] <213> Artificial Sequence
[2471] <220>
[2472] <223> Transfer nucleic acid border sequence
[2473] <400> 120
[2474] gggaagtaca tattggcggg taaac 25
[2475] <210> 121
[2476] <211> 25
[2477] <212> DNA

[2478]	<213> Artificial Sequence
[2479]	<220>
[2480]	<223> Transfer nucleic acid border sequence
[2481]	<400> 121
[2482]	ttacaggata tattaatatg tatga 25
[2483]	<210> 122
[2484]	<211> 25
[2485]	<212> DNA
[2486]	<213> Artificial Sequence
[2487]	<220>
[2488]	<223> Transfer nucleic acid border sequence
[2489]	<400> 122
[2490]	taacatgata tattcccttg taaat 25
[2491]	<210> 123
[2492]	<211> 25
[2493]	<212> DNA
[2494]	<213> Artificial Sequence
[2495]	<220>
[2496]	<223> Transfer nucleic acid border sequence
[2497]	<400> 123
[2498]	tgacaggata tatggtaatg taaac 25
[2499]	<210> 124
[2500]	<211> 25
[2501]	<212> DNA
[2502]	<213> Artificial Sequence
[2503]	<220>
[2504]	<223> Transfer nucleic acid border sequence
[2505]	<400> 124
[2506]	tggcaggata tataccgatg taaac 25
[2507]	<210> 125
[2508]	<211> 1395
[2509]	<212> DNA
[2510]	<213> Artificial Sequence
[2511]	<220>
[2512]	<223> Codon-optimized nucleotide sequence for production in plants of
[2513]	SEQ ID NO:7
[2514]	<400> 125
[2515]	atgtgtcctc ctaagaccga tggaagatct tctcctagat ctctctcac caggtctaag 60
[2516]	tcatctgctg aggctcttga tgctaaggat gcttctaccg ctctgttga tcttaagacc 120
[2517]	cttgagcctc atgaacttgc tgctaccttc gagactagat gggttagggt tgaggatggt 180
[2518]	gagtacgacg tgaccaactt caaacatcct ggtggaagcg tgatcttcta catgcttgct 240
[2519]	aacctgggtg ctgatgctac tgaggctttc aaagaatttc acatgcgtag cctcaaggct 300

[2520] tggaagatgc ttagagcttt gccttctaga cctgctgaga tcaagagatc tgagtctgag 360
 [2521] gatgctccta tgcttgagga tttcgctagg tggagagctg aacttgagag ggacggattc 420
 [2522] ttcaagcctt ctatcaccca tgttgcttac cgtcttttgg agcttcttgc tactttcget 480
 [2523] cttggaaccg ctcttatgta cgctggatac cctatcattg ctagecgttg gtacgggtgc 540
 [2524] ttcttcggag ctagatgtgg atgggttcaa catgagggtg gacacaactc tcttaccgga 600
 [2525] tctgtgtacg tggataagag acttcaggct atgacttgcg gattcggact ttctaccagc 660
 [2526] ggagagatgt ggaaccagat gcataacaag caccatgcta cccctcagaa agttagacac 720
 [2527] gacatggatc ttgataccac tctgtctgtg gctttcttca acaccgctgt ggaggataat 780
 [2528] agacctaggg gattctctag agcttgggct agacttcaag cttggacctt cgttctgtt 840
 [2529] acttctggac ttctcgttca ggctttctgg atctacgttc tccatcctag acaggtgctc 900
 [2530] aggaagaaga actacagga agcttcttgg atgctcgttt ctcacgttgt tagaaccgct 960
 [2531] gttatcaagc ttgctaccgg atactcttgg cctgttgctt actggtggtt cactttcgga 1020
 [2532] aactggatcg cttacatgta cctcttcgct cacttctcta cttctcacac tcacctccct 1080
 [2533] gttgttccat ctgacaagca ccttagctgg gttaactacg ctgttgatca caccgttgac 1140
 [2534] atcgatcctt ctctggata cgtttaactg cttatgggat accttaactg ccaggttatc 1200
 [2535] caccatctct tcctgatat gcctcaattc agacagcctg aggtgtcaag aagattcgtc 1260
 [2536] cctttccta agaagtggg actcaactac aaggtgctct cttactacgg tgcttgaag 1320
 [2537] gctactttca gcaacctga caaagtggga cagcactact acgttaacgg aaaggctgag 1380
 [2538] aaggctcact gatga 1395
 [2539] <210> 126
 [2540] <211> 870
 [2541] <212> DNA
 [2542] <213> Artificial Sequence
 [2543] <220>
 [2544] <223> Codon-optimized nucleotide sequence for production in plants of
 [2545] SEQ ID NO:3
 [2546] <400> 126
 [2547] atggaatttg ctcaacctct cgttgctatg gctcaagagc agtacgctgc tatcgatgct 60
 [2548] gttgttgctc ctgctatctt ctctgctacc gactctattg gatggggact caagcctatc 120
 [2549] tcttctgcta ctaaggatct cctctcgtt gaatctccta cccctcttat cctttctctc 180
 [2550] ctgcttact tcgctatcgt tggttctgga ctggtttacc gtaaagtgtt ccctagaacc 240
 [2551] gttaagggac aggatccttt ccttctcaag gctcttatgc tcgctcaca cgttttcctt 300
 [2552] atcggactca gcctttacat gtgcctcaag ctggtttacg aggccttacgt gaacaagtac 360
 [2553] tccttctggg gaaacgctta caacctgct caaaccgaga tggctaaggt gatctggatc 420
 [2554] ttctacgtgt ccaagatcta cgagttcatg gacaccttca tcatgcttct caagggaac 480
 [2555] gttaaccagg tttccttctt ccatgtttac caccacggat ctatctctgg aatctggtgg 540
 [2556] atgacactt atgctgctcc aggtggagat gcttacttct ctgctgctct caactcttgg 600
 [2557] gttcatgtgt gcatgtacac ctactacttc atggctgctg ttcttcttaa ggacgaaaag 660
 [2558] accaagagaa agtacctttg gtggggaaga taccttacc agatgcaaat gttccagttc 720
 [2559] ttcatgaacc ttctccaggc tgtttacctc ctctactctt cttctccta ccctaagttc 780
 [2560] attgctcaac tcctcgttgt ttacatgggt accctctca tgcttttcgg aaacttctac 840
 [2561] tacatgaagc accacgttc taagtataa 870

[2562]	<210> 127
[2563]	<211> 1281
[2564]	<212> DNA
[2565]	<213> Artificial Sequence
[2566]	<220>
[2567]	<223> Codon-optimized nucleotide sequence for production in plants of
[2568]	SEQ ID NO:25
[2569]	<400> 127
[2570]	atgcctccta gggactctta ctcttacgct gctcctcctt ctgctcaact tcacgaggtt 60
[2571]	gacactcctc aagagcacga caagaaagag cttgttatcg gagatagggc ttacgatgtg 120
[2572]	accaacttcg ttaagagaca ccctgggtga aagatcattg cttaccaagt gggaactgat 180
[2573]	gctaccgatg cttacaagca gttccatgtg agatctgcta aggctgacaa gatgctcaag 240
[2574]	tctctccctt ctagacctgt tcacaaggga tactctccta gaagagctga tcttatcgct 300
[2575]	gacttccaag agttcactaa gcaacttgag gctgaggga tggttcgaacc ttctctccct 360
[2576]	catgttgctt accgtcttgc tgaggttatc gctatgcatg ttgctgggtc tgctcttctc 420
[2577]	tggcacgat acactttcgc tggaatcgct atgcttgag ttgttcaggg aagatgcgga 480
[2578]	tggcttatgc atgagggtgg aactactct cttaccggaa acattgcttt cgatagggtc 540
[2579]	atccaagttg cttgttacgg acttgatgc ggaatgtctg gtgcttggtg gagaaccag 600
[2580]	cataacaagc accatgctac tcctcaaaag ctccagcacg atgttgatct tgataccctc 660
[2581]	cctctcggtg ctttccatga gagaatcgct gctaaggta agtctcctgc tatgaaggct 720
[2582]	tggctctcca tgcaagctaa actcttcgct cctgttacca ctcttcttgt tgctcttgga 780
[2583]	tggcagcttt acctcaccc tagacacatg ctcaagaacta agcactacga cgagcttgct 840
[2584]	atgcttggtg tcagatacgg acttggtgga taccttgctg ctaactacgg tgctggatac 900
[2585]	gttcttgctt gctaccttct ctacgttcag cttggagcta tgtacatctt ctgcaacttc 960
[2586]	gctgtttctc aactcatct cctgttggt gagcctaacg agcatgctac ttgggttgag 1020
[2587]	tacgtgcta accacactac taactgctct ccatcttggt ggtgtgattg gtggatgagc 1080
[2588]	tacctcaact accagatcga gcatcacctt tacccttcta tgcctcagtt caggcatcct 1140
[2589]	aagatcgctc ctagagtga gcaactcttc gagaagcacg gacttcaacta cgatgtgcgt 1200
[2590]	ggatacttcg aggtatggc tgatacttct gtaacctcg ataactgtgc tcatgctcct 1260
[2591]	gagaagaaaa tgcaatgatg a 1281
[2592]	<210> 128
[2593]	<211> 807
[2594]	<212> DNA
[2595]	<213> Artificial Sequence
[2596]	<220>
[2597]	<223> Codon-optimized nucleotide sequence for production in plants of
[2598]	SEQ ID NO:5
[2599]	<400> 128
[2600]	atggcttcta tcgctatccc tgctgctctt gctggaactc ttggatacgt gacctacaac 60
[2601]	gtggctaacc ctgatattcc tgcttctgag aaggttcag cttacttcat gcaagtggag 120
[2602]	tactggggac ctactatcgg aactatcggc tacctcctct tcatctactt cggaagcgt 180
[2603]	atcatgcaaa acagaagcca gcctttcgga cttaagaacg ctatgctcgt gtacaacttc 240

[2604]	taccagacct tcttcaacag ctactgcac	tacctcttcg ttacctctca tagggctcag	300
[2605]	ggacttaaaag tttggggaaa catccctgat	atgaccgcta actcttgggg aatctctcag	360
[2606]	gttatctggc tccactacaa caacaagtac	gtggagcttc tcgatacctt cttcatgggtg	420
[2607]	atgaggaaga agttcgacca gctttctttc	cttcacatct accaccacac tcttctcacc	480
[2608]	tggtcatggg tcgtgggttat gaagctcgag	cctgttggag attgctactt cggatctagc	540
[2609]	gttaacacct tcgtgcacgt gatcatgtac	tcttactacg gacttgctgc tcttggagtt	600
[2610]	aactgcttct ggaagaagta catcaccag	atccagatgc ttcagttctg tatctgcgct	660
[2611]	tctcactcta tctacaccgc ttacgttcag	aacactgctt tctggttcc ttaccttcag	720
[2612]	ctctgggtga tggtaacat gttcgtgctc	ttcgctaact tctaccgtaa aaggtacaag	780
[2613]	agcaagggtg ctaagaagca gtgataa		807
[2614]	<210>	129	
[2615]	<211>	1347	
[2616]	<212>	DNA	
[2617]	<213>	Artificial Sequence	
[2618]	<220>		
[2619]	<223>	Codon-optimized nucleotide sequence for production in plants of	
[2620]	SEQ ID NO:72		
[2621]	<400>	129	
[2622]	atgcctccat ctgctgctaa acagatggga	gcttctactg gtgttcacgc tgggtgttacc	60
[2623]	gattcttctg ctttcaccag aaaggatgtg	gctgatagac ctgatcttac catcggttgg	120
[2624]	gactctgtgt acgatgctaa ggctttcaga	tctgagcacc ctgggtgtgc tcatctcgtt	180
[2625]	tcactcttcg gaggaagaga tgctactgag	gctttcatgg aataccacag aagagcttgg	240
[2626]	cctaagtcta ggatgtctag gttccatgtt	ggatctcttg cttctaccga ggaacctgtt	300
[2627]	gctgctgatg agggatacct tcagctttgt	gctaggatcg ctaagatggg gccttctgtg	360
[2628]	tcatctggat tcgtccagc ttcttactgg	gttaaggctg gacttatcct cggatctgct	420
[2629]	atcgctcttg aggcttacat gctctacgct	ggaaagagac ttctcccttc tatcgttctt	480
[2630]	ggatggctct tcgctcttat cggacttaac	atccagcatg acgctaacca tgggtgcttg	540
[2631]	tctaagtctg ctagcgtaa cttgtctctt	ggactttgtc aggattggat cggaggatct	600
[2632]	atgatccctt ggctccaaga gcatgttgtt	atgcaccacc tccacacca cgaatgttgat	660
[2633]	aaggacctg atcaaaaggc tcatgggtgt	cttagactca agcctaccga tgcttggta	720
[2634]	cctatgcatt ggcttcagca cttttacctt	ctccctgggtg aaactatgta cgctttcaag	780
[2635]	ctcctcttcc tcgatatctc tgagcttgtg	atgtggagat gggagggtga acctatctct	840
[2636]	aagctcgtg gatactctt catgcttctt	cttctctca agcttacctt ctgggctaga	900
[2637]	ttcgttgcct ttctcttta cctcgtcctt	tctgttcata ctgctgtgtg tategctgct	960
[2638]	actgttatga ccggaagctt ctacctgtct	ttcttcttct tcatcagcca caacttcgag	1020
[2639]	ggtgttgctt ctgttgacc tgatggatct	atcacctcta tgaccagggg agcttcttct	1080
[2640]	cttaagaggc aggtgagac ttcttcta	atgtggaggac ctcttcttgc tactcttaac	1140
[2641]	ggtggactca actacaaat cgagcaccac	cttttcccta gagttcacca cggattctac	1200
[2642]	cctagacttg ctctcttgt gaaggtgaa	cttgaggcta gaggaatcga gtacaagcac	1260
[2643]	tacctacca tctggtctaa cctcgttctt	accctcagac atatgtacgc tcttgaaga	1320
[2644]	aggcctagat ctaaggctga gtgatga		1347

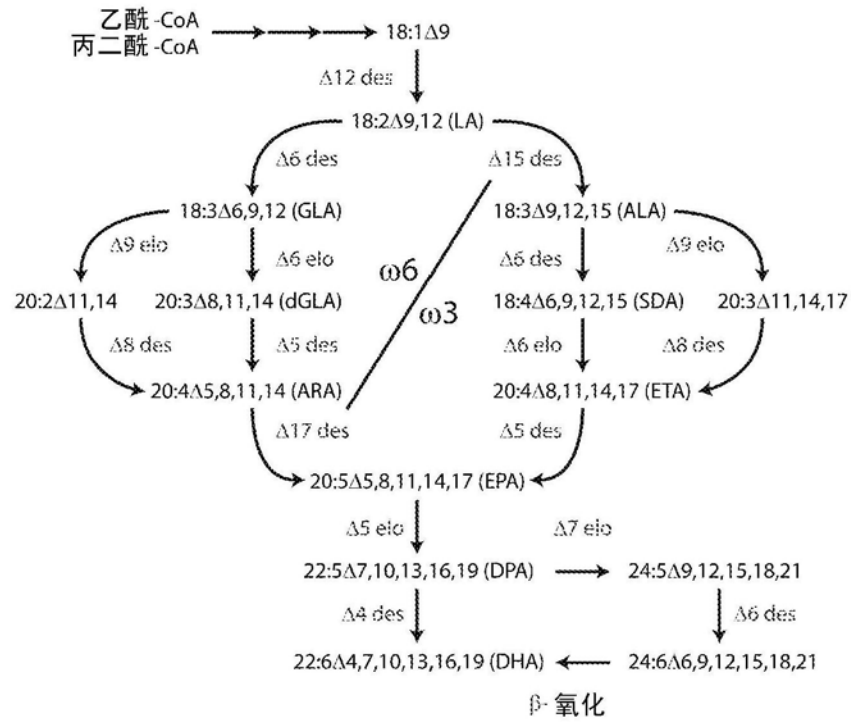


图1

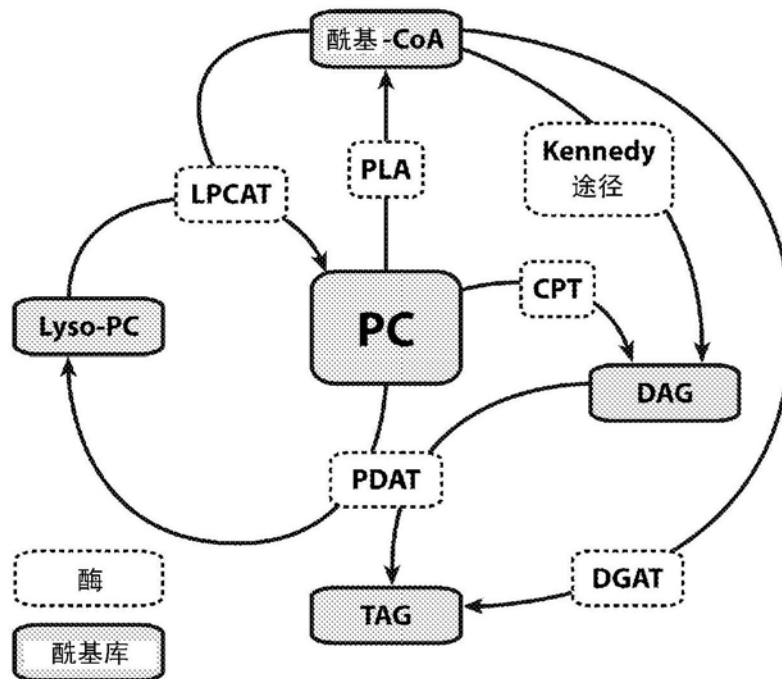


图2

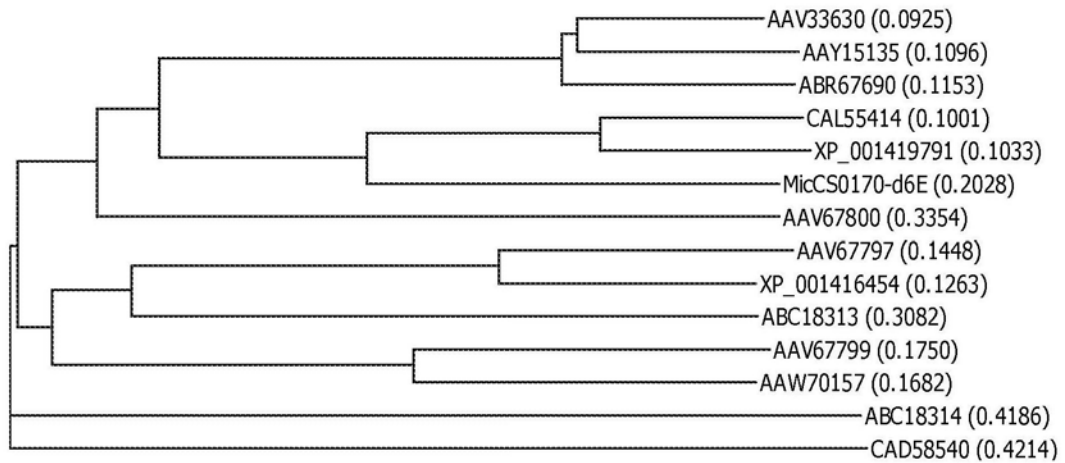


图3

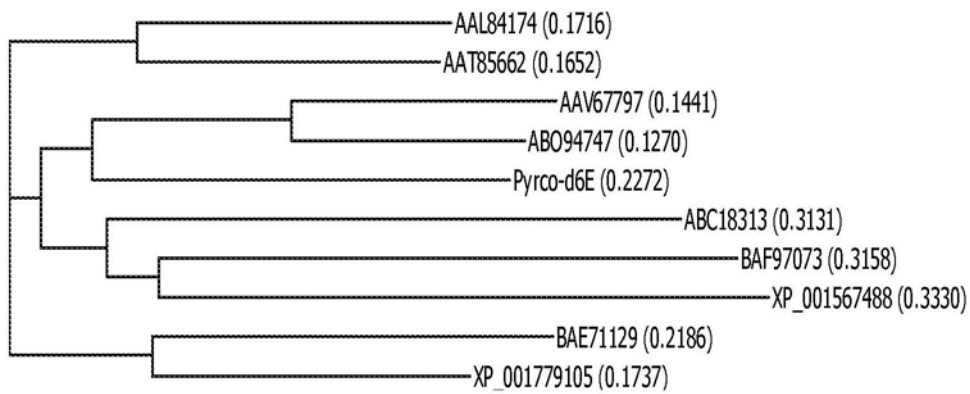


图4

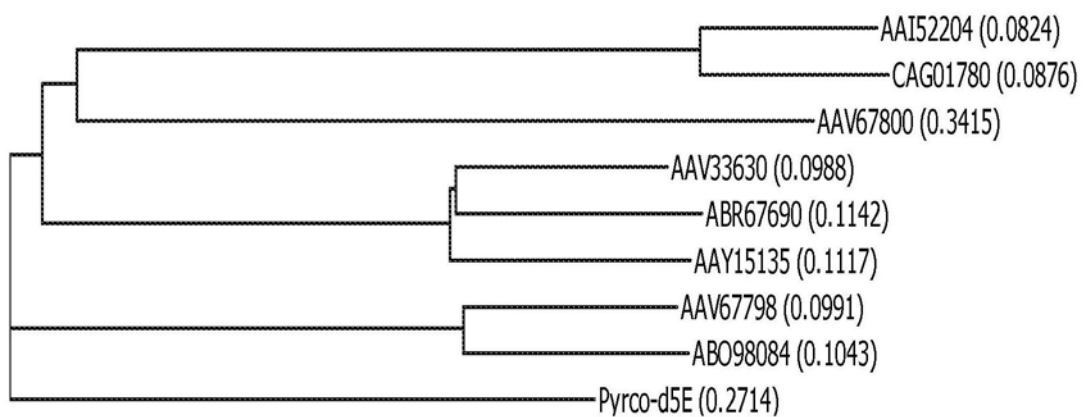


图5

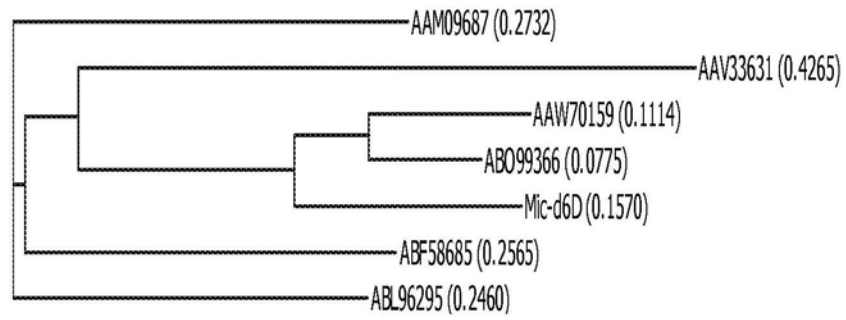


图6

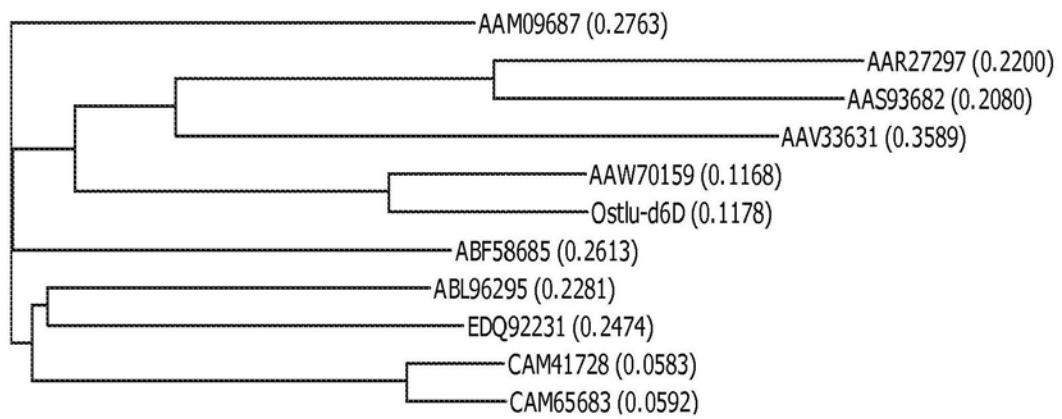


图7

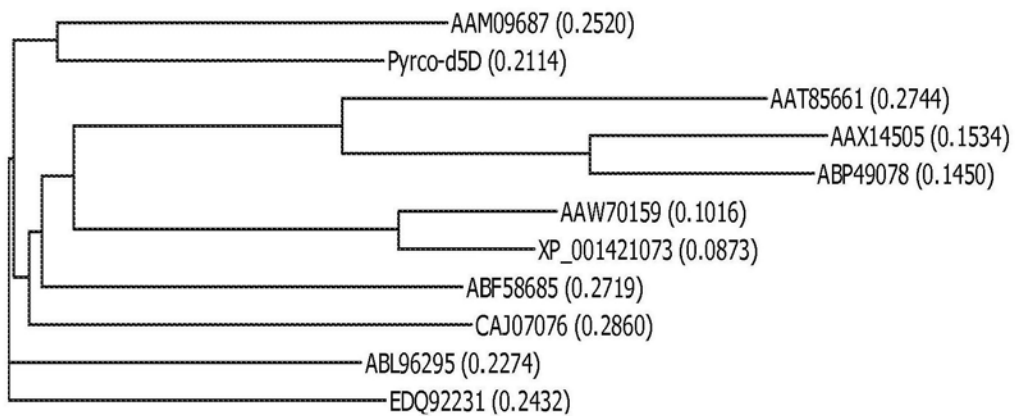


图8

M	(1)	MPPKTDG-----RSPRSPLTRSKSSAEALDAKDASTAPVDILKTLEPHELAAATFETRWVRVED	150
Ol	(1)	MVETEGTSRTMANERTSSSSSLSEGGTPVTVMGSEDAGKKTRNASVTAWTKELEPHALAKTFERRRYVTIEG	150
ot	(1)	MVETEN-----N-DGIP TVEIAFDGE---RERAEANVKLSAEKMEPAALAKTFARRRYVVIIEG	150
M	(60)	VEYDVTNFKHPGGSVIFYMLANTGADATEAFKEEFHRSILKAWKMLRALPSEPAEIKRSE-SEDAPMLEDFAEWRA	225
Ol	(76)	VEYDVTBFKHPGGSVIFYMLSNTGADATEAFKEEFHRSKKARKALAAPLHPKPVDAATREP IEDEAMLKDFAQWRK	225
ot	(55)	VEYDVTBFKHPGGTVIFYALSNTGADATEAFKEEFHRSKKARKALAAPLSEPAKTAK---VDDAEMLQDFAKWRK	225
M	(134)	ELERDGFCKPSITHVAYRLELLELALFALGTALMYAGYPIIASVYVYGAFFGARCGWVQHEGGHNSLTGSVYVDKRL	300
Ol	(151)	ELEREGFFKPSPAHVAYRFAELARMFALGTALMHARWHVASVIVYSCFFGARCGWVQHEGGHNSLTGNINWDKRI	300
ot	(127)	ELERDGFCKPSPAHVAYRFAELAAHYALGTILMYARYVVS SVLVYACFFGARCGWVQHEGGHNSLTGNINWDKRI	300
M	(209)	QAMTCGFGLSTSGEMNNQMHKHHATPQKVRHDMDLDTTPAVAFFNTAVE DNRPRGFSRAWARLQAWTFVPVTS	375
Ol	(226)	QAFAGFGGLASSGTMNNMHKHHATPQKVRHDMDLDTTPVAFFNSAVEENRPRGFSKLWLRLQAWTFVPVTS	375
ot	(202)	QAFTAGFGLAGSGTMNNSMHKKHHATPQKVRHDMDLDTTPAVAFFNTAVE DNRPRGFSKYWLRLQAWTFIPVTS	375
M	(284)	LI VQAFWIYVLHPROVLKKNYEEASMMELVSHVVRTAVIKLATGYSWPVAYWMTFGNWLAYMYLFAHFST SHTH	450
Ol	(301)	-MVLFFWMFVLHPRHALRKSFEAAWMFSAHVIRTAVIKAVTGYSWLASYGLEFAATMWSGCCYLF AHFST SHTH	450
ot	(277)	-LVLLFWMEFFLHP SKALKGGKYEELVWMLAAHVIRTWIKAVTGFTAAQS YGLEFLATSWVSGCYLF AHFST SHTH	450
M	(359)	LPVVP SDKHL SWVNYAVDHTVDIDP SRGYVNWLMGYLNCQV IHHLFPDMPQFRQPEVSRRFVPPFAKKWGLNYKVL	483
Ol	(375)	LDVVP SDKHL SWVRYAVDHTIDIMPNN SVVNWLMGYLNCQV IHHLFPDMPQFRQPEVSRRFVPPFAKKWGLNYKVL	483
ot	(351)	LDVVP ADEHL SWVRYAVDHTIDIDP SQGWVNWLMGYLNCQV IHHLFPDMPQFRQPEVSRRFVPPFAKKWGLNYKVM	483
M	(434)	SYGAWKATF SNLDKVGQHYVVGKAEKAH---	483
Ol	(450)	TYGAWKATF GNLDVGHYVVGHSQRVKSASA	483
ot	(426)	TYGAWKATLGNLDNVGKHYYVVGQHSKTA---	483

图9

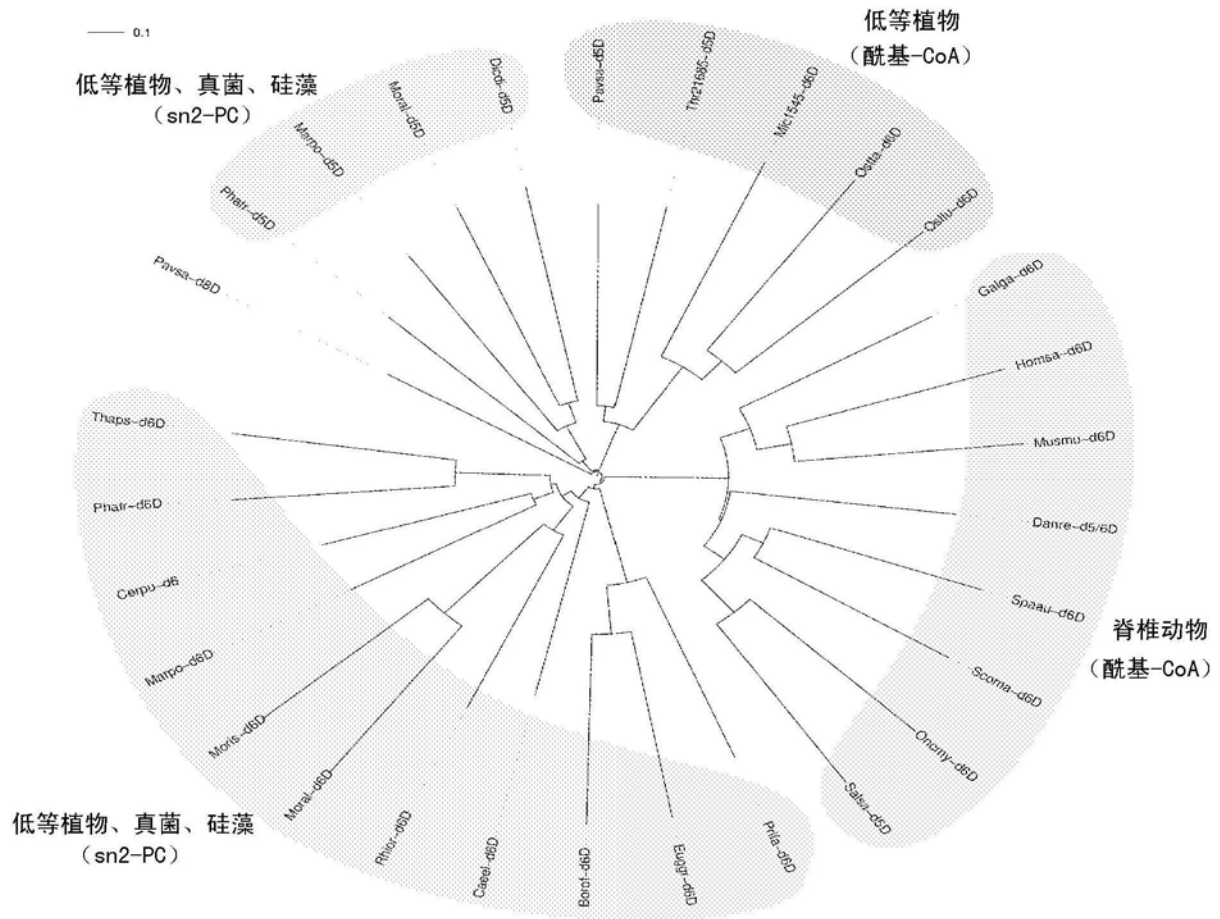


图10

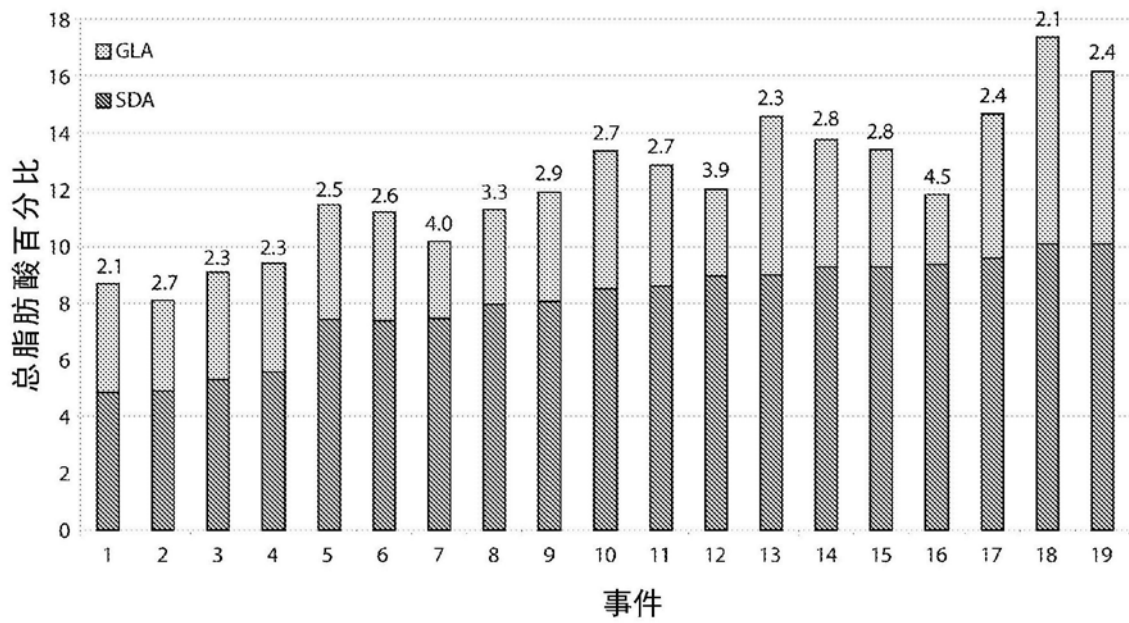


图11

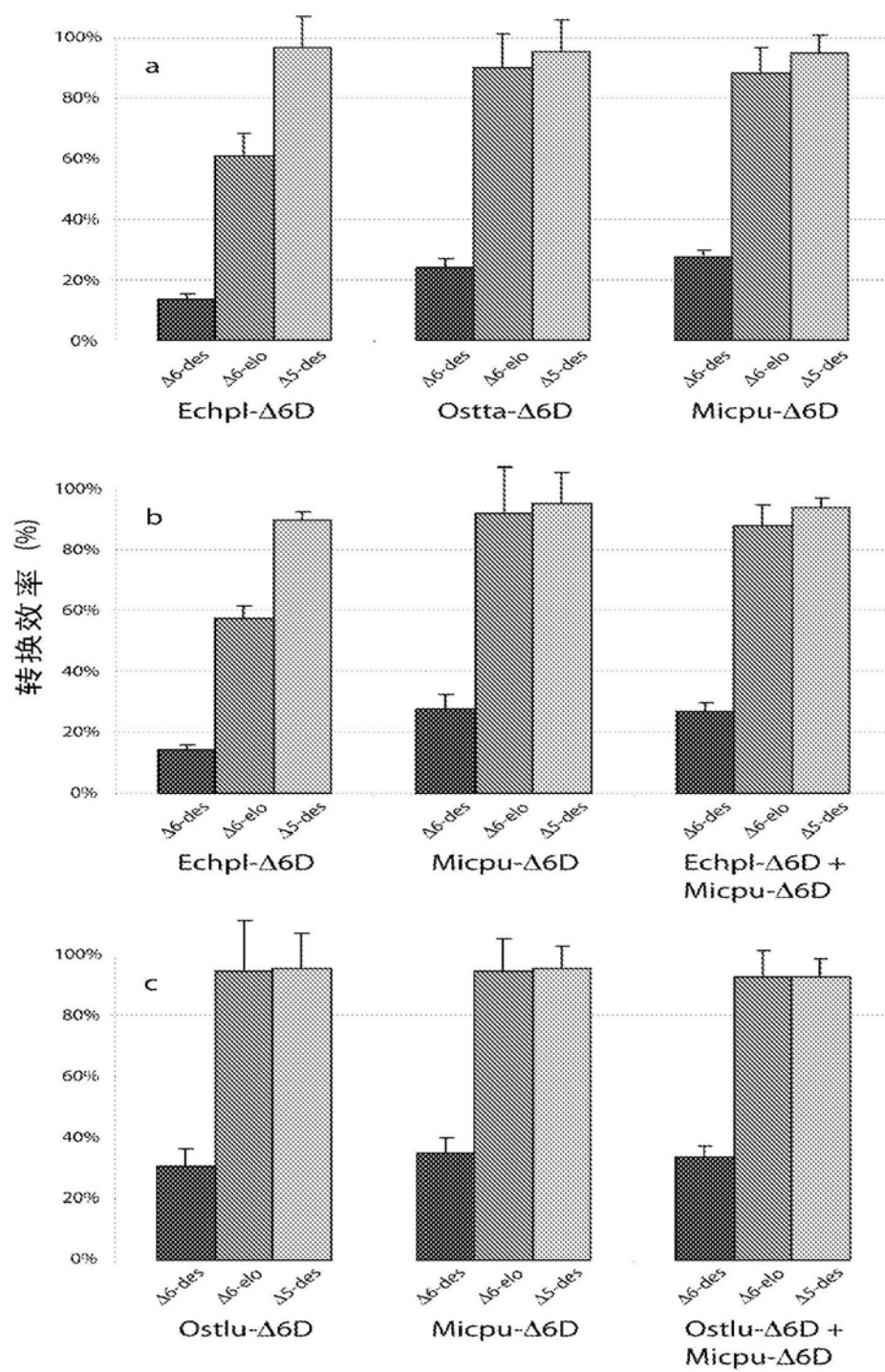


图12

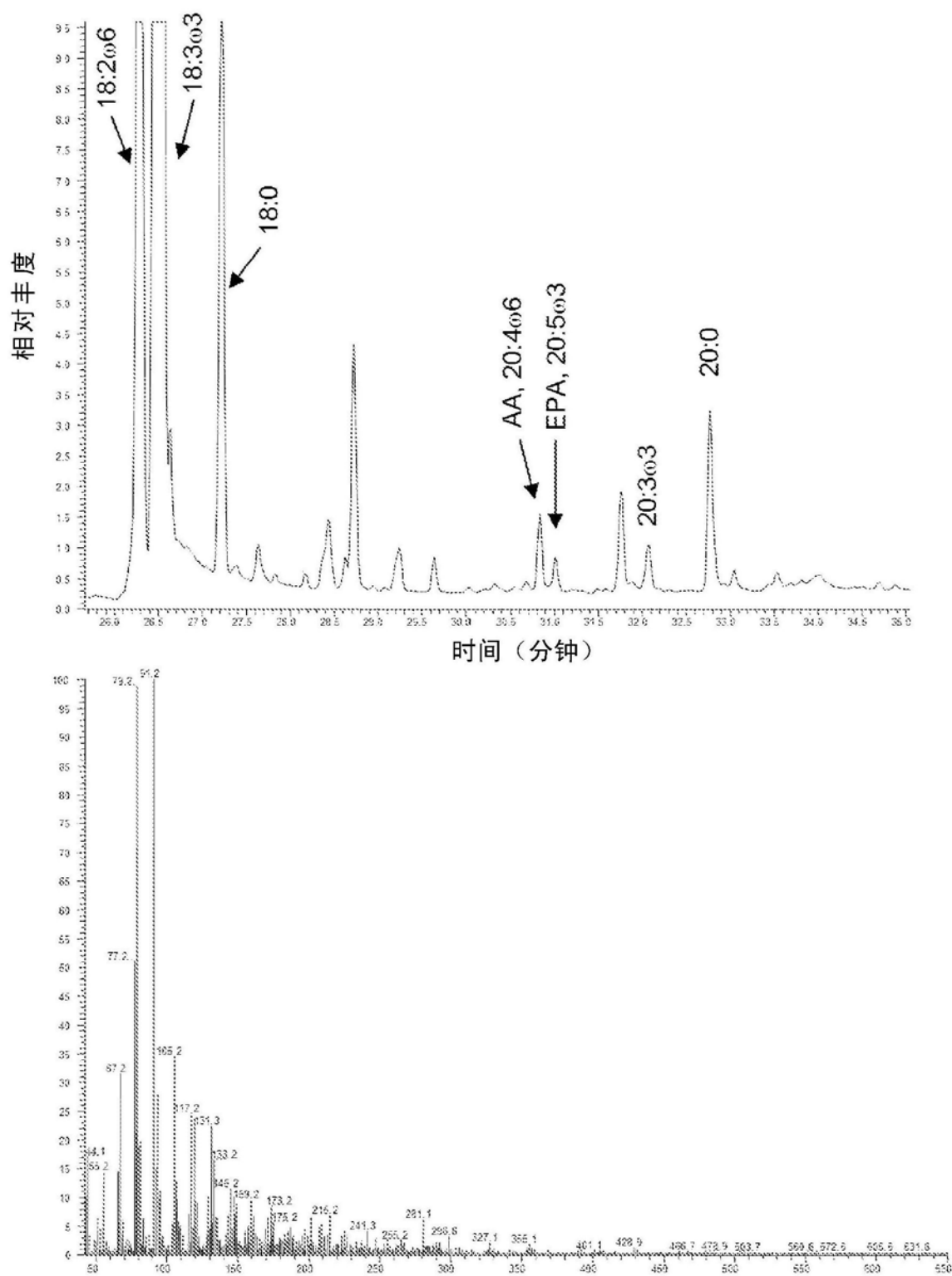


图13

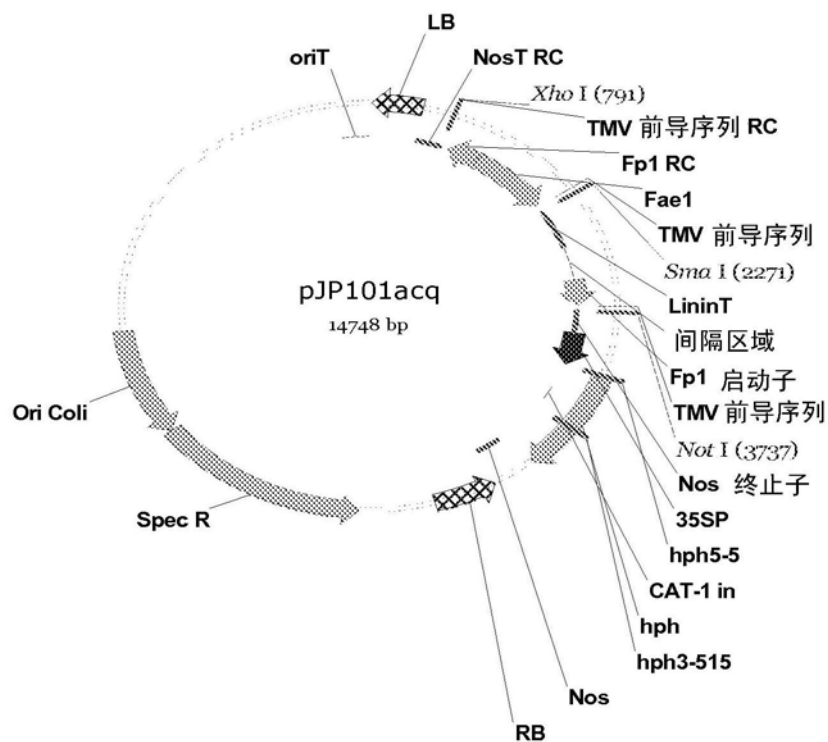


图14

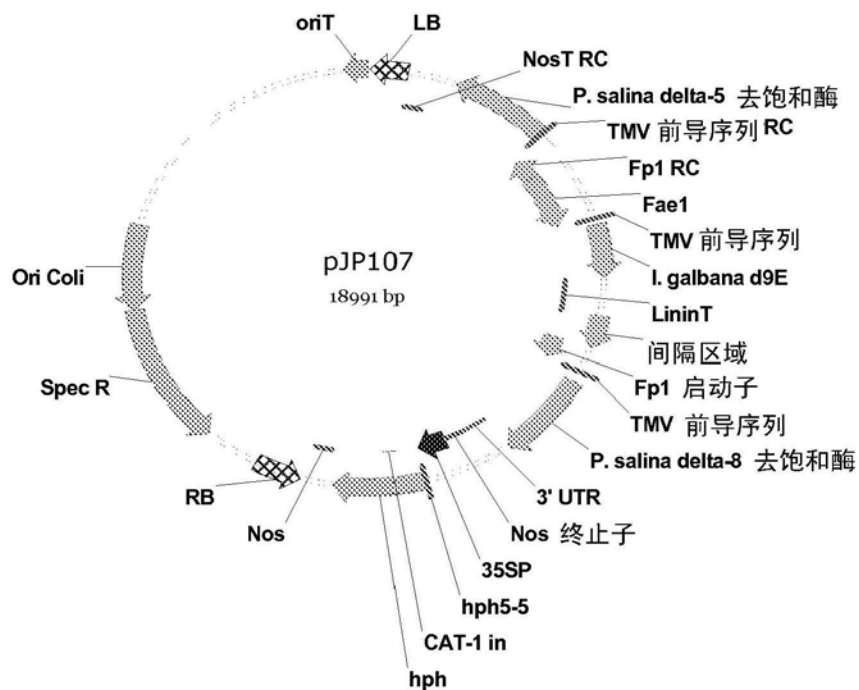


图15

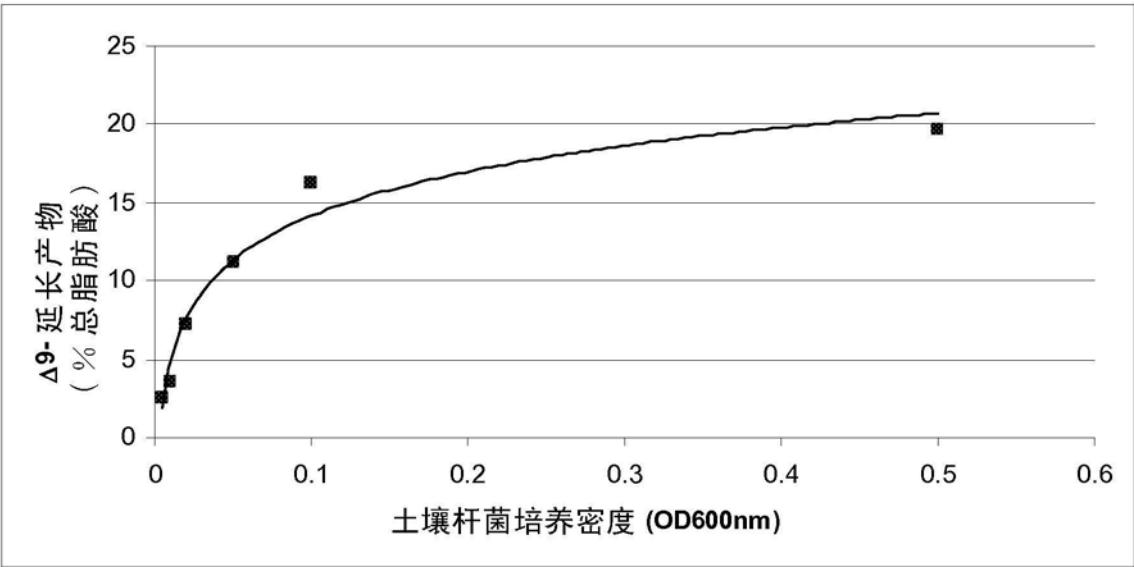


图16

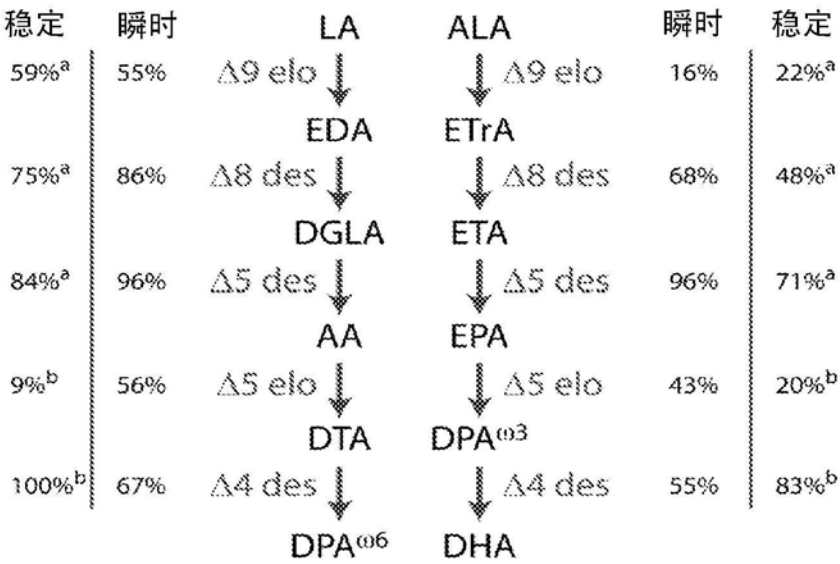


图17

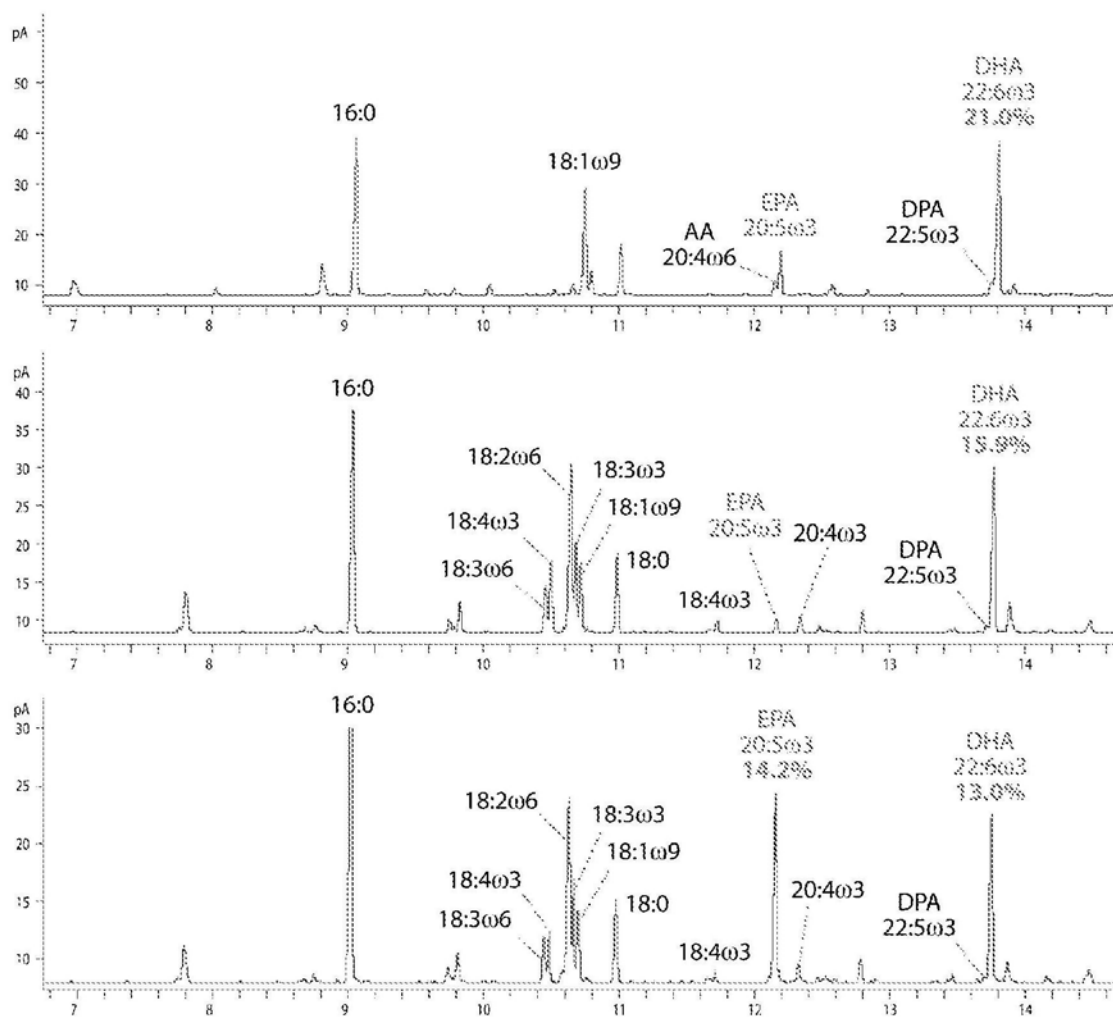


图18

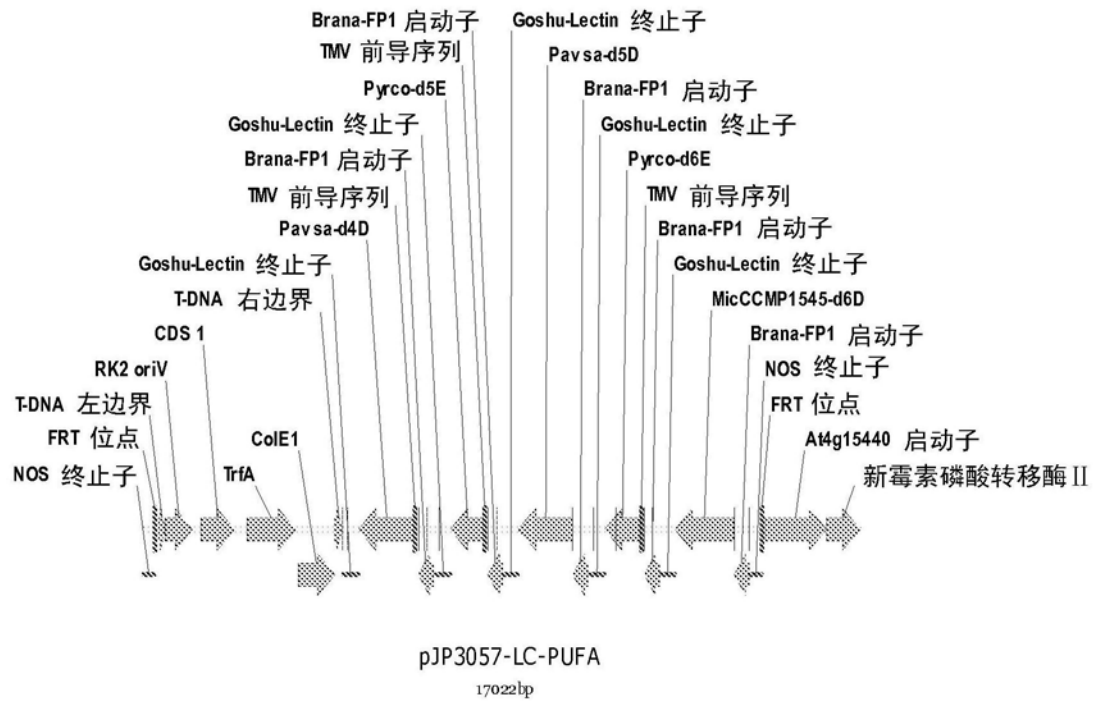


图19

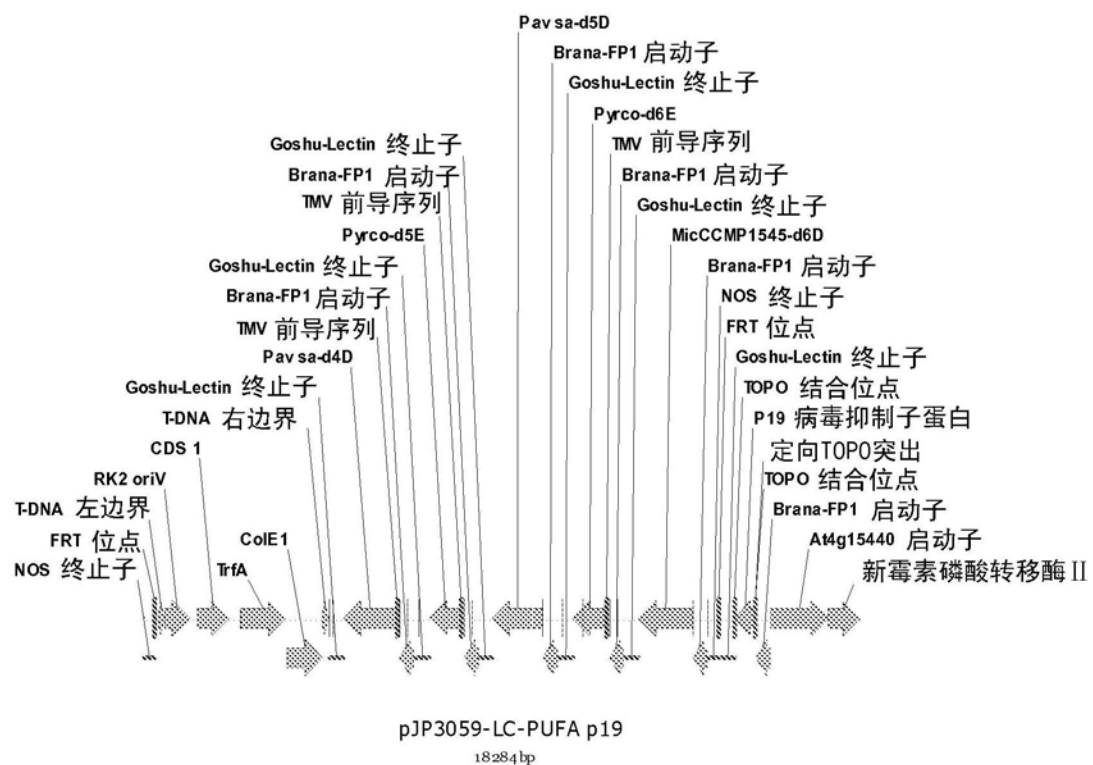


图20



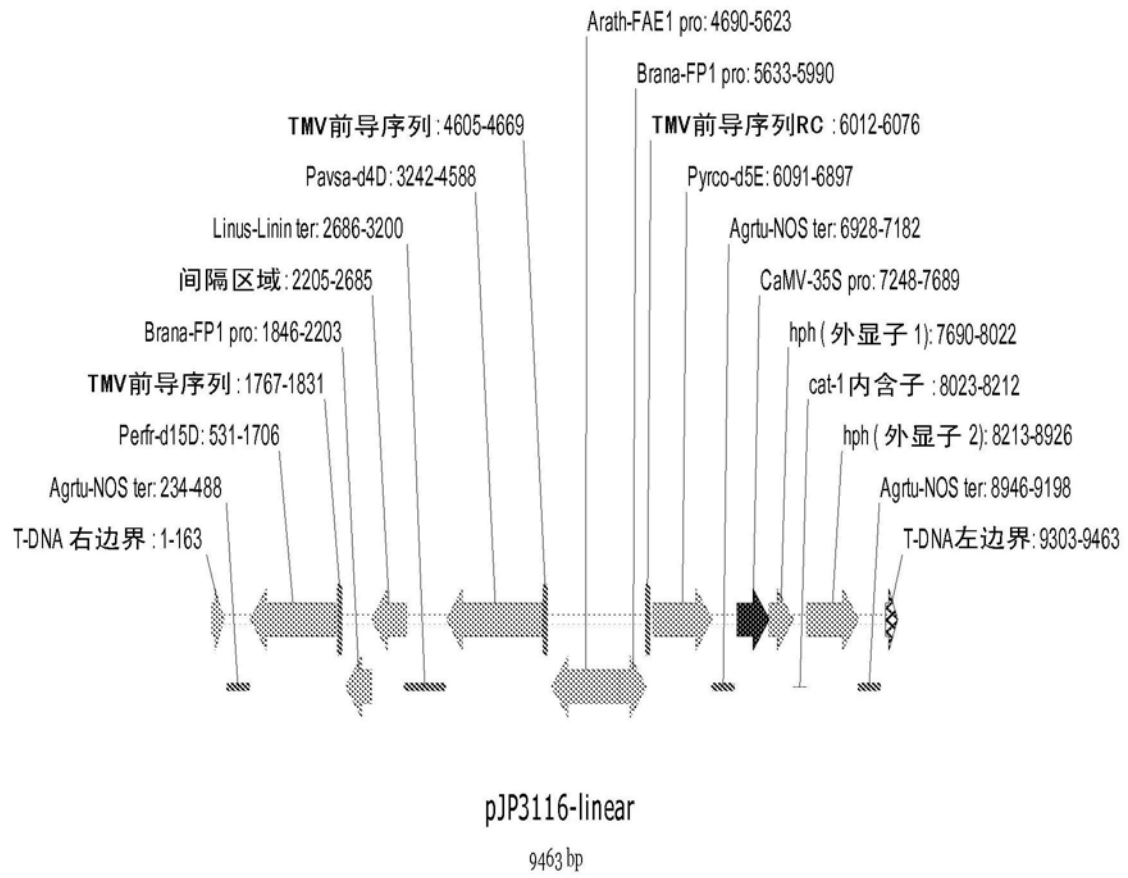


图23