(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 110066836 A
(43)申请公布日 2019.07.30

(21)申请号 201910164460.0

(22)申请日 2012.02.02

(30)优先权数据
61/438,969 2011.02.02 US
61/476,691 2011.04.18 US
61/484,458 2011.05.10 US
61/548,616 2011.10.18 US

(62)分案原申请数据
201280007593.3 2012.02.02

(71)申请人 柯碧恩生物技术公司
地址 美国加利福尼亚州

(72)发明人 S·富兰克林 A·索曼奇 J·维 G·鲁登科 J·莫瑟利 W·莱基特斯基 X·赵 R·巴特

(74)专利代理机构 北京市中咨律师事务所 11247
代理人 凌立 黄革生

(51)Int.Cl.
C12P 7/64(2006.01)
C12N 1/13(2006.01)
C10L 1/02(2006.01)
C12R 1/89(2006.01)

权利要求书2页 说明书193页
序列表304页 附图3页

(54)发明名称

产自重组产油微生物的定制油

(57)摘要

在此提供了用于在重组微生物中产生油、燃料、油脂化学品以及其他化合物的方法和组合物，包括含油微生物以及这样的微生物的低成本培养方法。包含编码例如脂肪酶、蔗糖转运蛋白、蔗糖转化酶、果糖激酶、多糖降解酶、酮脂酰-ACP合成酶、脂肪酰基-ACP硫酯酶、脂肪酰基辅酶A/醛还原酶、脂肪酰基辅酶A还原酶、脂肪醛还原酶、脂肪酸羟化酶、去饱和酶、脂肪醛脱羧酶、和/或酰基载体蛋白的外源基因的微藻细胞在制造运输燃料(如可再生柴油、生物柴油、以及可再生喷气燃料)、以及油脂化学品(如功能性流体、表面活性剂、肥皂以及润滑剂)中是有用的。

CN 110066836 A

1.一种用于生产包含三酰甘油酯的油或产自所述油的产品的方法,所述方法包括以下步骤:

培养无绿藻属或小球藻(Chlorella)属重组微藻细胞,所述细胞包含一种或多种外源基因,所述一种或多种外源基因可用于表达由编码β-酮脂酰-ACP合成酶II的第一外源基因编码的酶,任选其中所述细胞还包含编码外源酰基-ACP硫酯酶或外源去饱和酶的第二外源基因;和

从这种细胞回收所述油,并且任选地进一步加工所述油以生产食物、燃料或化学产品,

其中所述油由于所述一种或多种重组核酸而具有改变的脂肪酸谱。

2.一种无绿藻属或小球藻属重组微藻细胞,所述细胞包含一种或多种外源基因,所述细胞包含一种或多种外源基因,所述一种或多种外源基因可用于表达由编码β-酮脂酰-ACP合成酶II的第一外源基因编码的酶,任选其中所述细胞还包含编码外源酰基-ACP硫酯酶或外源去饱和酶的第二外源基因,其中当与非重组微藻细胞的脂肪酸谱相比时,由于所述一种或多种重组基因的表达,所述细胞具有改变的脂肪酸谱。

3.一种包含通过下述方法生产的三酰甘油酯的油,所述方法包括以下步骤:

培养无绿藻属或小球藻属重组微藻细胞,所述细胞包含一种或多种外源基因,所述一种或多种外源基因可用于表达由编码β-酮脂酰-ACP合成酶II的第一外源基因编码的酶,任选其中所述细胞还包含编码外源酰基-ACP硫酯酶或外源去饱和酶的第二外源基因;和从这种细胞回收所述油,并且任选地进一步加工所述油以生产食物、燃料或化学产品,其中所述油由于所述一种或多种重组核酸而具有改变的脂肪酸谱。

4.权利要求1的方法,权利要求2的细胞,或权利要求3的油,其中所述外源去饱和酶是硬脂酰-ACP去饱和酶(SAD)或脂肪酸去饱和酶(FAD)。

5.权利要求1或4的方法,权利要求2或4的细胞,或权利要求3或4的油,其中所述一种或多种外源基因整合到细胞的核基因组的同源重组位点中,并且其中整合减少或消除由所述同源重组位点编码的内源基因产物的表达,或其中所述内源基因的表达通过RNA诱导的机制降低。

6.权利要求5的方法、细胞或油,其中所述同源重组位点编码内源脂肪酰基-ACP硫酯酶。

7.权利要求6的方法、细胞或油,其中内源性脂肪酰基-ACP硫酯酶是FATA。

8.权利要求5的方法、细胞或油,其中核基因组中的同源重组位点编码内源性去饱和酶。

9.权利要求8的方法、细胞或油,其中内源性去饱和酶是硬脂酰-ACP去饱和酶(SAD)或脂肪酸去饱和酶(FAD)。

10.权利要求1或4至9中任一项的方法,权利要求2或4至9中任一项的细胞,或权利要求3至9中任一项的油,其中所述微藻选自无绿藻属。

11.如权利要求10所述的方法、细胞或油,其中所述微藻选自魏氏无绿藻、雍滞无绿藻、Prototheca portoricensis、Prototheca moriformis或祖菲无绿藻。

12.如权利要求11所述的方法、细胞或油,其中所述微藻是Prototheca moriformis。

13.权利要求1或4至9中任一项的方法,权利要求2或4至9中任一项的细胞,或权利要求3至9中任一项的油,其中所述微藻选自小球藻属。

14.如权利要求13所述的方法、细胞或油，其中所述微藻选自凯氏小球藻、Chlorella luteoviridis、原始小球藻或普通小球藻。

15.如权利要求14所述的方法、细胞或油，其中所述微藻是原始小球藻。

16.一种产品，其通过进一步加工权利要求3-15中任一项的油制备，以生产食品、燃料或化学产品。

产自重组产油微生物的定制油

[0001] 本申请是申请人于2012年2月2日提交的题为“产自重组产油微生物的定制油”的中国专利申请201280007593.3的分案申请。

[0002] 相关申请的交叉引用

[0003] 本申请要求于2011年2月2日提交的美国临时专利申请号61/438,969;2011年4月18日提交的美国临时专利申请号61/476,691;2011年5月10日提交的美国临时专利申请号61/484,458;以及2011年10月18日提交的美国临时专利申请号61/548,616的权益。出于所有目的将这些申请的每一篇通过引用以其全文结合在此。

[0004] 序列表引用

[0005] 本申请包括在详细说明的最后所显示的序列表。

发明领域

[0006] 本发明涉及从微生物制成的油、燃料、以及油脂化学品的生产。具体地说,本披露涉及含油微藻、培养它们用于生产有用的化合物(包括脂质、脂肪酸酯、脂肪酸、醛、醇以及链烷)的方法、以及用于在基因上改变它们以提高生产效率以及改变由它们产生的油的类型和组成的方法和试剂。

[0007] 发明背景

[0008] 化石燃料是用于有机物质的埋藏的可燃地质沉积物的一般术语,其从腐败的植物和动物形成,这些有机物质已经通过在地壳中暴露于热和压力经过数百万年而转化成原油、煤炭、天然气或重油。化石燃料是有限的不可再生资源。全球经济对能源的增加的需求也对烃的成本施加了日益增长的压力。除了能源之外,许多工业,包括塑料和化学品制造商,严重依赖于作为其生产工艺原料的烃的可获得性。针对目前供应源的节约成本的替代物会有助于缓和对于能源和这些原材料成本的上升的压力。

[0009] PCT公开号2008/151149描述了用于培养微藻以生产油的方法和材料,并且具体示例了从微藻原始小球藻所产生的油生产柴油燃料。对于改进的用于生产针对燃料、化学品、食物以及其他用途的油的方法仍然存在着需要,尤其是生产具有较短链长以及较高饱和度并且无色素的油、具有较高产率以及效率的方法。本发明满足了这种需要。

[0010] 聚氨酯是一种包含氨基甲酸酯(氨基甲酸乙酯)键的化合物。典型地,聚氨酯是一种有机单元聚合物。该聚合物是通过一种包含异氰酸酯部分的第一有机单元(C(O)N-R_1-NC(O))与一种包含羟基基团的第二有机单元(HO-R_2-OH)的反应而制备的。聚氨酯可以表示为-[C(O)NH-R_1-NHC(O)-O-R_2-O]$_m$-,其中下标m是一个数字,表示该在聚合物中包含的单体的数目。R_1以及R_2可以是相同的或不同的,但典型地是不同的。聚氨酯用于许多不同的应用,包括柔性材料和刚性材料两者。聚氨酯用于鞋、汽车、飞机、套管、垫圈、胶粘剂、地毯、斯潘德克斯弹纤维、电子产品外壳等等。

[0011] 发明概述

[0012] 本发明的说明性实施例提供了产生改变的甘油酯谱的产油细胞以及从这些细胞生产的产品。产油细胞的实例包括具有II型脂质生物合成途径的微生物细胞。实施例还以

使用这些细胞可获得的天然油为特征。实施例包括表达编码例如脂肪酰基-ACP硫酯酶等蛋白质的外源基因的重组细胞。本发明还提供了从这样的细胞制造脂质和油基产品(包括燃料,例如生物柴油、可再生柴油以及喷气燃料、食用油以及化学品)的方法。

[0013] 在一个第一方面,本发明提供了一种具有至少为3%C8:0脂质谱的微藻细胞。在一些情况下,该脂质谱是至少12%C8:0。在一些实施例中,这种细胞是一种重组细胞。在一些情况下,这种重组细胞包含编码酰基-ACP硫酯酶蛋白的外源基因,该酰基-ACP硫酯酶蛋白对于链长C8的脂肪酰基-ACP底物具有水解活性。在一些实施例中,该外源基因编码湿地萼距花酰基-ACP硫酯酶。在一些情况下,该细胞是一种无绿藻细胞。在一些情况下,这种细胞属于微藻属或选自在表1中鉴定的微藻种类。

[0014] 在一个第二方面,本发明提供了一种具有至少为4%C10:0脂质谱的微藻细胞。在一些情况下,该脂质谱是至少18%C10:0。在一些情况下,该脂质谱是至少20%C10:0。在一些情况下,该微藻细胞的脂质成分进一步包含C12:0。在一些情况下,C10:0与C12:0的比率是至少3:1。在一些情况下,该微藻细胞的脂质成分进一步包含C14:0。在一些情况下,C10:0与C14:0的比率是至少10:1。在一些实施例中,这种细胞是一种重组细胞。在一些情况下,这种重组细胞包含编码脂肪酰基-ACP硫酯酶蛋白的外源基因,该脂肪酰基-ACP硫酯酶蛋白对于链长C10的脂肪酰基-ACP底物具有水解活性。在一些实施例中,该外源基因编码来自选自下组的种类的脂肪酰基-ACP硫酯酶蛋白,该组由萼距花和美洲榆组成。在一些实施例中,该脂肪酰基-ACP硫酯酶蛋白来自选自下组的种类,该组由萼距花和美洲榆组成。在一些情况下,该脂肪酰基-ACP硫酯酶基因选自下组,该组由以下各项组成:来自萼距花和美洲榆的脂肪酰基-ACP硫酯酶基因,其对于链长C10的脂肪酰基-ACP底物具有水解活性。

[0015] 在一些情况下,这种细胞是一种无绿藻细胞。在一些实施例中,这种细胞属于微藻属或选自在表1中鉴定的微藻种类。

[0016] 在一个第三方面,本发明提供了一种具有至少为13%C12:0脂质谱的微藻细胞。在一些情况下,这种细胞是一种重组细胞。在一些实施例中,这种重组细胞包含编码脂肪酰基-ACP硫酯酶蛋白的外源基因,该脂肪酰基-ACP硫酯酶蛋白对于链长C12的脂肪酰基-ACP底物具有水解活性。在一些实施例中,该脂肪酰基-ACP硫酯酶蛋白来自选自下组的种类,该组由加州月桂和樟树组成。在一些情况下,该脂肪酰基-ACP硫酯酶基因选自下组,该组由以下各项组成:来自加州月桂和樟树的脂肪酰基-ACP硫酯酶基因,其对于链长C12的脂肪酰基-ACP底物具有水解活性。在一些实施例中,这种细胞是一种无绿藻细胞。

[0017] 在一个第四方面,本发明提供了一种具有至少为10%C14:0脂质谱的微藻细胞。在一些情况下,该脂质谱是至少35%C14:0。在一些情况下,该微藻细胞的脂质成分进一步包含C12:0。在一些情况下,C14:0与C12:0的比率是至少3:1。在一些情况下,这种细胞是一种重组细胞。在一些实施例中,这种重组细胞包含编码脂肪酰基-ACP硫酯酶蛋白的外源基因,该脂肪酰基-ACP硫酯酶蛋白对于链长C14的脂肪酰基-ACP底物具有水解活性。在一些实施例中,该脂肪酰基-ACP硫酯酶蛋白来自选自下组的种类,该组由樟树和美洲榆组成。在一些情况下,该脂肪酰基-ACP硫酯酶基因选自下组,该组由以下各项组成:来自樟树和美洲榆的脂肪酰基-ACP硫酯酶基因,其具有针对链长C14的脂肪酰基-ACP底物的水解活性。在一些情况下,这种细胞是一种无绿藻细胞。在一些实施例中,这种细胞属于微藻属或选自在表1中鉴定的微藻种类。

[0018] 在一个第五方面，本发明提供了一种具有至少为15％C16:0脂质谱的微藻细胞。在一些情况下，该脂质谱是至少39％C16:0。在一些情况下，该脂质谱是至少67％C16:0。在一些情况下，这种细胞是一种重组细胞。在一些实施例中，这种重组细胞包含编码脂肪酰基-ACP硫酯酶蛋白的外源基因，该脂肪酰基-ACP硫酯酶蛋白对于链长C16的脂肪酰基-ACP底物具有水解活性。在一些实施例中，该脂肪酰基-ACP硫酯酶蛋白来自一个选自蓖距花和美洲榆的种类。在一些实施例中，这种重组细胞包含编码脂肪酰基-ACP硫酯酶蛋白的外源基因，该脂肪酰基-ACP硫酯酶蛋白来自选自下组的种类，该组由蓖距花和美洲榆组成，其对于链长C16的脂肪酰基-ACP底物具有水解活性。在一些情况下，这种细胞是一种无绿藻细胞。在一些实施例中，该微藻细胞进一步包含一种内源去饱和酶基因，其中该内源去饱和酶基因已经突变以编码一种无活性的或比非突变去饱和酶活性低的去饱和酶，或者其中所述内源去饱和酶已经从该微藻细胞基因组中缺失。

[0019] 在一个第六方面，本发明提供了一种具有至少为60％饱和脂肪酸的脂质谱的微藻细胞。在一些情况下，该微藻细胞具有至少为85％饱和脂肪酸的脂质谱。在一些情况下，这种细胞是一种重组细胞。在一些实施例中，这种重组细胞包含编码脂肪酰基-ACP硫酯酶蛋白的外源基因，该脂肪酰基-ACP硫酯酶蛋白对于链长C10-C16的脂肪酰基-ACP底物具有水解活性。在一些情况下，这种细胞是一种无绿藻细胞。

[0020] 在一个第七方面，本发明提供了一种具有至少为19％C18:0脂质谱的微藻细胞。在一些情况下，该脂质谱是至少27％C18:0。在一些情况下，这种细胞是一种重组细胞。在一些实施例中，这种重组细胞包含编码脂肪酰基-ACP硫酯酶蛋白的外源基因，该脂肪酰基-ACP硫酯酶蛋白对于链长C18的脂肪酰基-ACP底物具有水解活性。在一些实施例中，该脂肪酰基-ACP硫酯酶蛋白来自一个选自甘蓝型油菜的种类。在一些实施例中，这种重组细胞包含编码来自甘蓝型油菜的脂肪酰基-ACP硫酯酶蛋白的外源基因，该脂肪酰基-ACP硫酯酶蛋白对于链长C16的脂肪酰基-ACP底物具有水解活性。

[0021] 在一个第八方面，本发明提供了一种包含编码选自下组的脂肪酰基-ACP硫酯酶蛋白的外源基因的微藻细胞，该组由以下各项组成：蓖距花、加州月桂、樟树、湿地蓖距花、披针叶蓖距花(Cuphea lanceolata)、德国鸢尾(Iris germanica)、肉豆蔻(Myristica fragrans)、山竹(Garcinia mangostana)、油棕(Elaeis guiniensis)、甘蓝型油菜、蓖麻(Ricinus communis)以及美洲榆。

[0022] 在一个第九方面，本发明提供了一种包含表达构建体的微藻细胞，其中该表达构建体下调一种选自下组的方法的内源基因的表达，该组由以下各项组成：该内源基因已经突变以编码一种无活性的或比非突变基因产物活性低的基因产物，该内源基因已经从该微藻细胞基因组中缺失，以及通过RNA诱导机制。在一些情况下，该方法是RNA诱导机制，如RNAi、反义和/或dsRNA。在一些情况下，该内源基因是一种去饱和酶基因。在一些实施例中，该去饱和酶基因是Δ12脂肪酸去饱和酶基因。在一些情况下，这种细胞是一种无绿藻细胞。

[0023] 在一个第十方面，本发明提供了一种如任何以上方面所述的微藻细胞，其中该微藻细胞是使用水苏糖、棉子糖或蜜二糖作为碳源而培养的。

[0024] 在一个第十一方面，本发明提供了一种具有不多于2％18:2脂质谱的微藻细胞。在一些情况下，本发明提供了一种具有不多于7％18:2的脂质谱的微藻细胞。

[0025] 在一个第十二方面，本发明提供了一种制造脂质的方法。在一个实施例中，该方法

包括(a)培养如以上所讨论的一种微藻细胞直到这种细胞具有按干重计至少20％的脂质；以及(b)从水溶性生物质组分中分离该脂质。

[0026] 在一个第十三方面，本发明提供了另一种制造脂质的方法。在一个实施例中，该方法包括(a)培养一种含有编码两种不同酰基-ACP硫酯酶的两种不同外源基因的微藻细胞，以及(b)从水溶性生物质组分中分离该脂质。在一些情况下，这些外源基因中的至少一种编码选自下组的脂肪酰基-ACP硫酯酶，该组由在表4中鉴定的硫酯酶组成。

[0027] 在一个第十四方面，本发明提供了一种制造油基产品的方法。在一个实施例中，该方法包括(a)培养如以上所讨论的一种微藻细胞直到这种细胞具有按干重计至少10％的脂质；(b)从这种微藻细胞中分离该脂质；(c)使该脂质经历至少一种选自下组的化学反应，该组由以下各项组成：皂化；复分解；酸水解；碱水解；酶水解；催化水解；加压热水水解；催化水解反应，其中该脂质被裂解为甘油和脂肪酸；胺化反应，以产生脂族氮化合物；臭氧分解反应，以产生一元酸和二元酸；甘油三酯裂解反应，选自由酶裂解和压力裂解组成的组；继水解反应之后的缩合反应；加氢处理反应；加氢处理反应以及在加氢处理反应之前或与其同时的脱氧反应和/或缩合反应；脱气反应；选自下组的脱氧反应，该组由以下各项组成：氢解反应、氢化、连续的氢化-氢解反应、连续的氢解-氢化反应、以及组合的氢化-氢解反应；继脱氧反应之后的缩合反应；酯化反应；酯交换反应(interestification reaction)；酯交换反应(transesterification reaction)；羟化反应；以及继羟化反应之后的缩合反应；以及(d)任选地将该反应的产物与其他组分分离，由此产生一种油基产品。

[0028] 在一些情况下，该油基产品选自肥皂或一种燃料产品。在一些实施例中，该油基产品是一种选自下组的燃料产品，该组由以下各项组成：生物柴油、可再生柴油、以及喷气燃料。在一些情况下，该燃料产品是生物柴油，具有以下一种或多种属性：(i)0.025-0.3mcg/g，优选0.05-0.244mcg/g总类胡萝卜素；(ii)少于0.005mcg/g，优选少于0.003mcg/g番茄红素；(iii)少于0.005mcg/g，优选少于0.003mcg/gβ胡萝卜素；(iv)0.025-0.3mcg/g，优选0.045-0.268mcg/g叶绿素A；(v)35-175mcg/g，优选38.3-164mcg/gγ生育酚；(vi)少于0.25％菜子甾醇、菜油甾醇、豆甾醇或β谷固醇；(vii)225-350mcg/g，优选249.6-325.3mcg/g总生育三烯酚；(viii)0.0025-0.05mcg/g，优选0.003-0.039mcg/g叶黄素；或(ix)50-300mcg/g，优选60.8-261.7mcg/g生育酚。在一些情况下，该燃料产品是具有至少20℃、40℃或60℃的T10-T90的可再生柴油。在一些情况下，该燃料产品是符合HRJ-5和/或ASTM规范D1655的喷气燃料。

[0029] 在一个第十五方面，本发明提供了一种甘油三酯油，该甘油三酯油包括(a)一种具有至少1％-5％、优选至少3％的C8:0，至少2.5％、优选至少4％的C10:0，至少10％、优选至少13％的C12:0，至少10％的C14:0、和/或至少60％的饱和脂肪酸的脂质谱；以及(b)以下一种或多种属性：(i)0.025-0.3mcg/g，优选0.05-0.244mcg/g总类胡萝卜素；(ii)少于0.005mcg/g，优选少于0.003mcg/g番茄红素；(iii)少于0.005mcg/g，优选少于0.003mcg/gβ胡萝卜素；(iv)0.025-0.3mcg/g，优选0.045-0.268mcg/g叶绿素A；(v)35-175mcg/g，优选38.3-164mcg/gγ生育酚；(vi)少于0.25％菜子甾醇、菜油甾醇、豆甾醇或β谷固醇；(vii)225-350mcg/g，优选249.6-325.3mcg/g总生育三烯酚；(viii)0.0025-0.05mcg/g，优选0.003-0.039mcg/g叶黄素；或(ix)50-300mcg/g，优选60.8-261.7mcg/g生育酚。

[0030] 在一个第十六方面，本发明提供了一种从微藻分离的甘油三酯油，该甘油三酯油

具有至少5:1的C8:C10脂肪酸比率。在一些实施例中，甘油三酯油分离自微藻细胞(例如，无绿藻属的微藻细胞)，其中该微藻细胞包含一种外源基因。在一个相关方面，本发明提供了一种从微藻分离的具有至少60％饱和脂肪酸的甘油三酯油。

[0031] 在另一个相关方面，本发明提供了一种从微藻分离的具有大约2:1的C16:14脂肪酸比率的脂质谱的甘油三酯油。在另一个相关方面，本发明提供了一种由微藻细胞产生的甘油三酯油，其中该微藻细胞包含一种外源基因。在一些情况下，该微藻属于无绿藻属。

[0032] 在另一个相关方面，本发明提供了一种从微藻分离的甘油三酯油，该甘油三酯油具有大约3:1的C12:14脂肪酸比率的脂质谱。在另一个相关方面，本发明提供了一种由微藻细胞产生的甘油三酯油，其中该微藻细胞包含一种外源基因。在一些情况下，该微藻属于无绿藻属。

[0033] 在一个第十七方面，本发明提供了一种甘油三酯油，该甘油三酯油包括(a)一种具有少于1％的<C12；在1％-10％之间，优选2％-7％的C14:0；在20％-35％之间，优选23％-30％的C16:0；在5％-20％之间，优选7％-15％的C18:0；在35％-60％之间，优选40％-55％的C18:1；以及在1％-20％之间，优选2％-15％的C18:2脂肪酸的脂质谱；以及(b)以下一种或多种属性：(i)0.025-0.3mcg/g，优选0.05-0.244mcg/g总类胡萝卜素；(ii)少于0.005mcg/g，优选少于0.003mcg/g番茄红素；(iii)少于0.005mcg/g，优选少于0.003mcg/gβ胡萝卜素；(iv)0.025-0.3mcg/g，优选0.045-0.268mcg/g叶绿素A；(v)35-175mcg/g，优选38.3-164mcg/gγ生育酚；(vi)少于0.25％菜子甾醇、菜油甾醇、豆甾醇或β谷固醇；(vii)225-350mcg/g，优选249.6-325.3mcg/g总生育三烯酚；(viii)0.0025-0.05mcg/g，优选0.003-0.039mcg/g叶黄素；或(ix)50-300mcg/g，优选60.8-261.7mcg/g生育酚。

[0034] 在一些情况下，该甘油三酯油是分离自包含一种或多种外源基因的微生物。在一些实施例中，该一种或多种外源基因编码脂肪酰基-ACP硫酯酶。在一些情况下，该脂肪酰基-ACP硫酯酶对于链长C14的脂肪酰基-ACP底物具有水解活性。在一些实施例中，该微生物进一步包含表达构建体，其中该表达构建体下调一种选自下组的方法的内源基因的表达，该组由以下各项组成：该内源基因已经突变以编码一种无活性的或比非突变基因产物活性低的基因产物，该内源基因已经从该微藻细胞基因组中缺失，以及通过RNA诱导机制。在一些实施例中，该内源基因编码一种去饱和酶。在一些情况下，该去饱和酶是一种硬脂酰-酰基载体蛋白去饱和酶(SAD)或一种脂肪酸去饱和酶(FAD)。

[0035] 在一个第十八方面，本发明提供了生产具有一种具有低于1％的<C12、在2％-7％之间的C14:0、在23％-30％之间的C16:0、在7％-15％之间的C18:0、在40％-55％之间的C18:1、以及在2％-15％之间的C18:2脂肪酸的脂质谱的甘油三酯油的方法，其中该甘油三酯油分离自包含一种或多种外源基因的微生物。在一些情况下，该甘油三酯油包含一种具有3％-5％C14:0、25％-27％C16:0、10％-15％C18:0；以及40％-45％C18:1的脂质谱。在一些实施例中，该一种或多种外源基因编码脂肪酰基-ACP硫酯酶。在一些情况下，该脂肪酰基-ACP硫酯酶对于链长C14的脂肪酰基-ACP底物具有水解活性。

[0036] 在一个第十九方面，本发明提供了一种产生蓖麻油酸的微生物细胞。在一些情况下，该微生物细胞是一种微藻细胞。在一些情况下，该微生物细胞和该微藻细胞包含一种编码脂肪酸羟化酶的外源基因。在一些实施例中，该脂肪酸羟化酶是一种油酸12-羟化酶。在一些情况下，该脂肪酸羟化酶来自蓖麻。在一些情况下，该微生物细胞属于无绿藻属，例如

像Prototheca moriformis。

[0037] 在一个第二十方面，本发明提供了一种包含编码α半乳糖苷酶的外源基因的微藻细胞。在一些情况下，该微藻细胞包含编码α半乳糖苷酶的外源基因，其中该α半乳糖苷酶是分泌的。在一些实施例中，该编码α半乳糖苷酶的外源基因来自一个选自下组的属，该组由以下各项组成：酵母菌属、曲霉属以及瓜儿豆属。

[0038] 在一个第二十一方面，本发明提供了一种产生脂质组合物的方法，该方法包括以下步骤：(a)在一种固定碳源的存在下、在异养条件下培养一种微藻细胞，其中该微藻细胞包含一种编码α半乳糖苷酶的外源基因并且该固定碳源选自下组，该组由以下各项组成：棉子糖、水苏糖以及蜜二糖；(b)将该脂质与非脂质组分分离；由此产生一种脂质组合物。在一些情况下，该微藻细胞属于无绿藻属。

[0039] 在一个第二十二方面，本发明提供了一种包含编码THIC酶的外源基因的微藻细胞。在一些情况下，该THIC酶来自一个选自下组的属，该组由以下各项组成：胶球藻属、拟南芥属、以及集胞藻属。

[0040] 在另一方面，本发明提供了一种在不存在硫胺素下培养微藻细胞的方法，该方法包括表达一种编码THIC酶的外源基因。在一些情况下，该THIC酶来自一个选自下组的属，该组由以下各项组成：胶球藻属、拟南芥属、以及集胞藻属。

[0041] 在其他方面，本发明提供了一种如在此所述的微藻细胞，其中所述微藻细胞进一步包含另一个选自下组的外源基因，该组由以下各项组成：蔗糖转化酶、α半乳糖苷酶、以及THIC酶。

[0042] 本发明的另一个方面提供了一种分离自微生物细胞的羟化油。在一个方面，该羟化油分离自一种微生物细胞，其中该微生物细胞是一种微藻细胞(例如，无绿藻属的微藻细胞)。在一个另外的方面，该羟化油分离自Prototheca moriformis细胞。在一个实施例中，该羟化油是一种羟化甘油三酯。该羟化甘油三酯在化学上可以是与蓖麻油类似或相同的。

[0043] 本发明的一个另外的方面是羟化脂肪酸。该羟化脂肪酸的一个实施例是蓖麻油酸。

[0044] 在又另一个方面，该微生物羟化油或羟化脂肪酸被进一步被羟化。当蓖麻油酸被羟化时，提供了一种含有两个羟基基团的脂肪酸。

[0045] 本发明的又另一个方面提供了一种通过使一种羟化油和/或一种羟化脂肪酸与一种含有异氰酸酯部分的化合物进行反应以形成一种聚氨酯而制备的组合物。

[0046] 本发明的另一个方面提供了一种具有至少为20％C18:2脂质谱的微藻细胞。在一些情况下，该微藻细胞具有至少为30％C18:2的脂质谱。在一些情况下，该微藻细胞具有至少为40％C18:2的脂质谱。在一些情况下，该微藻细胞具有至少为50％C18:2的脂质谱。

[0047] 本发明的另一个方面提供了一种制造脂质的方法，该方法包括：(a)在培养基中培养微藻细胞并监测糖浓度；(b)当培养基的糖浓度达到低于大约每升1克时，以大约每升每小时2克至大约每升每小时10克之间的连续速率将一种第一糖溶液添加至培养基，持续大约2至大约24小时；(c)将一种第二糖溶液添加至培养基以维持培养基每升大约15克至大约20克的糖浓度；以及(d)将该脂质与该微藻生物质分离。在一些情况下，该第一糖溶液是以大约每升每小时4克蔗糖至大约每升每小时6克蔗糖的速率添加至培养基。在一些情况下，该第一糖溶液是以大约每升每小时5.25克蔗糖的速率添加至该培养基。在一些情况下，该

糖是蔗糖或葡萄糖。

[0048] 本发明的另一个方面提供了一种具有至少为10％C18:3的脂质谱。在一些情况下，该微藻细胞具有至少为20％C18:3的脂质谱。在一些情况下，该微藻细胞具有至少为30％C18:3的脂质谱。在一些情况下，该微藻细胞具有至少为40％C18:3的脂质谱。在一些情况下，该微藻细胞具有至少为50％C18:3的脂质谱。

[0049] 本发明的另一个方面提供了一种产生包括亚油酸或亚麻酸的甘油三酯的微生物，其中该微生物包含一种编码β-酮脂酰-ACP合成酶II(KAS II)的重组核酸。在一些情况下，该微生物进一步包含一种编码硬脂酰ACP去饱和酶(SAD)的重组核酸。在一些情况下，该微生物进一步包含一种编码油酸特异性硫酯酶的重组核酸。在一些情况下，该微生物进一步包含一种编码脂肪酸去饱和酶(FAD)的重组核酸。在一些情况下，该微生物进一步包含一种编码甘油酯去饱和酶的重组核酸。

[0050] 在以上讨论的工程化微生物中，该KAS II酶可以包含一个选自下组的氨基酸序列，该组由以下各项组成：SEQ ID NO:175、SEQ ID NO:176、SEQ ID NO:177、SEQ ID NO:178以及SEQ ID NO:179。在一些情况下，该SAD酶可以包含一个选自下组的氨基酸序列，该组由以下各项组成：SEQ ID NO:172、SEQ ID NO:196、SEQ ID NO:197、SEQ ID NO:198、SEQ ID NO:199、SEQ ID NO:200、以及SEQ ID NO:201。在一些情况下，该油酸特异性硫酯酶可以包含一个选自下组的氨基酸序列，该组由以下各项组成：SEQ ID NO:195、SEQ ID NO:202、SEQ ID NO:203、SEQ ID NO:204、SEQ ID NO:205、以及SEQ ID NO:206。在一些情况下，该FAD酶可以包含一个选自下组的氨基酸序列，该组由以下各项组成：SEQ ID NO:181、SEQ ID NO:182、SEQ ID NO:183、SEQ ID NO:184以及SEQ ID NO:185。在一些情况下，该FAD酶是一种Δ12FAD酶，该Δ12FAD酶包含一个选自下组的氨基酸序列，该组由以下各项组成：SEQ ID NO:207、SEQ ID NO:208、SEQ ID NO:209、SEQ ID NO:210、SEQ ID NO:211、以及SEQ ID NO:212。在一些情况下，该FAD酶是一种Δ15FAD酶，该Δ15FAD酶包含一个选自下组的氨基酸序列，该组由以下各项组成：SEQ ID NO:213、SEQ ID NO:214、SEQ ID NO:215、SEQ ID NO:216、SEQ ID NO:217、SEQ ID NO:218、SEQ ID NO:219、SEQ ID NO:220、以及SEQ ID NO:221。在一些情况下，该甘油酯去饱和酶是一种ω-6脂肪酸去饱和酶、ω-3脂肪酸去饱和酶、或ω-6油酸去饱和酶。

[0051] 以上讨论的工程化微生物中的任何之一可以进一步包含一种编码蔗糖利用途径酶的重组核酸。在一些情况下，该蔗糖利用途径酶是一种蔗糖转化酶。

[0052] 在以上讨论的工程化微生物的任何之一中，该微生物可以进一步被工程化以增加亚油酸或亚麻酸相对于其他脂肪酸的比例。

[0053] 在以上讨论的工程化微生物的任何之一中，该微生物可以进一步被工程化以过表达对C18底物特异或优先的硫酯酶。

[0054] 在以上讨论的工程化微生物的任何之一中，该微生物可以进一步被工程化以减少对C8-C16底物特异或优先的硫酯酶的表达。

[0055] 在以上讨论的工程化微生物的任何之一中，该微生物可以是一种微藻细胞。在一些情况下，该微藻细胞属于无绿藻属。在一些情况下，该微藻细胞是Prototheca moriformis细胞。

[0056] 本发明的另一个方面提供了一种由以上讨论的工程化微生物中的任何之一所产

生的油。

[0057] 本发明的另一个方面提供了通过将一种或多种重组核酸引入这些微生物来产生以上讨论的工程化微生物以生产包含亚油酸或亚麻酸的甘油三酯的方法。

[0058] 本发明的另一个方面提供了一种用于生产包含三酰甘油酯的天然油或产自该天然油的产品的方法，其中该方法包括(i)培养一种重组微生物细胞，这种细胞包含可操作的重组核酸以(a)减少或消除由编码β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、脂肪酸去饱和酶、或酰基-ACP硫酯酶的一种或多种基因编码的酶的表达，并且任选地其中这种细胞包含可操作的重组核酸以减少或消除由编码β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶、或酰基-ACP硫酯酶的基因的两个拷贝的表达；或(b)表达编码活性β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶、或酰基-ACP硫酯酶的外源基因的产物；或(c)减少或消除由编码β-酮脂酰-ACP合成酶I或β-酮脂酰-ACP合成酶II的一种或多种基因编码的酶的表达，并且表达编码活性硬脂酰ACP去饱和酶、脂肪酸去饱和酶、或酰基-ACP硫酯酶的外源基因的产物；或(d)减少或消除由编码硬脂酰ACP去饱和酶或脂肪酸去饱和酶的一种或多种基因编码的酶的表达，并且表达编码活性β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、或酰基-ACP硫酯酶的外源基因的产物；以及(ii)从这种细胞中回收该天然油，并且任选地进一步对该天然油进行加工以生产食物、燃料或化学产品，其中该天然油由于这些重组核酸而具有改变的脂肪酸谱。

[0059] 在一些情况下，该微生物通过II型脂肪酸生物合成途径来合成脂肪酸。在一些情况下，该微生物是一种微藻。在一些情况下，该微藻是一种专性异养生物。在一些情况下，该微藻是无绿藻属或小球藻属的一个种类。在一些情况下，该微藻是魏氏无绿藻(Prototheca wickerhamii)、雍滞无绿藻(Prototheca stagnora)、Prototheca portoricensis、Prototheca moriformis、或祖菲无绿藻(Prototheca zopfii)。在一些情况下，该微藻是凯氏小球藻(Chlorella kessleri)、Chlorella luteoviridis、原始小球藻、或普通小球藻(Chlorella vulgaris)。在一些情况下，这种细胞是一种表达活性蔗糖转化酶的重组细胞。在一些情况下，该培养是异养的。在一些情况下，该脂肪酸去饱和酶是ω-6脂肪酸去饱和酶、ω-3脂肪酸去饱和酶、或ω-6油酸去饱和酶、或Δ12脂肪酸去饱和酶中的一种或多种。在一些情况下，将这种细胞进行培养以便包含按干细胞重量计在至少50％、至少60％、至少70％、或50％至90％之间的甘油三酯。在一些情况下，该油包含少于500、50、或5ppm的着色分子。在一些情况下，这些重组核酸被稳定地整合。在一些情况下，这些重组核酸被稳定地整合到该微生物的染色体中。在一些情况下，这种细胞进一步包含至少一种选择标记。

[0060] 在一些情况下，这种细胞包含可操作的重组核酸以通过反义、RNAi、或dsRNA的表达来减少或消除一种酶的表达，这些反义、RNAi、或dsRNA靶向编码该酶的基因的转录本。在一些情况下，由编码β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶的一种或多种基因编码的酶的表达的减少或消除是由于用一种或多种编码活性β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶、或酰基-ACP硫酯酶的基因对这一种或多种基因的中断或替换。在一些情况下，这种重组细胞进一步包含一种编码油酸12-羟化酶的外源基因，以便合成蓖麻油酸。在一些情况下，这种重组细胞包含可操作的核酸以减少或消除由KASII基因编码的β-酮脂酰-ACP合成

酶II的表达，并且表达编码酰基-ACP硫酯酶的外源基因的产物。在一些情况下，这种细胞产生一种具有特征在于具有至少40％、50％、60％、70％、或80％C16脂肪酸的脂肪酸谱的油。在一些情况下，这种细胞产生一种具有特征在于具有至少50％-75％C16:0的脂肪酸谱的油。在一些情况下，这种细胞产生一种具有特征进一步在于具有至少20％-40％C18:1的脂肪酸谱的油。在一些情况下，该编码酰基-ACP硫酯酶的外源基因产生一种活性酰基-ACP硫酯酶，该活性酰基-ACP硫酯酶在C8-C16脂肪酰基链的水解中具有比这种细胞的天然酰基-ACP硫酯酶更强的活性。在一些情况下，该编码酰基-ACP硫酯酶的外源基因中断该KASII基因。在一些情况下，这种重组细胞包含可操作的核酸以减少或消除由编码β-酮脂酰-ACP合成酶I的一种或多种基因编码的酶的表达。

[0061] 在一些情况下，所产生的油具有一种其特征在于由于这些重组核酸而更短的平均脂肪酸链长的脂肪酸谱。在一些情况下，这种重组细胞包含可操作的核酸以减少或消除由至少一种FAD基因编码的脂肪酸去饱和酶的表达，并且表达编码活性硬脂酰ACP去饱和酶的硬脂酰-ACP去饱和酶外源基因的产物。在一些情况下，核酸是可操作的以减少或消除由脂肪酸去饱和酶基因的多个拷贝编码的脂肪酸去饱和酶的表达。在一些情况下，该硬脂酰-ACP去饱和酶外源基因被重组进入该脂肪酸去饱和酶基因的编码区之内的一个基因座。在一些情况下，所产生的油具备一种具有升高的油酸的脂肪酸谱。在一些情况下，油酸构成这些脂肪酸的至少50％、60％、70％、80％、或90％。在一些情况下，这种重组细胞包含可操作的核酸以表达编码活性β-酮脂酰-ACP合成酶II的β-酮脂酰-ACP合成酶II外源基因的产物。在一些情况下，所产生的油的特征在于一种由于这些重组核酸所致的C18:1脂肪酸升高和C16脂肪酸降低的脂肪酸谱。在一些情况下，这种重组细胞包含可操作的核酸以通过RNA干扰来减少或消除由编码硬脂酰ACP去饱和酶的一种或多种基因编码的酶的表达。在一些情况下，所产生的油具有一种其特征在于C18:0脂肪酸增加的脂肪酸谱。在一些情况下，所产生的油的特征在于一种具有至少50％、60％、70％、80％、或90％C18:0的脂肪酸谱。在一些情况下，所产生的油的特征在于一种具有至少50％-75％C18:0脂肪酸谱。在一些情况下，所产生的油的特征另外地在于一种具有至少20％-40％C18:1的脂肪酸谱。在一些情况下，这种细胞包含可操作的重组核酸以减少或消除编码β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶的基因的两个拷贝的表达。在一些情况下，这些核酸是可操作的以表达编码活性ω-3脂肪酸去饱和酶和/或ω-6油酸去饱和酶的脂肪酸去饱和酶外源基因的产物。在一些情况下，所产生的油具有一种其特征在于水平提高的亚油酸、亚麻酸、或两者的脂肪酸谱。在一些情况下，该油的脂肪酸谱的特征在于具有至少10％、20％、30％、40％、或50％的亚油酸、亚麻酸、或两者。在一些情况下，该油的进一步的加工包括精炼、漂白、脱臭、复分解、酯交换(transesterification)、氢化、水解、氢化、脱氧、加氢裂解、异构化、羟基化、酯交换(interesterification)、酰胺化、磺化、以及硫化中的一种或多种。在一些情况下，该油被加工以产生食用油、脂肪酸、脂肪醇、润滑剂、肥皂、脂肪酸酯、脂肪酸乙氧基化物、脂肪胺、烷基氯、脂肪醇乙氧基化物、脂肪醇硫酸酯、脂肪酸烷醇酰胺、磺化油、硫化油、柴油燃料、喷气燃料、汽油、燃料混合原料(fuel blendstoc)、燃料添加剂、润滑添加剂、或涂层。

[0062] 本发明的另一个方面提供了通过以上讨论的方法而可获得的天然油。

[0063] 本发明的另一个方面提供了一种从以上讨论的天然油制成的产品。在一些情况

下，该产品包括食用油、脂肪酸、脂肪醇、润滑剂、肥皂、脂肪酸酯、脂肪酸乙氧基化物、脂肪胺、烷基氯、脂肪醇乙氧基化物、脂肪醇硫酸酯、脂肪酸烷醇酰胺、磺化油、硫化油、柴油燃料、喷气燃料、汽油、燃料混合原料、燃料添加剂、化学添加剂、或涂层。

[0064] 本发明的另一个方面提供了一种重组细胞，这种重组细胞包含可操作的重组核酸以(a)减少或消除由编码β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、脂肪酸去饱和酶、或酰基-ACP硫酯酶的一种或多种基因编码的酶的表达，并且任选地其中这种细胞包含可操作的重组核酸以减少或消除由编码β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶、或酰基-ACP硫酯酶的基因的两个拷贝的表达；或(b)表达编码活性β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶、或酰基-ACP硫酯酶的外源基因的产物；或(c)减少或消除由编码β-酮脂酰-ACP合成酶I或β-酮脂酰-ACP合成酶II的一种或多种基因编码的酶的表达，并且表达编码活性硬脂酰ACP去饱和酶、脂肪酸去饱和酶、或酰基-ACP硫酯酶的外源基因的产物；或(d)减少或消除由编码硬脂酰ACP去饱和酶或脂肪酸去饱和酶的一种或多种基因编码的酶的表达，并且表达编码活性β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、或酰基-ACP硫酯酶的外源基因的产物。

[0065] 在一些情况下，该微生物通过II型脂肪酸生物合成途径来合成脂肪酸。在一些情况下，该微生物是一种微藻。在一些情况下，该微藻是一种专性异养生物。在一些情况下，该微藻是无绿藻属或小球藻属的一个种类。在一些情况下，该微藻是魏氏无绿藻、雍滞无绿藻、Prototheca portoricensis、Prototheca moriformis、或祖菲无绿藻。在一些情况下，该微藻是凯氏小球藻、Chlorella luteoviridis、原始小球藻、或普通小球藻。在一些情况下，该细胞是一种表达活性蔗糖转化酶的重组细胞。在一些情况下，该细胞能够异养生长。在一些情况下，该脂肪酸去饱和酶是ω-6脂肪酸去饱和酶、ω-3脂肪酸去饱和酶、或ω-6油酸去饱和酶、或Δ12脂肪酸去饱和酶中的一种或多种。在一些情况下，这种细胞能够进行培养以便包含按干细胞重量计在至少50％、至少60％、至少70％、或50％至90％之间的甘油三酯。在一些情况下，这些重组核酸被稳定地整合。在一些情况下，这些重组核酸被稳定地整合到该微生物的染色体中。在一些情况下，这种细胞进一步包含至少一种选择标记。在一些情况下，这种细胞包含可操作的重组核酸以通过反义、RNAi或dsRNA的表达来减少或消除一种酶的表达，这些反义、RNAi或dsRNA靶向编码该酶的基因的转录本。在一些情况下，由编码β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、脂肪酸去饱和酶的一种或多种基因编码的酶的表达的减少或消除是由于用一种或多种编码活性β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶、或酰基-ACP硫酯酶的基因对这一种或多种基因的中断或替换。

[0066] 在一些情况下，这种重组细胞进一步包含一种编码油酸12-羟化酶的外源基因，以合成蓖麻油酸。在一些情况下，这种重组细胞包含可操作的核酸以减少或消除由KASII基因编码的β-酮脂酰-ACP合成酶II的表达，并且表达编码酰基-ACP硫酯酶的外源基因的产物。在一些情况下，这种细胞产生一种具有其特征在于具有至少40％、50％、60％、70％、或80％C16脂肪酸的脂肪酸谱的油。在一些情况下，这种细胞产生一种具有其特征在于具有至少50％-75％C16:0的脂肪酸谱的油。在一些情况下，这种细胞产生一种具有其特征进一步地在于具有至少20％-40％C18:1的脂肪酸谱的油。在一些情况下，该编码酰基-ACP硫酯酶的

外源基因产生一种活性酰基-ACP硫酯酶，该活性酰基-ACP硫酯酶在C8-C16脂肪酰基链的水解中具有比该细胞的天然酰基-ACP硫酯酶更强的活性。在一些情况下，该编码酰基-ACP硫酯酶的外源基因中断该KASII基因。在一些情况下，这种重组细胞包含可操作的核酸以减少或消除由编码β-酮脂酰-ACP合成酶I的一种或多种基因编码的酶的表达。在一些情况下，所产生的油具有一种其特征在于由于这些重组核酸而更短的平均脂肪酸链长的脂肪酸谱。在一些情况下，这种重组细胞包含可操作的核酸以减少或消除由至少一种FAD基因编码的脂肪酸去饱和酶的表达，并且表达编码活性硬脂酰ACP去饱和酶的硬脂酰-ACP去饱和酶外源基因的产物。在一些情况下，这些核酸是可操作的以减少或消除由脂肪酸去饱和酶基因的多个拷贝编码的脂肪酸去饱和酶的表达。在一些情况下，该硬脂酰-ACP去饱和酶外源基因是被重组进入该脂肪酸去饱和酶基因的编码区之内的一个基因座。

[0067] 在一些情况下，所产生的油具备一种具有升高的油酸的脂肪酸谱。在一些情况下，油酸构成这些脂肪酸的至少50％、60％、70％、80％、或90％。在一些情况下，这种重组细胞包含可操作的核酸以表达编码活性β-酮脂酰-ACP合成酶II的β-酮脂酰-ACP合成酶II外源基因的产物。在一些情况下，所产生的油的特征在于一种由于这些重组核酸所致的C18:1脂肪酸升高和C16脂肪酸降低的脂肪酸谱。在一些情况下，这种重组细胞包含可操作的核酸以通过RNA干扰来减少或消除由编码硬脂酰ACP去饱和酶的一种或多种基因编码的酶的表达。在一些情况下，所产生的油具有一种其特征在于C18:0脂肪酸增加的脂肪酸谱。在一些情况下，所产生的油的特征在于一种具有至少50％、60％、70％、80％、或90％C18:0的脂肪酸谱。在一些情况下，所产生的油的特征在于一种具有至少50％-75％C18:0的脂肪酸谱。在一些情况下，所产生的油的特征另外地在于一种具有至少20％-40％C18:1的脂肪酸谱。在一些情况下，这种细胞包含可操作的重组核酸以减少或消除编码β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶的基因的两个拷贝的表达。在一些情况下，这些核酸是可操作的以表达编码活性ω-3脂肪酸去饱和酶和/或ω-6油酸去饱和酶的脂肪酸去饱和酶外源基因的产物。在一些情况下，所产生的油具有一种其特征在于水平提高的亚油酸、亚麻酸、或两者的脂肪酸谱。在一些情况下，该油的脂肪酸谱的特征在于具有至少10％、20％、30％、40％、或50％的亚油酸、亚麻酸、或两者。

[0068] 本发明的另一个方面提供了一种从以上讨论的细胞所产生的天然油或含油产品。

[0069] 本发明的另一方面提供了一种用于生产包含三酰甘油酯(包括蓖麻油酸)的天然油或产自该天然油的产品的方法，该方法包括培养一种重组微生物细胞，这种细胞包含可操作的重组核酸以表达编码活性油酸12-羟化酶的外源基因的产物，以便合成该蓖麻油酸。

[0070] 在一些情况下，该微生物具有II型脂肪酸生物合成途径。在一些情况下，该微生物是一种微藻。在一些情况下，该微藻是一种专性异养生物。在一些情况下，该微藻是无绿藻属的一个种类。在一些情况下，该微藻是魏氏无绿藻、雍滞无绿藻、Prototheca portoricensis、Prototheca moriformis、或祖菲无绿藻。在一些情况下，该微藻是凯氏小球藻、Chlorella luteoviridis、原始小球藻、或普通小球藻。在一些情况下，这种细胞是一种表达活性蔗糖转化酶的重组细胞。在一些情况下，该培养是异养。在一些情况下，这种细胞在缺乏重组核酸的情况下产生至少40％、50％、60％、70％、80％、或90％的油酸，这些重组核酸可操作以表达编码活性油酸12-羟化酶的外源基因的产物。在一些情况下，这种细胞进一步包含可操作的重组核酸以增强油酸的产生，从而提高油酸12-羟化酶的底物水平。在

一些情况下,这种细胞包含可操作的重组核酸以(a)表达编码活性硬脂酰ACP去饱和酶的外源基因的产物并且减少或消除由编码脂肪酸去饱和酶的一种或多种基因编码的酶的表达;或(b)表达编码活性β-酮脂酰-ACP合成酶I的外源基因的产物并且表达编码活性酰基-ACP硫酯酶的外源基因的产物。

[0071] 本发明的另一个方面提供了根据以上讨论的任何方法所生产的产品。

[0072] 本发明的另一个方面提供了一种微生物细胞,这种微生物细胞包含可操作的重组核酸以表达编码活性油酸12-羟化酶的外源基因的产物,以便合成蓖麻油酸。

[0073] 在一些情况下,该微生物具有II型脂肪酸生物合成途径。在一些情况下,该微生物是一种微藻。在一些情况下,该微藻是一种专性异养生物。在一些情况下,该微藻是无绿藻属的一个种类。在一些情况下,该微藻是魏氏无绿藻、雍滞无绿藻、Prototheca portoricensis、Prototheca moriformis、或祖菲无绿藻。在一些情况下,该微藻是凯氏小球藻、Chlorella luteoviridis、原始小球藻、或普通小球藻。在一些情况下,这种细胞是一种表达活性蔗糖转化酶的重组细胞。在一些情况下,该细胞能够异养生长。在一些情况下,细胞在缺乏重组核酸的情况下产生至少40%、50%、60%、70%、80%、或90%的油酸,这些重组核酸可操作以表达编码活性油酸12-羟化酶的外源基因的产物。在一些情况下,这种细胞进一步包含可操作的重组核酸以增强油酸的产生,以便提高油酸12-羟化酶的底物水平。在一些情况下,该细胞包含可操作的重组核酸以(a)表达编码活性硬脂酰ACP去饱和酶的外源基因的产物并且减少或消除由编码脂肪酸去饱和酶的一种或多种基因编码的酶的表达;或(b)表达编码活性β-酮脂酰-ACP合成酶I的外源基因的产物并且表达编码活性酰基-ACP硫酯酶的外源基因的产物。

[0074] 本发明的另一个方面提供了一种包含如以上讨论的油的食品。

[0075] 本发明的这些以及其他方面与实施例描述于附图(紧随其后是简要说明)、以下对本发明的详细描述中,并且在以下实例中进行了示例。上文讨论和贯穿本申请的任何或所有特征可以在本发明的多种实施例中组合。

[0076] 附图简要说明

[0077] 图1显示了产自无绿藻属甘油三酯油的可再生柴油的色谱图。

[0078] 图2显示了与蓖麻油酸标准品和野生型对照相比较的代表性阳性转基因克隆的GC保留时间。

[0079] 图3显示了具有或不具有靶向基因破坏的Prototheca moriformis FADc等位基因的PstI限制酶切图谱,如在实例18中所描述。

[0080] 图4显示了在实例18中所描述的Southern印迹的结果。

[0081] 发明的详细说明

[0082] 本发明的说明性实施例的特征是产生改变的甘油酯谱的产油细胞以及从这些细胞生产的产品。产油细胞的实例包括具有II型脂质生物合成途径的微生物细胞。实施例包括重组细胞,这些重组细胞表达一种或多种编码例如脂肪酰基-ACP硫酯酶、脂肪酸去饱和酶、酮脂酰合成酶的蛋白质的外源基因并且任选地具有编码拥有类似活性的蛋白质的内源基因的一种或多种敲低。结果,一些实施例的特征在于之前从未获得的天然油。本发明还提供了从这样的细胞制造脂质和油基产品(包括燃料,例如生物柴油、可再生柴油以及喷气燃料、食用油以及化学品)的方法。

[0083] 根据本发明的实施例所生产的油可以除其他应用之外用在运输燃料、油脂化学品、和/或食品以及化妆品工业中。例如，脂质的酯交换可以产生作为生物柴油有用的长链脂肪酸酯。可以调节其他酶促和化学过程以产生脂肪酸、醛、醇、链烷、和烯。在一些应用中，产生了可再生柴油、喷气燃料、或其他烃化合物。本发明还提供了培养微藻的方法，该微藻用于增加生产力以及增加脂质产率，和/或用于更节约成本地生产在此所述的组合物。

[0084] 本发明的一个实施例提供了一种用于生产包含三酰甘油酯的天然油或用于生产产自该天然油的产品的方法。该天然油可以是一种非植物油或非种子油。该方法包括培养一种重组微生物细胞以生产一种定制油；即，由于在这种细胞中的这些重组核酸的存在而导致具有改变的脂肪酸谱的一种油。然后该天然油可以进一步被加工以生产一种食品、燃料、或化学产品。在这种细胞中的这些重组核酸是可操作的以(a)减少或消除由编码β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、脂肪酸去饱和酶、或酰基-ACP硫酯酶的一种或多种基因编码的酶的表达。任选地该细胞包含可操作的重组核酸以减少或消除编码β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶、或酰基-ACP硫酯酶的基因的两个拷贝(例如，二倍体生物中的两个等位基因)的表达；或(b)表达编码活性β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶、或酰基-ACP硫酯酶的外源基因的产物；或(c)减少或消除由编码β-酮脂酰-ACP合成酶I或β-酮脂酰-ACP合成酶II的一种或多种基因编码的酶的表达，并且表达编码活性硬脂酰ACP去饱和酶、脂肪酸去饱和酶、或酰基-ACP硫酯酶的外源基因的产物；或(d)减少或消除由编码硬脂酰ACP去饱和酶或脂肪酸去饱和酶的一种或多种基因编码的酶的表达，并且表达编码活性β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、或酰基-ACP硫酯酶的外源基因的产物。

[0085] 在编码一种或多种脂肪酸去饱和酶的重组核酸存在于这种细胞中的情况下，该核酸可以编码ω-6脂肪酸去饱和酶、ω-3脂肪酸去饱和酶、或ω-6油酸去饱和酶、或Δ12脂肪酸去饱和酶中的一种或多种。

[0086] 在这种细胞包含可操作的重组核酸以减少或消除一种酶的表达的情况下，这可以通过反义、RNAi、或dsRNA(这些反义、RNAi或dsRNA靶向编码该酶的基因的转录本)的表达而发生，或者通过其他适合的手段而发生，包括定向突变、完全缺失、或部分缺失。因此，由编码β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶的一种或多种基因编码的酶的表达的减少或消除可以是由于用一种或多种编码活性β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶、或酰基-ACP硫酯酶的基因对这一种或多种基因的中断或替换。

[0087] 优选地，这些重组核酸是被稳定地整合到这种细胞中；例如，进入细胞的染色体中，或游离基因中。可以使用一种选择标记来辅助选择具有稳定整合的核酸的细胞，选择标记如在此所述的蔗糖转化酶、抗生素抗性基因、或硫胺素营养缺陷型互补。

[0088] 优选地，该微生物可以是通过II型脂肪酸生物合成途径来合成脂肪酸的一种微生物。例如，该微生物可以是微藻，但还可以是使用基因工程技术向其引入了二型遗传机制的通常拥有I型脂肪酸生物合成途径的微生物(例如，一种产油酵母)。该微生物可以是异养生物，并且在一个具体实施例中，可以是专性异养生物。在使用微藻的情况下，该微藻可以是无绿藻属或小球藻属的一个种类。说明性的种类包括魏氏无绿藻、雍滞无绿藻、Prototheca

portoricensis、Prototheca moriformis、祖菲无绿藻、凯氏小球藻、Chlorella luteoviridis、原始小球藻、以及普通小球藻。为了能够使用蔗糖原料，例如甘蔗汁以及在此所述的其他原料，这种重组细胞可以包括包含蔗糖转化酶基因的重组核酸以表达活性蔗糖转化酶。蔗糖转化酶可以由该微生物分泌到培养基。

[0089] 培养可以是异养的；例如在使用固定碳源例如葡萄糖或蔗糖的生物反应器中进行。该培养可以是持续的直到该细胞达到按干细胞重量计的至少50％、至少60％、至少70％、或50％至90％的甘油三酯。这需要使用限制氮的培养，如下文所述。

[0090] 由这种细胞所产生的油可以从这种细胞中提取。在一个实施例中，该油包含少于500、50、或5ppm的着色分子。任选地，分析该油的脂肪酸谱，例如通过LC-MS。该油可以具有实例19，表60-63中的油的特性中的一种或多种。

[0091] 在一个具体实施例中，这种重组细胞包含可操作的核酸以减少或消除由KASII基因编码的β-酮脂酰-ACP合成酶II的表达，并且表达编码酰基-ACP硫酯酶的外源基因的产物。结果，这种细胞可以产生具有一种其特征在于具有至少40％、50％、60％、或70％C16脂肪酸(例如，棕榈酸)的脂肪酸谱的油。因此，该油可以具有朝向较短链长移动的脂肪酸分布。在脂肪酸分布方面的移动的特征可以在于减少的平均脂肪酸长度或该分布的其他统计特征。例如，为了计算平均脂肪酸长度，将构成这些甘油三酯的每一可检测的脂肪酸的百分数乘以该脂肪酸中的碳数目并且将乘积的和除以100。该编码酰基-ACP硫酯酶的外源基因可以产生一种活性酰基-ACP硫酯酶，该酰基-ACP硫酯酶在C8-C16脂肪酰基链的水解中具有比该细胞中的天然酰基-ACP硫酯酶更强的活性。该编码酰基-ACP硫酯酶的外源基因可以中断该KASII基因。以这种方式，该酰基-ACP硫酯酶的插入还可以一步消除β-酮脂酰-ACP合成酶II的表达。参见实例15和16。

[0092] 在另一个具体实施例中，这种重组细胞包含可操作的核酸以减少或消除由编码β-酮脂酰-ACP合成酶I的一种或多种基因编码的酶的表达。结果，所产生的油具有一种其特征在于与缺乏这些重组核酸的可比细胞相比更短的脂肪酸链长分布的脂肪酸谱。这可以表示为减少的平均脂肪酸链长。这种重组细胞可以包括可操作的核酸以减少或消除由至少一种FAD基因编码的脂肪酸去饱和酶的表达，并且表达编码活性硬脂酰ACP去饱和酶的硬脂酰-ACP去饱和酶外源基因的产物。任选地，这些核酸是可操作的以减少或消除由脂肪酸去饱和酶基因的多个拷贝(例如，等位基因)编码的脂肪酸去饱和酶的表达。在一个具体实施例中，该硬脂酰-ACP去饱和酶外源基因被重组进入该脂肪酸去饱和酶基因的编码区之内的一个基因座。结果，所产生的油可以具有与由缺乏这些核酸的可比细胞所产生的油相比的水平提高的油酸。油酸构成脂肪酸谱中的脂肪酸的至少50％、60％、70％、80％、或90％。参见实例10。

[0093] 在另一个具体实例中，这种重组细胞包含可操作的核酸以表达编码活性β-酮脂酰-ACP合成酶II的β-酮脂酰-ACP合成酶II外源基因的产物。结果，所产生的油的特征可以在于一种由于这些重组核酸所致的C18:1脂肪酸升高和C16脂肪酸降低的脂肪酸谱。参见实例13，其中KASII基因的过表达将C18脂肪酸的百分比从未经转化的细胞中的大约68％增加至大约84％。在相关实施例中，该增加大于70％(从75％到85％，或从70％到90％)。

[0094] 在另一个具体实施例中，这种重组细胞包含可操作的核酸以通过RNA干扰来减少或消除由编码硬脂酰ACP去饱和酶的一种或多种基因编码的酶的表达。结果，所产生的油可

以具有一种其特征在于18:0脂肪酸增加的脂肪酸谱。这些18:0脂肪酸可以是该谱中的脂肪酸的至少50％、60％、70％、80％、或90％。参见实例12。

[0095] 在另一个具体实施例中，该细胞包含可操作的重组核酸以减少或消除编码β-酮脂酰-ACP合成酶I、β-酮脂酰-ACP合成酶II、硬脂酰ACP去饱和酶、或脂肪酸去饱和酶的基因的两个拷贝(例如，两个等位基因)的表达。参见实例14，其中内源KASI等位基因在无绿藻属中被敲除。结果，观察到由于内源KASI的破坏而式总C14脂肪酸的百分比增加了大约35％至400％并且C16脂肪酸的百分比增加了大约30％至50％。

[0096] 在另一个具体实施例中，这种细胞包含可操作的重组核酸以表达编码活性ω-3脂肪酸去饱和酶和/或ω-6油酸去饱和酶的脂肪酸去饱和酶外源基因的产物。结果，可以由这种细胞产生水平提高的亚油酸、亚麻酸、或两者，并且其在该细胞脂质的脂肪酸谱中是可检测的。例如，这种细胞可以具有至少10％、20％、30％、40％、或50％的亚油酸、亚麻酸、或两者。例如，可以如在实例11中表达重组Δ15去饱和酶。结果，C18:3脂肪酸(即，亚麻酸)可以增加从大约2倍至17倍，或更多。

[0097] 在另一个实施例中，培养了重组微生物细胞。这种细胞包括重组核酸(操作这些重组核酸以表达编码活性油酸12-羟化酶的外源基因的产物)以便合成蓖麻油酸。这个基因可以存在于任何前述实施例中。参见实例7。12-羟化酶的一个优选底物是油酸。因此，在一个优选实施例中，可以通过在该细胞中包含操作以增加油酸产生的重组核酸而获得蓖麻油酸的更高产率。在没有限制的情况下，该细胞包含可操作的重组核酸以表达编码活性硬脂酰ACP去饱和酶的外源基因的产物并且减少或消除由编码脂肪酸去饱和酶的一种或多种基因编码的酶的表达；或表达编码活性β-酮脂酰-ACP合成酶I的外源基因的产物并且表达编码活性酰基-ACP硫酯酶的外源基因的产物。

[0098] 按照本发明的任何实施例，该油可以被提取并且进一步通过以下项的一种或多种进行加工：精炼、漂白、脱臭、复分解、酯交换(transesterification)、氢化、水解、氢化、脱氧、加氢裂解、异构化、羟基化、酯交换(interesterification)、酰胺化、磺化、以及硫化。该油可以被加工例如以产生食用油、脂肪酸、脂肪醇、润滑剂、肥皂、脂肪酸酯、脂肪酸乙氧基化物、脂肪胺、烷基氯、脂肪醇乙氧基化物、脂肪醇硫酸酯、脂肪酸烷醇酰胺、磺化油、或硫化油、柴油、喷气发动机汽油、或混合原料(blendstock)或添加剂、润滑剂、或涂料。

[0099] 在此提及的任何实施例作为食品或食用油可以是有用的。可以将整个生物结合到食品中。该生物可以是完整的、部分溶解的、大部分溶解的或完全溶解的。用于制备以及在食品中使用产油生物的方法传授于WO2011/150411、WO2010/12093、WO2011130578、以及WO2011/130576中。可替代地，从该生物提取的并且任选地纯化的油可以用作食用油，包括用作预制食品(如涂抹食品、酱汁、甜食以及冰冻甜食)中的食用油成分。在一个具体实施例中，该产油细胞或食用油包含50％-70％C18:0和20％-40％18:1(例如，油酸)。在另一个具体实施例中，这种产油细胞或食用油包含50％-70％C16:0和20％-40％18:1(例如，油酸)。

[0100] 为了方便读者，将本发明的详细说明分成多个部分。第I部分提供了在此所使用的术语的定义。第II部分提供了在本发明方法中有用的培养条件的说明。第III部分提供了基因工程方法和材料的说明。第IV部分提供了允许蔗糖利用的基因工程的说明。第V部分提供了改变脂质生物合成的基因工程的说明。第VI部分描述了用于制造燃料和化学品的方法。第VII部分披露了本发明的实例和实施例。本发明的详细说明之后为实例，这些实例展示了

本发明的各个方面和实施例。

[0101] I.定义

[0102] 除非另有定义,在此所使用的所有技术与科学术语具有如本发明所属领域的技术人员通常理解的含义。如在此所使用的,除非另外指明,以下术语具有归属于它们的含义。

[0103] “在微藻中有活性”指在微藻中有功能的核酸。例如,已经用来驱动抗生素抗性基因以便对转基因微藻赋予抗生素抗性的启动子在微藻中有活性。

[0104] “面积百分比”是指使用其中样品内每种脂肪酸在检测之前转化成脂肪酸甲酯(FAME)的FAME GC/FID检测方法观察到的峰面积。例如,与任何其他脂肪酸如C14:1相比,观察到没有不饱和的14碳原子脂肪酸(C14:0)的独立峰。每种种类FAME的峰面积与其在混合物中的组成百分比组成成正比并且基于样品中存在的所有峰的总和而被计算(即[特定峰下的面积/所有测量峰的总面积]X 100)。当提到本发明的油和细胞的脂质谱时,“至少4%C8-C14”表示细胞中或提取的甘油脂组合物中至少4%的总脂肪酸具有包括8、10、12或14个碳原子的链长。

[0105] “无菌的”是没有被其他活生物污染的生物培养物。

[0106] “生物柴油”是适合用作柴油发动机中的燃料的以生物方式产生的脂肪酸烷基酯。

[0107] “生物质”是通过细胞的生长和/或增殖产生的物质。生物质可以含有细胞和/或胞内内容物以及细胞外物质,包括但不限于由细胞分泌的化合物。

[0108] “生物反应器”是其中培养(任选地以悬浮方式培养)细胞的密闭罩或部分密闭罩。

[0109] “纤维素材料”是包含纤维素以及任选地半纤维素的生物材料。像这样它可以被消化为糖,如葡萄糖和木糖,并且任选地可以包含另外的化合物如二糖、寡糖、木质素、糠醛和其他化合物。纤维素材料的来源的非限制性实例包括甘蔗渣、甜菜浆、玉米秸秆、木屑、锯屑和柳枝稷。

[0110] “共培养”及其变体如“共培养”和“共发酵”是指在同一个生物反应器中培养两种或更多种类型的细胞。两种或更多种类型的细胞可以两者都是微生物,如微藻,或可以是与一种不同的细胞类型一起培养的微藻细胞。

[0111] 在此使用的“着色分子”或“产生颜色的杂质”是指将颜色赋予给所提取的油的任何化合物。“着色分子”或“产生颜色的杂质”包括例如叶绿素a、叶绿素b、番茄红素、生育酚、菜油甾醇、生育三烯酚以及类胡萝卜素,如β胡萝卜素、叶黄素、玉米黄质、虾青素。这些分子优选地以不超过500ppm、不超过250ppm、不超过100ppm、不超过75ppm、或不超过25ppm的浓度存在于该微生物生物质或所提取的油中。在其他实施例中,存在于该微生物生物质或所提取的油中的叶绿素的量少于500mg/kg、少于100mg/kg、少于10mg/kg、少于1mg.kg、少于0.5mg/kg、少于0.1mg/kg、少于0.05mg/kg、或少于0.01mg/kg。

[0112] “培养”(cultivated)及其变体如“培养”(cultured)和“发酵”是指通过使用选择和/或受控的条件有意地促进一个或多个细胞生长(在细胞大小、细胞内容物、和/或细胞活性方面增加)和/或增殖(增加细胞数)。生长和增殖的组合可以称作“繁殖”。选择和/或受控的条件的实例包括使用确定成分培养基(具有已知特征如pH、离子强度、和碳源)、指定的温度、氧张力、二氧化碳水平、以及在生物反应器中的生长。“培养”不是指微生物在自然界中或以别的方式在没有人类干预下的生长或增殖;例如,最终变成化石以产生地质原油的生物的自然生长不是培养。

[0113] “去饱和酶”是指在脂质合成途径中负责将双键(不饱和)引入三酰甘油酯分子的脂肪酸链的酶。实例包括但不限于硬脂酰-酰基载体蛋白去饱和酶(SAD)和脂肪酸去饱和酶(FAD),也称为脂肪酰基去饱和酶。

[0114] “表达载体”或“表在构建体”或“质粒”或“重组DNA构建体”是用于将一种核酸引入宿主细胞的载体。该核酸可以是已经通过人类干预(包括通过重组手段或直接化学合成)产生的一种核酸,其带有允许特定核酸的转录和/或翻译的一系列指定核酸元件。该表达载体可以是质粒的一部分、病毒、或核酸片段、或其他适合的载体。典型地,表达载体包含与启动子可操作地连接的待转录核酸。

[0115] “外源基因”是指令已经导入到细胞中(例如通过转化/转染)的RNA和/或蛋白质的表达的核酸,并且还被称为“转基因”。包含外源基因的细胞可以被称为重组细胞,可以向其中导入另外的外源基因。相对于正在转化的细胞,外源基因可以来自不同的种类(因此是异源的),或来自相同的种类(因此是同源的)。因此,外源基因可以包括相对于该基因的内源拷贝占据细胞基因组中不同位置或在不同控制之下的同源基因。外源基因可以在细胞中以一个拷贝以上存在。外源基因可以在细胞中作为基因组(核或质粒)中的插入片段或作为游离型分子而维持。

[0116] “外源提供的”是指提供到细胞培养物的培养基的分子。

[0117] 取决于上下文,“脂肪酸”应当表示游离脂肪酸、脂肪酸盐、或甘油脂中的脂肪酰基部分。

[0118] “固定碳源”是在环境温度和压力以固体或液体形式在培养基中存在的含碳分子,典型地是有机分子,它可以由其中所培养的微生物利用。因此,二氧化碳不是固定碳源。

[0119] 如涉及培养条件时,“异养的”是在基本上不存在光下,同时利用或代谢固定碳源的培养。

[0120] “匀浆”是已经被物理破坏的生物质。

[0121] “氢:碳比”是分子中基于原子对原子的氢原子对碳原子的比率。这个比率可以用来指代烃分子中碳原子和氢原子的数目。例如,具有最高比率的烃是甲烷CH_4(4:1)。

[0122] “诱导型启动子”是响应于特定刺激介导可操作地连接的基因的转录的启动子。这样的启动子的实例是在改变pH或氮水平的条件下被诱导的启动子序列。

[0123] “处于可操作连接”是在两个核酸序列(如控制序列(典型地是启动子)和连接的序列(典型地是编码蛋白质的序列,也称作编码序列))之间的功能性连接。如果一种启动子可以介导外源基因的转录,则该启动子与该基因处于可操作连接。

[0124] “脂质修饰酶”指改变脂质的共价结构或以别的方式可以导致细胞中的改变的脂肪酸谱的酶。脂质修饰酶的实例包括脂肪酶、脂肪酰基-ACP硫酯酶、脂肪酰基辅酶A/醛还原酶、脂肪酰基辅酶A还原酶、脂肪醛还原酶、去饱和酶,包括硬脂酰-酰基载体蛋白去饱和酶(SAD)和脂肪酰基去饱和酶(FAD)、以及脂肪醛脱羧酶。

[0125] “脂质途径酶”是在脂质代谢(即,脂质合成、修饰或降解)中发挥作用的任何酶和以化学方式修饰脂质的任何蛋白质以及载体蛋白。

[0126] “脂质谱”或“甘油脂谱”是指就链长和/或饱和模式而言,细胞或衍生自细胞的油中的脂肪酸的分布。在这个背景下,饱和模式可以包括饱和酸相对于不饱和酸的量度或者细胞的多种脂肪酸中的双键的分布位置的更为详细的分析。

[0127] “溶解”(lysis)是破坏生物生物的质膜和任选地细胞壁足以释放至少一些细胞内内容物,这通常是通过损害细胞完整性的机械、化学、病毒或渗透机制。

[0128] “溶解(Lysing)”是进行溶解(lysis)的过程。

[0129] “微藻”是微生物生物,其含有叶绿体或质体并且任选地能够进行光合作用,或能够进行光合作用的原核微生物生物。微藻包括不能代谢固定碳源作为能量的专性光能自养生物,以及能够仅仅以固定碳源为生的异养生物。微藻包括在细胞分裂之后不久与姐妹细胞分开的单细胞生物,如衣藻(Chlamydomonas),以及微生物,例如像,团藻(Volvox),它是两种不同细胞类型的简单多细胞光合生物。微藻包括细胞,如小球藻、杜氏藻(Dunaliella)和无绿藻属。微藻还包括展现出细胞-细胞粘附的其他微生物光合生物,如阿格门氏藻(Agmenellum)、鱼腥藻(Anabaena)和桑葚藻(Pyrobotry)。“微藻”还包括专性异养微生物,它们已经丧失进行光合作用的能力,如某些甲藻藻类和无绿藻种类。

[0130] 如在此在脂肪酸的背景下所使用的“中链”是指C10-C16脂肪酸。如在这个背景下所使用的“短链”是指C6-C10脂肪酸,而“长链”是指C17或更长的脂肪酸。这些界限并非旨在精确地限定,除非另外说明。

[0131] “天然油”应当表示主要从生物获得的甘油三酯油,其中该油未经历与另外的天然油或合成油掺合或分提以实质上改变该甘油三酯的脂肪酸谱。在此,术语“分提”表示以相对于由生物所产生的谱而改变其脂肪酸谱的方式来从该油去除物质,不管用什么方法完成。天然油包括如从生物获得的油,其中该油已经经历最低限度的不实质性改变其甘油三酯谱的加工,包括精炼、漂白和/或脱胶。天然油还可以是“非酯交换的天然油”,表示该天然油未经历一种其中脂肪酸在其酰基连接中被重分布到甘油的加工并且基本上保持如从生物中回收时的相同构型。

[0132] 关于两种蛋白质或基因的“天然共表达”表示这些蛋白质或它们的基因在它们从其衍生的组织或生物中天然地共表达,例如,由于编码这两种蛋白质的基因处于共同调节序列的控制之下或它们响应于相同的刺激而表达。

[0133] “渗透压休克”是渗透压急剧下降后细胞在溶液中破裂。有时诱导渗透压休克以便使这样的细胞的细胞组分释放到溶液中。

[0134] “多糖降解酶”是能够催化任何多糖水解或糖化的任何酶。例如,纤维素酶催化纤维素的水解。

[0135] “多糖”或“聚糖”是由糖苷键连接在一起的单糖构成的糖。纤维素是构成某些植物细胞壁的多糖。纤维素可以通过酶解聚以产生单糖如木糖和葡萄糖,以及较大的二糖和寡糖。

[0136] “启动子”是指导核酸转录的核酸控制序列。如在此所使用的,启动子包括转录起始位点附近必需的核酸序列,如,在聚合酶II型启动子的情况下的TATA元件。启动子还任选地包括远端增强子或阻抑元件,它们可以距转录起始位点多达几千碱基对而定位。

[0137] “重组体”是因导入外源核酸或改变天然核酸而已经被修饰的细胞、核酸、蛋白质或载体。因而,例如,重组细胞可以表达没有发现于天然(非重组)形式的细胞内部的基因,或以与非重组细胞表达这些基因不同的方式表达天然基因。重组细胞可以但不限于包括编码基因产物或抑制元件(如降低细胞中的活性基因产物的水平的突变、敲除、反义、干扰RNA(RNAi)或dsRNA)的重组核酸。“重组核酸”是最初在体外形成(通常,通过操作核酸,例如,使

用聚合酶、连接酶、核酸外切酶和核酸内切酶)的核酸或另外地处于通常未发现于自然界中的形式。可以产生重组核酸,例如,以便将两个或更多个核酸安排为可操作连接。因此,出于本发明的目的,分离核酸或通过连接通常在自然界中不接合的DNA分子在体外形成的表达载体均被认为是重组体。出于本发明的目的,一旦产生重组核酸并将其导入宿主细胞或生物中,这种核酸可以利用宿主细胞的体内细胞机器(cellular machinery)进行复制;然而一旦以重组的方式产生,虽然这样的核酸随后在细胞内复制,其仍将被认为是重组体。同样,"重组蛋白"是使用重组技术,即,通过表达重组核酸产生的蛋白质。

[0138] "可再生柴油"是从天然油产生的链烷(如C10:0、C12:0、C14:0、C16:0和C18:0)的混合物;例如通过脂质的氢化和脱氧。

[0139] "糖化"是将生物质(通常为纤维素生物质或木质纤维素生物质)转化成单体糖(如葡萄糖和木糖)的过程。"糖化"或"解聚"的纤维素材料或生物质是指已经通过糖化作用转化成单体糖的纤维素材料或生物质。

[0140] "糠醛种类"是2-呋喃甲醛或保留相同基本结构特征的衍生物。

[0141] 根据在此的实施例(并且不必讨论现有技术),关于用来产生一个的重组株系的株系转化,"稳定的"或"稳定地整合的"应当表示这些重组核酸被该株系的细胞保留至少10代。例如,在重组株系具有一个使培养能够在一种选择压力存在下进行的选择标记的情况下,这些重组核酸在该选择压力不存在下在培养10代之后被保留。

[0142] "蔗糖利用基因"是在表达时辅助细胞利用蔗糖作为能源的能力的基因。由蔗糖利用基因编码的蛋白质在此称为"蔗糖利用酶"并且包括蔗糖转运蛋白、蔗糖转化酶、和己糖激酶如葡萄糖激酶和果糖激酶。

[0143] II.培养

[0144] 本发明总体上涉及微生物并且尤其是具有II型脂肪酸生物合成途径的产油微生物(如产生甘油三酯的微藻)的培养。在一个实施例中,这些微生物是专性异养生物。这些微生物可以是基于例如在下文披露的基因工程方法的重组微生物。为了方便读者,将这个部分细分成多个小节。小节1描述了微生物的种类和株系。小节2描述了用于培养的生物反应器。小节3描述了用于培养的培养基。小节4描述了根据本发明的说明性培养方法的油的生产。

[0145] 1.微生物种类与株系

[0146] 虽然以下呈现的说明性实施例适用于众多微生物,无绿藻属是用于在脂质生产中使用的优选微生物。重要的是,在此描述的以无绿藻属作为实例的基因工程方法适用于其他微生物(例如,Chlorella sorokiniana、普通小球藻、椭圆小球藻(Chlorella ellipsoidea)、凯氏小球藻、杜氏盐藻(Dunaliella tertiolecta)、强壮团藻(Volvox carteri)、雨生红球藻(Haematococcus pluvialis)、极锐新月藻-瘦狭新月藻-沿岸新月藻复合体(Closterium peracerosum-strigosum-littorale complex)、绿色杜氏藻(Dunaliella viridis)、盐生杜氏藻(Dunaliella salina)、胸状盘藻(Gonium pectorale)、三角褐指藻(Phaeodactylum tricornutum)、角毛藻属、梭形筒柱藻(Cylindrotheca fusiformis)、前沟藻种类(Amphidinium sp.)、Symbiodinium microadriacticum、微拟球藻属(Nannochloropsis)、隐秘小环藻(Cyclotella cryptica)、腐生舟形藻(Navicula saprophila)、或假微型海链藻(Thalassiosira pseudonana))。

[0147] 从专性异养微藻(如无绿藻属)获得的脂质或油通常可以是低色素的(例如,低至叶绿素以及某些类胡萝卜素的不可检测水平,例如少于500ppm、50ppm或5ppm的着色分子、产生颜色的杂质、或者叶绿素与类胡萝卜素浓度的总计)并且在任何情况下含有比来自其他微藻的脂质少得多的色素。而且,相对于从其他微生物产生脂质,由本发明提供的重组无绿藻细胞可以用来以更高的产量和效率并且以降低的成本产生脂质。用于在本发明的方法中使用的说明性无绿藻株包括另外,这种微藻异养生长并且可以作为魏氏无绿藻(Prototheca wickerhamii)、雍滞无绿藻(Prototheca stagnora)(包括UTEX 327)、Prototheca portoricensis(Prototheca portoricensis)、Prototheca moriformis(Prototheca moriformis)(包括UTEX株1441、1435)、以及祖菲无绿藻(Prototheca zopfii)而被基因工程化。无绿藻种类是专性异养生物。

[0148] 除了产生适合的脂质或烃用于生产油、燃料和油脂化学品之外,影响选择用于在本发明的实施例中使用的微生物的考虑事项包括:(1)作为细胞重量百分比的高脂质含量;(2)易于生长;(3)易于进行基因工程;和(4)易于进行生物质加工。在特定的实施例中,野生型或基因工程化的微生物产生了含有至少40%、至少45%、至少50%、至少55%、至少60%、至少65%或至少70%或更多脂质的细胞。在其他特定的实施例中,野生型或基因工程化的微生物产生了含有在40%与80%之间或50%与90%之间的甘油三酯的细胞。优选的生物以异养方式(在不存在光时依赖糖)生长。

[0149] 可以用来实施本发明的藻类的实例包括,但不限于在表1中列出的以下藻类。

[0150] 表1.藻类的实例。

[0151]

东方曲壳藻(Achnanthes orientalis)、阿格门氏藻属(Agmenellum)、Amphiprora hyaline、咖啡形双眉藻(Amphora coffeiformis)、Amphora coffeiformis linea、Amphora coffeiformis punctata、Amphora coffeiformis taylori、Amphora coffeiformis tenuis、优美双眉藻(Amphora delicatissima)、Amphora delicatissima capitata、双眉藻种类(Amphora sp.)、鱼腥藻属(Anabaena)、纤维藻属(Ankistrodesmus)、镰形纤维藻(Ankistrodesmus falcatus)、黄金色藻(Boekelovia hooglandii)、Borodinella sp.、布朗葡萄藻(Botryococcus braunii)、Botryococcus sudeticus、四鞭藻属(Carteria)、纤细角毛藻(Chaetoceros gracilis)、牟氏角毛藻(Chaetoceros muelleri)、Chaetoceros

[0152]

muelleri subsalsum、*角毛藻种类*（*Chaetoceros sp.*）、*Chlorella anitrata*、*Chlorella Antarctica*、*Chlorella aureoviridis*、*Chlorella candida*、*Chlorella capsulate*、*Chlorella desiccate*、*椭圆小球藻*（*Chlorella ellipsoidea*）、*浮水小球藻*（*Chlorella emersonii*）、*Chlorella fusca*、*Chlorella fusca var. vacuolata*、*Chlorella glucotropha*、*Chlorella infusionum*、*Chlorella infusionum var. actophila*、*Chlorella infusionum var. auxenophila*、*凯氏小球藻*（*Chlorella kessleri*）、*Chlorella lobophora*（SAG 37.88 株）、*Chlorella luteoviridis*、*Chlorella luteoviridis var. aureoviridis*、*Chlorella luteoviridis var. lutescens*、*Chlorella miniata*、*极微小球藻*（*Chlorella minutissima*）、*Chlorella mutabilis*、*Chlorella nocturna*、*Chlorella parva*、*Chlorella photophila*、*Chlorella pringsheimii*、*原始小球藻*（*Chlorella protothecoides*）（包括以下任何株：UTEX 株 1806、411、264、256、255、250、249、31、29、25 和 CCAP 株 211/17 和 211/8d）、*Chlorella protothecoides var. acidicola*、*普通小球藻*（*Chlorella regularis*）、*Chlorella regularis var. minima*、*Chlorella regularis var. umbricata*、*Chlorella reisiglii*、*Chlorella saccharophila*、*Chlorella saccharophila var. ellipsoidea*、*海水小球藻*（*Chlorella salina*）、*Chlorella simplex*、*Chlorella sorokiniana*、*小球藻种类*（*Chlorella sp.*）、*Chlorella sphaerica*、*Chlorella stigmatophora*、*Chlorella vanniellii*、*普通小球藻*（*Chlorella vulgaris*）、*普通小球藻*（*Chlorella vulgaris*）、*Chlorella vulgaris f. tertia*、*Chlorella vulgaris var. autotrophica*、*Chlorella vulgaris var. viridis*、*Chlorella vulgaris var. vulgaris*、*Chlorella vulgaris var. vulgaris f. tertia*、*Chlorella vulgaris var. vulgaris f. viridis*、*Chlorella xanthella*、*Chlorella zofingiensis*、*Chlorella trebouxioides*、*普通小球藻*（*Chlorella vulgaris*）、*水溪绿球藻*（*Chlorococcum infusionum*）、*绿球藻种类*（*Chlorococcum sp.*）、*绿梭藻*（*Chlorogonium*）、*蓝隐藻种类*（*Chroomonas sp.*）、*金球藻种类*（*Chrysosphaera sp.*）、*球钙板藻种类*（*Cricosphaera sp.*）、*隐藻种类*（*Cryptomonas sp.*）、*隐秘小环藻*（*Cyclotella cryptica*）、*梅尼小环藻*（*Cyclotella meneghiniana*）、*小环藻种类*（*Cyclotella sp.*）、*杜氏藻种类*（*Dunaliella sp.*）、*拜尔代维勒杜氏藻*（*Dunaliella bardawil*）、*Dunaliella bioculata*、*Dunaliella granulate*、*Dunaliella maritime*、*Dunaliella minuta*、*巴夫杜氏藻*（*Dunaliella parva*）、*Dunaliella peircei*、*海产性绿藻*（*Dunaliella primolecta*）、*盐生杜氏藻*（*Dunaliella salina*）、*Dunaliella terricola*、*杜氏盐藻*（*Dunaliella tertiolecta*）、*绿色杜氏藻*（*Dunaliella viridis*）、*杜氏盐藻*（*Dunaliella tertiolecta*）、*绿色独球藻*（*Eremosphaera viridis*）、*独球藻种类*（*Eremosphaera sp.*）、*Ellipsoidon sp.*、*眼虫藻*（*Euglena*）、*被刺藻种类*（*Franceia sp.*）、*克罗脆杆藻*（*Fragilaria crotonensis*）、*脆杆藻*（*Fragilaria sp.*）、*粘球藻*（*Gleocapsa sp.*）、*Gloeothamnion sp.*、*膜胞藻种类*（*Hymenomonas sp.*）、*等鞭金藻*（*Isochrysis aff. galbana*）、*球等鞭金藻*（*Isochrysis galbana*）、*鳞孔藻*（*Lepocinclis*）、*微芒藻*（*Micractinium*）、*微*

[0153]

芒藻（Micractinium）（UTEX LB 2614）、单壳缝藻（Monoraphidium minutum）、单针藻种类（Monoraphidium sp.）、微绿球藻种类（Nannochloris sp.）、盐生微绿球藻（Nannochloropsis salina）、微绿球藻种类（Nannochloropsis sp.）、适意舟形藻（Navicula acceptata）、Navicula biskanterae、Navicula pseudotenelloides、小皮舟形藻（Navicula pelliculosa）、腐生舟形藻（Navicula saprophila）、舟形藻种类（Navicula sp.）、肾鞭藻种类（Nephrochloris sp.）、肾藻种类（Nephroselmis sp.）、Nitschia communis、亚历山大菱形藻（Nitzschia alexandrina）、普通菱形藻（Nitzschia communis）、细端菱形藻（Nitzschia dissipata）、碎片菱形藻（Nitzschia frustulum）、汉氏菱形藻（Nitzschia hantzschiana）、Nitzschia inconspicua、Nitzschia intermedia、小头菱形藻（Nitzschia microcephala）、微小菱形藻（Nitzschia pusilla）、Nitzschia pusilla elliptica、Nitzschia pusilla monoensis、Nitzschia quadrangular、菱形藻种类（Nitzschia sp.）、棕鞭藻种类（Ochromonas sp.）、小卵囊藻（Oocystis parva）、极小卵囊藻（Oocystis pusilla）、卵囊藻种类（OOocystis sp.）、沼泽颤藻（Oscillatoria limnetica）、颤藻种类（Oscillatoria sp.）、拟短形颤藻（Oscillatoria subbrevis）、Pascheria acidophila、巴夫藻种类（Pavlova sp.）、扁裸藻（Phagus）、席藻（Phormidium）、扁藻种类（Platymonas sp.）、颗石藻（Pleurochrysis carterae）、Pleurochrysis dentate、颗石藻种类（Pleurochrysis sp.）、魏氏无绿藻（Prototheca wickerhamii）、雍滞无绿藻（Prototheca stagnora）、Prototheca portoricensis、Prototheca moriformis、祖菲无绿藻（Prototheca zopfii）、塔胞藻种类（Pyramimonas sp.）、桑葚藻（Pyrobotrys）、Sarcinoid chrysophyte、被甲栅藻（Scenedesmus armatus）、水绵（Spirogyra）、钝顶螺旋藻（Spirulina platensis）、裂丝藻种类（Stichococcus sp.）、聚球藻种类（Synechococcus sp.）、四角藻（Tetraedron）、融合微藻种类（Tetraselmis sp.）、四鞭片藻（Tetraselmis suecica）、威氏海链藻（Thalassiosira weissflogii）和 Viridiella fridericiana

[0154] 2.生物反应器

[0155] 出于实施基因操作和生产烃(例如,脂质、脂肪酸、醛、醇和链烷)两者的目的而培养微生物。前一类型的培养按小规模实施并且至少起初在起始微生物可以生长的条件下进行。出于烃生产目的的培养通常按大规模(例如,10,000L、40,000L、100,000L或更大的生物反应器)在生物反应器中实施。典型地按本发明的方法在生物反应器内的液体培养基中培养微藻(包括无绿藻种类)。典型地,该生物反应器不允许光进入。

[0156] 生物反应器或发酵罐用来培养微藻细胞经过其生理周期的多个阶段。生物反应器为用于异养生长和增殖方法提供许多优点。为了产生生物质,微藻优选在液体(如在悬浮培养物中)中大量发酵。生物反应器(如钢发酵罐)可以容纳十分大的培养体积(本发明的各种实施例中使用40,000升和更大容量的生物反应器)。生物反应器典型地也允许控制培养条件如温度、pH、氧张力、和二氧化碳水平。例如,生物反应器典型地是可配置的,例如,使用与管路连接的端口以允许气体组分(如氧或氮)鼓泡穿过液体培养物。也可以使用生物反应器

更轻易地操纵其他培养参数，如培养基的pH、微量元素的一致性和浓度、以及其他培养基组分。

[0157] 配备多种装置(如旋转叶片和叶轮、摆动机构、搅拌棒、加压气体输注装置)的生物反应器可以用来混合培养物。混合可以是连续或间歇的。例如，在一些实施例中，为繁殖细胞，不维持气体进入和培养基进入的湍流态直到已经实现所希望的所述细胞数目增加。

[0158] 生物反应器端口可以用来将气体、固体、半固体、和液体导入或提取到含有微藻的生物反应器腔室中。虽然许多生物反应器具有一个以上的端口(例如，一个用于培养基进入，并且另一个用于采样)，仅一种物质进入或离开端口是不必要的。例如，一个端口可以用来使培养基流入生物反应器并且稍后用于采样、气体进入、气体离开或其他目的。优选地，采样端口可以反复使用，而不改变或损害培养物的无菌性质。采样端口可以配置有允许样品流停止和开始或提供连续采样手段的阀门或其他装置。生物反应器典型地具有允许接种培养物的至少一个端口，并且这种端口也可以用于其他目的，如培养基或气体进入。

[0159] 生物反应器端口允许操纵微藻培养物的气体含量。为了举例说明，生物反应器的部分体积可以是气体，而不是液体，并且生物反应器的气体入口允许将气体泵送入生物反应器。可以被有益地泵送入生物反应器的气体包括空气、氧、空气/CO_2混合物、稀有气体(如氩气)、和其他气体。生物反应器可以被配备为使得用户能够控制气体进入生物反应器的速率。如上文指出，增加进入生物反应器的气体流量可以用来增进培养物的混合。

[0160] 增加的气体流量也影响培养物的浊度。可以通过安置气体进口低于含水培养基液位实现紊流，从而进入生物反应器的气体鼓泡至培养物表面。一个或多个气体出口允许气体逃逸，从而防止生物反应器中的压力累积。优选地，气体出口通向防止污染微生物进入生物反应器的“单向”阀门。

[0161] 3.培养基

[0162] 微藻培养基典型地含有多种组分，如固定氮源、固定碳源、微量元素、任选地用于pH维持的缓冲剂、和磷酸盐(典型地作为磷酸盐提供)。其他组分可以包括盐，如氯化钠，尤其对于海水微藻而言。氮源包括有机和无机氮源，例如包括而不限于分子氮、硝酸盐、硝酸盐、氨(纯氨或盐形式，如$(NH_4)_2SO_4$和NH_4OH)、蛋白质、大豆粕、玉米浆、以及酵母提取物。微量元素的实例包括锌、硼、钴、铜、锰、和钼，例如处于对应的形式：$ZnCl_2$、H_3BO_3、$CoCl_2 \cdot 6H_2O$、$CuCl_2 \cdot 2H_2O$、$MnCl_2 \cdot 4H_2O$和$(NH_4)_6Mo_7O_{24} \cdot 4H_2O$。

[0163] 根据本发明的方法使用的微生物发现于世界各处的各个地点和环境。由于它们与其他种类的隔离和因此产生的进化趋异，可能难以预测用于最佳生长和产生脂质和/或烃组分的特定生长培养基。在一些情况下，微生物的某些菌株可能在特定生长培养上不能生长，原因在于存在某种抑制性组分或不存在特定微生物菌株需要的某种必需营养需求。

[0164] 固体和液体生长培养基通常可从广泛来源获得，并且可以找到制备适合于多种微生物株系的特定培养基的说明书，例如在www.utex.org/在线获得，该站点由奥斯汀的德克萨斯大学(1University Station A6700，奥斯汀，德克萨斯，78712-0183)，为其藻类培养物保藏中心(UTEX)维护。例如，多种淡水和咸水培养基包括在通过引用结合在此的PCT公开号2008/151149中描述的那些。

[0165] 在一个具体的实例中，月示培养基(Proteose Medium)适于无菌培养，并且可以通过添加1g月示蛋白胨至1升Bristol培养基来制备1L体积的培养基(pH约6.8)。Bristol培养

基包含在水溶液中的2.94mM $NaNO_3$、0.17mM $CaCl_2 \cdot 2H_2O$、0.3mM $MgSO_4 \cdot 7H_2O$、0.43mM、1.29mM KH_2PO_4、和1.43mM NaCl。对于1.5％琼脂培养基，可以将15g琼脂添加至1L的这种溶液。将溶液盖好并高压灭菌，并且然后在使用之前贮存在冷藏温度。另一个实例是无绿藻分离培养基(PIM)，其包含10g/L邻苯二甲酸氢钾(KHP)、0.9g/L氢氧化钠、0.1g/L硫酸镁、0.2g/L磷酸二氢钾、0.3g/L氯化铵、10g/L葡萄糖、0.001g/L盐酸硫胺素、20g/L琼脂、0.25g/L 5-氟胞嘧啶，pH在5.0至5.2的范围内(参见Pore,1973,App.Microbiology,26:648-649)。可以通过咨询以上鉴定的URL或通过咨询维持微生物培养物的其他机构(如SAG、CCAP、或CCALA)而容易地确定供本发明的方法使用的其他适合的培养基。SAG是指在哥廷根大学的藻类培养物保藏中心(哥廷根，德国)，CCAP是指苏格兰海洋科学协会管理的藻类和原虫培养物保藏中心(苏格兰，英国)，并且CCALA是指在植物研究所(Institute of Botany)的藻类实验室培养物保藏中心(Třeboň，捷克共和国)。另外，美国专利号5,900,370描述了适合于无绿藻种类异养发酵的培养基制品和条件。

[0166] 为了生产油，固定碳源的选择是重要的，因为固定碳源的成本必须充分低从而使油生产是经济的。因此，虽然适合的碳源可以包括例如乙酸盐、弗罗里多苷、果糖、半乳糖、葡糖醛酸、葡萄糖、甘油、乳糖、甘露糖、N-乙酰葡糖胺、鼠李糖、棉子糖、水苏糖、蔗糖和/或木糖，包含这些化合物的原料的选择是本发明实施例的方法的一个重要方面。根据本发明的方法有用的适合的原料包括例如黑液、玉米淀粉、解聚的纤维素材料、乳清、糖蜜、马铃薯、高粱、蔗糖、甜菜、甘蔗、水稻、和小麦。碳源也可以作为混合物如蔗糖和解聚甜菜浆的混合物而被提供。一种或多种碳源可以按一种或多种外源提供的固定碳源的至少大约50μM、至少大约100μM、至少大约500μM、至少大约5mM、至少大约50mM、以及至少大约500mM的浓度而供应。高度浓缩的碳源作为用于发酵的原料是优选的。例如，在一些实施例中，在培养步骤之前向补料分批培养中添加原料的至少300g/L、至少400g/L、至少500g/L、或至少600g/L或更高的葡萄糖水平，其中当细胞生长并且积累脂质时，高度浓缩的固定碳源随时间推移而被供应给细胞。在其他实施例中，在培养之前向补料分批培养中添加至少500g/L、至少600g/L、至少700g/L、至少800g/L蔗糖水平的蔗糖或更多，其中当细胞生长并且积累脂质时，高度浓缩的固定碳源随时间推移而被供应给细胞。高度浓缩的固定碳源(如蔗糖)的非限制性实例包括稠甘蔗汁、甘蔗汁、甜菜汁和糖蜜。为了本发明的目的，特别感兴趣的碳源包括纤维素(处于解聚形式)、甘油、蔗糖、和高粱，下文更详细地讨论每种碳源。

[0167] 根据本发明，可以使用解聚的纤维素生物质作为原料培养微生物。纤维素生物质(例如，秸秆，如玉米秸秆)是便宜并且可以容易地获得的；然而，已经发现此类原料抑制酵母生长，并且酵母不能利用从纤维素材料中产生的5-碳糖(例如，来自半纤维素的木糖)。相反，至少一些微藻可以在加工的纤维素材料上生长。纤维素材料通常包括大约40％-60％的纤维素、大约20％-40％的半纤维素、以及10％-30％的木质素。

[0168] 纤维素材料包括来自草本和木质能源作物以及农业作物的残余物，即，植物部分，主要是茎和叶，它们不随主要食物或纤维产物一起从田间去除。实例包括农业废弃物如甘蔗渣、稻壳、玉米纤维(包括茎、叶、玉米皮、和玉米芯)、麦秸、稻草、甜菜浆、柑橘浆、柑橘皮；林业废弃物如硬木和软木间伐材(thinnings)、和来自木材操作的硬木和软木残余物；木材废弃物如锯木厂废弃物(木屑、锯屑)和纸浆厂废弃物；城市废弃物如市政固体废弃物的纸部分、城市木质废弃物和城市绿色废弃物如市政碎草(grass clipping)；和木结构废弃物。

另外的纤维素材料包括专用的纤维素作物如柳枝稷、杂交杨木、和芒草、纤维甘蔗、和纤维高粱。从这些材料产生的五碳糖包括木糖。

[0169] 可以处理纤维素材料以增加微生物能够藉此利用包含在这些材料内的糖的效率。本发明的实施例提供了用于在酸爆炸之后处理纤维素材料的方法，从而这些材料适合用于微生物(例如，微藻和产油酵母)的异养培养。如上文讨论，木质纤维素生物质由多种部分组成，这些部分包括纤维素(β-1,4连接的葡萄糖(六碳糖)的结晶聚合物)、半纤维素(更松散结合的聚合物，主要地由木糖(五碳糖)并且较少程度地由甘露糖、半乳糖、阿拉伯糖组成)、木质素(芥子醇及其衍生物组成的复杂芳香族聚合物)、以及果胶(是α-1,4连接的多聚半乳糖醛酸的直链)。因为纤维素和半纤维素的聚合结构，在它们中的糖(例如，单体葡萄糖和木糖)不是呈可以被许多微生物有效利用(代谢)的形式。对于这些微生物，进一步加工纤维素生物质以产生构成聚合物的单体糖对于确保有效利用纤维素材料作为原料(碳源)可能是十分有益的。

[0170] 细胞或纤维素生物质经受一个称作“爆炸”的过程，其中所述生物质用稀硫酸(或其他酸)在升高的温度和压力下处理。这个过程将生物质调理，使得生物质可以有效地经受纤维素部分和半纤维素部分的酶促水解，变成葡萄糖和木糖单体。所生成的单体糖被称作纤维素糖。纤维素糖可以随后被微生物利用以产生多种代谢物(例如，脂质)。酸爆炸步骤导致半纤维素部分部分水解为组成单糖。这些糖可以随进一步处理从生物质完全释放。在一些实施例中，进一步处理是包括用热水洗涤爆裂材料的水热处理(去除污染物，例如盐)。这个步骤对纤维素乙醇发酵不是必要的，原因在于在这类工艺中使用更加稀释的糖浓度。在其他实施例中，进一步处理是另外的酸处理。在再另外的实施例中，进一步处理是爆裂材料的酶促水解。这些处理也可以按任何组合方式使用。处理的类型可以影响释放的糖的类型(例如，五碳糖与六碳糖)和它们在这种过程中释放的阶段。因此，可以产生不同的糖流，无论它们是主要为五碳还是六碳。因此可以将这些富集的五碳流或六碳流导向到具有不同碳利用能力的特定微生物。

[0171] 本发明的方法可以涉及发酵至比典型地在乙醇发酵中实现的更高的细胞密度。由于用于异养纤维素油生产的培养物的密度更高，固定碳源(例如，源自纤维素材料的糖流)优选地处于浓缩形式。解聚纤维素材料的葡萄糖水平在培养步骤之前优选至少300g/升、至少400g/升、至少500g/升或至少600g/升，所述培养步骤任选地是在细胞生长并且积累脂质时将这种材料随时间推移而被供应给细胞的补料分批培养。因此，为了在生产木质纤维素油期间产生和维持很高的细胞密度，可以将碳原料以高度浓缩的形式输送至异养培养物中。然而，在进料流中不是产油微生物的底物并且不被其代谢的任何组分将在生物反应器中积累，如果这种组分有毒或抑制所希望的终产物的产生，这可能导致问题。虽然木质素和木质素衍生的副产物、碳水化合物衍生的副产物如糠醛与羟甲基糠醛、以及源自纤维素材料的产生的盐(在爆炸过程和后续中和过程中均存在)、甚至未代谢的戊糖/己糖可能在乙醇发酵中带来问题，在这些物质的浓度在初始原料中高的过程中，这些影响被显著放大。为了实现针对六碳糖(它们可以在本发明所述的木质纤维素油的大规模生产中使用)的300g/L范围(或更高)的糖浓度，这些有毒材料的浓度可以比典型地存在于纤维素生物质的乙醇发酵中的浓度高20倍。

[0172] 纤维素材料的爆炸加工处理利用显著量的硫酸、热和压力，从而释放碳水化合物

副产物,即糠醛和羟甲基糠醛。在半纤维素水解期间通过木糖脱水成糠醛和水,产生糠醛和羟甲基糠醛。在本发明的一些实施例中,将这些副产物(例如,糠醛和羟甲基糠醛)在导入生物反应器之前从糖化的木质纤维素材料中去除。在本发明的某些实施例中,用于去除碳水化合物的副产物的方法是爆裂纤维素材料的水热处理。另外,本发明提供了其中能够耐受化合物如糠醛或羟甲基糠醛的株系被用于木质纤维素油生产的方法。在另一个实施例中,本发明还提供了多种方法和微生物,这些微生物不仅能够耐受发酵培养基中的糠醛,而且实际上能够在木质纤维素油生产期间代谢这些副产物。

[0173] 爆炸过程还产生显著水平的盐。例如,当爆裂的纤维素生物质以10:1的水:固形物(干重)比率再悬浮时,用于爆炸的典型条件可以导致超过5mS/cm的电导率。在本发明的某些实施例中,稀释的爆裂的生物质经受酶促糖化,并且将所生成的上清液浓缩直至25倍以便用于生物反应器中。浓缩糖流中的盐水平(如由电导率量度的)可以不可接受地高(高达1.5M Na^+当量)。当中和用于后续酶促糖化过程的爆裂的材料时,也产生另外的盐。本发明的实施例提供了用于去除这些盐的方法,从而可以将生成的浓缩纤维素糖流用于产生木质纤维素油的异养过程中。在一些实施例中,去除这些盐的方法为用树脂(例如,但不限于,DOWEX Marathon MR3)去离子化。在某些实施例中,用树脂去离子化的步骤在糖浓缩之前或在糖化之前的生物质的pH调节和水热处理或前述步骤的任何组合之前进行;在其他实施例中,在这些加工过程中的一个或多个之后实施该步骤。在其他实施例中,爆炸过程本身被改变,从而避免以不可接受高的水平产生盐。例如,纤维素生物质的硫酸(或其他酸)爆炸过程的替代方案是机械制浆以使纤维素生物质易于接受酶促水解(糖化)。在再另外的实施例中,使用抵抗高水平盐的天然微生物菌株或抵抗高水平盐的基因工程菌株。

[0174] 如下实施制备爆裂纤维素生物质的方法的优选实施例,其中爆裂纤维素生物质用于利用产油微生物生产异养木质纤维素油。第一个步骤包括将再悬浮的爆裂纤维素生物质的pH调节至5.0-5.3范围,随后将纤维素生物质洗涤3次。可以通过多种手段完成这个洗涤步骤,包括使用脱盐和离子交换树脂、反渗透、水热处理(如上文所述)、或仅在去离子水中反复再悬浮和离心。这个洗涤步骤产生其电导率在100-300μS/cm之间的纤维素流以及显著量的糠醛和羟甲基糠醛的去除。可以保存来自这个洗涤步骤的滤析液以浓缩从半纤维素部分释放的五碳糖。第二个步骤包括酶促糖化洗涤的纤维素生物质。在一个优选实施例中,使用Accellerase(Genencor)。第三个步骤包括通过离心或滤析和漂洗糖化的生物质回收糖。所生成的生物质(固体)是一种能量密集、富含木质素的组分,它可以用作燃料或送到废弃点。收集在离心/滤析和淋洗过程中回收的糖流。第四个步骤包括微过滤以去除染固形物,同时回收渗透物。第五个步骤包括可以使用真空蒸发器完成的浓缩步骤。这个步骤可以任选地包括添加消泡剂如P'2000(Sigma/Fluka),由于所生成的糖原料的蛋白质含量,有时它是必要的。

[0175] 在本发明方法的另一个实施例中,碳源是甘油,包括来自生物柴油酯交换的酸化和未酸化的甘油副产物。在一个实施例中,碳源包括甘油和至少一种其他碳源。在一些情况下,将所有甘油和至少一种其他固定碳源在发酵开始时提供给微生物。在一些情况下,将甘油和至少一种其他固定碳源以预定比率同时提供给微生物。在一些情况下,将甘油和至少一种其他固定碳源在发酵过程期间以预定的速率供应给微生物。

[0176] 一些微藻在甘油存在下经历比葡萄糖存在下更快的细胞裂解(参见PCT公开号

2008/151149)。在这些情况下,其中首先将甘油供应给细胞以快速增加细胞密度并且然后供应葡萄糖以积累脂质的两阶段生长过程可以提高产生脂质的效率。当返回该生产工艺时,酯交换法的甘油副产物的使用可以提供显著的经济优势。还提供了其他的进料方法,如甘油和葡萄糖的混合物。供应这类混合物还获得相同的经济益处。另外,本发明提供了向微藻供应替代性糖(如与甘油不同地组合的蔗糖)的方法。

[0177] 在本发明方法的另一个实施例中,碳源是转化糖。转化糖与蔗糖相比较不易结晶并且因此可以为储存和补料分批发酵提供优点,其中在异养培养微生物(包括微藻)的情况下,对于浓缩碳源存在着需要。在一个实施例中,该碳源是转化糖,在该培养步骤(其任选地是补料分批培养)之前优选处于浓缩形式,优选至少800g/升、至少900g/升、至少1000g/升或至少1100g/升。当细胞生长并且积累脂质时,优选地处于浓缩形式的转化糖随时间推移被供应给细胞。

[0178] 在本发明方法的另一个实施例中,碳源是蔗糖,包括含有蔗糖的复杂原料,如来自甘蔗加工的稠甘蔗汁。由于用于异养油生产的培养物的密度更高,固定碳源(例如,蔗糖、葡萄糖,等等)在培养步骤之前优选处于浓缩形式,优选至少500g/升、至少600g/升、至少700g/升或至少800g/升的固定碳源,所述培养步骤任选地是在细胞生长并且积累脂质时将这种材料随时间推移供应给细胞的补料分批培养。在一些情况下,碳源是在培养步骤之前处于稠甘蔗汁形式的蔗糖,优选处于浓缩形式,优选至少60%固形物或大约770g/升糖、至少70%固形物或大约925g/升糖、或至少80%固形物或大约1125g/升糖,所述培养步骤任选地是补料分批培养。当细胞生长并且积累脂质时,浓缩的稠甘蔗汁随时间推移被供应给细胞。

[0179] 在一个实施例中,培养基进一步包括至少一种蔗糖利用酶。在一些情况下,蔗糖利用酶是蔗糖转化酶。蔗糖转化酶是由外源蔗糖转化酶基因编码的可分泌性蔗糖转化酶,所述外源蔗糖转化酶基因由微生物群体表达。该可分泌性蔗糖转化酶可以由这些微生物分泌到培养基中以将培养基中的蔗糖转变为被微生物利用的葡萄糖和果糖。如以下所述,该蔗糖转化酶可以是重组体,由此赋予微生物使用纯的或复杂的蔗糖原料作为固定碳源用于生长或油生产的能力。在一些情况下,如下文第IV部分中更详细地描述,该微生物已经被基因工程化以表达蔗糖利用酶,如蔗糖转运蛋白、蔗糖转化酶、己糖激酶、葡萄糖激酶、或果糖激酶。

[0180] 含有蔗糖的复杂原料包括来自甘蔗加工的废弃糖蜜;这种低价值的甘蔗加工废弃产物的使用在生产烃和其他油中能够提供显著的成本节约。在本发明的方法中有用的另一种含有蔗糖的复杂原料是高粱,包括高粱饴和纯高粱。高粱饴是从甜高粱秆的汁中产生的。其糖谱主要由葡萄糖(右旋糖)、果糖和蔗糖组成。

[0181] 4.油生产

[0182] 为了根据本发明的方法生产油,优选的是在黑暗下培养细胞,例如,当使用不允许光冲击培养物的极大(40,000升和更大)发酵罐时,情况就是如此。可以在含有固定碳源的培养基中并且在没有光的情况下使异养生物种类生长和增殖用于产生油;这种生长称作异养生长。

[0183] 作为实例,将产生脂质的微藻细胞的接种物导入培养基中;在细胞开始增殖之前存在一个延迟期(延滞期)。在延迟期后,增殖速率稳定增加并且进入对数期或指数期。指数

期之后转而是增殖减慢，这是由于营养素(如氮)减少、有毒物质增加、以及群体感应机制的缘故。在这种减慢之后，增殖停止，并且细胞进入静止阶段或稳定生长状态，这取决于提供给细胞的特定环境。为了获得富含脂质的生物质，典型地正好在指数期结束之后收获培养物，可以通过允许氮或另一种关键营养素(除了碳之外)变得耗尽，从而迫使细胞将过量存在的碳源转化成脂质(并且具体地是甘油三酯)而提前终止该阶段。可以操纵培养条件参数来优化总油生产量、所产生的脂质种类的组合、和/或特定油的产生。

[0184] .借助于在此披露的细胞的脂质生产可以发生在对数期或其后(包括静止期)，在静止期中营养素被供应或仍然是可获得的，以允许在不存在细胞分裂下的脂质产生的继续。

[0185] 优选地，使用在此所述和/或本领域已知的条件培养的微生物包含按干重计至少大约20％-30％、30％-40％、40％-50％、50％-60％、60％-70％、或80％-90％的甘油三酯。可以调节工艺条件以增加适于特定用途的脂质的产量和/或降低生产成本。例如，在某些实施例中，在限制性浓度的一种或多种营养素(例如像，氮、磷、或硫)存在同时提供过量的固定碳能量如葡萄糖的情况下培养微藻。氮限制倾向于增加微生物脂质产量(每克干细胞重产生的脂质的量的度量)，超过其中过量提供氮的培养物中的微生物脂质产量。在特定的实施例中，脂质产量的增加是至少大约：10％、50％、100％、200％、或500％。可以将微生物在限制量的营养素存在的情况下培养持续总培养期的一部分或持续整个时期。在特定的实施例中，在总培养期期间，营养素浓度在限制性浓度与非限制性浓度之间循环至少2次。可以通过继续培养持续增加的时间段同时提供过量的碳但是限制氮或没有氮而增加细胞的脂质含量。

[0186] 在另一个实施例中，通过在脂质途径酶(例如，脂肪酸合成酶的辅酶或辅基)的一种或多种辅因子存在的情况下培养产生脂质的微生物(例如，微藻)而增加脂质产量。通常，辅因子的浓度足以增加微生物脂质(例如，脂肪酸)产量，其超过在辅因子不存在情况下的微生物脂质产量。在一个特定的实施例中，通过在培养物中包含微生物(例如，微藻)将辅因子提供给培养物，其中该微生物含有编码这种辅因子的外源基因。可替代地，可以通过包括含有外源基因的微生物(例如，微藻)将辅因子提供给培养物，其中该外源基因编码参与这种辅因子的合成的蛋白质。在某些实施例中，适合的辅因子包括脂质途径酶所需要的任何维生素(例如像，生物素或泛酸盐)。编码适合用于本发明的方法中的辅因子或参与这类辅因子的合成的基因是熟知的，并且可以使用构建体和如上文描述的那些技术将这些基因导入微生物中(例如，微藻)。

[0187] 在此所述的生物反应器、培养条件、以及异养生长和增殖方法的具体实例可以按任何适合的方式组合以提高微生物生长和脂质和/或蛋白质产生的效率。

[0188] 使用本领域内已知的不同的培养方法已经产生了具有按干重计高百分比的油/脂质积累的微藻生物质(参见PCT公开号2008/151149)。通过在此所述的培养方法产生的并且根据本发明使用的微藻生物质包含按干重计至少10％的微藻油。在一些实施例中，微藻生物质包含按干重计至少25％、50％、60％、70％或至少80％的微藻油。在一些实施例中，微藻生物质含有按干重计10％-90％的微藻油、25％-75％的微藻油、40％-75％的微藻油、75％-85％或50％-70％的微藻油。

[0189] 在此所述的或从用于本发明的方法和组合物中的生物质中提取的生物质微藻油

可以包含具有一条或多条不同的脂肪酸酯侧链的甘油脂。甘油脂是由与一个、二个或三个脂肪酸分子酯化的甘油分子组成，所述脂肪酸分子可以具有不同的长度并且具有不同的饱和度。如下文第IV部分中更详细地描述的，可以通过培养条件或通过脂质途径工程操纵脂肪酸分子(和微藻油)的长度特征和饱和特征来改变在本发明的实施例的微藻油中的脂肪酸分子的特性或比例。特性与比例的具体修饰包括改变微生物甘油三酯的脂肪酸分布，如改变链长谱、饱和谱、以及脂肪酸的羟化。如此产生的油可以包含天然油。可替代地，可以在单一藻种类中通过将源自两个或更多个微藻种类的生物质或藻油混合，或通过将本发明的藻油与源自其他来源的油(如大豆油、菜籽油、卡诺拉油、棕榈油、棕榈仁油、椰子油、玉米油、废弃植物油、乌桕油、橄榄油、向日葵油、棉籽油、鸡脂肪、牛脂、猪脂、微藻油、巨藻油、微生物油、正蓖距花油、亚麻油、花生油、精选白脂膏、猪油、亚麻荠(Camelina sativa)油、芥子油、腰果油、燕麦油、羽扇豆油、洋麻油、金盏花油、海草(help)油、咖啡油、亚麻油、榛果油、大戟油、南瓜子油、芫荽油、茶油、芝麻油、红花油、米糠油、油桐树油、可可油、干椰子肉油、鸦片罂粟油、蓖麻油、山核桃油、霍霍巴油、澳洲坚果油、巴西坚果油、鳄梨油、石油、或任何前述油的馏分)混合而制备微藻油的特定共混物。

[0190] 也可以通过组合来自至少两个不同微藻种类的生物质或油而操纵油的组成，即，甘油脂的脂肪酸组分的特性和比例。在一些实施例中，至少两个不同的微藻种类具有不同的甘油脂谱。不同的微藻种类可以如在此所述那样一起或分开培养，优选地在异养条件下培养，以便产生对应的油。不同的微藻种类可以在细胞的甘油脂中含有不同百分比的不同脂肪酸组分。

[0191] 通常，无绿藻株具有极少或不具有链长C8-C14的脂肪酸。例如，Prototheca moriformis(UTEX 1435)、Prototheca krugani(UTEX 329)、雍滞无绿藻(UTEX1442)以及祖菲无绿藻(UTEX 1438)不含有(或含有不可检测量的)C8脂肪酸、0-0.01％的C10脂肪酸、0.03％-2.1％的C12脂肪酸以及1.0％-1.7％的C14脂肪酸。

[0192] 在一些情况下，这些含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于链长C8或C8-10的脂肪酰基-ACP底物具有活性)的转基因的微生物株系具有至少1.5％、至少3.0％、至少10％、至少12％或更多的链长C8的脂肪酸。在其他情况下，这些含有编码脂肪酰基ACP硫酯酶(该硫酯酶对于链长C10的脂肪酰基-ACP底物具有活性)的转基因的微生物株系具有至少5.0％、至少10.0％、至少24％、至少29％或更多的链长C10的脂肪酸。在其他情况下，这些含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于链长C12的脂肪酰基-ACP底物具有活性)的转基因的微生物株系具有至少5％、至少15％、至少34％、至少50％或更多的链长C12的脂肪酸。在其他情况下，这些含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于链长C14的脂肪酰基-ACP底物具有活性)的转基因的微生物株系具有至少2.0％、至少7％、至少10％、至少15％、至少30％、至少43％或更多的链长C14的脂肪酸。在其他情况下，这些含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于链长C16的脂肪酰基-ACP底物具有活性)的转基因的微生物株系具有至少30％、至少40％、至少66％或更多的链长C16的脂肪酸。仍然在其他情况下，这些含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于链长C18并且尤其是C18:0的脂肪酰基-ACP底物具有活性)的转基因的微生物株系具有至少5％、至少10％、至少26％、至少40％或更多的C18:0脂肪酸水平。在任何这些实例中，该微生物可以是一种微藻，如无绿藻属。

[0193] 在非限制性的实例中，含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于链长C8的脂

肪酰基-ACP底物具有活性)的微生物株系具有1%-20%、优选1.8%-13%的链长C8的脂肪酸。在其他非限制性的实例中,含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于链长C10的脂肪酰基-ACP底物具有活性)的微生物株系具有1%-40%、优选1.91%-30%的链长C10的脂肪酸。在其他非限制性的实例中,含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于链长C12的脂肪酰基-ACP底物具有活性)的微生物株系具有10%-60%、优选13.55%-55%的链长C12的脂肪酸。在其他非限制性的实例中,含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于链长C14的脂肪酰基-ACP底物具有活性)的微生物株系具有1%-50%、优选2.59%-43.27%的链长C14的脂肪酸。在其他非限制性的实例中,含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于碳链长度变化的脂肪酰基-ACP底物具有广谱特异性)的微生物株系具有高达70%的链长C16的脂肪酸。在其他情况下,含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于链长C16的脂肪酰基-ACP底物具有活性)的微生物株系具有高达75%、优选高达67.42%的链长C16的脂肪酸。在一些情况下,这些含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于链长C8与C14之间的脂肪酰基-ACP底物具有活性)的转基因的微生物株系具有1%-790%、或大约2%-80%的(C8-C14)脂肪酸。在一些情况下,这些含有编码脂肪酰基-ACP硫酯酶(该硫酯酶对于链长C12与C14之间的脂肪酰基-ACP底物具有活性)的转基因的微生物株系具有至少50%或60%的(C12-C14)脂肪酸。在一些情况下,由于具有特定链长的所希望的脂肪酸的增加,在恒定的并且高的选择性压力下保持这些转基因微生物株系以保留外源基因是有利的。还可以通过使用同源重组载体以及在此披露的方法将外源基因插入这些细胞的核染色体中来出现高水平的外源基因保留。含有整合到核染色体中的外源基因的重组细胞是本发明的目的。在任何这些实例中,该微生物可以是一种微藻,如无绿藻属。

[0194] 任选地,该微生物油还可以包含由微藻产生或从培养基中结合到微藻油中的其他组分。这些其他组分可以按不同的量存在,这取决于用于培养微藻的培养条件、微藻的种类、用来从生物质回收微藻油的提取方法和可能影响微藻油组成的其他因素。此类组分的非限制性实例包括以油的0.025-0.3mcg/g,优选从0.05至0.244微克/克存在的类胡萝卜素;以油的0.025-0.3mcg/g,优选从0.045至0.268微克/克存在的叶绿素A;少于油的0.03mcg/g,优选少于0.025微克/克的总叶绿素;以油的35-175mcg/g,优选从38.3-164微克/克存在的γ-生育酚;以油的50-300mcg/g,优选从60.8至261.7微克/克存在的总生育酚;少于0.5%,优选少于0.25%的菜子甾醇、菜油甾醇、豆甾醇或β谷固醇;少于油的300微克/克的总生育三烯酚;以及以油的225-350mcg/g,优选从249.6微克/克至325.3微克/克存在的总生育三烯酚。

[0195] 其他组分可以包括而不限于磷脂、生育酚、生育三烯酚、类胡萝卜素(例如,α-胡萝卜素、β-胡萝卜素、番茄红素,等等)、叶黄素(xanthophyll)(例如,叶黄素(lutein)、玉米黄质、α-隐黄质和β-隐黄质)、以及各种有机或无机化合物。在一些情况下,从无绿藻种类提取的油包含每克油0.001至0.05微克之间,优选0.003至0.039微克之间的叶黄素,每克油少于0.005微克、优选少于0.003微克的番茄红素;以及每克油少于0.005微克的、优选少于0.003微克的β胡萝卜素。

[0196] III.基因工程方法和材料

[0197] 本发明提供了用于对在本发明的方法中有用的无绿藻细胞以及重组宿主细胞进行基因修饰的方法和材料,这些细胞包括但不限于重组Prototheca moriformis、祖菲无绿

藻、Prototheca krugani、和雍滞无绿藻宿主细胞。为了方便读者，将这些方法和材料的描述内容分成多个小节。在小节1中，描述了转化方法。在小节2中，描述了使用同源重组的基因工程方法。在小节3中，描述了表达载体和组件。

[0198] 1.工程化方法-转化

[0199] 可以通过任何适合的技术转化细胞，所述技术例如包括生物射弹(biolistics)、电穿孔(参见Maruyama等人(2004)，Biotechnology Techniques 8:821-826)、玻璃珠转化(glass bead transformation)和碳化硅晶须转化(silicon carbide whisker transformation)。另一种可以使用的方法涉及形成原生质体和使用$CaCl_2$和聚乙二醇(PEG)以将重组DNA导入微藻细胞中(参见Kim等人(2002)，Mar.Biotechnol.4:63-73，其报道了这种椭圆小球藻(Chorella ellipsoidea)转化方法的用途)。微藻的共转化可以用来将两种不同的载体分子同时导入细胞中(参见例如Protist 2004Dec;155(4):381-93)。

[0200] 生物射弹法(参见例如，Sanford,Trends In Biotech(《生物技术趋势》).(1988)6:299 302，美国专利号4,945,050)；电穿孔(Fromm等人，Proc.Nat'l.Acad.Sci(《美国科学院院刊》)(美国).(1985)82:5824 5828)、使用激光束、显微注射或能够将DNA导入微藻中的任何其他方法也可以用于转化无绿藻细胞。

[0201] 2.工程化方法-同源重组

[0202] 同源重组是互补DNA序列对齐和交换同源区的能力。将含有与被靶向的基因组序列(“模板”)同源的序列的转基因DNA(“供体”)导入生物中并且然后在相应的基因组同源序列的位点处经历重组到基因组中。

[0203] 对于在分子遗传水平可以进行的事项而言，在宿主生物中实施同源重组的能力具有许多实际意义，并且在生成能够产生定制油的产油微生物中是有用的。就其本质而言，同源重组是一个精确的基因打靶事件；因此，用相同的靶向序列产生的大部分转基因系就表型而言将是基本上相同的，从而迫使筛选少得多的转化事件。同源重组还将基因插入事件靶向到宿主染色体中，潜在地产生优良的遗传稳定性，甚至在没有遗传选择的情况下也是如此。由于不同的染色体基因座将可能影响基因表达(甚至从异源启动子/UTR)，同源重组可以是一种在陌生基因组环境中查找基因座并评定这些环境对基因表达影响的方法。

[0204] 一个利用同源重组的特别有用的基因工程方法是共同选择(co-opt)特异性宿主调节元件如启动子/UTR以高度特异性方式驱动异源基因表达。例如，用编码选择性标记的异源基因消融或敲除去饱和酶基因/基因家族可能被期望增加宿主细胞中产生的饱和脂肪酸的总百分比。实例6描述了同源重组靶向构建体和在Prototheca moriformis中产生的这类去饱和酶基因消融或敲除的工作实例。减少内源基因表达的另一种方法是使用RNA诱导的基因表达下调或沉默，包括但不限于RNAi或反义方法、以及dsRNA方法。反义法、RNAi法、dsRNA法是本领域熟知的并且包括导入当表达为mRNA时会导致发夹RNA形成的表达构建体或含有靶基因的一部分的表达构建体，其中所述部分会以反义方向转录。所有三种方法均导致减少的靶基因表达。实例6还描述了通过RNAi或反义方法下调内源Prototheca moriformis Δ12去饱和酶基因(FADc)的表达构建体和工作实例。

[0205] 由于同源重组是一个精确的基因打靶事件，它可以用来精确地修饰在感兴趣的基因或区域内的任何核苷酸，只要已经鉴定足够的侧翼区即可。因此，可以使用同源重组作为修饰影响RNA和/或蛋白质的基因表达的调节序列的手段。它还可以用来修饰蛋白质编码区

以调节酶活性如底物特异性、亲和力、和Km，并且从而实现所希望的宿主细胞代谢变化。同源重组提供了强有力的操纵宿主基因组的手段，导致基因打靶、基因转换、基因缺失、基因重复、和基因倒位以及交换基因表达调节元件如启动子、增强子和3'UTR。

[0206] 可以通过使用靶向构建体实现同源重组，所述靶向构建体含有"靶向"宿主细胞内源基因组内的感兴趣的基因或区域的内源序列片段。这样的靶向序列可以位于感兴趣的基因或区域的5'、感兴趣的基因/区域的3'或甚至位于感兴趣的基因/区域的侧翼。可以将这样的靶向构建体作为具有另外的载体骨架的超螺旋质粒DNA、不具有载体骨架的PCR产物或作为线性化分子转化到宿主细胞中。在一些情况下，可能有利的是，首先用限制性酶暴露转基因DNA(供体DNA)内部的同源序列。这个步骤可以增加重组效率并且减少不希望的事件的发生。增加重组效率的其他方法包括使用PCR产生转化用转基因DNA，其含有与正在被靶向的基因组序列同源的线性末端。

[0207] 出于非限制性说明的目的，用于同源重组的供体DNA序列的区域包括在Prototheca moriformis中的DNA的KE858区域。KE858是一个1.3kb的基因组片段，它包含蛋白质的编码区的一部分，其中所述蛋白质与转移RNA(tRNA)蛋白质家族共享同源性。Southern印迹已经显示KE858序列在Prototheca moriformis(UTEX 1435)基因组中以单拷贝存在。已经在PCT申请号PCT/US2009/066142中描述了这个区域和使用这个区域进行同源重组打靶的实例。另一种有用的供体DNA区域是在此表示为"6S"的基因组序列(在SEQ ID NO:82、SEQ ID NO:84的供体序列)。注意该6S序列不是6S rRNA序列。这个序列在Prototheca moriformis中的同源重组中的用途描述于以下实例中。

[0208] 3.载体和载体组件

[0209] 就在此的披露内容而言，可以通过本领域技术人员熟悉的已知技术制备用于根据本发明的微生物的转化的载体。载体典型地含有一种或多种基因，其中每种基因指令所希望的产物(基因产物)的表达并与一个或多个控制序列可操作地连接，所述控制序列调节基因表达或将该基因产物靶向到重组细胞中的特定位置。为了方便读者，将这个小节分成多个小节。小节A描述了典型地包含在载体上的控制序列以及由本发明提供的新颖的控制序列。小节B描述了典型地包含在载体中的基因以及新颖的密码子优化方法和由本发明提供的使用它们所制备的基因。

[0210] A.控制序列

[0211] 控制序列是调节编码序列表达或指导基因产物至细胞内或细胞外特定位置的核酸。调节表达的控制序列例如包括调节编码序列转录的启动子和终止编码序列转录的终止子。另一个控制序列是位于编码序列末端的编码多聚腺苷化信号的3'非翻译序列。将基因产物导向至特定位置的控制序列包括编码信号肽的那些，这些信号肽将它们连接到其上的蛋白质导向至细胞内或细胞外的特定位置。

[0212] 因此，用于在微藻中表达外源基因的示例性载体设计含有所希望的基因产物(例如，选择标记、脂质途径修饰酶、或蔗糖利用酶)的编码序列，该编码序列与微藻中有活性的启动子可操作地连接。可替代地，如果这种载体不含与感兴趣的编码序列可操作地连接的启动子，则可以将该编码序列转化到细胞中，使得它在载体整合时变成与内源启动子可操作地连接。已经证明无启动子的转化方法在微藻(参见例如Plant Journal 14:4，(1998)，第441-447页)中有效。

[0213] 许多启动子在微藻中有活性,包括相对于正在转化的藻类为内源的启动子,以及相对于正在转化的藻类不是内源的启动子(即,来自其他藻类的启动子、来自高等植物的启动子、以及来自某些植物病毒或藻类病毒的启动子)。PCT公开号2008/151149和其中引用的参考文献中描述了在微藻中有活性的说明性外源和/或内源启动子(以及在微藻中有功能的抗生素抗性基因)。

[0214] 用来表达外源基因的启动子可以是与这个基因天然连接的启动子或可以是异源启动子。一些启动子在一个以上的微藻种类类中是有活性的。其他启动子是种类特异性的。说明性的启动子包括在下文实例中使用的启动子,如来自莱茵衣藻(Chlamydomonas reinhardtii)β-微管蛋白的启动子,和已经显示在多个微藻种类中有活性的病毒启动子,如花椰菜花叶病毒(CMV)和小球藻病毒(参见例如Plant Cell Rep(《植物细胞报告》).2005Mar;23(10-11):727-35;J Microbiol(《微生物学杂志》).2005Aug;43(4):361-5;Mar Biotechnol(《海洋生物技术》)(NY).2002Jan;4(1):63-73)。适用于在无绿藻中表达外源基因的另一种启动子是Chlorella sorokiniana谷氨酸脱氢酶启动子/5'UTR。任选地,使用含有启动子的这些序列的至少10、20、30、40、50、或60个或更多个核苷酸。在本申请的序列表中列出了用于无绿藻中表达外源基因的说明性启动子,如小球藻HUP1基因的启动子(SEQ ID NO:1)和椭圆小球藻硝酸盐还原酶启动子(SEQ ID NO:2)。小球藻病毒启动子也可以用来在无绿藻中表达基因,如美国专利6,395,965的SEQ ID NOs:1-7。另外的在无绿藻中有活性的启动子可以例如在Biochem Biophys Res Commun.(《生物化学与生物物理学研究通讯》)1994Oct 14;204(1):187-94;Plant Mol Biol(《植物分子生物学》).1994Oct;26(1):85-93;Virology(《病毒学》).2004Aug 15;326(1):150-9;和Virology(《病毒学》).2004Jan 5;318(1):214-23中找到。其他有用的启动子详细描述于以下实例中。

[0215] 启动子通常可以表征为组成型或诱导型。组成型启动子通常具有在所有时间(或在细胞生活周期中的某些时间)以相同水平驱动表达的活性或功能。相反,诱导型启动子仅仅在响应于刺激时有活性(或变得无活性)或显著上调或下调。两种类型的启动子均用于本发明的方法中。用于本发明中的诱导型启动子包括这些启动子,它们介导可操作地连接的基因响应于刺激(如外源提供的小分子(例如,葡萄糖,如在SEQ ID NO:1中)、温度(热或冷)、培养基中氮的缺乏,等等)的转录。适合的启动子可以激活基本上沉默的基因的转录或上调、优选地显著上调以低水平转录的可操作地连接的基因的转录。以下实例描述了在无绿藻细胞中有用的另外的诱导型启动子。

[0216] 包含终止区控制序列是任选的,并且如果使用,则选择主要在于方便,因为终止区是相对可互换的。终止区可以相对于转录起始区(启动子)是天然的,可以相对于感兴趣的DNA序列是天然的,或可以是从另一个来源可获得的。参见例如,Chen和Orozco,Nucleic Acids Res.(1988)16:8411。

[0217] 本发明还提供了控制序列和重组基因以及含有它们的载体,它们被提供为将感兴趣的基因产物导向特定的细胞区室,如叶绿体、质体、线粒体、或内质网。另外,本发明的实施例包括控制序列和重组基因以及含有它们的载体,它们被提供为蛋白质在细胞外部分泌。

[0218] 可以使用质体靶向信号,将无绿藻的核基因组中表达的蛋白质靶向到质体。相对于小球藻为内源的质体靶向序列是已知的,如在小球藻核基因组中编码被靶向至质体的蛋

白质的基因;参见例如GenBank登录号AY646197和AF499684,并且在一个实施例中,这样的控制序列用于本发明的载体中将蛋白质表达靶向到无绿藻质体。

[0219] 以下实例描述了藻质体靶向序列将异源蛋白质靶向到宿主细胞中的正确区室的用途。使用Prototheca moriformis和原始小球藻细胞产生cDNA文库,并且在PCT申请号PCT/US2009/066142中描述了这些文库。

[0220] 在本发明的另一个实施例中,将在无绿藻属中的多肽的表达靶向到内质网。在表达载体中包含适当的滞留或分选信号确保蛋白质留在内质网(ER)中并且不顺流进入高尔基体。例如,来自Wageningen UR-Plant Research International(瓦赫宁根UR-国际植物研究所)的IMPACTVECTOR1.3载体包含熟知的KDEL滞留或分选信号。用这种载体,ER滞留具有实用优点,因为已经报道它提高表达水平5倍或更多。这种现象的主要原因似乎是相比于在细胞质中存在的负责表达的蛋白质的翻译后降解的蛋白酶,ER含有较低浓度和/或不同的蛋白酶。在绿色微藻中有功能的ER滞留信号是已知的。例如,参见Proc Natl Acad Sci U S A.2005Apr 26;102(17):6225-30。

[0221] 在本发明的另一个实施例中,将在细胞外部分泌的多肽靶向到培养基中。根据本发明的方法可以在无绿藻属中使用的在小球藻中有活性的分泌信号的实例参见Hawkins等人,Current Microbiology(《当代微生物学》)第38卷(1999),第335-341页。

[0222] B.基因和密码子优化

[0223] 典型地,基因包括启动子、编码序列、和终止控制序列。当通过重组DNA技术装配时,基因可以称作表达盒并且可以在侧翼具有方便插入载体中的限制性酶切位点,所述载体被用来将重组基因导入宿主细胞中。这个表达盒可以在侧翼具有来自基因组或其他核酸靶的DNA序列,以便促进表达盒通过同源重组稳定整合到基因组中。可替代地,载体及其表达盒可以保持未整合(例如,游离基因),在这种情况下,载体典型地包含能够提供异源载体DNA复制的复制起点。

[0224] 在载体上存在的常见基因是编码蛋白质的基因,所述蛋白质的表达允许含有这种蛋白质的重组细胞区别于不表达这种蛋白质的细胞。这种基因以及其相应的基因产物称作选择标记(selectable marker)或选择标记(selection marker)。在对于转化无绿藻属有用的转基因构建体中可以使用广泛的多种选择标记中的任一种。适合的选择标记的实例包括G418抗性基因、硝酸盐还原酶基因(参见Dawson等人(1997),Current Microbiology(《当代微生物学》)35:356-362)、潮霉素磷酸转移酶基因(HPT;参见Kim等人(2002),Mar.Biotechnol(《海洋生物技术》).4:63-73)、新霉素磷酸转移酶基因和赋予腐草霉素抗性的ble基因(Huang等人(2007),Appl.Microbiol.Biotechnol(《应用微生物学与生物技术》).72:197-205)。确定微藻对抗生素敏感性的方法是熟知的。例如,Mol Gen Genet(《分子基因遗传学》),19960ct 16;252(5):572-9,如在此所述的蔗糖转化酶、以及也如在此所述的硫胺素营养缺陷型互补。

[0225] 不是基于抗生素的其他选择标记也可以在用于转化微藻(通常,包括无绿藻种类)的转基因构建体中使用。也可以使用赋予利用某些碳源的能力的基因作为选择标记,所述碳源以前是不能由该微藻利用的。以实例说明,Prototheca moriformis株典型地在蔗糖上不良地生长(如果生长的话)。使用含有蔗糖转化酶基因的构建体可以赋予阳性转化子在作为底物的蔗糖上生长的能力。有关将蔗糖利用作为选择标记连同其他选择标记的另外的细

节在以下部分IV中讨论。

[0226] 出于本发明的目的,用来制备本发明的重组宿主细胞的表达载体将包括至少2个并且经常是3个基因,如果这些基因之一是选择标记的话。例如,可以通过用本发明的载体转化产生本发明的基因工程化无绿藻属,其中除了选择标记之外,所述载体还包含一个或多个外源基因,例如像,蔗糖转化酶基因或酰基ACP-硫酯酶基因。一个或两个基因可以使用诱导型启动子表达,这允许控制这些基因表达的相对定时以增强脂质生产和转化成脂肪酸酯。两个或更多个外源基因的表达可以在相同的诱导型启动子控制之下或者在不同的诱导型(或组成型)启动子控制之下。在后一种情况下,可以诱导第一外源基因表达持续第一时间段(在此期间,可以诱导或可以不诱导第二外源基因的表达),并且可以诱导第二外源基因表达持续第二时间段(在此期间,可以诱导或可以不诱导第一外源基因的表达)。

[0227] 在其他实施例中,两个或更多个外源基因(除了任何选择标记之外)是脂肪酰基-ACP硫酯酶和脂肪酰基辅酶A/醛还原酶,它们的组合作用产生醇产物。进一步提供的是外源基因的其他组合,包括而不限于,脂肪酰基-ACP硫酯酶和脂脂肪酰基辅酶A还原酶,以产生醛。在一个实施例中,载体提供脂肪酰基-ACP硫酯酶、脂脂肪酰基辅酶A还原酶和脂肪醛脱羧酶的组合,以产生链烷。在这些实施例的每一个中,可以使用诱导型启动子表达一个或多个外源基因。

[0228] 表达两个或更多个外源基因的本发明的其他说明性载体包含编码蔗糖转运蛋白和蔗糖转化酶两者的那些载体以及编码选择标记和分泌型蔗糖转化酶的那些载体。由于使用甘蔗(和甘蔗衍生糖)作为碳源的工程化能力,用任一类型载体转化的重组无绿藻属以较低的制造成本产生脂质。上文描述的两种外源基因的插入可以与通过定向和/或随机诱变破坏多糖生物合成相组合,这甚至使更多的碳流转向进入脂质生产。单独和组合地,营养型转换、改变脂质产生的工程化、以及用外源酶处理改变了由微生物产生的脂质组成。这种改变可以是以下方面的变化:所产生脂质的量、所产生的一种或多种烃种类相对于其他脂质的量、和/或在微生物中产生的脂质种类的类型。例如,可以将微藻工程化以产生更高量和/或百分比的TAG(三酰甘油酯)。

[0229] 为了最佳表达重组蛋白,采用这样的编码序列是有益的:这些编码序列用优选被有待转化的宿主细胞使用的密码子产生mRNA。因此,转基因的正确表达可能需要转基因的密码子使用匹配于其中正在表达这种转基因的生物的特定密码子偏倚。作为这种作用基础的精确机制是多样的,但是包括将可获得的氨基酰tRNA池与细胞中正在合成的蛋白质正确平衡,当满足这种需要时,这伴随转基因信使RNA(mRNA)的更高效翻译。当转基因中的密码子使用时未优化时,可获得的tRNA池并不足以允许异源mRNA的高效翻译,从而导致核糖体失速和终止以及转基因mRNA可能不稳定。

[0230] 本发明提供了用于在无绿藻中成功表达重组蛋白的密码子优化的核酸。通过研究从Prototheca moriformis分离的cDNA序列分析了无绿藻种类中的密码子使用。这项分析代表对超过24,000个密码子的查询,并且在下表2中显示结果。

[0231] 表2.在无绿藻株中优选的密码子使用

[0232]

Ala	GCG	345 (0.36)	Asn	AAT	8 (0.04)
	GCA	66 (0.07)		AAC	201 (0.96)
	GCT	101 (0.11)			
	GCC	442 (0.46)	Pro	CCG	161 (0.29)
				CCA	49 (0.09)
Cys	TGT	12 (0.10)		CCT	71 (0.13)
	TGC	105 (0.90)		CCC	267 (0.49)
Asp	GAT	43 (0.12)	Gln	CAG	226 (0.82)
	GAC	316 (0.88)		CAA	48 (0.18)
Glu	GAG	377 (0.96)	Arg	AGG	33 (0.06)
	GAA	14 (0.04)		AGA	14 (0.02)
				CGG	102 (0.18)
Phe	TTT	89 (0.29)		CGA	49 (0.08)
	TTC	216 (0.71)		CGT	51 (0.09)
				CGC	331 (0.57)
Gly	GGG	92 (0.12)			
	GGA	56 (0.07)	Ser	AGT	16 (0.03)
	GGT	76 (0.10)		AGC	123 (0.22)
	GGC	559 (0.71)		TCG	152 (0.28)
				TCA	31 (0.06)
His	CAT	42 (0.21)		TCT	55 (0.10)
	CAC	154 (0.79)		TCC	173 (0.31)

[0233]

Ile	ATA	4 (0.01)	Thr	ACG	184 (0.38)
	ATT	30 (0.08)		ACA	24 (0.05)
	ATC	338 (0.91)		ACT	21 (0.05)
				ACC	249 (0.52)
Lys	AAG	284 (0.98)			
	AAA	7 (0.02)	Val	GTG	308 (0.50)
				GTA	9 (0.01)
Leu	TTG	26 (0.04)		GTT	35 (0.06)
	TTA	3 (0.00)		GTC	262 (0.43)
	CTG	447 (0.61)			
	CTA	20 (0.03)	Trp	TGG	107 (1.00)
	CTT	45 (0.06)			
	CTC	190 (0.26)	Tyr	TAT	10 (0.05)
				TAC	180 (0.95)
Met	ATG	191 (1.00)			
			终止	TGA/TAG/TAA	

[0234] 在其他实施例中，参考除了无绿藻株之外的微藻株系，已经将重组载体中的基因进行密码子优化。例如，在美国专利号7,135,290中描述了再编码用于在微藻中表达的基因的方法。另外的用于密码子优化的信息是可获得的，例如，在GenBank的密码子使用数据库处获得。

[0235] 关于具有密码子优化的基因的实施例，这些密码子优化的基因优选被优化以增加该优化的基因的基因产物的表达至少10％并且更优选至少20％、40％、60％、80％、100％或200％。

[0236] 虽然本发明的方法和材料允许将任何外源基因导入无绿藻中，涉及蔗糖利用和脂质途径修饰的基因是特别感兴趣的，如在以下部分中所讨论的。

[0237] IV.选择标记

[0238] 1.蔗糖利用

[0239] 在一个实施例中，本发明的重组细胞进一步包含一个或多个外源蔗糖利用基因。在各种实施例中，一种或多种基因编码选自下组的一种或多种蛋白质，该组由以下各项组成：果糖激酶、葡萄糖激酶、己糖激酶、蔗糖转化酶、蔗糖转运蛋白。例如，蔗糖转运蛋白和蔗糖转化酶的表达允许无绿藻将蔗糖从培养基转运至细胞中并且水解蔗糖以产生葡萄糖和果糖。任选地，在内源己糖激酶活性不足以导致果糖最大磷酸化的情况下，也可以表达果糖激酶。适合的蔗糖转运蛋白的实例是Genbank登录号CAD91334、CAB92307、和CAA53390。适合的果糖激酶的实例是Genbank登录号P26984、P26420和CAA43322。

[0240] 在一个实施例中，本发明提供了一种分泌蔗糖转化酶的宿主细胞。蔗糖转化酶的分泌避免需要表达可以将蔗糖转运至细胞中的转运蛋白。这是因为分泌型转化酶催化一分子蔗糖转化成一分子葡萄糖和一分子果糖，两者均可以被本发明提供的微生物转运和利用。例如，带有分泌信号(如SEQ ID NO:4的分泌信号(来自酵母)、SEQ ID NO:5(来自高等植物)、SEQ ID NO:6(真核共有分泌信号)和SEQ ID NO:7(来自高等植物和真核共有信号序列的组合)的蔗糖转化酶(如SEQ ID NO:3)的表达在细胞外产生转化酶活性。这种蛋白质的表达(如通过在此披露的基因工程方法学能够实现)允许已经能够利用细胞外葡萄糖作为能源的细胞利用蔗糖作为细胞外能源。

[0241] 在含有蔗糖的培养基中表达转化酶的无绿藻种类是用于生产油的优选微藻种类。这种充分有活性的蛋白质的表达及其细胞外靶向允许所得到的宿主细胞在蔗糖上生长，而它们未转化的对应物则不能。因此，本发明的实施例提供了具有一种密码子优化的转化酶基因的重组微藻(包括无绿藻属)细胞，该转化酶基因包括但不限于酵母转化酶基因，其整合到它们的基因组中，使得该转化酶基因被表达为通过转化酶活性和蔗糖水解进行评定。将转化酶基因用作重组细胞中的选择标记，因为这类细胞能够在蔗糖上生长，而它们未转化的对应物则不能；并且使用转化酶选择重组宿主细胞的方法是藻类分子遗传学的强有力选择标记。

[0242] 蔗糖转化酶在无绿藻中的成功表达也表明了本发明的另一个方面，因为它证明异源(重组)蛋白可以在藻细胞中表达并且成功地以完全活性和功能性形式运送至细胞外并且进入培养基。因此，本发明的实施例提供了用于在微藻中表达广泛和多样类型的异源蛋白质并且将它们分泌于宿主细胞外部的方法和试剂。这种蛋白质包括例如工业用酶，例如像，脂肪酶、蛋白酶、纤维素酶、果胶酶、淀粉酶、酯酶、氧化还原酶、转移酶、乳糖酶、异构酶、以及转化酶，以及治疗性蛋白，例如像生长因子、细胞因子、包含两轻链与两重链的全长抗体、Fab、scFv(单链可变片段)、新型骆驼抗体(camellid-type antibody)、抗体片段、抗体片段融合物、抗体-受体融合物、胰岛素、干扰素、以及胰岛素样生长因子。

[0243] 蔗糖转化酶在无绿藻中的成功表达还表明了本发明的另一个方面，因为它提供了

用于在藻类中使用真菌转运肽引导在无绿藻中的蛋白质分泌的方法和试剂;以及用于确定一种肽在无绿藻细胞中作为转运肽是否起作用以及其起作用的能力的方法和试剂。本发明的这些方法和试剂可以用作工具与平台来鉴定其他的可以成功地将蛋白质运输到细胞之外的转运肽,并且酵母转化酶在这些方法中具有很好的效用。如在这一实例中证明的,内源酵母转化酶转运肽的去除以及其被其他转运肽(相对于宿主藻类是内源的或来自其他来源(真核的、原核的以及病毒的))的替换可以鉴定任何感兴趣的肽在引导蛋白质从细胞流出中是否可以作为转运肽起作用。

[0244] 适合的蔗糖转化酶的实例包括Genbank登录号CAB95010、NP_012104和CAA06839鉴定的那些。下文表3中列出了适合的转化酶的非限制性实例。下文序列表中包括每种列出的转化酶的氨基酸序列。在一些情况下,适合用于本发明的方法和载体中的外源蔗糖利用基因编码与选自表3的蔗糖转化酶具有至少40%、50%、60%、75%、或90%或更高的氨基酸一致性的蔗糖转化酶。

[0245] 表3.蔗糖转化酶

[0246]

描述	生物	GenBank 登录号	SEQ ID NO:
转化酶	*菊苣(Chicorium intybus)*	Y11124	SEQ ID NO:20
转化酶	*粟酒裂殖酵母(Schizosaccharomyces pombe)*	AB011433	SEQ ID NO:21
β-呋喃果糖苷酶（转化酶）	*异常毕赤酵母(Pichia anomala)*	X80640	SEQ ID NO:22
转化酶	*西洋德巴利酵母(Debaryomyces occidentalis)*	X17604	SEQ ID NO:23
转化酶	*水稻 (Oryza sativa)*	AF019113	SEQ ID NO:24
转化酶	*洋葱(Allium cepa)*	AJ006067	SEQ ID NO:25
转化酶	*栽培甜菜(Beta vulgaris subsp. Vulgaris)*	AJ278531	SEQ ID NO:26
β-呋喃果糖苷酶（转化酶）	*短双歧杆菌(Bifidobacterium breve) UCC2003*	AAT28190	SEQ ID NO:27
转化酶	*酿酒酵母(Saccharomyces cerevisiae)*	NP_012104	SEQ ID NO:8 (核苷酸)
			(SEQ ID NO:28(氨基酸)
转化酶 A	*运动发酵单胞菌(Zymomonas mobilis)*	AAO38865	SEQ ID NO:29

[0247] 转化酶由无绿藻分泌至培养基也能够使细胞在来自甘蔗加工的废弃糖蜜上生长，如同它们在纯试剂级的葡萄糖上生长那样；这种低价值的甘蔗加工废弃产物的使用可以在脂质和其他油的生产中提供显著的成本节省。因此，本发明提供了含有无绿藻属微生物群体的微生物培养物，以及包含(i)蔗糖和(ii)蔗糖转化酶的培养基。在各种实施例中，培养物中的蔗糖来自高粱、甜菜、甘蔗、糖蜜、或解聚的纤维素材料(它可以任选地含有木质素)。

在另一个方面，本发明的这些方法与试剂显著增加了可以被重组微藻或其他微生物利用的原料的数目和类型。虽然此处例举的微生物被改变使得它们可以利用蔗糖，可以应用本发明的方法和试剂，使得原料如纤维素材料是可以被本发明的具有分泌纤维素酶、果胶酶、异构酶等等的能力的工程化宿主微生物利用的，从而酶促反应的分解产物不再是仅仅简单地被耐受，而是被宿主用作碳源。这种实例描述在下文以及在微生物实例中，这些微生物被工程化以表达分泌型α-半乳糖苷酶，从而赋予水解寡糖中的α-半乳糖苷键(如那些包含在棉子糖和水苏糖中的那些)的能力，棉子糖和水苏糖是发现于农业废弃物流中的两种寡糖。

[0248] 2.α-半乳糖苷酶表达

[0249] 如上所述，虽然蔗糖转化酶的表达赋予无绿藻细胞更有效地利用蔗糖作为碳源的能力(通过该酶水解二糖蔗糖中的果糖与葡萄糖分子之间的α-连接)，其他的水解寡糖中的其他类型的α-连接的酶可以赋予无绿藻细胞利用其他碳源的能力。这些酶的表达(以及所得到的利用无绿藻属和其他微藻细胞通常不能利用的碳源的能力)可以通过允许能够在这些碳源上生长的阳性克隆的选择而作为用于这些转基因无绿藻细胞的选择标记。

[0250] 在一个实施例中，本发明的重组无绿藻细胞进一步含有一种或多种编码多糖降解酶的外源基因。在各种实施例中，这些编码多糖降解酶的一种或多种基因是编码分泌型α-半乳糖苷酶的基因。无绿藻细胞中的外源分泌型α-半乳糖苷酶的表达赋予这样的转化株系在含D-半乳糖苷键(如在半乳糖与葡萄糖单糖单元之间的α-连接)的糖(碳源)上生长的能力。表达外源分泌型α-半乳糖苷酶的无绿藻株将能够利用二糖，如蜜二糖(由α-D-半乳糖-葡萄糖组成的二糖)。

[0251] 糖，如棉子糖(由α-连接的半乳糖-葡萄糖-果糖构成的三糖)和水苏糖(由两个α-连接的D-半乳糖单元接着α-连接的葡萄糖与果糖构成的四糖)，在农业废弃物流(如甜菜浆(棉子糖)以及大豆粕(水苏糖))中以相当大比例存在。这样的农业残余物代表了大量未使用的碳源，可以用于通过能够利用它们的微生物(包括无绿藻属)而转变成油。

[0252] 无绿藻株不能够以任何显著的量利用寡糖或根本不能利用寡糖，如棉子糖和水苏糖。在棉子糖和水苏糖的情况下，尽管表达蔗糖转化酶的转基因株系(如上所述)具有水解在蔗糖的α-半乳糖苷衍生物中的果糖与葡萄糖之间的α-连接的能力，但是该寡糖的剩余部分仍未得以利用，因为蔗糖转化酶不会切割这种糖中的其余α-连接并且所产生的二糖是不可利用的。在另一个实施例中，本发明的重组无绿藻细胞包含编码蔗糖转化酶的外源基因以及编码α-半乳糖苷酶的外源基因两者。因此，表达蔗糖转化酶和α-半乳糖苷酶两者的株系将能够完全水解寡糖，如棉子糖和水苏糖，使这些组分单体能够消耗。另外，α-半乳糖苷酶编码基因可以被用作一种用于转化的选择标记。含有该外源α-半乳糖苷酶基因的克隆将具有在蜜二糖上生长的能力。适合的用于在无绿藻株中使用的α-半乳糖苷酶基因的实例包括来自卡尔酵母(Saccharomyces carlbergensis)的MEL1基因、来自黑曲霉的AglC基因。有趣的是，已经发现不是所有α-半乳糖苷酶基因都在无绿藻种类中具有功能，即使这些基因根据无绿藻株中的偏好密码子使用而被优化。以下实例显示了当用密码子优化的来自卡尔酵母(S.carlbergensis的)MEL1基因以及来自黑曲霉(A.niger)的AglC基因但不是来自高等植物Cyamopsis tetragonobola(瓜儿豆)的α-半乳糖苷酶编码基因转化时转基因无绿藻细胞在蜜二糖上生长的能力。

[0253] 3.硫胺素营养缺陷型互补

[0254] 包括Prototheca moriformis的无绿藻株已知是硫胺素营养缺陷的(参见,例如,Ciferri,O.(1956)Nature(《自然》),v.178,pp.1475-1476),这意味着这些株系在营养培养基中需要硫胺素用于生长。硫胺素营养缺陷可以是突变或硫胺素生物合成途径中的酶表达缺乏的结果。表达硫胺素生物合成途径中缺失的一种或多种酶的补偿的转基因株系然后可以不需要添加的硫胺素而生长,因此降低了营养培养基的成本,还致使所得到的微藻生物质更合乎需要地用作动物饲料。带有硫胺素生物合成途径酶的补偿还可以用作一种选择标记,因为该转基因赋予在不含硫胺素的平板/培养基上生长的能力。

[0255] 在一个实施例中,本发明的重组无绿藻细胞进一步含有一个或多个编码胺素生物合成途径酶的外源基因。在另一个实施例中,本发明的重组无绿藻细胞包含编码来自藻类、植物或蓝藻来源的羟甲基嘧啶磷酸合成酶(hydroxymethylpyrimidine phosphate synthase)的外源基因。仍然在其他实施例中,该羟甲基嘧啶磷酸合成酶由THIC基因编码。仍然在其他实施例中,该THIC基因是胶球藻属(Coccomyxa)C-169THIC、拟南芥(Arabidopsis thaliana)THIC、或集胞藻种类(Synechocystis sp.)PCC 6803thiC。以下实例详述了具有恢复的硫胺素原养型的Prototheca moriformis UTEX 1435的工程化。

[0256] V.脂质途径工程化

[0257] 除了改变微藻或其他微生物利用原料如含蔗糖原料的能力之外,本发明还提供了已经被修饰以改变所产生脂质的特性和/或比例的重组微藻或其他微生物。可以进一步或可替代地修饰这个途径以改变通过酶促加工脂肪酸途径中的脂质和中间体所产生的各种脂质分子的特性和/或比例。在各种实施例中,相对于其未转化的对应物,本发明的重组无绿藻细胞具有优化的每单位体积和/或每单位时间的脂质产量、碳链长度(例如,用于可再生柴油生产或用于需要脂质原料的工业化学品应用)、减少的或增加的双键数目和/或位置(任选地至零)、脂肪酸的羟化、以及增加的特定脂质种类或不同脂质群体的氢:碳比。

[0258] 在特定的实施例中,已经上调或下调控制针对脂肪酸合成的在代谢中的分支点的一种或多种关键酶来改进脂质产生。可以例如,通过用表达构建体转化细胞实现上调,其中编码感兴趣的酶的基因被表达,例如,使用增加转录的强启动子和/或增强子元件。这样的构建体可以包含选择标记,使得转化子可以经受选择,这可以导致基因保持,以及可能地扩增构建体并且增加所编码的酶的表达水平。适合于根据本发明的方法上调的酶的实例包括在将丙酮酸转化成乙酰辅酶A中发挥作用的丙酮酸脱氢酶(例如,一些来自微藻,包括Genbank登录号NP_415392;AAA53047;Q1XDM1和CAF05587)。丙酮酸脱氢酶的上调可以增加乙酰辅酶A的产生,并且从而增加脂肪酸合成。乙酰辅酶A羧化酶催化脂肪酸合成中的初始步骤。因此,可以上调这种酶以增加脂肪酸的产生(例如,一些来自微藻,包括Genbank登录号BAA94752;AAA75528;AAA81471;YP_537052;YP_536879;NP_045833和BAA57908)。也可以通过上调酰基载体蛋白(ACP)来增加脂肪酸产生,酰基载体蛋白在脂肪酸合成过程中携带正在增长的酰基链(例如,一些自微藻,包括Genbank登录号A0T0F8;P51280;NP_849041;YP_874433)。甘油-3-磷酸酰基转移酶催化脂肪酸合成的限速步骤。这种酶的上调可以增加脂肪酸产生(例如,一些来自微藻,包括Genbank登录号AAA74319;AAA33122;AAA37647;P44857;和ABO94442)。

[0259] 基因的上调和/或下调可以适用于控制脂肪酸生物合成途径的基因表达的全面调节子(global regulator)。因此,在适当的情况下,可以上调或下调脂肪酸合成的一种或多

种全面调节子以分别抑制或增强多个脂肪酸合成基因的表达并且最终增加脂质产生。实例包括固醇调节元件结合蛋白(SREBP),如SREBP-1a和SREBP-1c(例如,参见Genbank登录号NP_035610和Q9WTN3)。

[0260] 本发明还提供了已经被修饰以含有一个或多个外源基因的重组无绿藻细胞,其中所述外源基因编码脂质修饰酶如,例如像,脂肪酰基-ACP硫酯酶(参见表4)、脂肪酰基辅酶A/醛还原酶(参见表6)、脂肪酰基辅酶A还原酶(参见表7)、脂肪醛脱羧酶(参见表8)、脂肪醛还原酶、去饱和酶(如硬脂酰-ACP去饱和酶和脂肪酰基去饱和酶)和鲨烯合酶(参见GenBank登录号AF205791)。虽然脂肪酰基-ACP硫酯酶典型地不是直接地化学修饰这些脂质,根据本发明的实施例对它们的操作可以改变细胞的脂肪酸谱,尤其是在链长以及双键分布方面。在一些实施例中,将编码脂肪酰基-ACP硫酯酶和天然共表达的酰基载体蛋白的基因转化到无绿藻或其他微藻或微生物细胞中,任选地连同编码其他脂质修饰酶的一种或多种基因一起转化。在其他实施例中,ACP和脂肪酰基-ACP硫酯酶可以具有当这两种酶在本发明的微生物和方法中一起使用时赋予优势的相互亲和力,无论它们是否在特定组织或生物中天然共表达。因此,本发明的实施例考虑了这些酶的天然共表达对以及共享彼此相互作用的亲和力以促进从ACP切下特定长度碳链的那些酶。

[0261] 仍然在其他实施例中,将编码去饱和酶的外源基因连同编码其他脂质修饰酶的一种或多种基因一起转化到微藻或其他微生物细胞中,以便提供关于脂质饱和的修饰。在其他实施例中,内源去饱和酶基因在微藻或其他微生物细胞中过表达(例如,通过导入另外的基因拷贝)。硬脂酰-ACP去饱和酶(参见,例如,GenBank登录号AAF15308;ABM45911;和AAY86086)例如催化硬脂酰-ACP转化成油酰-ACP。上调这种基因可以增加由细胞产生的单不饱和脂肪酸比例;而下调可以降低单不饱和物的比例。出于说明性目的,硬脂酰-ACP去饱和酶(SAD)负责从C18:0前体合成C18:1脂肪酸。另一个去饱和酶家族是脂肪酰基去饱和酶(FAD),包括Δ12脂肪酸去饱和酶(Δ12FAD)。这些去饱和酶也提供关于脂质饱和的修饰。出于说明性目的,Δ12脂肪酸去饱和酶负责从C18:1前体合成C18:2脂肪酸。同样,可以控制一种或多种甘油脂去饱和酶(如ω-6脂肪酸去饱和酶、ω-3脂肪酸去饱和酶、或ω-6-油酸去饱和酶)的表达以改变不饱和脂肪酸对饱和脂肪酸的比率。在一些实施例中,可以参考所希望的碳链长度来选择去饱和酶,使得去饱和酶能够在指定碳长度的底物或具有指定范围内碳长度的底物内产生位置特异性修饰。在另一个实施例中,如果所希望的脂肪酸谱是单不饱和物(如C16:1和/或C18:1)增加,则SAD的过表达或异源性SAD的表达可以与脂肪酰基去饱和酶(FAD)或另一种去饱和酶基因的沉默或失活(例如,通过突变、RNAi、反义、敲除内源去饱和酶基因,等等)偶联。

[0262] 在其他实施例中,该微藻以及其他微藻细胞已经被修饰而具有突变的内源去饱和酶基因,其中该突变致使该基因或去饱和酶无活性。在一些情况下,该突变的内源去饱和酶基因是脂肪酸去饱和酶(FAD)。在一些情况下,该突变的内源去饱和酶基因是硬脂酰酰基载体蛋白去饱和酶(SAD)。以下实例6描述了无绿藻属中的硬脂酰-ACP去饱和酶以及Δ12脂肪酸去饱和酶的靶向消融或敲除。实例6还描述了RNAi或反义构建体用来减少内源去饱和酶基因的表达的用途。

[0263] 在一些情况下,可能有利的是将一种或多种基因工程技术进行配对以获得产生所希望的脂质谱的转基因细胞。在一个实施例中,微藻或其他微藻细胞包含突变的内源去饱

和酶基因以及一个或多个外源基因。在一个非限制性实例中，具有突变的内源去饱和酶基因的微藻或其他微藻细胞还可以表达外源脂肪酰基-ACP硫酯酶基因和/或蔗糖转化酶基因。以下实例6描述了含有内源SAD的靶向消融或敲除并且还表达樟树C14偏好型硫酯酶(C14-preferring thioesterase)和蔗糖转化酶的转基因无绿藻细胞。在这种情况下，这种转基因无绿藻细胞产生一种与发现于动物脂中的脂质谱非常接近的脂质谱。动物脂典型地是衍生自炼制的牛脂或羊脂，在室温是固体并且利用于食品、化妆品、以及化学品工业中的多种应用中。动物脂的脂肪酸谱是：4％C14:0;26％C16:0;3％C16:1;14％C18:0;41％C18:1;3％C18:2;以及1％C18:3。如在以下实例6中所显示的，具有内源SAD的靶向消融或敲除并且表达樟树C14偏好型硫酯酶的转基因无绿藻细胞的克隆具有这样的脂质谱：少于1％的C12以及更短碳链长度的脂肪酸;2.74％至6.13％的C14:0;23.07％至25.69％的C16:0;7.02％至11.08％的C18:0;42.03％至51.21％的C18:1;以及9.37％至13.45％的C18:2(以面积百分数表示)。在一些情况下，这些转基因无绿藻细胞具有这样的脂质谱：3％-5％C14:0;25％-27％C16:0;10％-15％C18:0;以及40％-45％C18:1。

[0264] 在特定的实施例中，本发明的微生物经基因工程化以表达选自以下的一种或多种外源基因：酰基-ACP硫酯酶、酰基辅酶A/醛还原酶、脂肪酰基辅酶A还原酶、脂肪醛还原酶、脂肪醛脱羧酶、或天然共表达的酰基载体蛋白。上文描述了适合的关于表达脂肪酶基因的表达方法，除了其他方法之外，包括诱导型表达和区室化表达。脂肪酰基-ACP硫酯酶在脂质合成过程中从酰基载体蛋白(ACP)切下脂肪酸。通过进一步酶促加工，切下的脂肪酸然后与辅酶组合以产生酰基辅酶A分子。这种酰基辅酶A是产生醛的脂肪酰基辅酶A还原酶以及产生醇的脂肪酰基辅酶A/醛还原酶的酶活性的底物。如上文通过脂肪酰基辅酶A还原酶的作用所产生的醛是产生醇的脂肪醛还原酶或产生链烷或烯的脂肪醛脱羧酶的进一步酶活性的底物。

[0265] 在一些实施例中，通过在此所述的方法产生的脂肪酸、甘油酯或相应的伯醇、醛、链烷或烯含有8、10、12或14个碳原子。用于生产柴油、生物柴油、可再生柴油或喷气燃料、或用于工业应用的相应的伯醇、醛、链烷以及烯的优选脂肪酸含有8至14个碳原子。在某些实施例中，以上脂肪酸、以及其他相应的烃分子是饱和的(没有碳碳双键或三键)；单不饱和的(单个双键)；多不饱和的(两个或更多个双键)；是直链(不是环状的)或支链的。对于燃料生产而言，较大的饱和度是优选的。

[0266] 上文直接描述的酶具有用于水解含有特定碳原子数的底物的优先特异性。例如，脂肪酰基-ACP硫酯酶可以优先从ACP切下具有12个碳原子的脂肪酸。在一些实施例中，ACP和长度特异性硫酯酶可以彼此具有亲和力，从而使它们作为组合特别有用(例如，外源ACP和硫酯酶基因可以在它们自其衍生的特定组织或生物中天然共表达)。因此，在不同实施例中，本发明的重组无绿藻细胞可以含有编码蛋白质的外源基因，所述蛋白质就底物中所含碳原子的数目而言具有用于催化酶活性(例如，从ACP切下脂肪酸、将酰基辅酶A还原成醛或醇、或将醛转化成链烷)的特异性。在各种实施例中，酶特异性可以针对具有从8个至34个碳原子，优选地从8个至18个碳原子，并且更优选从8至14个碳原子的底物。优选特异性是针对具有较少(即，12)而不是较多(即18)碳原子的底物。

[0267] 适合于供本发明的微生物和方法使用的其他脂肪酰基-ACP硫酯酶包括而不限于在表4中列出的那些。

[0268] 表4.脂肪酰基-ACP硫酯酶和GenBank登录号

[0269]

加州月桂（Umbellularia californica）脂肪酰基-ACP *硫酯酶*(GenBank #AAC49001)
樟树（Cinnamomum camphora）脂肪酰基-ACP *硫酯酶*(GenBank #Q39473)
加州月桂脂肪酰基-ACP *硫酯酶*(GenBank #Q41635)
肉豆蔻（Myristica fragrans）脂肪酰基-ACP *硫酯酶*(GenBank #AAB71729)
肉豆蔻脂肪酰基-ACP *硫酯酶*(GenBank #AAB71730)
油棕（Elaeis guineensis）脂肪酰基-ACP *硫酯酶*(GenBank #ABD83939)
油棕脂肪酰基-ACP *硫酯酶*(GenBank #AAD42220)
油棕脂肪酰基-ACP *硫酯酶*(GenBank#AAD42220.2)
毛白杨(Populus tomentosa)脂肪酰基-ACP *硫酯酶*(GenBank #ABC47311)
拟南芥（Arabidopsis thaliana）脂肪酰基-ACP *硫酯酶*(GenBank #NP_172327)
拟南芥脂肪酰基-ACP *硫酯酶*(GenBank #CAA85387)
拟南芥脂肪酰基-ACP *硫酯酶*(GenBank #CAA85388)
陆地棉（Gossypium hirsutum）脂肪酰基-ACP 硫酯酶(GenBank #Q9SQI3)
披针叶萼距花（Cuphea lanceolata）脂肪酰基-ACP *硫酯酶*(GenBank #CAA54060)
萼距花（cuphea hookeriana）*脂肪酰基*-ACP *硫酯酶*(GenBank #AAC72882)
Cuphea calophylla subsp. mesostemon 脂肪酰基-ACP *硫酯酶*(GenBank #ABB71581)
披针叶萼距花脂肪酰基-ACP *硫酯酶*(GenBank #CAC19933)

[0270]

油棕脂肪酰基-ACP 硫酯酶（GenBank #AAL15645）
蓼距花脂肪酰基-ACP 硫酯酶(GenBank #Q39513)
蓼距花脂肪酰基-ACP 硫酯酶(GenBank #Q39513.1)
陆地棉脂肪酰基-ACP 硫酯酶(GenBank #AAD01982)
葡萄（*Vitis vinifera*）脂肪酰基-ACP 硫酯酶(GenBank #CAN81819)
山竹（*Garcinia mangostana*）脂肪酰基-ACP 硫酯酶(GenBank #AAB51525)
山竹脂肪酰基-ACP 硫酯酶(GenBank #AAB51525.1)
芥菜（*Brassica juncea*）脂肪酰基-ACP 硫酯酶(GenBank #ABI18986)
长叶马府油树（*Madhuca longifolia*）脂肪酰基-ACP 硫酯酶(GenBank #AAX51637)
甘蓝型油菜（*Brassica napus*）脂肪酰基-ACP 硫酯酶(GenBank #ABH11710)
甘蓝型油菜脂肪酰基-ACP 硫酯酶(GenBank #CAA52070.1)
水稻（籼型栽培品种组）脂肪酰基-ACP 硫酯酶(GenBank #EAY86877)
水稻（粳型栽培品种组）脂肪酰基-ACP 硫酯酶(GenBank #NP_001068400)
水稻（籼型栽培品种组）脂肪酰基-ACP 硫酯酶(GenBank #EAY99617)
蓼距花脂肪酰基-ACP 硫酯酶(GenBank #AAC49269)
美洲榆（*Ulmus Americana*）脂肪酰基-ACP 硫酯酶(GenBank #AAB71731)
披针叶蓼距花脂肪酰基-ACP 硫酯酶(GenBank #CAB60830)
湿地蓼距花（*Cuphea palustris*）脂肪酰基-ACP 硫酯酶(GenBank #AAC49180)
德国鸢尾（*Iris germanica*）脂肪酰基-ACP 硫酯酶(GenBank #AAG43858)
德国鸢尾脂肪酰基-ACP 硫酯酶(GenBank #AAG43858.1)
湿地蓼距花脂肪酰基-ACP 硫酯酶(GenBank #AAC49179)
肉豆蔻脂肪酰基-ACP 硫酯酶(GenBank# AAB71729)
肉豆蔻脂肪酰基-ACP 硫酯酶(GenBank# AAB717291.1)
蓼距花脂肪酰基-ACP 硫酯酶(GenBank #U39834)
加州月桂脂肪酰基-ACP 硫酯酶(GenBank # M94159)
樟树脂肪酰基-ACP 硫酯酶(GenBank #U31813)
蓖麻（*Ricinus communis*）脂肪酰基-ACP 硫酯酶(GenBank#ABS30422.1)

[0271] 以下实例描述了来自蓼距花、加州月桂、樟树、湿地蓼距花、披针叶蓼距花、德国鸢尾、肉豆蔻、山竹、油棕、甘蓝型油菜、蓖麻以及美洲榆的异源脂肪酰基-ACP硫酯酶在无绿藻种类中的成功靶向和表达。另外地，证实了在表达这些异源脂肪酰基-ACP硫酯酶的宿主细胞中的脂肪酸谱的改变。如在这些实例中所显示的，这些异源硫酯酶在无绿藻属中的表达产生了一种转基因微藻，该转基因微藻能够产生具有真正独特的脂肪酸谱的油/脂质，这样的油/脂质目前从商业种子作物是不可获得的，即使通过将数种种子作物油掺混。表5显示了常见商业种子油的脂肪酸谱。下文的所有商业种子油数据均从美国药典食品和化学品代码，第7版(2010-2011)编纂。动物脂数据来自美国国家研究委员会：动物产品的脂肪含量和

组成(1976)(National Research Council:Fat Content and Composition of Animal Products)。

[0272] 表5.商业种子油的脂质谱

[0273]

	C8:0	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:0-diOH	C18:1-OH	C18:2	C18:3 α
蓖麻（蓖麻油）	0	0	0	0	0.9-1.6	1.0-1.8	3.7-6.7	0.4-1.3	83.6-89.0	0	0.2-0.6
椰子(C. nucifera)（椰子油）	5.0-9.0	4.0-8.0	44-52	15-21	8.0-11.0	1.0-4.0	5.0-8.0	0	0	0-2.5	0
玉米(Z. mays)（玉米油）	0	0	0	< 1.0	8.0-19.0	0.5-4.0	19-50	0	0	38-65	< 2.0
海岛棉(G. barbadense)（棉籽油）	0	0	< 0.1	0.5-2.0	17-29	1.0-4.0	13-44	0	0	40-63	0.1-2.1
芜青(B. rapa)，甘蓝型油菜(B.napus)，芥菜(B. juncea)（卡诺拉）	0	0	< 0.1	< 0.2	< 6.0	< 2.5	> 50	0	0	< 40	< 14
油橄榄(O. europea)（橄榄）	0	0	0	< 0.1	6.5-20.0	0.5-5.0	56-85	0	0	3.5-20.0	< 1.2
落花生(A. hypogaea)（花生）	0	0	< 0.1	< 0.2	7.0-16.0	1.3-6.5	35-72	0	0	13.0-43	< 0.6
油棕（棕榈仁）	3.0-5.0	2.5-6.0	40-52	14.0-18.0	7.0-10.0	1.0-3.0	11.0-19.0	0	0	0.5-4.0	0
油棕（棕榈）	0	0	0	0.5-5.9	32.0-47.0	2.0-8.0	34-44	0	0	7.2-12.0	0
红花(C. tinctorus)（红花）	0	0	< 0.1	< 0.1	2.0-10.0	1.0-10.0	7.0-16.0	0	0	72-81	< 1.5
向日葵(H. annus)（向日葵）	0	0	< 0.1	< 0.5	3.0-10.0	1.0-10.0	14-65	0	0	20-75	< 0.5
大豆(G. max)（大豆）	0	0	< 0.1	< 0.5	7.0-12.0	2.0-5.5	19-30	0	0	48-65	5.0-10.0
亚麻(L. usitatissimum) 低亚麻酸亚麻（Solin-Flax）	0	0	< 0.1	< 0.5	2.0-9.0	2.0-5.0	8.0-60	0	0	40-80	< 5.0
乳油木(*B. parkii*) 乳油木（Sheanut）	0	0	0	0	3.8-4.1	41.2-56.8	34.0-46.9	0	0	3.7-6.5	0
动物脂				4	26	14	41			3	1

[0274] 作为实例，这些常见种子油没有一种含有高量的C8或C10脂肪酸，其中椰子油和棕榈仁油是最大的来源，但是两者具有大约1:1(C8:C10脂肪酸)的比率。如在这些实例中所显

示的，用湿地蓼距花C:8偏好型硫酯酶转化的无绿藻属能够不仅达到超过12％的C8脂肪酸水平，而且C8:C10脂肪酸的比率为大约5:1。在脂肪酸水平方面的改变对于生产含有用于多种商业应用的定制脂肪酸谱的油是有用的。另外地，在不同脂肪酸链长之间的比率的改变在未曾经历进一步的高成本的化学过程(如酯化、蒸馏、分提、以及再酯化)的油中在某种程度上是不曾在商业上可得的。作为另一个实例，棕榈油是含C16:0脂肪酸最高(32％-47％)的油，但是棕榈油具有极少的C14:0脂肪酸。含有美洲榆硫酯酶的无绿藻属达到了大约33％-38％的C16:0脂肪酸以及大约10％-16％的C14:0脂肪酸(大约2:1的C16:0与C14:0的比率)。通过掺混商业水平的现有的油实现这个脂肪酸谱在商业上不切实际的，因为这些高16:0脂肪酸的种子油通常不含有很多14:0脂肪酸。

[0275] 以下实例还描述了至少两种脂肪酰基-ACP硫酯酶在一个克隆中的成功靶向和表达。证实在了这些克隆中的脂肪酸谱的改变并且这些改变取决于哪两种硫酯酶在一个克隆中共表达，这些脂肪酸谱以不同方式受到影响。作为一个实例，根据以上表5，椰子油合棕榈仁油两者都具有大致3:1的C12:C14比率。如在以下实例中所描述的，含有两个异源硫酯酶基因的无绿藻转化子能够产生大致5:1比率的C12:C14脂肪酸水平。这种C12:C14脂肪酸比率曾经在商业上是不切实际的(即，通过将种子油掺混)。

[0276] 由转基因微藻产生的油的另一个新颖的方面是脂肪酸的饱和度。棕榈油目前是饱和油的最大来源，具有52％比48％的总饱和物比不饱和物。如在以下实例中所显示的，具有来自美洲榆和樟树的异源硫酯酶的无绿藻属在其产生的油中达到了超过60％的总饱和物水平。还在以下实例中显示，具有来自美洲榆的异源硫酯酶的无绿藻属在其产生的油中达到了超过86％的总饱和物水平。

[0277] 适合于供本发明的微生物和方法使用的脂肪酰基辅酶A/醛还原酶包括而不限于在表6中列出的那些。

[0278] 表6.通过GenBank登录号列出的脂肪酰基辅酶A/醛还原酶

[0279]

AAC45217、YP_047869、BAB85476、YP_001086217、YP_580344、YP_001280274、YP_264583、YP_436109、YP_959769、ZP_01736962、ZP_01900335、ZP_01892096、ZP_01103974、ZP_01915077、YP_924106、YP_130411、ZP_01222731、YP_550815、YP_983712、YP_001019688、YP_524762、YP_856798、ZP_01115500、YP_001141848、NP_336047、NP_216059、YP_882409、YP_706156、YP_001136150、YP_952365、ZP_01221833、YP_130076、NP_567936、AAR88762、ABK28586、NP_197634、CAD30694、NP_001063962、BAD46254、NP_001030809、EAZ10132、EAZ43639、EAZ07989、NP_001062488、CAB88537、NP_001052541、CAH66597、CAE02214、CAH66590、CAB88538、EAZ39844、AAZ06658、CAA68190、CAA52019、以及BAC84377

[0280] 适合于供本发明的微生物和方法使用的脂肪酰基辅酶A还原酶包括而不限于在表7中列出的那些。

[0281] 表7.通过GenBank登录号列出的脂肪酰基辅酶A还原酶

[0282]

NP_187805、ABO14927、NP_001049083、CAN83375、NP_191229、EAZ42242、EAZ06453、CAD30696、BAD31814、NP_190040、AAD38039、CAD30692、CAN81280、NP_197642、NP_190041、AAL15288、以及NP_190042

[0283] 适合于供本发明的微生物和方法使用的脂肪醛脱羧酶包括而不限于在表8中列出的那些。

[0284] 表8.通过GenBank登录号列出的脂肪醛脱羧酶

[0285]

NP_850932、ABN07985、CAN60676、AAC23640、CAA65199、AAC24373、CAE03390、ABD28319、NP_181306、EAZ31322、CAN63491、EAY94825、EAY86731、CAL55686、XP_001420263、EAZ23849、NP_200588、NP_001063227、CAN83072、AAR90847、以及AAR97643

[0286] 天然共表达的脂肪酰基-ACP硫酯酶以及酰基载体蛋白的组合适合于供本发明的微生物和方法使用。

[0287] 烃或脂质修饰酶的另外的实例包括多种氨基酸序列,这些氨基酸序列包含于、引用于或编码自包含于或引用于任何以下美国专利中的核酸序列中:6,610,527、6,451,576、6,429,014、6,342,380、6,265,639、6,194,185、6,114,160、6,083,731、6,043,072、5,994,114、5,891,697、5,871,988、6,265,639,并且进一步描述于以下GenBank登录号中:AAO18435、ZP_00513891、Q38710、AAK60613、AAK60610、AAK60611、NP_113747、CAB75874、AAK60612、AAF20201、BAA11024、AF205791以及CAA03710。

[0288] 在脂质生物合成途径中的其他酶也适合供本发明的微生物和方法使用。例如,酮脂酰-ACP合成酶(Kas)与一些以上列出的酶在脂质生物合成途径中共同起作用。存在不同类别的Kas酶:Kas I参与在不断增长的酰基ACP链与丙二酰-ACP之间的连续缩合步骤。Kas II典型地参与引导从C16:0-ACP到掺入丙二酰-ACP的C18:0-ACP的最终缩合步骤。像这样,在主要合成C16-C18:0脂肪酸(以及它们的不饱和衍生物)的高等植物以及一些微藻种类/株系中,Kas II酶与FatA基因的产物(酰基-ACP硫酯酶)相互作用。

[0289] 酰基-ACP在脂肪酸生物合成过程中从ACP释放增长的脂肪酸链,并且在大多数植物种类中,这是由FatA基因家族成员实施的,其中所述FatA基因家族成员的作用是在C16:0至C18:0阶段终止延伸。在合成较短链脂肪酸的种类(如蓖距花、油棕、肉豆蔻或月桂)中,由FatB基因编码的不同组的酰基-ACP硫酯酶实施这个终止步骤。在Kas II酶与酰基-Acp硫酯酶之间的相互作用对于脂肪酸链延长的正确终止是重要的。结果,在已经进化出能够生物合成较短链脂质的FatB基因的高等植物种类(以及微藻种类)中,存在着相应的另外类别的Kas基因(称作Kas IV基因)的共进化。Kas IV基因负责特定大小范围的脂肪酸(在长度上为4-14个碳)的链长延伸。

[0290] 用于供本发明的微生物和方法使用的其他适合的酶包括与表4、6-8中列出的蛋白质之一具有至少70%氨基酸一致性并且展现出相应的所希望的酶活性(例如,从酰基载体

蛋白切下脂肪酸、将酰基辅酶A还原成醛或醇、或将醛转化成链烷)的那些酶。在另外的实施例中,酶活性存在于与上述序列之一具有至少大约75%、至少大约80%、至少大约85%、至少大约90%、至少大约95%、或至少大约99%一致性的序列中,特此将所有序列通过引用进行结合,如同本文完全阐述。

[0291] 通过选择有待表达的外源基因的所希望的组合,人们可以定制由微生物产生的产物,然后可以从含水生物质提取这种产物。例如,这种微生物可以含有以下各项中的一种或多种:(i)编码脂肪酰基-ACP硫酯酶的外源基因;以及任选地,(ii)天然共表达的酰基载体蛋白或另外地对脂肪酰基-ACP硫酯酶具有亲和力的酰基载体蛋白(或相反地);以及任选地,(iii)编码脂肪酰基辅酶A/醛还原酶或脂肪酰基辅酶A还原酶的外源基因;以及任选地,(iv)编码脂肪醛还原酶或脂肪醛脱羧酶的外源基因。这种微生物还可以包含以下各项中的一种或多种:外源硬脂酰ACP去饱和酶、脂肪酸去饱和酶、β-酮脂酰-ACP合成酶I(例如,由KASI基因编码的)、β-酮脂酰-ACP合成酶II(例如,由KASII基因编码的)、或油酸-12羟化酶。这种微生物在此处所述的培养条件下合成连接至ACP的脂肪酸并且脂肪酰基-ACP硫酯酶催化从ACP切下脂肪酸,从而通过进一步的酶促加工而产生脂肪酰基辅酶A分子。当存在时,脂肪酰基辅酶A/醛还原酶催化酰基辅酶A还原为醇。同样,脂肪酰基辅酶A还原酶当存在时催化酰基辅酶A还原为醛。在其中一种编码脂肪酰基辅酶A还原酶的外源基因存在并且被表达以产生醛产物的那些实施例中,由第三外源基因编码的脂肪醛还原酶催化醛还原为醇。同样,脂肪醛脱羧酶当存在时催化醛转变为链烷或烯。

[0292] 在另一个实施例中,这种微生物可以含有:(i)编码脂肪酰基-ACP硫酯酶的外源基因;(ii)任选地,天然共表达的酰基载体蛋白或对脂肪酸酰基-ACP硫酯酶具有亲和力的酰基载体蛋白;(iii)突变的内源去饱和酶基因,其中所述突变使得去饱和酶基因或去饱和酶蛋白质无活性,如去饱和酶敲除或去饱和酶抑制元件如靶向RNAi、反义或dsRNA构建体;(iv)内源硬脂酰酰基载体蛋白去饱和酶过表达或异源SAD表达;和(v)前述情况的任何组合。

[0293] 编码这类酶(如脂肪酰基ACP硫酯酶)的基因可以从已经已知展现出显著脂质生产的细胞如原始小球藻获得。可以将已经已知在脂质产生中发挥作用的基因,例如,编码使双键饱和的酶的基因,单独地转化到受体细胞中。然而,为了实施本发明,没有必要做出关于需要何种基因的先验假设。在PCT公开号2008/151149中描述了用于鉴定可以改变(改进)微藻中的脂质产生的基因的方法。

[0294] 因此,本发明提供了已经被基因工程化从而与相同种类的野生型细胞相比以改变的水平表达脂质途径酶的无绿藻细胞。在一些情况下,当两种细胞均在相同条件下生长时,与野生型细胞相比,这种细胞产生更多的脂质。在一些情况下,细胞已经被基因工程化和/或被选择为与野生型细胞相比以更高水平或更低水平表达脂质途径酶。在一些情况下,脂质途径酶选自下组,该组由以下各项组成:丙酮酸脱氢酶、乙酰辅酶A羧化酶、酰基载体蛋白、以及甘油-3磷酸酰基转移酶。在一些情况下,细胞已经被基因工程化和/或被选择为与野生型细胞相比以更低的水平表达脂质途径酶。在细胞以较低水平表达脂质途径酶的至少一个实施例中,脂质途径酶包括柠檬酸合酶。

[0295] 在一些实施例中,细胞已经被基因工程化和/或被选择为与野生型细胞相比以改变的水平表达脂肪酸合成的全面调节子,由此与野生型细胞相比,多个脂肪酸合成基因的

表达水平被改变。在一些情况下，脂质途径酶包括修饰脂肪酸的酶。在一些情况下，脂质途径酶选自硬脂酰-ACP去饱和酶以及甘油脂去饱和酶。在一些情况下，细胞已经被基因工程化和/或被选择为表达较低水平的脂质途径酶或根本不表达特定的脂质途径酶(即，其中脂质途径酶已经被敲除或替换为外源基因，或者表达已经使用RNAi或反义方法而被降低)。在另一个实施例下，脂质途径酶是去饱和酶基因的异源表达，包括但不限于硬脂酰-ACP去饱和酶或脂肪酸去饱和酶(FAD)。实例6描述了来自油橄榄的异源硬脂酰-ACP在Prototheca moriformis遗传背景中的表达。

[0296] 在其他实施例中，本发明是针对含有一种或多种外源基因的产油微生物，其中所述外源基因编码选自下组的蛋白质，该组由以下各项组成：脂肪酰基-ACP硫酯酶、脂肪酰基辅酶A还原酶、脂肪醛还原酶、脂肪酰基辅酶A/醛还原酶、脂肪醛脱羧酶、去饱和酶、以及酰基载体蛋白。在另一个实施例中，内源去饱和酶基因在含有一个或多个以上外源基因的微生物中过表达。在一个实施例中，外源基因与一个启动子可操作地连接，所述启动子是响应于刺激可诱导或可抑制的。在一些情况下，这种刺激选自下组，该组由以下各项组成：外源提供的小分子、热、冷、以及在培养基中限制氮或无氮。在一些情况下，外源基因在细胞区室中表达。在一些实施例中，细胞区室选自叶绿体、质体和线粒体。在一些实施例中，微生物是Prototheca moriformis、Prototheca krugani、雍滞无绿藻或祖菲无绿藻。

[0297] 在一个实施例中，外源基因编码脂肪酸酰基-ACP硫酯酶。在一些情况下，由外源基因编码的硫酯酶催化从酰基载体蛋白(ACP)切下8至18碳脂肪酸。在一些情况下，由外源基因编码的硫酯酶催化从ACP切下10至14碳脂肪酸。在一个实施例中，由外源基因编码的硫酯酶催化从ACP切下12碳脂肪酸。

[0298] 在一个实施例中，外源基因编码脂肪酰基辅酶A/醛还原酶。在一些情况下，由外源基因编码的还原酶催化8至18碳脂肪酰基辅酶A还原成相应的伯醇。在一些情况下，由外源基因编码的还原酶催化10至14碳脂肪酰基辅酶A还原成相应的伯醇。在一个实施例中，由外源基因编码的还原酶催化12碳脂肪酰基辅酶A还原成相应的十二醇。

[0299] 本发明还提供了含有两种外源基因的重组无绿藻或其他细胞，其中第一外源基因编码脂肪酰基-ACP硫酯酶并且第二外源基因编码选自下组的蛋白质，该组由以下各项组成：脂肪酰基辅酶A还原酶、脂肪酰基辅酶A/醛还原酶、以及酰基载体蛋白。在一个实施例中，这两个外源基因各自与一个启动子可操作地连接，所述启动子是响应于刺激可诱导的。在一些情况下，每个启动子是响应于相同刺激(如在培养基中限制氮或无氮)可诱导的。培养基中限制氮或完全缺乏氮在一些微生物如无绿藻种类中刺激油产生，并且可以用作诱导油产生至高水平的触发物。当与在此披露的基因工程方法组合使用时，作为干细胞重量百分比的脂质可以被推向高水平，如至少30％、至少40％、至少50％、至少60％、至少70％、至少75％、至少80％、至少85％或在75％至90％之间；在此披露的方法为细胞提供了这些水平的脂质，其中该脂质是至少4％C8-C14、至少0.3％C8、至少2％C10、至少2％C12、以及至少2％C14。在一些实施例中这些细胞具有按干细胞重计超过25％的脂质并且含有至少为10％C8-C14、至少20％C8-C14、至少30％C8-C14、10％-30％C8-C14以及20％-30％C8-C14的脂质。

[0300] 在此披露的这些新颖的油与其他天然产生的中链脂肪酸含量高的油(如棕榈油、棕榈仁油、以及椰子油)不同。例如，在棕榈油和棕榈仁油中的污染物(如类胡萝卜素)水平

远远高于本发明的油中的污染物水平。棕榈油和棕榈仁油特别地含有α与β胡萝卜素和番茄红素，其量远远高于本发明的油中的α与β胡萝卜素和番茄红素。另外，在棕榈油和棕榈仁油中发现了超过20种的类胡萝卜素，而这些实例证明本发明的油含有很少的类胡萝卜素种类并且水平很低。另外，在棕榈油、棕榈仁油、以及椰子油中的维生素E化合物(如生育三烯酚)的水平远远高于在本发明的油中的水平。

[0301] 在一个实施例中，由该第一外源基因编码的硫酯酶催化从ACP切下8至18碳脂肪酸。在一些实施例中，该第二外源基因编码一种脂肪酰基辅酶A/醛还原酶，该脂肪酰基辅酶A/醛还原酶催化8-18碳脂肪酰基辅酶A还原为一种相应的伯醇。在一些情况下，由该第一外源基因编码的硫酯酶催化从ACP切下10-14碳脂肪酸，并且由该第二外源基因编码的还原酶催化10-14碳脂肪酰基辅酶A还原为相应的伯醇，其中该硫酯酶和该还原酶作用于相同的碳链长度。在一个实施例中，由该第一外源基因编码的硫酯酶催化从ACP切下12碳脂肪酸，并且由该第二外源基因编码的还原酶催化12碳脂肪酰基辅酶A还原为十二醇。在一些实施例中，该第二外源基因编码一种脂肪酰基辅酶A还原酶，该脂肪酰基辅酶A还原酶催化8-18碳脂肪酰基辅酶A还原为一种相应的醛。在一些实施例中，该第二外源基因编码一种与脂肪酰基-ACP硫酯酶天然共表达的酰基载体蛋白。

[0302] 在一些实施例中，该第二外源基因编码一种脂肪酰基辅酶A还原酶，并且该微生物进一步包含一种编码脂肪醛脱羧酶的第三外源基因。在一些情况下，由该第一外源基因编码的硫酯酶催化从ACP切下8-18碳脂肪酸，由该第二外源基因编码的还原酶催化8-18碳脂肪酰基辅酶A还原为相应的脂肪醛，并且由该第三外源基因编码的脱羧酶催化8-18碳脂肪醛转变为相应的链烷，其中该硫酯酶、该还原酶、以及该脱羧酶作用于相同的碳链长度。

[0303] 在一些实施例中，该第二外源基因编码一种酰基载体蛋白，并且该微生物进一步含有一种编码选自下组的蛋白质的第三外源基因，该组由以下各项组成：脂肪酰基辅酶A还原酶和脂肪酰基辅酶A/醛还原酶。在一些情况下，该第三外源基因编码一种脂肪酰基辅酶A还原酶，并且该微生物进一步含有一种编码脂肪醛脱羧酶的第四外源基因。

[0304] 本发明还提供了用于生产醇的方法，这些方法包括将重组无绿藻细胞群体在培养基中培养，其中这些细胞含有(i)一种编码脂肪酰基-ACP硫酯酶的第一外源基因，以及(ii)一种编码脂肪酰基辅酶A/醛还原酶的第二外源基因，并且这些细胞合成连接至酰基载体蛋白(ACP)的脂肪酸，该脂肪酰基-ACP硫酯酶催化从ACP切下该脂肪酸以通过进一步的加工而产生脂肪酰基辅酶A，并且该脂肪酰基辅酶A/醛还原酶催化该酰基辅酶A还原为醇。

[0305] 本发明还提供了在无绿藻细胞中产生脂质分子的方法。在一个实施例中，该方法包括将无绿藻细胞群体在培养基中培养，其中这些细胞含有(i)一种编码脂肪酰基-ACP硫酯酶的第一外源基因，以及(ii)一种编码脂肪酰基辅酶A还原酶的第二外源基因，并且这些微生物合成连接至酰基载体蛋白(ACP)的脂肪酸，该脂肪酰基-ACP硫酯酶催化从ACP切下该脂肪酸以通过进一步的加工而产生脂肪酰基辅酶A，并且该脂肪酰基辅酶A还原酶催化该酰基辅酶A还原为醛。

[0306] 本发明还提供了在无绿藻细胞中产生具有特定碳链长度的脂肪酸分子的方法。在一个实施例中，该方法包括将产生脂质的无绿藻细胞群体在培养基中培养，其中这些微生物含有编码一种脂肪酰基-ACP硫酯酶的外源基因，该脂肪酰基-ACP硫酯酶具有针对某一碳链长度(如8、10、12或14个碳原子)的特异性的或优先的活性，并且其中这些微生物合成连

接至酰基载体蛋白(ACP)的脂肪酸并且当该脂肪酸已经被合成到特定碳链长度时该硫酯酶催化从ACP切下该脂肪酸。

[0307] 在上文所述的各种实施例中,无绿藻细胞可以含有至少一个编码脂质途径酶的外源基因或抑制基因产物表达的抑制元件(如RNA干扰元件)。在一些情况下,该脂质途径酶选自下组,该组由以下各项组成:硬脂酰-ACP去饱和酶、甘油脂去饱和酶、丙酮酸脱氢酶、乙酰辅酶A羧化酶、酰基载体蛋白、以及甘油-3磷酸酰基转移酶。在其他情况下,无绿藻细胞含有选自下组的脂质修饰酶,该组由以下各项组成:脂肪酰基-ACP硫酯酶、脂肪酰基辅酶A/醛还原酶、脂肪酰基辅酶A还原酶、脂肪醛还原酶、脂肪醛脱羧酶、和/或酰基载体蛋白。

[0308] 本发明还提供了以一种含有异源基因的微生物细胞,该异源基因编码一种产生羟化脂肪酸的羟化酶。这种微生物细胞可以包含II型脂肪酸合成途径。例如,这种微生物细胞可以是一种微藻细胞。在一些实施例中,这种微藻细胞选自在以上表1中列出的微藻细胞。在其他实施例中,这种微藻细胞属于无绿藻属。仍然在其他实施例中,这种微藻细胞是Prototheca moriformis。羟化酶是将羟基基团(-OH)添加至底物上的酶类。脂肪酸羟化酶是在一些高等植物中发现的天然存在的酶。在高等植物中发现的天然存在的羟化酶的一个非限制性实例是负责油酸的产生的来自蓖麻的油酸12-羟化酶。实例7描述了在无绿藻细胞中的羟化酶的异源表达的一个实例,具体地是蓖麻油酸12-羟化酶在Prototheca moriformis细胞中的表达。

[0309] VI.燃料和化学品生产

[0310] 为了根据本发明的方法生产燃料,通过任何方便的手段收获或以别的方式收集由本发明的细胞产生的脂质。可以通过全细胞提取分离脂质。首先破坏细胞,并且然后可以例如通过使用如上文所述的离心从细胞团中分离细胞内脂质和细胞膜/细胞壁结合的脂质以及细胞外烃。在一些实施例中,在溶解微生物细胞之后提取产生在微生物中的细胞内脂质。一旦被提取,将这些脂质进一步精炼以便生产油、燃料、或油脂化学品。

[0311] 在完成培养之后,从发酵液中分离微生物。任选地,通过离心实现分离以产生浓缩膏。离心并不从这些微生物中去除显著量的细胞内水分并且并不是一个干燥步骤。这种生物质可以然后任选地用洗涤液(例如,蒸馏水)洗涤以除去发酵液和碎片。任选地,也可以在破坏细胞之前干燥(烘干、冻干,等等)洗涤的微生物生物质。可替代地,当发酵结束时,可以在不与一些或所有发酵液分开的情况下溶解细胞。例如,当溶解细胞时,细胞可以处于小于1:1v:v细胞对胞外液体的比率。

[0312] 可以溶解含有脂质的微生物以产生溶解产物。如在此详述,可以通过任何方便的手段实现溶解微生物(也称为细胞溶解)的步骤,所述手段包括热诱导的溶解、添加碱、添加酸、使用酶如蛋白酶和多糖降解酶(如淀粉酶)、使用超声波、机械溶解、使用渗透压休克、用溶解病毒感染和/或表达一个或多个溶解基因。进行溶解以释放已经由微生物产生的细胞内分子。这些用于溶解微生物的方法的每一种可以作为单一方法或以组合方式同时或依次地使用。可以通过显微镜分析来观察细胞破坏的程度。使用一种或多种在此所述的方法,典型地观察到超过70%的细胞破碎。优选地,细胞破裂超过80%、更优选超过90%并且最优选大约100%。

[0313] 在特定的实施例中,在生长之后溶解微生物,例如以便增加用于提取或进一步加工的细胞脂质和/或烃的暴露。可以调节脂肪酶表达(例如,通过诱导型启动子)或细胞溶解

的定时，以便优化脂质和/或烃的产生。下文描述众多溶解技术。这些技术可以单独或组合地使用。

[0314] 在本发明的一个实施例中，溶解微生物的步骤包括加热含有该微生物的细胞悬液。在这个实施例中，加热含有该微生物的发酵液(或从发酵液中分离的微生物悬液)直到微生物(即微生物的细胞壁和细胞膜)降解或分解。典型地，应用的温度是至少50℃。更高的温度，如至少30℃、至少60℃、至少70℃、至少80℃、至少90℃、至少100℃、至少110℃、至少120℃、至少130℃或更高用于更高效的细胞溶解。可以通过煮沸微生物通过热处理来溶解细胞。可替代地，可以在高压釜中进行热处理(不沸腾)。可以将热处理的溶解产物冷却，用于进一步处理。也可以通过蒸汽处理，即，通过添加加压蒸汽进行细胞破碎。例如在美国专利号6,750,048中描述了用于细胞破碎的微藻蒸汽处理。在一些实施例中，可以通过将蒸汽鼓泡到发酵罐中并且将培养液维持在所希望的温度持续小于大约90分钟、优选小于大约60分钟、并且更优选小于大约30分钟而实现蒸汽处理。

[0315] 在本发明的另一个实施例中，溶解微生物的步骤包括将碱添加至含有微生物的细胞悬液。该碱应当强到足以水解所使用的微生物的蛋白质化合物的至少一部分。用于增溶蛋白质的碱是化学领域已知的。用于本发明方法的示例性碱包括但不限于锂、钠、钾、钙的氢氧化物、碳酸盐和碳酸氢盐，及其混合物。优选的碱是KOH。例如在美国专利号6,750,048中描述了用于破碎细胞的微藻碱处理。

[0316] 在本发明的另一个实施例中，溶解微生物的步骤包括将酸添加至含有微生物的细胞悬液。可以使用浓度为10mN-500mN或优选40nM-160nM的酸实现酸溶解。优选在室温以上(例如，在40°-160°，并且优选是50°-130°的温度)进行酸溶解。对于中等温度(例如，室温至100℃，并且尤其是室温至65°)，酸处理可以有用地与超声处理或其他细胞破碎方法结合。

[0317] 在本发明的另一个实施例中，溶解微生物的步骤包括通过使用酶溶解微生物。用于溶解微生物的优选酶是蛋白酶和多糖降解酶，如半纤维素酶(例如，来自黑曲霉(Aspergillus niger)的半纤维素酶，Sigma Aldrich,St.Louis,MO,#H2125)；果胶酶(例如，来自根霉种类(Rhizopus sp.)的果胶酶，Sigma Aldrich,St.Louis,MO,#P2401)；Mannaway 4.0L(Novozymes)；纤维素酶(例如，来自绿色木霉(Trichoderma viride)的纤维素酶，Sigma Aldrich,St.Louis,MO,#C9422)；以及崩溃酶(例如，来自担子菌纲种类(Basidiomycetes sp.)的崩溃酶，Sigma Aldrich,St.Louis,MO,#D9515)。

[0318] 在本发明的其他实施例中，使用酶完成溶解，所述酶例如是纤维素酶如多糖降解酶，其任选地来自小球藻或小球藻病毒，或蛋白酶，如灰色链霉菌(Streptomyces griseus)蛋白酶、胰凝乳蛋白酶、蛋白酶K、在Degradation of Polylactide by Commercial Proteases(通过商业蛋白酶降解聚乳酸)，Oda Y等人，Journal of Polymers and the Environment(《聚合物与环境杂志》)，第8卷，第1期，2000年1月，第29-32页(4)中所列出的蛋白酶、Alcalase 2.4FG(Novozymes)、以及Flavourzyme 100L(Novozymes)。也可以使用蛋白酶和多糖降解酶的任何组合，包括前述蛋白酶和多糖降解酶的任何组合。

[0319] 在另一个实施例中，可以使用螺旋榨机进行溶解。在这个过程中，以高压迫使生物质穿过螺杆型装置，溶解细胞并引起细胞内脂质释放并且与细胞中的蛋白质和纤维(和其他组分)分开。

[0320] 在本发明的另一个实施例中，通过使用超声波(即，超声处理)进行溶解微生物的

步骤。因此，也可以用高频声波溶解细胞。可以电子地产生声波并且将其通过金属尖端传输至适当浓缩的细胞悬液。这种超声处理(或超声波处理)基于在细胞悬液中产生空腔而破坏细胞完整性。

[0321] 在本发明的另一个实施例中，通过使用机械溶解进行溶解微生物的步骤。可以将细胞机械地溶解并且任选地均质化以促进烃(例如，脂质)收集。例如，压力破碎仪可以用来泵送含有细胞的浆料穿过限流孔阀(restricted orifice valve)。施加高压(高达1500巴)，随后穿过出口喷嘴瞬间膨胀。通过三种不同的机制完成细胞破碎：冲击在阀门上、在孔口中的高液体剪切和释放时压力骤降，从而引起细胞爆裂。这种方法释放细胞内分子。可替代地，可以使用球磨机。在球磨机中，在悬浮液中将细胞与小磨粒(例如珠子)一起搅动。细胞因剪切力、在珠子之间的碾磨以及与珠子碰撞而破裂。珠子破坏细胞以释放细胞内容物。也可以通过剪切力破坏细胞，如使用共混(例如用高速或瓦林混碎机(Waring blender)作为实例)、弗氏压碎器、或在薄弱细胞壁的情况下甚至离心来破坏细胞。

[0322] 在本发明的另一个实施例中，通过应用渗透压休克进行溶解微生物的步骤。

[0323] 在本发明的另一个实施例中，溶解微生物的步骤包括用溶解病毒感染微生物。已知多种病毒溶解适合用于本发明中的微生物，并且针对特定的微生物选择和使用特定的溶解病毒是在本领域的技术水平范围之内的。例如，草履虫小球藻病毒(PBCV-1)是一组巨大的形成蚀斑的二十面体双链DNA病毒的原型(藻DNA病毒科(Phycodnaviridae)，绿藻病毒属(Chlorovirus))，这些病毒在某些单细胞真核小球藻样绿藻中复制并将其溶解。因此，可以通过用适合的小球藻病毒感染培养物溶解任何易感性微藻。用小球藻病毒感染小球藻种类的方法是已知的。参见例如Adv.Virus Res(《病毒研究进展》).2006;66:293-336;Virology(《病毒学》),1999Apr 25;257(1):15-23;Virology(《病毒学》),2004Jan 5;318(1):214-23;Nucleic Acids Symp.Ser(《核酸研讨会论文集丛刊》).2000;(44):161-2;J.Virol(《病毒杂志》).2006Mar;80(5):2437-44;以及Annu.Rev.Microbiol(《微生物学年度评论》).1999;53:447-94。

[0324] 在本发明的另一个实施例中，溶解微生物的步骤包括自溶。在这个实施例中，一种根据本发明的微生物经基因工程化以产生溶解蛋白，这种蛋白将溶解这种微生物。这种溶解基因可以使用诱导型启动子表达，使得可以首先使细胞在发酵罐中生长至所希望的密度，随后诱导启动子以表达溶解基因来溶解细胞。在一个实施例中，溶解基因编码多糖降解酶。在某些其他实施例中，溶解基因来自溶解病毒的基因。因此，例如，来自小球藻属病毒的溶解基因可以在藻细胞中表达；参见Virology(《病毒学》)260,308-315(1999)；FEMS Microbiology Letters(《FEMS微生物通讯》)180(1999)45-53;Virology(《病毒学》)263,376-387(1999)；和Virology(《病毒学》)230,361-368(1997)。优选地使用由刺激(如存在小分子、光、热和其他刺激)诱导的诱导型启动子，如微藻中有活性的启动子完成溶解基因的表达。

[0325] 多种方法可用于从通过以上方法产生的细胞溶解产物中分离脂质。例如，脂质和脂质衍生物如脂肪醛、脂肪醇和烃(如链烷)可以用疏水性溶剂如己烷提取(参见Frenz等人1989,Enzyme Microb.Technol.(《酶与微生物技术》),11:717)。脂质和脂质衍生物也可以使用液化(参见例如Sawayama等人1999,Biomass and Bioenergy(《生物质与生物能源》)17:33-39和Inoue等人1993,Biomass Bioenergy(《生物质与生物能源》)6(4):269-274)；油

液化(参见例如Minowa等人1995,Fuel(《燃料》)74(12):1735-1738);和超临界CO_2提取(参见例如Mendes等人2003,Inorganica Chimica Acta(《无机化学学报》)356:328-334)提取。Miao和Wu描述了从原始小球藻培养物回收微藻脂质的方案,其中将细胞通过离心收获、用蒸馏水洗涤并且通过冷冻干燥法干燥。在研钵中粉碎所得到的细胞粉末并且然后将其用正己烷提取。Miao和Wu,Biosource Technology(《生物资源技术》)(2006)97:841-846。

[0326] 因此,可以用有机溶剂提取回收由本发明的微生物产生的脂质、脂质衍生物和烃。在一些情况下,优选的有机溶剂是己烷。典型地,将有机溶剂直接添加至溶解产物,而不事先分离溶解产物组分。在一个实施例中,将通过上文所述的一种或多种方法产生的溶解产物与有机溶剂接触足以允许脂质和/或烃组分与有机溶剂形成溶液的一段时间。在一些情况下,可以然后进一步精炼溶液以回收特定的所希望的脂质或烃组分。己烷提取法是本领域熟知的。

[0327] 可以通过使用一种或多种酶(包括脂肪酶)修饰如在此所述的细胞产生的脂质和脂质衍生物,如脂肪醛、脂肪醇、以及烃(如链烷),如上所述。当烃处于细胞的细胞外环境中时,可以将一种或多种酶在酶修饰烃或从烃前体中完成其合成的条件下添加至这个环境。可替代地,在添加一种或多种催化剂(例如酶)之前,可以从细胞物质中部分或完全地分离烃。这样的催化剂是外源添加的,并且它们的活性在细胞外部或在体外发生。

[0328] 因此,如在此所述的在体内由细胞产生或在体外酶促修饰的脂质和烃可以任选地通过常规手段进一步加工。加工可以包括“裂解”以减小烃分子的大小,并且因此增加烃分子的氢:碳比。催化裂解和热裂解方法常规地用于烃和甘油三酯油加工中。催化方法涉及使用催化剂,如固体酸催化剂。这种催化剂可以是二氧化硅-氧化铝或沸石,这引起碳-碳键的异裂断裂或非对称断裂,产生碳正离子和氢负离子(hydride anion)。这些活性中间体然后经历重排或与另一种烃发生氢负离子转移。反应可以因此再生这些中间体以导致自增殖链机制(self-propagating chain mechanism)。也可以加工烃以减少其中的碳-碳双键或三键的数目,任选地减少至零。也可以加工烃以除去或消除其中的环或环状结构。也可以加工烃以增加氢:碳比。这可以包括加氢(“氢化”)和/或烃“裂解”成更小的烃。

[0329] 热方法包括使用升高的温度和压力来减小烃的大小。可以使用大约800℃的升高的温度以及大约700kPa的升高的压力。这些条件产生“轻质(light)”烃,这是一个有时用来指富氢烃分子的术语(与光子通量区别),同时还通过缩合产生更重质的氢相对耗尽的烃分子。该方法论提供了均裂断裂或对称断裂并且产生烯,其可以是任选地如以上所述酶促饱和的。

[0330] 在工厂中用于烃加工和炼油的催化方法和热方法是标准的。因此可以通过常规手段收集并且加工或精炼由在此所述的细胞产生的烃。参见Hillen等人(Biotechnology and Bioengineering(《生物技术与生物工程》),Vol.XXIV:193-205(1982))关于微藻产生的烃的加氢裂解的报道。在可替代的实施例中,将该馏分用另一种催化剂处理,如有机化合物、热、和/或无机化合物。为了将脂质加工为生物柴油,在这一节中如以下描述使用酯交换方法。

[0331] 通过本发明的方法产生的烃在多种工业应用中是有用的。例如,生产直链烷基苯磺酸盐(LAS)(一种在几乎所有类型的去污剂以及清洁用品中使用的阴离子表面活性剂)利用了通常包括10-14个碳原子链的烃。参见,例如,美国专利号6,946,430、5,506,201、6,

692,730、6,268,517、6,020,509、6,140,302、5,080,848、以及5,567,359。表面活性剂，例如LAS，可以用在个人护理组合物和去污剂的制造中，如描述于美国专利号5,942,479、6,086,903、5,833,999、6,468,955、以及6,407,044中的那些。

[0332] 日益增长的兴趣指向生物起源的烃组分在燃料(如生物柴油、可再生柴油、以及喷气燃料)中的用途，因为可以替代衍生自化石燃料的起始材料的可再生生物起始材料是可获得的，并且其使用是令人希望的。对用于从生物材料生产烃组分的方法存在迫切需要。本发明通过以下满足了这一需要：提供了使用通过在此描述的方法产生的脂质作为生物材料生产生物柴油、可再生柴油、以及喷气燃料的方法，从而生产生物柴油、可再生柴油、以及喷气燃料。

[0333] 传统的柴油燃料是富含石蜡烃的石油馏出物。它们具有宽至370°到780°F的沸程，适合在压燃式发动机中燃烧，如柴油发动机车辆。根据该沸程，连同可允许范围的其他燃料特性，如十六烷值、浊点、闪点、粘度、苯胺点、含硫量、含水量、灰分含量、铜片腐蚀、以及碳渣，美国试验与材料协会(American Society of Testing and Materials)(ASTM)建立了柴油的等级。技术上，衍生自生物质或以别的方式符合适当的ASTM规范的任何烃馏出物材料可以定义为柴油燃料(ASTM D975)、喷气燃料(ASTM D1655)，或如果它是脂肪酸甲酯则定义为生物柴油(ASTM D6751)。

[0334] 在提取之后，可以使从在此描述的微生物生物质回收的脂质和/或烃组分经受化学处理以制造在柴油车辆或喷气发动机中使用的燃料。

[0335] 生物柴油是一种颜色在金色与深棕色之间变化的液体，这取决于生产原料。它与水几乎不混溶，具有高沸点和低蒸汽压。生物柴油是指供在柴油发动机车辆中使用的等同柴油的加工燃料。生物柴油是可生物降解的并且是无毒的生物柴油胜过常规柴油燃料的一个另外的益处是较低的发动机磨损。典型地，生物柴油包含C14-C18烷基酯。多种方法将生物质或如在此描述地产生并分离的脂质转变为柴油燃料。生产生物柴油的一个优选方法是借助于如在此描述的脂质的酯交换反应。用于用作生物柴油的一种优选烷基酯是甲酯或乙酯。

[0336] 在大多数现有的柴油发动机车辆中，通过在此描述的方法生产的生物柴油可以单独使用或以任何浓度与常规柴油燃料掺混使用。当与常规柴油燃料(石化柴油)掺混使用时，生物柴油可以从大约0.1％至大约99.9％存在。世界上许多地方使用被称为“B”因子的系统来说明生物柴油在任何燃料混合物中的量。例如，含有20％生物柴油的燃料被标识为B20。纯生物柴油被称为B100。

[0337] 生物柴油还可以在家用或商用锅炉中用作加热燃料。现有的燃油锅炉可能含有橡胶部件并且可能需要转变以便靠生物柴油来运行。转变方法通常相对简单，包括将橡胶部件更换为合成部件，因为生物柴油是一种强溶剂。由于其强溶解能力，燃烧的生物柴油将增加锅炉的效率。生物柴油可以在柴油的配制品中用作添加剂以增加纯的超低硫柴油(ULSD)燃料的润滑性，超低硫柴油是有利的，因为它几乎不具有含硫量。生物柴油是一种比石化柴油更好的溶剂并且可以用来分解先前靠石化柴油运行的车辆的燃料管线中的残渣沉积物。

[0338] 生物柴油可以通过包含在富油生物质中的甘油三酯的酯交换来生产。因此，本发明的另一个方面提供了一种用于生产生物柴油的方法。在一个优选实施例中，用于生产生物柴油的方法包括以下步骤：(a)使用在此披露的方法培养含脂质微生物，(b)溶解含脂质

微生物以产生一种溶解产物，(c)从该溶解的微生物分离脂质，并且(d)将该脂质组合物酯交换，由此产生了生物柴油。用于微生物生长、溶解微生物以产生溶解产物、在包含有机溶剂的介质中处理该溶解产物以形成异质混合物以及将该处理的溶解产物分离为脂质组合物的方法已经描述于上文并且也可以在生产生物柴油的方法中使用。

[0339] 生物柴油的脂质谱通常与原料油的脂质谱高度类似。可以使通过这些方法提供的其他油以及本发明的组合物经受酯交换以产生具有多种脂质谱的生物柴油，这些脂质谱包括(a)至少4%C8-C14；(b)至少0.3%C8；(c)至少2%C10；(d)至少2%C12；以及(3)至少30%C8-C14。

[0340] 脂质组合物可以经受酯交换以产生作为生物柴油有用的长链脂肪酸酯。优选的酯交换反应概述于下文并且包括碱催化的酯交换和使用重组脂肪酶的酯交换。在碱催化的酯交换方法中，在碱性催化剂(典型地是氢氧化钾)的存在下使三酰甘油酯与醇(如甲醇或乙醇)反应。这一反应形成甲酯或乙酯以及作为副产物的丙三醇(甘油)。

[0341] 动物或植物油典型地是由甘油三酯构成的，甘油三酯是游离脂肪酸与三元醇甘油的酯。在酯交换中，用一种短链醇(如甲醇或乙醇)来替代三酰甘油酯(TAG)中的甘油。典型的反应方案如下：

[0342]



[0343] 在这一反应中，用碱使醇去质子化以使其成为一种更强的亲核体。通常，使用非常过量(高达50倍)的乙醇或甲醇。通常，这一反应将极慢地进行或不进行。可以使用热，以及酸或碱来帮助该反应更快地进行。该酸或碱不被该酯交换反应消耗，因此它们不是反应物而是催化剂。使用碱催化技术已经生产了几乎所有的生物柴油，因为它仅要求低温低压并且产生了超过98%的转换率(其条件是起始油水分低并且不含脂肪酸)。

[0344] 如以上所讨论，还已经使用一种酶(如脂肪酶)代替碱来进行酯交换。可以在例如室温至80℃之间的温度、以及大于1:1(优选大约3:1)的TAG与低级醇的摩尔比下进行脂肪酶催化的酯交换。适合在酯交换中使用的脂肪酶包括但不限于在表9中列出的那些。对于酯交换有用的脂肪酶的其他实例发现于例如美国专利号4,798,793、4,940,845、5,156,963、5,342,768、5,776,741以及WO89/01032中。这样的脂质肪酶包括但不限于由以下微生物产生的脂肪酶：根霉属、曲霉属、假丝酵母属、毛霉属、假单胞菌属、根毛霉属、假丝酵母属以及腐质霉属，以及胰脂肪酶。

[0345] 表9.适合供酯交换中使用的脂肪酶。

[0346]

*黑曲霉*脂肪酶 ABG73614、*南极假丝酵母（Candida antarctica）*脂肪酶 B

[0347]

（novozym-435） CAA83122、*圆柱形假丝酵母（Candida cylindracea）*脂肪酶AAR24090、*解脂假丝酵母（Candida lipolytica）*脂肪酶（脂肪酶L；天野制药株式会社（Amano Pharmaceutical Co., Ltd.））、*皱褶假丝酵母（Candida rugosa）*脂肪酶（例如，脂肪酶-OF；名糖产业株式会社（Meito Sangyo Co., Ltd.））、*米黑毛霉（Mucor miehei）*脂肪酶（Lipozyme IM 20）、*萤光假单胞菌（Pseudomonas fluorescens）*脂肪酶AAA25882、*日本根霉（Rhizopus japonicas）*脂肪酶（Lilipase A-10FG）Q7M4U7_1、*米黑根毛霉（Rhizomucor miehei）*脂肪酶B34959、*米根霉（Rhizopus oryzae）脂肪酶*（脂肪酶F） AAF32408、*粘质沙雷菌（Serratia marcescens）*脂肪酶（SM酶） ABI13521、*绵毛嗜热丝孢菌（Thermomyces lanuginosa）*脂肪酶CAB58509、脂肪酶P（长濑化成株式会社（Nagase ChemteX Corporation））、以及脂肪酶QLM（*名糖产业株式会社,名古屋,日本*）

[0348] 使用脂肪酶用于生产适合于生物柴油的脂肪酸酯的一个挑战是:脂肪酶的价格比强碱法所使用的氢氧化钠(NaOH)的价格高得多。已经通过使用可以再循环的固定化脂肪酶来解决这一挑战。然而,在循环持续一个最小的循环数之后必须维持固定化脂肪酶的活性,以允许基于脂肪酶的方法在生产成本方面与强碱法相竞争。固定化脂肪酶会遭受典型地在酯交换中使用的低级醇的毒害。美国专利号6,398,707(2002年6月4日颁布给Wu等人)描述了用于增强固定化脂肪酶活性以及再生具有降低活性的固定化脂肪酶的方法。一些适合的方法包括将固定化脂肪酶在具有不少于3的碳原子数目的醇中浸渍持续一段时间,优选从0.5-48小时,并且更优选从0.5-1.5小时。一些适合的方法包括将失活的固定化脂肪酶用具有不少于3的碳原子数目的醇进行洗涤并且然后将该失活的固定化脂肪酶浸渍在植物油中持续0.5-48小时。

[0349] 在特定实施例中,在产生重组脂肪酶作用于其上的脂质的相同微生物中表达该重组脂肪酶。适合的重组脂肪酶包括在表9中列出的那些和/或具有在以上表9中列出的GenBank登录号的那些,或一种与在以上表9中列出的脂肪酶之一具有至少70％氨基酸一致性并且展现出脂肪酶活性的多肽。在另外的实施例中,酶活性存在于与上述序列之一具有至少大约75％、至少大约80％、至少大约85％、至少大约90％、至少大约95％、或者至少大约99％一致性的序列中,特此将所有序列通过引用进行结合,如同本文完全阐述。编码脂肪酶和选择性标记的DNA优选是密码子优化的cDNA。再编码用于在微藻种类中表达的基因的方法描述于美国专利7,135,290中。

[0350] 针对生物柴油的国际通用标准是EN 14214。ASTM D6751是在美国和加拿大引用的最通用的生物柴油标准。德国使用DIN EN 14214而在英国要求遵循BS EN 14214。确定这些产品是否符合这些标准的基本工业测试典型地包括气相色谱、HPLC以及其他。符合质量标准的生物柴油是完全无毒的,具有大于50mL/kg的毒性度(LD_{50})。

[0351] 虽然符合ASTM标准的生物柴油必须是无毒的,但是可以存在倾向于作为沉积物从溶液中结晶和/或沉淀和沉降的污染物。当在较低温度下使用柴油时,沉积物形成尤其是个问题。沉积物或沉淀物可以引起问题,如降低燃料流动、阻塞燃料线管、阻塞过滤器等。具体

涉及去除生物柴油中的这些污染物和沉积物以便产生更高质量产品的方法是本领域熟知的。这样的方法的实例包括但不限于对油进行预处理以去除污染物(例如磷酸酯和游离脂肪酸)(例如脱胶、碱炼和二氧化硅吸附剂过滤)和冷过滤。冷过滤是一种特别研发以去除在生产之后存在于生物柴油中的任何微粒和沉积物的方法。这种方法冷却生物柴油并且过滤出当这种燃料在较低温度使用时可能形成的任何沉积物或沉淀物。这样一种方法在本领域内是熟知的并且描述于美国专利申请公开号2007-0175091中。适合的方法可以包括将生物柴油冷却至低于大约38℃的温度,使得杂质和污染物作为在生物柴油液体中的微粒沉淀出来。然后可以将硅藻土或其他过滤材料添加至冷却的生物柴油以形成浆料,然后可以将这种浆料通过压力叶(pressure leaf)滤机或其他类型的过滤器过滤以去除微粒。然后可以将过滤的生物柴油通过增泽过滤器(polish filter)以去除任何剩余的沉积物和硅藻土,从而产生最终的生物柴油产品。

[0352] 实例9描述了使用来自Prototheca moriformis的甘油三酯油生产生物柴油。根据ASTM D6751 A1方法的冷浸过滤性(Cold Soak Filterability),在实例9中生产的生物柴油是300ml体积120秒。这种测试包括过滤300ml的B100,冷却至40°F持续16小时,允许加温至室温,并且在真空下使用0.7微米的具有不锈钢支撑的玻璃纤维过滤器进行过滤。可以对本发明的油进行酯交换以产生具有少于120秒、少于100秒、以及少于90秒的冷浸时间的生物柴油。

[0353] 如果将在特别冷的温度中使用生物柴油,还可以使用后续过程。这样的过程包括冬化和分提。两种过程都被设计为通过降低浊点(生物柴油开始结晶的温度)来改进燃料的冷流和冬季性能。存在着冬化生物柴油的若干途径。一种途径是将生物柴油与石化柴油掺混。另一种途径是使用能够降低生物柴油的浊点的添加剂。另一种途径是通过在添加剂中混合并且允许饱和物结晶并且然后过滤出晶体而不加选择地去除饱和的甲酯。分提选择性地将甲酯分离成单独的组分或馏分,从而允许去除或者包含特定的甲酯。分提方法包括尿素分提、溶剂分提和热蒸馏。

[0354] 由本发明的方法提供的另一种有价值的燃料是可再生柴油,其包含链烷,如C10:0、C12:0、C14:0、C16:0以及C18:0并且因此是与生物柴油是可区别的。高质量的可再生柴油符合ASTM D975标准。由本发明的方法产生的脂质可以用作生产可再生柴油的原料。因此,本发明的另一个方面提供了一种用于生产可再生柴油的方法。可以通过至少三个过程来生产可再生柴油:氢热处理(氢化处理)、加氢处理、以及间接液化。这些方法产生非酯馏出物。在这些过程期间,如在此描述所产生和分离的三酰甘油酯被转化成链烷。

[0355] 在一个实施例中,用于生产可再生柴油的方法包括(a)使用在此披露的方法培养含脂质微生物,(b)将微生物溶解以产生溶解产物,(c)从该溶解的微生物分离脂质,以及(d)将该脂质脱氧和氢化处理以产生链烷,由此产生可再生柴油。适合于制造可再生柴油的脂质可以通过使用有机溶剂(如己烷)从微生物生物质提取,或通过其他方法(如描述于美国专利5,928,696中的那些)而获得。一些适合的方法可以包括机械压制和离心。

[0356] 在一些方法中,首先将微生物脂质裂解连同将其氢化处理以分别降低碳链长度度和减少饱和双键。然后将该材料异构化,也连同将其氢化处理。然后通过蒸馏去除石脑油(naptha)部分,随后另外地蒸馏以蒸发并且蒸馏柴油燃料中的所希望的组分以符合ASTM D975标准,同时留下比所希望的组分更重的组分以便符合D975标准。化学修饰油(包括甘油

三酯油)的氢化处理、加氢裂解、脱氧以及异构化方法在本领域内是熟知的。参见例如欧洲专利申请EP1741768(A1)、EP1741767(A1)、EP1682466(A1)、EP1640437(A1)、EP1681337(A1)、EP1795576(A1)、以及美国专利7,238,277、6,630,066、6,596,155、6,977,322、7,041,866、6,217,746、5,885,440、6,881,873。

[0357] 在用于生产可再生柴油的方法的一个实施例中,产生链烷的脂质处理是通过脂质组合物的氢化处理来进行的。在氢热处理中,典型地,生物质是在升高的温度和压力下在水中发生反应以形成油和残留固体。转化温度典型地是300°到660°F,伴随足以保持水主要为液态的压力,100到170标准大气压(atm)。反应时间是近似15分钟至30分钟。在反应完成之后,从水中分离有机物。由此产生了适合于柴油的馏出物。

[0358] 在制造可再生柴油的一些方法中,处理甘油三酯的第一步骤是加氢处理以使双键饱和,接着是在氢和催化剂存在下在升高的温度进行脱氧。在一些方法中,氢化和脱氧在同一个反应中发生。在其他方法中,脱氧在氢化之前发生。然后任选地进行异构化,同样是在氢和催化剂的存在下进行。优选通过蒸馏去除石脑油组分。例如,参见美国专利5,475,160(甘油三酯的氢化);5,091,116(脱氧、氢化以及脱气);6,391,815(氢化);以及5,888,947(异构化)。

[0359] 一种适合的用于甘油三酯的氢化的方法包括制备铜、锌、镁以及镧盐的水溶液以及另外的碱金属溶液,或优选碳酸铵溶液。这两种溶液可以被加热至大约20℃至大约85℃的温度并且以使得沉淀容器中的pH维持在5.5与7.5之间的速率一起计量进入该沉淀容器,以便形成一种催化剂。另外的水可以最初用于该沉淀容器中或者与该盐溶液和沉淀溶液同时添加。然后将得到的沉淀物充分洗涤、干燥、在大约300℃煅烧并且在从大约100℃至大约400℃的温度范围在氢中进行活化。然后可以在反应器中在上述催化剂的存在下使一种或多种甘油三酯与氢接触并反应。该反应器可以是滴流床反应器、固定床气固反应器、填充的鼓泡塔反应器、连续搅拌槽反应器、浆液反应器、或在本领域内已知的任何其他适合的反应器类型。该过程可以分批或以连续方式进行。反应温度典型地是处于从大约170℃至大约250℃的范围,同时反应压力典型地是处于从大约300psig至大约2000psig的范围。而且,在本发明的该过程中的氢与甘油三酯的摩尔比典型地是处于从大约20:1至大约700:1的范围。该过程典型地是在范围从大约0.1hr^{-1}至大约5hr^{-1}的重量时空速度(WHSV)下进行。本领域的普通技术人员将认识到,反应所需要的时间期限将根据所使用的温度、氢与甘油三酯的摩尔比、以及氢分压而变化。通过这样的氢化过程产生的产物包括脂肪醇、甘油、痕量的石蜡以及未反应的甘油三酯。典型地通过常规手段将这些产物分离,例如像蒸馏、萃取、过滤、结晶,等等。

[0360] 石油精炼厂使用加氢处理以便通过用氢处理进料来去除杂质。加氢处理转化温度典型地是300°至700°F。压力典型地是40至100个大气压。反应时间典型地是近似10分钟至60分钟。采用固体催化剂来增加某些反应速率、改进对于某些产物的选择性、以及优化氢消耗。

[0361] 适合的用于油的脱氧的方法包括将一种油加热至大约350°F至大约550°F范围的温度并且在至少范围从大气压至大气压以上的压力下使该加热的油连续地与氮接触持续至少大约5分钟。

[0362] 适合的用于异构化的方法包括使用碱异构化以及本领域内已知的其他油异构化。

[0363] 氢化处理和加氢处理最终导致甘油三酯进料的分子量的减小。甘油三酯分子在加氢处理条件下被还原为四个烃分子:一个丙烷分子以三个较重的烃分子,典型地在C8至C18的范围。

[0364] 因此,在一个实施例中,对本发明的脂质组合物进行的一种或多种化学反应的产物是包含ASTM D975可再生柴油的链烷混合物。通过微生物生产烃由Metzger等人Appl Microbiol Biotechnol(《应用微生物学与生物技术》)(2005)66:486-496以及A Look Back at the U.S.Department of Energy's Aquatic Species Program:Biodiesel from Algae(美国能源部水生生物计划回顾:来自藻类的生物柴油),NREL/TP-580-24190,John Sheehan,Terri Dunahay,John Benemann以及Paul Roessler(1998)综述。

[0365] 根据T10-T90(分别是在10%与90%蒸馏体积处的温度)描述柴油燃料的蒸馏特性。从Prototheca moriformis甘油三酯油生产可再生柴油并且将其描述于实例9中。在实例9中产生的材料的T10-T90是57.9℃。可以采用在此披露的油的氢化处理、异构化以及其他的共价修饰方法、以及在此披露的蒸馏和分提方法(如冷过滤),使用根据在此披露的方法生产的甘油三酯油来产生具有其他T10-T90范围(如20℃、25℃、30℃、35℃、40℃、45℃、50℃、60℃以及65℃)的可再生柴油组合物。

[0366] 在实例9中产生的材料的T10是242.1℃。可以采用在此披露的油的氢化处理、异构化以及其他的共价修饰方法、以及在此披露的蒸馏和分提方法(如冷过滤)来产生具有其他T10值(如在180与295之间、在190与270之间、在210与250之间、在225与245之间、以及至少290的T10)的可再生柴油组合物。

[0367] 在实例9中产生的材料的T90是300℃。可以采用在此披露的油的氢化处理、异构化以及其他的共价修饰方法、以及在此披露的蒸馏和分提方法(如冷过滤)来产生具有其他T90值(如在280与380之间、在290与360之间、在300与350之间、在310与340之间、以及至少290的T90)的可再生柴油组合物。

[0368] 在实例9中产生的材料的FBP是300℃。可以采用在此披露的油的氢化处理、异构化以及其他的共价修饰方法、以及在此披露的蒸馏和分提方法(如冷过滤)来产生具有其他FBP值(如在290与400之间、在300与385之间、在310与370之间、在315与360之间、以及至少300的FBP)的可再生柴油组合物。

[0369] 可以使通过本发明的方法和组合物提供的其他油经受氢化处理、异构化以及其他共价修饰的组合,这些油包括具有多种脂质谱的油,这些脂质谱包括(a)至少4%C8-C14;(b)至少0.3%C8;(c)至少2%C10;(d)至少2%C12;以及(3)至少30%C8-C14。

[0370] 可以通过两步法从任何形式的生物质生产传统的超低硫柴油。首先,将该生物质转化为一种合成气,合成气是一种富含氢和一氧化碳的气体混合物。然后,将该合成气催化转化为液体。典型地,使用费歇尔-托罗普希(Fischer-Tropsch)(FT)合成来完成液体的产生。这一技术适用于煤、天然气、以及重油。因此,在用于生产可再生柴油的方法的又另一个优选实施例中,产生链烷的脂质组合物处理是通过对该脂质组合物进行间接液化来进行的。

[0371] 本发明还提供了生产喷气燃料的方法。喷气燃料是澄明至稻草色的。最常见的燃料是一种基于无铅/石蜡油的、被分类为飞机A-1(Aeroplane A-1)的燃料,其按照一套国际标准化的规范生产。喷气燃料是大量的不同的烃的混合物,这些烃可能多达一千种或更多。

它们的大小范围(分子量或碳数)受到对产品的要求的限制,例如凝固点或烟点。煤油型(Kerosone-type)飞机燃料(包括喷气A(Jet A)以及喷气A-1(Jet A-1))具有在大于8与16碳数之间的碳数分布。宽馏分以及石脑油类型的飞机燃料(包括喷气B(Jet B))典型地具有在大约5与15个碳之间的碳数分布。

[0372] 两种飞机燃料(喷气A和喷气B)可以含有一些添加剂。有用的添加剂包括但不限于抗氧化剂、抗静电剂、腐蚀抑制剂、以及燃料系统结冰抑制剂(FSII)。抗氧化剂预防胶质化并且通常基于烷基化苯酚,例如AO-30、AO-31、或AO-37。抗静电剂消散静电并且预防发电花。Stadis 450是一个实例,其具有二壬基萘磺酸(DINNSA)作为活性成分。腐蚀抑制剂(例如DCI-4A)用于民用和军用燃料并且DCI-6A用于军用燃料。FSII剂包括例如Di-EGME。

[0373] 在本发明的一个实施例中,喷气燃料是通过将藻类燃料与现有喷气燃料掺混而生产的。由本发明的方法产生的这些脂质可以用作生产喷气燃料的原料。因此,本发明的另一个方面提供了一种用于生产喷气燃料的方法。同此提供了两种用于从这些通过本发明的方法产生的脂质来生产喷气燃料的方法:流化催化裂解(FCC)、以及加氢脱氧(HDO)。

[0374] 流化催化裂解(FCC)是一种用来从重质粗馏分生产烯烃尤其是丙烯的方法。通过本发明的方法生产的脂质可以转化为烯烃。该过程包括使脂质流动通过FCC区并且收集由烯烃组成的产物流,其作为喷气燃料是有用的。将所产生的脂质在裂解条件下与裂解催化剂相接触以提供作为喷气燃料有用的包含烯烃和烃的产物流。

[0375] 在一个实施例中,用于生产喷气燃料的方法包括(a)使用在此披露的方法培养含脂质微生物,(b)将该含脂质微生物溶解以产生溶解产物,(c)从该溶解产物分离脂质,并且(d)处理该脂质组合物,由此产生喷气燃料。在用于生产喷气燃料的方法的一个实施例中,可以使该脂质组合物流动通过一个流体催化裂解区,在一个实施例中,其包括使该脂质组合物在裂解条件下与裂解催化剂相接触以提供一种包含C_2-C_5烯烃的产物流。

[0376] 在这种方法的某些实施例中,可能令人希望的是去除任何可能存在于该脂质组合物中的污染物。因此,在使该脂质组合物流动通过液体催化裂解区之前,将该脂质组合物进行处理。预处理可以涉及使该脂质组合物与一种离子交换树脂相接触。该离子交换树脂是一种酸性离子交换树脂,如Amberlyst™-15,并且可以在反应器中用作该脂质组合物流动通过(向上流动或向下流动)的床。其他预处理可以包括通过使该脂质组合物与酸(如硫酸、乙酸、硝酸、或盐酸)相接触而进行的温和酸洗。接触通常是在环境温度和大气压用稀释酸溶液来完成的。

[0377] 任选地使预处理的脂质组合物流到FCC区,在此这些含烃组分被裂解为烯烃。催化裂解是通过使该脂质组合物在反应区内与由精细分散的微粒材料组成的催化剂相接触来完成的。与加氢裂解相反,该反应是催化裂解并且是在不添加氢或无氢消耗下进行的。随着该裂解反应的进行,大量的焦炭沉积在催化剂上。通过在一个再生区内燃烧来自该催化剂的焦炭在高温下再生该催化剂。含焦炭的催化剂(在此称为"积炭催化剂")被从反应区连续地输送至再生区以进行再生并且被来自再生区的基本上无焦炭的再生催化剂替代。不同气体流对这些催化剂微粒的流化允许催化剂在反应区与再生区之间输送。用于在流化的催化剂流中裂解烃(如在此描述的脂质组合物中的那些)、在反应区与再生区之间输送催化剂、以及在再生器中燃烧焦炭的方法对于FCC过程领域内的技术人员是熟知的。示例性的FCC应用以及对于裂解脂质组合物以产生C_2-C_5烯烃有用的催化剂描述于美国专利号6,538,169、

7,288,685中，通过引用以全文将其结合。

[0378] 适合的FCC催化剂通常包括至少至少两种组分，它们可以在或可以不在相同基质上。在一些实施例中，两种组分都可以循环遍及整个反应容器。第一组分通常包括在流化催化裂解领域内使用的任何熟知催化剂，如活性无定形粘土型催化剂和/或高活性结晶分子筛。分子筛催化剂可以是比无定形催化剂更优选的，这是因为它们的相对于所希望的产物的大大改进的选择性。在一些优选的实施例中，沸石可以用作FCC过程中的分子筛。优选地，该第一催化剂组分包括大孔沸石，如Y型沸石、活性氧化铝材料、粘合剂材料，包括二氧化硅或氧化铝以及惰性填料(如高岭土)。

[0379] 在一个实施例中，裂解本发明的脂质组合物发生在FCC区的提升管段(riser section)或可替代地提升器部分(lift section)。通过一个喷嘴将脂质组合物引入该提升管，导致脂质组合物的迅速蒸发。在接触催化剂之前，脂质组合物一般将具有大约149℃至大约316℃(300°F至600°F)的温度。催化剂从掺混容器流到提升管，在此它与脂质组合物接触大约2秒或更少的时间。

[0380] 然后将掺混的催化剂和反应的脂质组合物蒸气通过一个出口从提升管的顶部释放并且分离成包含烯烃的裂解产物蒸气流以及一批覆盖有大量焦炭并且通常被称为“积炭催化剂”的催化剂颗粒。在最小化脂质组合物与催化剂的接触时间的努力中(可以促进所希望的产物另外转化为不希望的其他产物)，可以使用任何分离器安排(如涡旋臂安排)以从产物流中迅速去除积炭催化剂。该分离器，例如涡旋臂分离器，被定位在一个腔的上部分，该腔具有一个位于该腔的下部分的汽提区(stripping zone)。通过该涡旋臂安排分离的催化剂落下到该汽提区中。包括裂解的烃(包括低碳烯烃)的裂解产物蒸气流和某一催化剂经由一个管道从该腔排出，该管道与旋风分离器连通。这些旋风分离器从该产物蒸气流去除剩余的催化剂颗粒以将颗粒浓度降低至极低水平。然后该产物蒸气流从分离容器的顶部排出。通过这些旋风分离器分离的催化剂返回到分离容器并且然后到达汽提区。汽提区通过与流的逆流接触将来自催化剂表面的吸附的烃去除。

[0381] 低的烃分压起作用以有利于低碳烯烃的产生。因此，提升管压力被设置在大约172至241kPa(25至35磅/平方英寸(绝对压力))，具有大约35至172kPa(5至25磅/平方英寸(绝对压力))的烃分压，具有优选的大约69至138kPa(10至20磅/平方英寸(绝对压力))的烃分压。这种相对低的烃分压是通过使用流作为稀释剂达到这种程度而实现的：该稀释剂是脂质组合物的10wt％至55wt％并且优选是脂质组合物的大约15wt％。其他稀释剂，如干气，可以用来达到等同的烃分压。

[0382] 在提升管出口的裂解流的温度将是大约510℃至621℃(950°F至1150°F)。然而，566℃(1050°F)以上的提升管出口温度产生更多的干气和更多的烯烃。然而，566℃(1050°F)以上的提升管出口温度产生更少的乙烯和丙烯。因此，在大约566℃至大约630℃的优选温度、在大约138kPa至大约240kPa(20至35磅/平方英寸(绝对压力))的优选压力下运行FCC过程是优选的。用于该过程的其他条件是催化剂与脂质组合物的比率，其可以从大约5至大约20并且优选是从大约10至大约15而变化。

[0383] 在用于生产喷气燃料的方法的一个实施例中，脂质组合物被引入FCC反应器的提升器部分。提升器部分的温度会非常热并且范围是从大约700℃(1292°F)至大约760℃(1400°F)，具有大约100至大约150的催化剂与脂质组合物比率。预期的是，将脂质组合物引

入提升器部分将产生大量的丙烯和乙烯。

[0384] 在使用脂质组合物或如在此描述所产生的脂质而生产喷气燃料的方法的另一个实施例中，该脂质组合物或脂质的结构是通过一种被称为加氢脱氧(HDO)的过程而破坏的。HDO表示借助于氢来去除氧，即，在破坏该材料的结构时氧被去除。烯属双键被氢化并且任何硫或氮化合物被去除。硫去除被称作加氢脱硫(HDS)。原料(脂质组合物或脂质)的预处理和纯化有助于催化剂的使用寿命。

[0385] 通常在HDO/HDS步骤中，将氢与原料(脂质组合物或脂质)混合并且然后使该混合物作为并向流、或者作为单相或两相原料通过催化剂床。在HDO/MDS步骤之后，将产物馏分分离并且传送到一个分开的异构化反应器。用于生物起始材料的异构化反应器作为一种并流反应器(co-current reactor)描述于文献(FI 100 248)中。

[0386] 用于通过氢化烃进料(例如，在此处的脂质组合物或脂质)而生产燃料的方法还可以通过以下来进行：使该脂质组合物或脂质作为与氢气的并流通过一个第一氢化区，并且其后通过使氢气作为相对于烃流出物的逆流经过一个第二氢化区而使该烃流出物进一步在该第二氢化区中氢化。示例性的HDO应用以及对于裂解脂质组合物以产生C_2-C_5烯烃有用的催化剂描述于美国专利号7,232,935中，通过引用以其全文进行结合。

[0387] 典型地，在加氢脱氧步骤中，生物学组分(如在此处的脂质组合物或脂质)的结构被分解，氧、氮、磷以及硫化合物、以及作为气体的轻烃被去除，并且烯属键被氢化。在该过程的第二个步骤中，即，在所谓的异构化步骤中，进行异构化以便支化烃链并且改进石蜡在低温下的性能。

[0388] 在该裂解过程的第一个步骤中，即HDO步骤，将氢气以及在此处的有待氢化的脂质组合物和脂质作为并流或逆流传送到HDO催化剂床系统，所述催化剂床系统包括一个或多个催化剂床，优选是1-3个催化剂床。该HDO步骤典型地是以并流方式运行的。在HDO催化剂床系统包括两个或更多个催化剂床的情况下，可以利用逆流原理对这些床中的一个或多个进行操作。在该HDO步骤中，压力在20巴与150巴之间、优选在50巴与100巴之间变化，并且温度是在200℃与500℃之间、优选在300℃-400℃的范围内变化。在该HDO步骤中，可以使用已知的、含有来自周期系统的VII族和/或VIB族金属的氢化催化剂。优选地，这些氢化催化剂是负载有Pd、Pt、Ni、NiMo或CoMo的催化剂，载体是氧化铝和/或二氧化硅。典型地，使用$NiMo/Al_2O_3$和$CoMo/Al_2O_3$催化剂。

[0389] 在该HDO步骤之前，可以任选地通过在较温和的条件下预氢化来处理在此处的脂质组合物和脂质，由此避免这些双键的副反应。这样的预氢化是在预氢化催化剂的存在下、在50℃-400℃的温度以及在1-200巴的氢压(优选在150℃与250℃之间的温度以及在10巴与100巴之间的氢压)进行的。该催化剂可以含有来自周期系统的VIII族和/或VIB族金属。优选地，该预氢化催化剂是负载有Pd、Pt、Ni、NiMo或CoMo的催化剂，载体是氧化铝和/或二氧化硅。

[0390] 将来自该HDO步骤的含有氢的气体流进行冷却并且然后从其中去除一氧化碳、二氧化碳、氮、磷以及硫化合物、气态轻烃以及其他杂质。在压缩之后，将经纯化的氢或再循环的氢返回至该第一催化剂床和/或这些催化剂床之间以补偿退出的气流。从凝聚的液体中去除水。将该液体传送到该第一催化剂床或这些催化剂床之间。

[0391] 在该HDO步骤之后，使产物经受异构化步骤。对于该过程而言重要的是在烃与异构

化催化剂接触之前尽可能完全去除杂质。该异构化步骤包括一个任选的汽提步骤，其中可以通过用水蒸气或适合的气体(如轻烃、氮或氢)进行汽提而纯化来自该HDO步骤的反应产物。该任选的汽提步骤是以逆流方式在异构化催化剂上游的一个单元中进行的，其中气体与液体彼此接触，或者在利用逆流原理的分开的汽提单元中的实际异构化反应器之前进行。

[0392] 在该汽提步骤之后，将氢气以及在此处的氢化脂质组合物或脂质、以及任选地n-石蜡混合物传送到包括一个或多个催化剂床的反应性异构化单元。该异构化步骤的这些催化剂床可以按并流或逆流方式进行操作。

[0393] 对于该过程而言重要的是逆流原理在该异构化步骤中适用。在该异构化步骤中，这是通过以逆流方式进行该任选的汽提步骤或该异构化反应步骤或两者来完成的。在该异构化步骤中，压力在20-150巴范围内、优选是在20-100巴范围内变化，温度是在200℃与500℃之间、优选是在300℃与400℃之间变化。在该异构化步骤中，可以使用在本领域内已知的异构化催化剂。适合的异构化催化剂含有分子筛和/或来自VII族的金属和/或载体。优选地，该异构化催化剂含有SAPO-11或SAPO41或ZSM-22或ZSM-23或镁碱沸石以及Pt、Pd或Ni以及Al_2O_3或SiO_2。典型的异构化催化剂是例如Pt/SAPO-11/Al_2O_3、Pt/ZSM-22/Al_2O_3、Pt/ZSM-23/Al_2O_3以及Pt/SAPO-11/SiO_2。该异构化步骤和该HDO步骤可以在同一个压力容器中或在分开的压力容器中进行。任选的预氢化可以在分开的压力容器中或在同一个压力容器中进行，如同HDO和异构化步骤。

[0394] 因此，在一个实施例中，一种或多种化学反应的产物是包含HRJ-5的链烷混合物。在另一个实施例中，一种或多种化学反应的产物是包含ASTM D1655喷气燃料的链烷混合物。在一些实施例中，符合ASTM 1655喷气燃料的规范的组合物具有低于10ppm的含硫量。在其他实施例中，符合ASTM 1655喷气燃料的规范的组合物具有低于205℃的蒸馏曲线的T10值。在另一个实施例中，符合ASTM 1655喷气燃料的规范的组合物具有低于300℃的终沸点(FBP)。在另一个实施例中，符合ASTM 1655喷气燃料的规范的组合物具有至少38℃的闪点。在另一个实施例中，符合ASTM 1655喷气燃料的规范的组合物具有在775K/M^3与840K/M^3之间的密度。在又另一个实施例中，符合ASTM 1655喷气燃料的规范的组合物具有低于-47℃的凝固点。在另一个实施例中，符合ASTM 1655喷气燃料的规范的组合物具有至少42.8MJ/K的燃烧热。在另一个实施例中，符合ASTM 1655喷气燃料的规范的组合物具有至少13.4质量%的氢含量。在另一个实施例中，符合ASTM 1655喷气燃料的规范的组合物具有在260℃的热稳定性，如通过定量重量分析JFTOT在低于3mm的Hg所测试的。在又另一个实施例中，符合ASTM 1655喷气燃料的规范的组合物具有低于7mg/dl的实际胶质。

[0395] 因此，本发明披露了多种方法，其中进行微藻脂质的化学修饰以产生在多种工业和其他应用中有用的产物。用于修饰由在此披露的方法产生的油的方法的实例包括但不限于，油的水解、油的加氢处理和油的酯化。微藻脂质的其他化学修饰包括而不限于包括而不限于环氧化、氧化、水解、硫酸盐化、磺化、乙氧基化、丙氧基化、酰胺化和皂化。微藻油的修饰产生基础油脂化学品，这些油脂化学品可以进一步修饰为针对所希望的功能而选择的衍生油脂化学品。以与关于燃料生产方法的以上描述类似的方式，也可以对从在此描述的微生物培养物中产生的油进行这些化学修饰。基础油脂化学品的实例包括但不限于肥皂、脂肪酸、脂肪酯、脂肪醇、脂族氮化合物、脂肪酸甲酯、以及甘油。衍生油脂化学品的实例包括

但不限于脂肪腈、酯类、二聚酸、季铵化合物、表面活性剂、脂肪酸烷醇酰胺、脂肪醇硫酸酯、树脂、乳化剂、脂肪醇、烯烃、钻井泥浆、多元醇、聚氨酯、聚丙烯酸酯、橡胶、蜡烛、化妆品、金属皂、肥皂、α-磺化甲酯、脂肪醇硫酸酯、脂肪醇乙氧基化物、脂肪醇醚硫酸盐、咪唑啉、表面活性剂、去污剂、酯类、季铵化合物、臭氧分解产物、脂肪胺、脂肪酸烷醇酰胺、乙氧基硫酸盐、甘油单酯、甘油二酯、甘油三酯(包括中链甘油三酯)、润滑剂、液压流体、润滑脂、介电流体、脱模剂、金属加工液、传热流体、其他功能性流体、工业化学品(例如,清洁剂、纺织加工助剂、增塑剂、稳定剂、添加剂)、表面涂层、涂料(paint)以及清漆(lacquer)、电气布线绝缘材料、以及较高级链烷。

[0396] 来自通过本发明的方法产生的甘油酯的脂肪酸组分的水解产生游离脂肪酸,这些脂肪酸可以被衍生化以产生其他有用的化学品。水解在水和催化剂(可以是酸或碱)的存在下发生。这些释放的游离脂肪酸可以被衍生化以产生多种产物,如在以下文献中报道的:美国专利号5,304,664(高度硫酸化的脂肪酸)、7,262,158(清洁组合物)、7,115,173(织物柔软剂组合物)、6,342,208(用于处理皮肤的乳剂)、7,264,886(拒水组合物)、6,924,333(涂料添加剂)、6,596,768(脂质富集的反刍动物原料)、以及6,380,410(用于去污剂和清洁剂的表面活性剂)。

[0397] 关于水解,在本发明的一个实施例中,甘油三酯油任选地首先在一种液体介质(如水或氢氧化钠)中水解以获得甘油和肥皂。存在着各种适合的甘油三酯水解方法,包括但不限于皂化、酸水解、碱水解、酶水解(在此称为裂解)、以及使用加压热水的水解。本领域内的技术人员将认识到,为了产生油脂化学品,甘油三酯油不需要水解;相反,这种油可以通过其他已知的过程直接转化为所希望的油脂化学品。例如,甘油三酯油可以通过酯化直接转化为脂肪酸甲酯。

[0398] 在一些实施例中,由在此披露的方法产生的油的催化水解通过将该油裂解为甘油和脂肪酸而发生。如以上所讨论的,然后可以通过一些其他修饰进一步加工这些脂肪酸以获得衍生油脂化学品。例如,在一个实施例中,这些脂肪酸可以经历胺化反应以产生脂族氮化合物。在另一个实施例中,这些脂肪酸可以经历臭氧分解以产生一元酸或二元酸。

[0399] 在其他实施例中,水解可以经由在此产生的油的裂解而发生,从而产生油脂化学品。在本发明的一些优选实施例中,甘油三酯油可以在其他过程之前裂解。本领域的技术人员将认识到,存在着许多适合的甘油三酯裂解方法,包括但不限于酶裂解以及加压裂解。

[0400] 通常,酶促的油裂解方法使用酶(脂肪酶)作为作用于水/油混合物的生物催化剂。然后酶裂解将油或脂肪分别裂解为甘油或游离脂肪酸。然后甘油迁移至水相中,而有机相富集有游离脂肪酸。

[0401] 这些酶裂解反应通常在有机相与水相之间的相界处发生,其中酶仅存在于相界处。然后遇到相界的甘油三酯促成或参与裂解反应。随着反应进行,与游离脂肪酸相比,仍然化学键合为甘油酯的脂肪酸的占有密度或浓度在相界处减少,使得反应减慢。在某些实施例中,酶裂解可以在室温发生。本领域的普通技术人员将知道适合的用于将油裂解为所希望的脂肪酸的条件。

[0402] 作为实例,可以通过增加界面边界表面来加速反应速度。一旦反应完成,则将游离脂肪酸从不含酶的有机相中分离,并且将仍然含有化学键合为甘油酯的脂肪酸的残余物进行反馈或再循环并与有待经受裂解的新鲜油或脂肪混合。以这种方式,然后将再循环的甘

油酯经受另外的酶裂解过程。在一些实施例中，以这样一种方式从部分裂解的油或脂肪中提取游离脂肪酸。以那种方式，如果将这些化学键合的脂肪酸(甘油三酯)返回或反馈进入该裂解过程，则酶消耗可以大幅减少。

[0403] 裂解程度测定为测量的酸值除以理论上可能的酸值的比率，对于一种给定的油或脂肪而言，这可以计算。优选地，酸值是根据通用标准方法通过滴定而测量的。可替代地，可以将含水甘油相的密度当作裂解程度的量度。

[0404] 在一个实施例中，在此描述的裂解过程还适合于裂解包含在来自这些产生的油的碱精炼过程的所谓的皂料中的甘油单酯、甘油二酯以及甘油三酯。以这种方式，可以将该皂料定量转化，无需事先将中性油皂化为脂肪酸。出于这个目的，优选地在裂解之前通过加入酸来释放化学键合在皂料中的脂肪酸。在某些实施例中，除了用于裂解过程的水和酶之外，使用一种缓冲溶液。

[0405] 在一个实施例中，还可以使根据本发明的方法所产生的油经受皂化作用(作为一种水解方法)。动物或植物油典型地是由三酰甘油(TAG)构成，三酰甘油是脂肪酸与三元醇(甘油)的酯。在碱水解反应中，TAG中的甘油被去除，留下三羧酸阴离子，这些三羧酸阴离子能与碱金属阳离子(如钠或钾)缔合而产生脂肪酸盐。在这种方案中，这些羧酸组分被从甘油部分上切割并且被用基基团替代。在该反应中使用的碱(例如，KOH)的量是由所希望的皂化度决定的。如果目的是例如生产一种含有一些最初存在于TAG组合物中的油类的肥皂产品，则将不足以把所有TAG转化为脂肪酸盐的碱量引入反应混合物中。通常，这种反应是在一种水溶液中进行的并且进展缓慢，但是可以通过加热而加快。可以通过向反应混合物中添加盐(如水溶性碱金属卤化物(例如，NaCl或KCl))来促进脂肪酸盐的沉淀。优选地，该碱是一种碱金属氢氧化物，如NaOH或KOH。可替代地，可以在该反应方案中使用其他碱，如烷醇胺，包括例如三乙醇胺以及氨甲基丙醇。在一些情况下，这些可替代物可以优先地产生一种透明肥皂产品。在一个实施例中，经受皂化的脂质组合物是如在此描述而产生的动物脂模拟物(即，类似于动物脂的脂质组合物)或动物脂模拟物与另一种甘油三酯油的掺混物。

[0406] 在一些方法中，化学修饰的第一步可以是加氢处理以使双键饱和，接着是在氢和催化剂存在下在升高的温度进行脱氧。在其他方法中，氢化和脱氧可以在同一反应中发生。仍然在其他方法中，脱氧在氢化之前发生。然后可以任选地进行异构化，同样在氢和催化剂的存在下进行。最后，如果希望的话，可以去除气体和石脑油。例如，参见美国专利5,475,160(甘油三酯的氢化)、5,091,116(脱氧、氢化以及脱气)、6,391,815(氢化)、以及5,888,947(异构化)。

[0407] 在本发明的一些实施例中，甘油三酯油被部分或完全脱氧。这些脱氧反应形成所希望的产物，包括但不限于脂肪酸、脂肪醇、多元醇、酮、以及醛。通常，不受任何特定的理论限制，这些脱氧反应涉及各种不同反应途径的组合，包括但不限于：氢解、氢化、连续的氢化-氢解、连续的氢解-氢化、以及组合的氢化-氢解反应，其导致氧从脂肪酸或脂肪酸酯中至少部分去除以产生可以通过进一步加工而容易地转化为所希望的化学品的反应产物，如脂肪醇。例如，在一个实施例中，可以通过FCC反应将脂肪醇转化为烯烃或通过缩合反应转化为高级链烷。

[0408] 一种这样的化学修饰是氢化，氢化是将氢加成至甘油酯的或游离脂肪酸的脂肪酸组分中的双键。氢化过程允许液体油转化成半固体或固体脂肪，其更适合于特定应用。

[0409] 由在此描述的方法产生的油的氢化可以与在此提供的一种或多种方法和/或材料结合进行，如在以下文献中报道：美国专利号7,288,278(食品添加剂或药剂)、5,346,724(润滑产品)、5,475,160(脂肪醇)、5,091,116(食用油)、6,808,737(用于人造黄油或涂抹食品的结构性脂肪)、5,298,637(卡路里降低的脂肪代用品)、6,391,815(氢化催化剂乙基以及硫吸附剂)、5,233,099和5,233,100(脂肪醇)、4,584,139(氢化催化剂)、6,057,375(泡沫抑制剂)、以及7,118,773(可食用乳液涂抹食品)。

[0410] 本领域的技术人员将认识到可以使用各种方法将碳水化合物氢化。一种适合的方法包括使碳水化合物与氢或与混合有适合的气体或催化剂的氢在氢化反应器中在足以形成氢化产物的条件下接触。氢化催化剂通常可以包括Cu、Re、Ni、Fe、Co、Ru、Pd、Rh、Pt、Os、Ir、以及其合金或任何组合，其可以是单独的或具有促进剂(如W、Mo、Au、Ag、Cr、Zn、Mn、Sn、B、P、Bi、以及其合金或任何组合)。其他有效的氢化催化材料包括负载型镍或用铼修饰的钌。在一个实施例中，取决于所希望的催化剂的官能度，该氢化催化剂还包括任何一种载体。可以通过本领域的普通技术人员已知的方法来制备这些氢化催化剂。

[0411] 在一些实施例中，氢化催化剂包括负载型VIII族金属催化剂和金属海绵材料(例如，海绵镍催化剂)。拉尼镍提供了一个适合于在本发明中使用的活化的海绵镍催化剂的实例。在其他实施例中，本发明的氢化反应是使用一种包含镍-铼催化剂或钨修饰的镍催化剂的催化剂来进行的。适合的用于本发明的氢化反应的催化剂的一个实例是碳负载型镍-铼催化剂。

[0412] 在一个实施例中，适合的拉尼镍催化剂可以通过以下来制备：用水性碱溶液(例如，含有大约25重量％的氢氧化钠)处理按重量计大致等量的镍和铝的合金。铝被水性碱溶液选择性地溶解，产生主要包含镍并带有少量铝的海绵状材料。最初的合金包括一定量的促进剂金属(即，钼或铬)，该量使得大约1重量％至2重量％保留在所形成的海绵镍催化剂中。在另一个实施例中，使用浸透适合的载体材料的在水中的亚硝酰硝酸钌(III)、氯化钌(III)溶液来制备该氢化催化剂。然后将该溶液干燥以形成一种具有低于按重量计大约1％的含水量的固体。然后可以将该固体在大气压、在氢气流中、在300℃(非煅烧的)或400℃(煅烧的)、在旋转球炉(rotary ball furnace)中还原4小时。在冷却并用氮气致使该催化剂具有惰性之后，使氮气中按体积计5％的氧在催化剂上流过持续2小时。

[0413] 在某些实施例中，描述的催化剂包括一种催化剂载体。该催化剂载体稳定并负载该催化剂。所使用的催化剂载体的类型取决于所选择的催化剂和反应条件。适合的用于本发明的载体包括但不限于碳、二氧化硅、二氧化硅-氧化铝、氧化锆、二氧化钛、二氧化铈、氧化钒、氮化物、一氮化硼、杂多酸、羟基磷灰石、氧化锌、氧化铬、沸石、碳纳米管、碳富勒烯以及其任何组合。

[0414] 可以使用本领域已知的那些常规方法来制备在本发明中使用的催化剂。适合的方法可以包括但不限于初期润湿、蒸发浸透、化学气相沉积、喷涂式包衣(wash-coating)、磁控溅射技术，等等。

[0415] 用于进行氢化反应的条件将根据起始材料的类型和所希望的产物而变化。得益于本披露，本领域的普通技术人员将认识到适当的反应条件。通常，氢化反应是在80℃至250℃、并且优选在90℃至200℃、并且最优选在100℃至150℃的温度下进行的。在一些实施例中，氢化反应是在500KPa至14000KPa的压力下进行的。

[0416] 在在本发明的氢解反应中使用的氢可以包括外部氢、再循环氢、原位产生的氢、以及其任何组合。如在此使用的，术语“外部氢”是指不是源自生物质反应本身而是从另一种来源加入到系统中的氢。

[0417] 在本发明的一些实施例中，令人希望的是将起始碳水化合物转化为一种较小的分子，该分子更容易转化为所希望的较高级烃。一种适合的用于这种转化的方法是借助于氢解反应。已知多种方法可以用于进行碳水化合物的氢解。一种适合的方法包括使碳水化合物与氢或与混合有适合的气体或氢解催化剂的氢在氢解反应器中在足以形成包含较小的分子或多元醇的反应产物的条件下接触。在此，术语“较小的分子或多元醇”包括具有较小的分子量的任何分子，这些分子可以包含比起始碳水化合物更少数目的碳原子或氧原子。在一个实施例中，这些反应产物包括包含多元醇和醇的较小的分子。本领域的普通技术人员将能够选择适当的藉此进行氢解反应的方法。

[0418] 在一些实施例中，可以使用一种氢解催化剂将5和/或6碳糖或糖醇转化为丙二醇、乙二醇、以及甘油。氢解催化剂可以包括Cr、Mo、W、Re、Mn、Cu、Cd、Fe、Co、Ni、Pt、Pd、Rh、Ru、Ir、Os、以及其合金或任何组合，其可以是单独的或具有促进剂(如Au、Ag、Cr、Zn、Mn、Sn、Bi、B、O、以及其合金或任何组合)。氢解催化剂还可以包括一种含有过渡金属(例如，铬、钼、钨、铼、锰、铜、镉)或VIII族金属(例如，铁、钴、镍、铂、钯、铑、钌、铱、以及锇)的含碳焦化聚合物催化剂(carbonaceous pyropolymer catalyst)。在某些实施例中，氢解催化剂可以包括与碱土金属氧化物结合的或附着至催化活性载体上的任何以上金属。在某些实施例中，在氢解反应中的所描述的催化剂可以包括一种如在以上描述的用于氢化反应的催化剂载体。

[0419] 用于进行氢解反应的条件将根据起始材料的类型和所希望产物而变化。得益于本披露，本领域的普通技术人员将认识到用来进行该反应的适当条件。通常，氢解反应是在110℃至300℃、并且优选在170℃至220℃、并且最优选在200℃至225℃的温度下进行的。在一些实施例中，氢解反应是在碱性条件下、优选在8至13的pH、并且甚至更优选在10至12的pH下进行的。在一些实施例中，氢解反应是在60KPa与16500KPa之间的范围、并且优选在1700KPa与14000KPa之间的范围、并且甚至更优选在4800KPa与11000KPa之间的范围的压力下进行的。

[0420] 在本发明的氢解反应中使用的氢可以包括外部氢、再循环氢、原位产生的氢、以及其任何组合。

[0421] 在一些实施例中，可以在缩合反应器中通过缩合反应将以上讨论的反应产物转化为高级烃。在这样的实施例中，这些反应产物的缩合在一种能够形成高级烃的催化剂的存在下发生。虽然不意图受理论限制，但是认为高级烃的产生通过逐步加成反应而进行，包括碳-碳或碳-氧键的形成。所得到的反应产物包括含有这些部分的任何数目的化合物，如在以下更详细描述。

[0422] 在某些实施例中，适合的缩合催化剂包括酸催化剂、碱催化剂、或酸/碱催化剂。如在此所使用的，术语“酸/碱催化剂”是指具有酸和碱官能度两者的催化剂。在一些实施例中，缩合催化剂可以包括但不限于沸石、碳化物、氮化物、氧化锆、氧化铝、二氧化硅、铝硅酸盐、磷酸盐、氧化钛、氧化锌、氧化钒、氧化镧、氧化钇、氧化钪、氧化镁、氧化铈、氧化钡、氧化钙、氢氧化物、杂多酸、无机酸、酸改性树脂、碱改性树脂、以及其任何组合。在一些实施例中，缩合催化剂还可以包括一种改性剂。适合的改性剂包括La、Y、Sc、P、B、Bi、Li、Na、K、Rb、

Cs、Mg、Ca、Sr、Ba、及其任何组合。在一些实施例中，缩合催化剂还可以包括一种金属。适合的金属包括Cu、Ag、Au、Pt、Ni、Fe、Co、Ru、Zn、Cd、Ga、In、Rh、Pd、Ir、Re、Mn、Cr、Mo、W、Sn、Os、合金、以及其任何组合。

[0423] 在某些实施例中，在缩合反应中所描述的催化剂可以包括一种如在以上描述的用于氢化反应的催化剂载体。在某些实施例中，缩合催化剂是自负载的。如在此所使用的，术语“自负载”表示该催化剂不需要另一种材料作为载体。在其他实施例中，该缩合催化剂与一种分开的适合于悬浮该催化剂的载体结合使用。在一个实施例中，该缩合催化剂载体是二氧化硅。

[0424] 缩合反应在其下发生的条件将根据起始材料的类型和所希望的产物而变化。得益于本披露，本领域的普通技术人员将认识到用来进行该反应的适当条件。在一些实施例中，该缩合反应是在对于所提出的反应的热力学是有利的一个温度下进行的。缩合反应的温度将根据特定的起始多元醇或醇而变化。在一些实施例中，缩合反应的温度是在从80℃至500℃、并且优选从125℃至450℃、并且最优选从125℃至250℃的范围内。在一些实施例中，缩合反应是在0KPa与9000KPa之间的范围、并且优选在0KPa与7000KPa之间的范围、并且甚至更优选在0KPa与5000KPa之间的范围的压力下进行的。

[0425] 由本发明形成的高级链烷包括但不限于具有从4至30个碳原子的支链或直链链烷、具有从4至30个碳原子的支链或直链烯、具有从5至30个碳原子的环烷、具有从5至30个碳原子的环烯、芳基、稠和芳基、醇、以及酮。适合的链烷包括但不限于丁烷、戊烷、戊烯、2-甲基丁烷、己烷、己烯、2-甲基戊烷、3-甲基戊烷、2,2,-二甲基丁烷、2,3-二甲基丁烷、庚烷、庚烯、辛烷、辛烯、2,2,4-三甲基戊烷、2,3-二甲基己烷、2,3,4-三甲基戊烷、2,3-二甲基戊烷、壬烷、壬烯、癸烷、癸烯、十一烷、十一烯、十二烷、十二烯、十三烷、十三烯、十四烷、十四烯、十五烷、十五烯、壬基癸烷、壬基癸烯、二十烷、二十烯、二十一烷、二十一烯、二十二烷、二十二烯、二十三烷、二十三烯、二十四烷、二十四烯、以及其异构体。这些产物中的一些可以适合用作燃料。

[0426] 在一些实施例中，这些环烷和环烯是未经取代的。在其他实施例中，这些环烷和环烯是单取代的。仍然在其他实施例中，这些环烷和环烯是多取代的。在包含取代的环烷和环烯的实施例中，这些取代的基团包括但不限于具有1至12个碳原子的支链或直链烷基、具有1至12个碳原子的支链或直链亚烷基、苯基、以及其任何组合。适合的环烷和环烯包括但不限于环戊烷、环戊烯、环己烷、环己烯、甲基-环戊烷、甲基-环戊烯、乙基-环戊烷、乙基-环戊烯、乙基-环己烷、乙基-环己烯、其异构体以及任何组合。

[0427] 在一些实施例中，所形成的芳基是未取代的。在另一个实施例中，所形成的芳基是单取代的。在包含取代的芳基的实施例中，这些取代的基团包括但不限于具有1至12个碳原子的支链或直链烷基、具有1至12个碳原子的支链或直链亚烷基、苯基、以及其任何组合。适合的用于本发明的芳基包括但不限于苯、甲苯、二甲苯、乙苯、对二甲苯、间二甲苯、以及其任何组合。

[0428] 在本发明中产生的醇具有从4至30个碳原子。在一些实施例中，这些醇是环状的。在其他实施例中，这些醇是支链的。在另一个实施例中，这些醇是直链的。适合的用于本发明的醇包括但不限于丁醇、戊醇、己醇、庚醇、辛醇、壬醇、癸醇、十一醇、十二醇、十三醇、十四醇、十五醇、十六醇、庚基癸醇、辛基癸醇、壬基癸醇、二十醇、二十一醇(uneicosanol)、二

十二醇(doeicosanol)、二十三醇(trieicosanol)、二十四醇(tetraeicosanol)、以及其异构体。

[0429] 本发明中产生的酮具有从4至30个碳原子。在一个实施例中,这些酮是环状的。在另一个实施例中,这些酮是支链的。在另一个实施例中,这些酮是直链的。适合的用于本发明的酮包括但不限于丁酮、戊酮、己酮、庚酮、辛酮、壬酮、癸酮、十一烷酮、十二烷酮、十三烷酮、十四烷酮、十五烷酮、十六烷酮、庚基癸酮、辛基癸酮、壬基癸酮、二十烷酮、二十一烷酮(uneicosanone)、二十二烷酮(doeicosanone)、二十三烷酮(trieicosanone)、二十四烷酮(tetraeicosanone)、以及其异构体。

[0430] 另一种这样的化学修饰是酯交换。天然产生的甘油脂不具有脂肪酸组分的均匀分布。在油的情况下,酯交换是指在不同甘油酯的两个酯之间的酰基的交换。该酯交换程序提供了一种机制,通过该机制可以将甘油酯混合物的脂肪酸组分重排以改变分布模式。酯交换是一种熟知的化学过程,并且通常包括在一种催化剂的存在下(如碱金属或碱金属烷基化物(例如,甲醇钠))将油的混合物加热(至大约200℃)持续一段时间(例如,30分钟)。这个过程可以用来使一种油混合物的脂肪酸组分的分布模式随机化,或者可以指向产生所希望的分布模式。可以对在此提供的材料(如具有至少20%细胞干重的脂质的微生物生物质)进行脂质的这种化学修饰方法。

[0431] 可以通过将油混合物维持在低于可能出现的一些TAG的熔点以下的温度来进行定向酯交换(其中寻求一种特定的脂肪酸分布模式)。这导致这些TAG的选择性结晶,从而有效地随着它们结晶而将其从反应混合物中去除。例如,该过程可以继续直到油中的大多数脂肪酸已经沉淀。定向酯交换可以用来例如通过用较短链对应物取代较长链脂肪酸来生产一种具有较低卡路里含量的产物。定向酯交换还可以用来生产一种具有脂肪混合物的产物,该产物可以提供在食品添加剂或产品(例如,人造黄油)中寻求的所希望的熔化特性和结构特征,无需借助于氢化,氢化可能产生不需要的反式异构体。

[0432] 由在此描述的方法产生的油的酯交换可以与一种或多种方法和/或材料结合进行,或产生如在以下文献中报道的产品:美国专利号6,080,853(不易消化的脂肪代用品)、4,288,378(花生酱稳定剂)、5,391,383(食用喷淋油(Edible spray oil))、6,022,577(用于食品的食用脂肪)、5,434,278(用于食品的食用脂肪)、5,268,192(低卡路里坚果产品)、5,258,197(卡路里降低的食用组合物)、4,335,156(食用脂肪产品)、7,288,278(食品添加剂或药剂)、7,115,760(分提方法)、6,808,737(结构脂肪)、5,888,947(发动机润滑剂)、5,686,131(食用油混合物)、以及4,603,188(固化的氨基甲酸乙酯组合物)。

[0433] 在根据本发明的一个实施例中,如以上描述的油的酯交换后继之为酯交换的产物与多元醇的反应(如在美国专利号6,465,642中报道)以产生多元醇脂肪酸聚酯。这样一种酯化与分离过程可以包括以下步骤:在肥皂的存在下使一种低级烷基酯与多元醇反应;从产物混合物中去除残余的肥皂;水洗并且干燥产物混合物以去除杂质;漂白产物混合物用于精炼;将至少一部分未反应的低级烷基酯与产物混合物中的多元醇脂肪酸聚酯分离;并且将分离的未反应的低级烷基酯进行再循环。

[0434] 还可以对具有短链脂肪酸酯的微生物生物质进行酯交换,如在美国专利6,278,006中报道的。通常,可以通过在一种适合的催化剂存在下将一种短链脂肪酸酯添加至一种油并且加热该混合物来进行酯交换。在一些实施例中,该油构成了按重量计该反应混合物

的大约5％至大约90％。在一些实施例中，该短链脂肪酸酯构成了按重量计该反应混合物的大约10％至大约50％。催化剂的非限制性实例包括碱催化剂；甲醇钠；酸催化剂，包括无机酸(如硫酸)与酸化粘土，有机酸，如甲磺酸、甲磺酸、以及甲苯磺酸，以及酸性树脂，如Amberlyst 15。金属(如钠与镁)、以及金属氢化物也是有用的催化剂。

[0435] 另一种这样的化学修饰是羟化，其包括将水加成至双键，从而导致饱和以及羟基部分的掺入。该羟化过程提供了一种用于将甘油脂的一种或多种脂肪酸组分转化为羟基脂肪酸的机制。羟化可以例如经由在美国专利号5,576,027中报道的方法来进行。羟化脂肪酸(包括蓖麻油及其衍生物)作为在若干工业应用(包括食品添加剂、表面活性剂、颜料润湿剂、消泡剂、防水添加剂、增塑剂、化妆品乳化和/或除臭剂)以及电子设备、药物、涂料、油墨、胶粘剂、以及润滑剂中的组分是有用的。可以如何进行甘油酯的羟化的一个实例如下：加热脂肪，优选与庚烷结合加热至大约30℃-50℃并且在此温度下保持三十分钟或更长；然后将乙酸添加至该混合物，接着添加硫酸水溶液，再接着以小的增量经过一小时添加过氧化氢水溶液至该混合物；在添加过氧化氢水溶液之后，然后可以将温度增加到至少大约60℃并且搅拌至少六小时；在搅拌之后，允许该混合物沉降并且可以将由该反应形成的下部水层去除，同时用具有大约60℃的温度的热水洗涤由该反应形成的上部庚烷层；然后可以将洗涤的庚烷层用氢氧化钾水溶液中和至大约5至7的pH并且然后在真空下通过蒸馏将其去除；然后可以将反应产物在100℃在真空下干燥并且将干燥的产物在真空条件下进行蒸汽除臭并且在大约50℃至60℃使用硅藻土过滤。

[0436] 由在此描述的方法产生的微生物油的羟化可以与一种或多种方法和/或材料结合进行，或生产如在以下文献中报道的产品：美国专利号6,590,113(油基涂层和油墨)、4,049,724(羟化过程)、6,113,971(橄榄油脂(Olive oil butter))、4,992,189(润滑剂和润滑油添加剂)、5,576,027(羟化奶)、以及6,869,597(化妆品)。蓖麻油酸的羟化提供了一种多元醇。

[0437] 羟化甘油酯可以被转化为交内酯。交内酯由其中羟化脂肪酸组分已经被酯化为另一种脂肪酸分子的甘油酯组成。羟化甘油酯至交内酯的转化可以通过以下来进行：加温甘油酯与脂肪酸的混合物并且将该混合物与一种矿物酸接触，如由Isbell等人，JAOCS 71(2):169-174(1994)所描述的。交内酯在多种应用中是有用的，包括但不限于在以下文献中报道的那些：美国专利号7,196,124(弹性材料与地板覆盖物)、5,458,795(用于高温应用的稠化油)、5,451,332(用于工业应用的流体)、5,427,704(燃料添加剂)、以及5,380,894(润滑剂、润滑脂、增塑剂以及印刷墨)。

[0438] 另一种这样的化学修饰链烯烃复分解。在烯烃复分解中，催化剂切断烯(烯烃)中的亚烷基碳并且通过使它们中的每一个与不同的亚烷基碳进行配对而形成新的烯。烯烃复分解反应提供了用于如以下过程的机制：通过乙烯醇分解截断烯上的不饱和脂肪酸烷基链，通过自我复分解经由烯键交联脂肪酸，以及通过与衍生烯烃的交叉复分解在脂肪酸上掺入新的官能团。

[0439] 结合其他反应，如酯交换和氢化，烯烃复分解可以将不饱和甘油酯转化为不同的终产物。这些产物包括用于蜡的甘油酯低聚物；用于润滑剂的短链甘油酯；用于化学品和聚合物的同-以及异-双功能烷基链；用于生物燃料的短链酯；以及用于喷气燃料的短链烃。可以对三酰甘油和脂肪酸衍生物进行烯烃复分解，例如使用在美国专利号7,119,216、美国专

利公开号2010/0160506、以及美国专利公开号2010/0145086中报道的催化剂和方法。

[0440] 生物油的链烯烃复分解通常包括在其他烯存在(交叉复分解)或不存在(自我复分解)下、在惰性条件下将Ru催化剂溶液以大约10至250ppm的装载量添加至不饱和脂肪酸酯。典型地允许这些反应进行从数小时至数日并且最终产生烯产物的一种分布。如何进行脂肪酸衍生物的烯烃复分解的一个实例如下:将第一代格拉布催化剂(二氯[2(1-甲基乙氧基-α-O)苯基]亚甲基-α-C](三环己基-膦))在甲苯中的溶液以222ppm的催化剂装载量添加至含有经脱气并干燥的油酸甲酯的容器中。然后将该容器用大约60磅/平方英寸(绝对压力)的乙烯气体进行加压并且在大约30℃或30℃以下维持3小时,由此产生大约50%产率的甲基9-癸烯酸酯。

[0441] 由在此描述的方法产生的油的烯烃复分解可以与一种或多种方法和/或材料结合进行,或生产如在以下文献中报道的产品:专利申请PCT/US07/081427(α-烯烃脂肪酸)以及美国专利申请号12/281,938(石油膏)、12/281,931(彩弹枪胶囊(paintball gun capsules))、12/653,742(增塑剂和润滑剂)、12/422,096(双功能有机化合物)、以及11/795,052(烛用蜡)。

[0442] 其他的可以对微生物油进行的化学反应包括使三酰甘油与环丙烷化剂(cyclopropanating agent)反应以增强流动性和/或氧化稳定性,如在美国专利6,051,539中报道;从三酰甘油制造蜡,如在美国专利6,770,104中报道;以及三酰甘油的环氧化,如在“脂肪酸组合物对环氧化的三酰甘油的酰化动力学的影响”("The effect of fatty acid composition on the acrylation kinetics of epoxidized triacylglycerols"),Journal of the American Oil Chemists'Society(美国油化学协会杂志),79:1,59-63,(2001)以及Free Radical Biology and Medicine(自由基生物学与医学),37:1,104-114(2004)中报道的。

[0443] 用于如以上描述的燃料和化学产品的含油微生物生物质的产生导致脱脂生物质粉的产生。脱脂粉是制备藻油的副产物并且作为农场动物(例如,反刍动物、家禽、猪以及水产养殖)的动物饲料是有用的。所得到的粉虽然具有降低的油含量,但仍然含有高品质的蛋白质、碳水化合物、纤维、灰分、残油以及其他适合用于动物饲料的营养素。由于这些细胞大部分被油分离过程溶解,该脱脂粉可以容易地被这些动物消化。脱脂粉可以任选地在动物饲料中与其他成分结合,如谷粒。由于脱脂粉具有粉状一致性(powdery consistency),可以使用可商业获得的挤出机或膨胀器或另一种类型的机器将它压制成小丸。

[0444] 蓖麻油是一种从蓖麻子中分离的天然存在的油。蓖麻油的水解产生蓖麻油酸。从蓖麻子中生产蓖麻油是困难的,因为蓖麻子含有大量的篦麻毒素。蓖麻毒素是一种在化学武器公约(Chemical Weapons Convention)中列为附表1化合物的极其危险的毒素。因此,在从蓖麻子中生产蓖麻油时必须非常小心。由本发明的一个实施例提供了一种从微藻细胞分离的羟化油。以这种方式,可以生产蓖麻油酸。在一个实施例中,该羟化油是一种羟化甘油三酯。本发明的羟化甘油三酯与蓖麻油可以是在化学上类似的。如在实例7中所显示的,本发明提供了一种羟化微生物油。当通过GC/MS进行分析时,实例7的油显示诸位发明人已经生产出了蓖麻油酸(12-羟基-9-顺式-十八碳烯酸)。

[0445] 根据本发明的一个实施例的一种脂肪酸是羟化脂肪酸。该羟化脂肪酸的一个实施例是蓖麻油酸。

[0446] 该微生物羟化油或羟化脂肪酸可以进一步被羟化。当蓖麻油酸被进一步羟化时，提供了一种含有两个羟基基团的脂肪酸(一种多元醇)。

[0447] 本发明提供了一种通过使一种多元醇(例如，羟化油和/或羟化脂肪酸)与一种含有异氰酸酯部分的化合物反应而制备的组合物。已经使用蓖麻油和异氰酸酯生产了聚氨酯。聚氨酯在我们当今使用的各种产品中普遍存在。聚氨酯存在于汽车、玩具、体育用品、消费电子产品、鞋、床垫、垫子、胶粘剂、建筑材料，等等。目前，由蓖麻油制成的聚氨酯可从巴斯夫公司(BASF)、伊藤制油公司(Itoh Oil)、以及其他公司商购。由羟化大豆油制得的聚氨酯可从嘉吉公司(Cargill)、陶氏化学公司(Dow)、拜耳公司(Bayer)、以及其他公司商购。

[0448] 在一个实施例中，可以根据本领域已知的方法将由微生物细胞产生的蓖麻油酸进一步加工成油脂化学品，包括蓖麻油酸酯、蓖麻油酸酰胺、聚氨酯、聚氨酯泡沫、或聚氨酯部分。参见，例如，美国专利号6194475、4266617、6403664、以及4058492、以及美国专利申请号20100227151。

[0449] 已经如上详述的本发明在以下实例中进行了举例说明，这些实例被提供为旨在说明而不是限制要求保护的发明。

[0450] VII.实例

[0451] 实例1:用于培养无绿藻的方法

[0452] 培养无绿藻株以实现按干细胞重量计高百分比的油。将冷冻的细胞在室温融化，并且将500μl细胞添加至加有2％葡萄糖的4.5ml的培养基(4.2g/L K_2HPO_4、3.1g/L NaH_2PO_4、0.24g/L $MgSO_4 \cdot 7H_2O$、0.25g/L柠檬酸一水合物、0.025g/L $CaCl_2 2H_2O$、2g/L酵母提取物)，并且在28℃在6孔平板中在搅拌(200转/分钟)下培养7天。通过以14,000转/分钟在预称重的Eppendorf管中将1ml的培养物离心5分钟，确定干细胞重量。弃去培养上清液并且将所得到的细胞沉淀用1ml的去离子水洗涤。将培养物再次离心，弃去上清液，并且将细胞沉淀置于-80℃直到冷冻。然后将样品冻干24小时并且计算干细胞重量。为了确定培养物中的总脂质，移出3ml培养物并且使用Ankom系统(Ankom公司，马其顿，纽约)根据制造商的方案进行分析。将样品用Amkom XT10提取器，根据制造商的操作方案进行溶剂提取。将总脂质确定为在酸水解的干燥样品与溶剂提取的干燥样品之间的质量差。在表10中显示了干细胞重量的油百分比测量值。

[0453] 表10.按干细胞重量计的油百分比

[0454]

种类	株	％油
雍滞无绿藻	UTEX 327	13.14
Prototheca moriformis	UTEX 1441	18.02
Prototheca moriformis	UTEX 1435	27.17

[0455] 将来自无绿藻属多个株的微藻样品进行基因分型。如下从藻生物质分离基因组DNA。以14,000x g持续5分钟将细胞(大约200mg)从液体培养物中离心出来。然后将细胞再悬浮于无菌蒸馏水中，以14,000x g离心5分钟，并且弃去上清液。将直径约2mm的单个玻璃珠添加至生物质，并且将管置于-80℃持续至少15分钟。移出样品并添加150μl研磨缓冲液(1％Sarkosyl，0.25M蔗糖，50mM NaCl，20mM EDTA，100mM Tris-HCl，pH 8.0，RNA酶A 0.5μg/μl)。通过短暂涡旋将沉淀再悬浮，随后添加40μl的5M NaCl。将样品短暂涡旋，随后添加

66μl的5％CTAB(十六烷基三甲基溴化铵)并且进行最终短暂涡旋。接下来,将样品在65℃孵育10分钟,此后将它们以14,000x g离心10分钟。将上清液转移至一支新管并且用300μl酚:氯仿:异戊醇(12:12:1)提取一次,随后以14,000x g离心5分钟。将所得到的水相转移至一支含有0.7体积异丙醇(约190μl)的新管,通过倒置进行混合并且在室温孵育30分钟或在4℃孵育过夜。通过以14,000x g离心10分钟回收DNA。然后将所得到的沉淀用70％乙醇洗涤2次,随后用100％乙醇最终洗涤。将沉淀在室温晾干20-30分钟,随后再悬浮于50μl的10mM TrisCl、1mM EDTA(pH 8.0)中。

[0456] 将如上文所述那样制备的5μl总藻DNA按1:50稀释于10mM Tris(pH 8.0)中。如下设置PCR反应,终体积20μl。将10μl 2x iProof HF主混合物(BIO-RAD)以10mM储存浓度添加至0.4μl引物SZ02613(5'-TGTTGAAGAATGAGCCGGCGAC-3'(SEQ ID NO:9))。这个引物序列范围在Gen Bank登录号L43357中的位置567-588并且在高等植物和藻类质体基因组中是高度保守的。随后添加至处于10mM储存浓度的0.4μl引物SZ02615(5'-CAGTGAGCTATTACGCACTC-3'(SEQ ID NO:10))。这个引物序列与Gen Bank登录号L43357中的位置1112-1093互补并且在高等植物和藻质体基因组中是高度保守的。接下来,添加5μl稀释的总DNA和3.2μl dH_2O。如下运行PCR反应:98℃,45";98℃,8";53℃,12";72℃,20",持续35个循环,随后72℃持续1分钟,并且保持在25℃。为了纯化PCR产物,将20μl的10mM Tris(pH 8.0)添加至每个反应,随后用40μl酚:氯仿:异戊醇(12:12:1)提取,涡旋并且以14,000x g离心5分钟。将PCR反应物施加至S-400柱(GE Healthcare)并且以3,000x g离心2分钟。将纯化的PCR产物随后TOPO克隆到PCR8/GW/TOPO中,并且在LB/Spec平板上选择阳性克隆。使用M13正向和反向引物,将纯化的质粒DNA按两个方向测序。总共选择十二个无绿藻株对其23S rRNA DNA测序,并且在序列表中列出这些序列。下文包括株的编号和序列表编号的简述。针对来自UTEX 1435(SEQ ID NO:15)序列的总趋异(overall divergence)进行序列分析。出现了作为最趋异的两对(UTEX 329/UTEX 1533和UTEX 329/UTEX 1440)。在两种情况下,成对比对导致75.0％的成对序列一致性。下文还包括针对UTEX 1435的序列一致性百分比:

[0457]

种类	株	% nt 一致性	SEQ ID NO.
Prototheca kruegani	UTEX 329	75.2	SEQ ID NO: 11
魏氏无绿藻	UTEX 1440	99	SEQ ID NO: 12
雍滞无绿藻	UTEX 1442	75.7	SEQ ID NO: 13
Prototheca moriformis	UTEX 288	75.4	SEQ ID NO: 14
Prototheca moriformis	UTEX 1439; 1441; 1435; 1437	100	SEQ ID NO: 15
魏氏无绿藻	UTEX 1533	99.8	SEQ ID NO: 16
Prototheca moriformis	UTEX 1434	75.9	SEQ ID NO: 17
祖菲无绿藻	UTEX 1438	75.7	SEQ ID NO: 18
Prototheca moriformis	UTEX 1436	88.9	SEQ ID NO: 19

[0458] 使用HPLC分析来自上文列出的株的亚组中的脂质样品的脂质谱。结果显示于表11中。可替代地,可以使用实例11中概述的程序来确定脂质谱。

[0459] 表11.无绿藻种类中脂质链的多样性

[0460]

株	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3	C20:0	C20:1
UTEX 327	0	12.01	0	0	50.33	17.14	0	0	0
UTEX 1441	1.41	29.44	0.70	3.05	57.72	12.37	0.97	0.33	0
UTEX 1435	1.09	25.77	0	2.75	54.01	11.90	2.44	0	0

[0461] 分析从Prototheca moriformis UTEX 1435(通过溶剂提取或使用螺旋榨机)提取的油的类胡萝卜素、叶绿素、生育酚、其他固醇和生育三烯酚。结果汇总于下表12中。

[0462] 表12.从Prototheca moriformis(UTEX 1435)提取的油中的类胡萝卜素、叶绿素、生育酚/固醇和生育三烯酚分析

	压榨的油 (mcg/ml)	溶剂提取的油 (mcg/ml)
顺式叶黄素	0.041	0.042
反式叶黄素	0.140	0.112
反式玉米黄质	0.045	0.039
顺式玉米黄质	0.007	0.013
t-α-隐黄质	0.007	0.010
t-β-隐黄质	0.009	0.010
t-α-胡萝卜素	0.003	0.001
c-α-胡萝卜素	未检测到	未检测到
t-β-胡萝卜素	0.010	0.009
9-顺式-β-胡萝卜素	0.004	0.002
番茄红素	未检测到	未检测到
总类胡萝卜素	**0.267**	**0.238**
叶绿素	<0.01 mg/kg	<0.01 mg/kg

[0463]

生育酚和固醇		
	压榨的油 (mg/100g)	溶剂提取的油 (mg/100g)
γ生育酚	0.49	0.49
菜油甾醇	6.09	6.05
豆甾醇	47.6	47.8
β-谷甾醇	11.6	11.5
其他固醇	445	446
生育三烯酚		
	压榨的油 (mg/g)	溶剂提取的油 (mg/g)
α生育三烯酚	0.26	0.26
β生育三烯酚	< 0.01	< 0.01
γ生育三烯酚	0.10	0.10
δ生育三烯酚	< 0.01	< 0.01
总生育三烯酚	**0.36**	**0.36**

[0464] 使用标准植物油加工方法精炼和漂白来自4个独立批次的从Prototheca moriformis提取的油。简而言之，在水平滗析器(horizontal decanter)中澄清从Prototheca moriformis提取的粗制油，其中将固形物与油分离。然后将澄清的油转移至含有柠檬酸和水的罐并且让其沉降大约24小时。在24小时之后，罐中的混合物形成2个分离层。底层由水和胶组成，然后通过滗析移除所述胶，之后将脱胶油转移至漂白罐中。然后将油与另一个剂量的柠檬酸一起加热。然后添加漂白粘土至漂白罐，并且将混合物在真空下进一步加热，以便蒸发掉存在的任何水。然后将混合物泵送通过叶滤机以移除漂白粘土。然后使过滤的油通过最终5μm增泽过滤器(polishing filter)并且然后收集储存直至使用。然后分析精炼和漂白的(RB)油的类胡萝卜素、叶绿素、固醇、生育三烯酚、和生育酚。在以下表13中汇总了这些分析的结果。“Nd”表示未检测到，并且下文列出检测灵敏度：

[0465] 检测灵敏度

[0466] 类胡萝卜素(mcg/g)nd＝<0.003mcg/g

[0467] 叶绿素(mcg/g)nd＝<0.03mcg/g

[0468] 固醇(%)nd＝0.25%

[0469] 生育酚(mcg/g);nd＝3mcg/g

[0470] 表13.来自精炼和漂白的Prototheca moriformis油的类胡萝卜素、叶绿素、固醇、生育三烯酚和生育酚分析

[0471]

	批次A		批次B	批次C	批次D
总类胡萝卜素（mcg/g）					
叶黄素	0.025		0.003	nd	0.039
玉米黄质	nd		nd	nd	nd
顺式叶黄素/玉米黄质	nd		nd	nd	nd
反式-α-隐黄质	nd		nd	nd	nd
反式-β-隐黄质	nd		nd	nd	nd
反式-α-胡萝卜素	nd		nd	nd	nd
顺式-α-胡萝卜素	nd		nd	nd	nd
反式-β-胡萝卜素	nd		nd	nd	nd
顺式-β-胡萝卜素	nd		nd	nd	nd
番茄红素	nd		nd	nd	nd
未鉴定	0.219		0.066	0.050	0.026
总类胡萝卜素	**0.244**		**0.069**	**0.050**	**0.065**
叶绿素（mcg/g）					
叶绿素A	0.268		0.136	0.045	0.166
叶绿素B	nd		nd	nd	nd
总叶绿素	**0.268**		**0.136**	**0.045**	**0.166**
固醇（%）					
菜子甾醇	nd		nd	nd	nd
菜油甾醇	nd		nd	nd	nd
豆甾醇	nd		nd	nd	nd
β-谷甾醇	nd		nd	nd	nd
总固醇	nd		nd	nd	nd
生育酚（mcg/g）					
α-生育酚	23.9		22.8	12.5	8.2
β-生育酚	3.72		nd	nd	nd
γ-生育酚	164		85.3	43.1	38.3
δ-生育酚	70.1		31.1	18.1	14.3
总生育酚	**262**		**139.2**	**73.7**	**60.8**
生育三烯酚（mcg/g）					
α-生育三烯酚	190		225	253	239
β-生育三烯酚	nd		nd	nd	nd
γ-生育三烯酚	47.3		60.4	54.8	60.9
δ-生育三烯酚	12.3		16.1	17.5	15.2
总生育三烯酚	**250**		**302**	**325**	**315**

[0472] 还分析了相同4个批次Prototheca moriformis油的微量元素,并且将结果汇总在以下表14中。

[0473] 表14.精炼和漂白的Prototheca moriformis油的元素分析

[0474]

	批次 A	批次 B	批次 C	批次 D
元素分析（ppm）				
钙	0.08	0.07	< 0.04	0.07
磷	< 0.2	0.38	< 0.2	0.33
钠	< 0.5	0.55	< 0.5	< 0.5
钾	1.02	1.68	< 0.5	0.94
镁	< 0.04	< 0.04	< 0.04	0.07
锰	< 0.05	< 0.05	< 0.05	< 0.05
铁	< 0.02	< 0.02	< 0.02	< 0.02
锌	< 0.02	< 0.02	< 0.02	< 0.02
铜	< 0.05	< 0.05	< 0.05	< 0.05
硫	2.55	4.45	2.36	4.55
铅	< 0.2	< 0.2	< 0.2	< 0.2
硅	0.37	0.41	0.26	0.26
镍	< 0.2	< 0.2	< 0.2	< 0.2
有机氯化物	< 1.0	< 1.0	< 1.0	2.2
无机氯化物	< 1.0	< 1.0	< 1.0	< 1.0
氮	4.4	7.8	4.2	6.9
锂	< 0.02	< 0.02	< 0.02	< 0.02
硼	0.07	0.36	0.09	0.38
铝	--	< 0.2	< 0.2	< 0.2
钒	< 0.05	< 0.05	< 0.05	< 0.05
洛维邦得（Lovibond）色（°L）				
红	5.0	4.3	3.2	5.0
黄	70.0	70.0	50.0	70.0
经 HPLC 的甘油单酯和甘油二酯（%）				
甘油二酯	1.68	2.23	1.25	1.61
甘油单酯	0.03	0.04	0.02	0.03
游离脂肪酸(FFA)	1.02	1.72	0.86	0.83
肥皂	0	0	0	
氧化和聚合的甘油三酯				
氧化甘油三酯（%）	3.41	2.41	4.11	1.00
聚合甘油三酯（%）	1.19	0.45	0.66	0.31
过氧化值（meg/kg）	0.75	0.80	0.60	1.20
p-茴香胺值(无量纲的)	5.03	9.03	5.44	20.1
水和其他杂质（%）				
卡尔-费休（Karl Fisher）水分	0.8	0.12	0.07	0.18
总极性化合物	5.02	6.28	4.54	5.23
不可皂化物	0.92	1.07	0.72	1.04

[0475]

不溶性杂质	< 0.01	< 0.01	0.01	< 0.01
总油（%）				
中性油	98.8	98.2	99.0	98.9

[0476] 实例2:用于生物射弹转化无绿藻的一般方法

[0477] 根据来自制造商的方案制备Seashell金微载体550纳米。质粒(20μg)与50μl的结合缓冲液和60μl(30mg)的S550d金载体混合并且在冰中孵育1分钟。添加沉淀缓冲液(100μl),并且将混合物在冰中孵育另外1分钟。在涡旋之后,通过在Eppendorf 5415C微量离心机中以10,000转/分钟旋转10秒,使DNA包被的颗粒沉淀。将金沉淀物用500μl冷100％乙醇洗涤一次,通过在离心管中短暂旋转沉淀,并且用50μl冰冷的乙醇再悬浮。在短暂(1-2秒)超声处理之后,将10μl DNA包被的颗粒立即转移至载体膜。

[0478] 将无绿藻株在转动摇床上含有2％葡萄糖的月示培养基(2g/L酵母提取物,2.94mM NaNO3,0.17mM CaCl2·2H2O,0.3mM MgSO4·7H2O,0.4mM K2HPO4,1.28mM KH2PO4,0.43mM NaCl)中培养直到它达到2x 10^6个细胞的细胞密度。收获细胞,用无菌蒸馏水洗涤一次并且再悬浮于50μl的培养基中。将1x 10^7个细胞铺展在非选择性月示培养基平板的中心三分之一(center third)。用PDS-1000/He生物射弹粒子递送体系(Bio-Rad)轰击细胞。使用裂碟片(1350psi),并且将平板置于筛选/巨载体组件下方6cm。允许细胞在25℃恢复12-24小时。回收时,将细胞用橡胶抹刀从板上刮下,与100μl培养基混合并且铺展在含有适当抗生素选择的平板上。在25℃孵育7-10天之后,在平板上可见代表转化细胞的集落。将集落挑出并点在选择性(抗生素或碳源)琼脂平板上用于第二轮选择。

[0479] 实例3:在无绿藻属中的各种硫酯酶的表达

[0480] 先前已经在PCT申请号PCT/US2009/066412中描述了用于在无绿藻种类中表达异源硫酯酶基因的方法和作用,特此将该专利通过引用进行结合。进一步研究了来自高等植物种类的其他硫酯酶基因/基因产物的作用。这些硫酯酶包括来自以下高等植物的硫酯酶:

[0481]

种类	GenBank 登录号	特异性	SEQ ID NO:
樟树	Q39473	C14	SEQ ID NOs: 30-31
加州月桂	Q41635	C10-C12	SEQ ID NOs: 34-35
荨距花	AAC49269	C8-C10	SEQ ID NOs: 32-33
湿地荨距花	AAC49179	C8	SEQ ID NOs: 36-37
披针叶荨距花	CAB60830	C10	SEQ ID NOs: 38-39
德国鸢尾	AAG43858.1	C14	SEQ ID NOs: 40-41
肉豆蔻	AAB717291.1	C14	SEQ ID NOs: 42-43
湿地荨距花	AAC49180	C14	SEQ ID NOs: 44-45
美洲榆	AAB71731	广谱	SEQ ID NOs: 46-47

[0482]

肉豆蔻	AAB71729	*广谱*	SEQ ID NOs: 145-146
山竹	AAB51525.1	C16	SEQ ID NOs: 147-148
蓼距花	Q39513.1	C16	SEQ ID NOs: 149-150
油棕（Elaeis guiniensis）	AAD42220.2	C16	SEQ ID NO: 151-152
甘蓝型油菜	CAA52070.1	C18	SEQ ID NO: 153-154
蓖麻	ABS30422.1	C18:1	SEQ ID NO: 155-156

[0483] 在所有情况下，使用生物射弹粒子轰击将以上硫酯酶构建体中的每一个转化到Prototheca moriformis(UTEX 1435)中。其他转化方法，包括如在PCT申请号PCT/US2009/066142中所描述的同源重组，也适合用于感兴趣的基因的异源表达。使用在实例2中描述的方法用以上硫酯酶构建体中的每一个进行Prototheca moriformis的(UTEX 1435)转化。每一个构建体包含NeoR基因并且使用100μg/ml G418进行阳性克隆选择。所有编码区经密码子优化以反映在Prototheca moriformis UTEX 1435(见表2)核基因中固有的密码子偏倚。在序列一致性表中列出了所使用的构建体的氨基酸序列和cDNA序列两者。除非另外说明，用来自Prototheca moriformis Δ12脂肪酸去饱和酶(SEQ ID NO:48)或来自原始小球藻硬脂酰ACP去饱和酶(SEQ ID NO:49)的藻类密码子优化的转运肽替代每一高等植物硫酯酶的转运肽。所有硫酯酶构建体是由莱茵衣藻β-微管蛋白启动子/5'UTR驱动的。将所选择的阳性克隆的生长和脂质生产与野生型(未转化的)Prototheca moriformis(UTEX 1435)进行比较。使野生型与所选择的阳性克隆在2％葡萄糖的G418平板上生长。对于每一构建体的所选择的阳性克隆的脂质谱分析总结在以下表15中(以面积％表示)。

[0484] 表15.表达不同异源硫酯酶的Prototheca moriformis的脂质谱。

[0485]

脂肪酸	UTEX 1435 wt	硫酯酶								
		加州月桂	*樟树*	*德国蓼尾*	*肉豆蔻*	*湿地蓼距花 C8:0*	*蓼距花*	*披针叶蓼距花*	*湿地蓼距花 C14:0*	*美洲椿*
C8:0	0	0	0	0		3.1	1.8	0	0	.09
C10:0	0.02	.07	.02	.01	.09	.56	6.85	1.91	.01	2.85
C12:0	0.05	14	1.82	.09	.05	.25	.2	.29	.06	.74
C14:0	1.65	3	17.3	2.59	5.31	1.45	1.8	1.83	2.87	10.45
C16:0	28.0	21.4	24.3	26.52	31.08	22.84	23.9	25.55	27.23	33.3
C18:0	2.9	2.9	2.7	3.11	2.71	3.24	2.8	3.26	3.62	3.47
C18:1	53.8	45.2	41.3	49.96	39.77	56.62	49.8	55.43	51.04	38.71
C18:2	10.95	10	9.7	11.86	14.17	8.24	9.7	8.17	10.81	7.38
C18:3 α	0.8	.86	.8	.40	.64	.61	.9	.58	.97	.52
总饱和物（面积%）	32.62	44.97	46.14	32.32	39.24	31.44	37.35	32.84	33.79	50.9

[0486] 这些结果显示所有表达的硫酯酶都在一定水平上影响脂肪酸谱。参看“总饱和物”行，饱和度受这些硫酯酶中的几种的表达的深刻影响，包括来自加州月桂、樟树的那些，并

且最显著地来自美洲榆的那些。在总饱和物百分比方面的这些变化是意料之外的，因为来自高等植物的硫酯酶的异源表达可以明显影响不只是脂质链长；它还可以影响由微藻产生的脂质谱的其他属性，即脂肪酸的饱和度。

[0487] 用湿地萼距花C8硫酯酶、萼距花硫酯酶、加州月桂和樟树硫酯酶转化的选定克隆进一步在变化量的G418(从25mg/L至50mg/L)中以及在变化的温度(从22℃至25℃)下生长并且测定这些克隆的脂质谱。表16总结了含有各硫酯酶的代表性克隆的脂质谱(表示为面积%)。构建了含有美洲榆硫酯酶的一个第二构建体并且使用以上描述的生物射弹方法将其转化到Prototheca moriformis(UTEX 1435)中。将这个第二构建体经由同源重组引入细胞中。在无绿藻种类中的同源重组的方法先前描述于PCT申请号PCT/US2009/66142中。所使用的同源DNA来自自Prototheca moriformis UTEX 1435的6S基因组DNA(以SEQ ID 92和SEQ ID 84给出的供体序列)。选择剂是使用来自酿酒酵母的、由莱茵衣藻β微管蛋白启动子驱动的密码子优化的suc2基因在蔗糖上生长的能力。用原始小球藻(UTEX 250)硬脂酰ACP去饱和酶转运肽替代天然美洲榆转运肽。这种构建体的cDNA在序列表中作为SEQ ID NO:50列出。在2%蔗糖平板上进行阳性克隆选择并且使产生的用于脂质谱测定的培养物也在含有2%蔗糖的培养基上生长。含有同源重组的异源美洲榆硫酯酶的这个Prototheca moriformis株的代表性脂质谱总结于表16中。

[0488] 表16.含有异源硫酯酶基因的Prototheca moriformis株的脂质谱。

[0489]

	湿地萼距花C8	*萼距花*	*樟树*	*美洲榆2*
C8:0	12.28	2.37	0	0
C10:0	2.17	12.09	0.02	4.69
C12:0	0.34	0.33	3.81	1.02
C14:0	1.59	2.08	32.73	16.21
C16:0	15.91	20.07	24.03	38.39
C18:0	1.59	1.57	1.21	2.83
C18:1	50.64	41.80	18.64	27.22
C18:2	13.02	16.37	16.57	7.65
C18:3α	1.52	1.75	1.66	0.74
总饱和物	33.88	38.51	61.80	63.14

[0490] 正如以上描述的克隆，与野生型(未转化)Prototheca moriformis相比，所有含有异源硫酯酶基因的转化子显示出在一定水平上受影响的脂肪酸谱，并且饱和脂肪酸的总百分比也被改变。含有通过同源重组引入的美洲榆硫酯酶的Prototheca moriformis在总饱和物方面具有最大的增长。

[0491] 另外，针对新颖的脂质谱评定了含有外源萼距花、樟树、加州月桂或美洲榆硫酯酶的转基因克隆。当在22℃以2%葡萄糖、25mg/ml G418培养时，含有萼距花硫酯酶的克隆实现了以下脂质谱：5.10%C8:0、18.28%C10:0、0.41%C12:0、1.76%C14:0、16.31%C16:0、1.40%C18:0、40.49%C18:1、以及13.16%C18:2。当在25℃以2%蔗糖培养时，含有樟树硫

酯酶的克隆(还含有外源蔗糖转化酶)实现了以下脂质谱:0.04%C10:0、6.01%C12:0、35.98%C14:0、19.42C16:0、1.48%C18:0、25.44%C18:1、以及9.34%C18:2。当在22℃以2%葡萄糖、25-100mg/ml G418培养时,含有加州月桂硫酯酶的克隆实现了以下脂质谱:0%C8:0、0.11%C10:0、34.01%C12:0、5.75%C14:0、14.02%C16:0、1.10%C18:0、28.93%C18:1、以及13.01%C18:2。当在28℃以2%葡萄糖培养时,含有美洲榆硫酯酶的克隆达到了以下脂质谱:1.54%C10:0、0.43%C12:0、7.56%C14:0、39.45%C16:0、2.49%C18:0、38.49%C18:1、以及7.88%C18:2。

[0492] 还将来自高等植物的另外的硫酯酶引入到Prototheca moriformis UTEX 1435遗传背景中,并且在如以上说明的序列表中列出密码子优化的cDNA序列和氨基酸序列。这些另外的硫酯酶包括来自肉豆蔻的广谱特异性硫酯酶(C14:0-C18:0)、来自山竹的C16:0偏好型硫酯酶、来自蓖距花的C16:0偏好型硫酯酶、来自油棕(Elaeis guiniensis)的C16:0偏好型硫酯酶、来自甘蓝型油菜的C18:0偏好型硫酯酶、以及来自蓖麻的C18:1偏好型硫酯酶。这些表达构建体的详情以及从以上各转基因/转化产生的转基因克隆描述于下文。

[0493] 使用以上描述的方法将来自肉豆蔻的广谱特异性硫酯酶(C14:0-C18:0)引入到Prototheca moriformis UTEX 1435遗传背景中。对两种不同的表构建体进行测试,每一种含有不同的质体靶向序列。在两种构建体中,使用酿酒酵母蔗糖转化酶基因suc2作为选择标记,从而赋予阳性转化子利用蔗糖作为唯一碳源在平板上生长的能力。两种表达构建体(pSZ1318和pSZ1317)均含有针对用于整合至核基因组中的6S基因组区域的5'(SEQ ID NO:82)和3'(SEQ ID NO:84)同源重组靶向序列(在构建体的侧翼)和在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出。pSZ1318包含肉豆蔻硫酯酶编码区,其中天然转运肽被来自在Prototheca moriformis Amt03启动子(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR控制下的Prototheca moriformis Δ12FAD(SEQ ID NO:48)的转运肽替换。具有来自Prototheca moriformis Δ12FAD的转运肽的、密码子优化的肉豆蔻硫酯酶作为SEQ ID NO:145列出。pSZ1317包含肉豆蔻硫酯酶编码区,其中天然转运肽被来自在Prototheca moriformis Amt03启动子(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR控制下的原始小球藻硬脂酰ACP去饱和酶(SEQ ID NO:49)的转运肽替换。具有来自原始小球藻硬脂酰ACP去饱和酶的转运肽的、密码子优化的肉豆蔻硫酯酶作为SEQ ID NO:158列出。两种表达构建体pSZ1318和pSZ1317均被转化到无绿藻细胞中并且在蔗糖作为唯一碳源的平板上进行选择。从各转化中选择阳性克隆并使其在氮限制条件下在蔗糖作为唯一碳源的培养基上培养用于生产脂质。使用以上描述的直接酯交换方法确定所选择的阳性克隆的亚组的脂质谱并且总结于表17中。

[0494] 表17.肉豆蔻广谱特异性硫酯酶转基因无绿藻细胞的脂质谱。

[0495]

株	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2

[0496]

野生型	0.01	0.03	1.17	25.86	2.84	58.33	9.16
pSZ1318 克隆 A	0.03	0.23	16.09	37.72	6.11	27.39	9.98
pSZ1318 克隆 B	0.03	0.22	15.74	37.17	6.23	28.16	9.94
pSZ1318 克隆 C	0.03	0.22	14.97	36.05	5.87	30.48	9.86
pSZ1317 克隆 A	0.02	0.21	15.23	36.62	5.11	31.83	8.76
pSZ1317 克隆 B	0.03	0.27	18.06	38.88	5.64	26.11	8.90
pSZ1317 克隆 C	0.02	0.24	16.19	37.02	5.61	29.52	9.19

[0497] 含有肉豆蔻硫酯酶转基因的阳性克隆显示出改变的脂质谱。然而,以上总结的结果显示出意料之外的结果;在高等植物中,肉豆蔻硫酯酶展现出显著的针对C16:0脂肪酰基-ACP的活性(Voelker等人,1997),然而,在无绿藻细胞中,肉豆蔻硫酯酶似乎具有分级为C14:0>C18:0>C16:0的影响并且有更广泛的基础,而不仅仅是针对C16:0。

[0498] 将来自山竹的C16:0偏好型硫酯酶引入到Prototheca moriformis UTEX 1435遗传背景中,并且在如以上说明的序列表中列出密码子优化的cDNA序列和氨基酸序列。该表达构建体含有针对用于整合至核基因组中的6S基因组区域的5'(SEQ ID NO:82)和3'(SEQ ID NO:84)同源重组靶向序列(在构建体侧翼)和在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。山竹编码区在Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR的控制下。还将山竹天然转运肽替换为来自原始小球藻硬脂酰去饱和酶(SEQ ID NO:49)的转运肽,并且将带有替换的转运肽的硫酯酶的cDNA序列作为SEQ ID NO:147列出。将整个山竹表达盒称作pSZ1452并且转化到Prototheca moriformis遗传背景中。在含有蔗糖作为唯一碳源的平板上筛选阳性克隆。选择阳性克隆的亚组并且使其在脂质产生条件下生长,并且使用如以上所述的直接酯交换法确定脂质谱。将所选择的克隆的脂质谱汇总在以下表18中。

[0499] 表18.山竹C16:0偏好型硫酯酶转基因无绿藻细胞的脂质谱。

[0500]

株	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2
野生型	0.01	0.03	1.17	25.86	2.84	58.33	9.16
pSZ1452 克隆 A	0.02	0.07	5.52	62.77	4.36	18.99	6.29
pSZ1452 克隆 B	0.02	0.08	5.69	61.66	4.76	19.28	6.54
pSZ1452 克隆 C	0.01	0.05	3.44	57.97	4.21	24.76	7.38

[0501] 结果显示,与野生型比较,具有山竹硫酯酶转基因的转化子具有显著受影响的

C16:0脂肪酸水平以及在较小程度上受影响的C14:0和C18:0脂肪酸水平、连同在C18:1脂肪酸水平上的急剧下降。

[0502] 将来自萼距花的C16:0偏好型硫酯酶引入到Prototheca moriformis UTEX 1435遗传背景中。建立了两个表达构建体，一个具有天然萼距花C16偏好型硫酯酶转运肽序列，称为pSZ1417，并且第二个称为pSZ1462，其中天然转运肽序列被来自原始小球藻硬脂酰-ACP去饱和酶(SEQ ID NO:49)的转运肽替代。将具有该天然转运肽的萼距花硫酯酶的编码序列作为SEQ ID NO:149列出并且将具有该替代转运肽的萼距花硫酯酶的编码序列作为SEQ ID NO:160列出。两种表达构建体均含有针对用于整合至核基因组中的6S基因组区域的5'(SEQ ID NO:82)和3'(SEQ ID NO:84)同源重组靶向序列(在构建体侧翼)和在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。在两种构建体中，萼距花编码区在Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR的控制下。将这两个构建体转化到Prototheca moriformis遗传背景中，并且在含有蔗糖作为唯一碳源的平板上筛选阳性克隆。选择阳性克隆的亚组并且使其在脂质产生条件下生长，并且使用如以上所述的直接酯交换法确定脂质谱。将所选择的克隆的脂质谱汇总在下表19中。

[0503] 表19.萼距花C16:0偏好型硫酯酶转基因无绿藻细胞的脂质谱。

[0504]

株	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2
野生型	0.01	0.03	1.17	25.86	2.84	58.33	9.16
pSZ1417克隆A	0.02	0.06	4.21	55.29	2.59	26.87	9.02
pSZ1417克隆B	0.02	0.06	4.12	54.57	2.31	26.43	10.45
pSZ1417克隆C	0.01	0.05	3.59	53.18	2.60	29.02	9.48
pSZ1462克隆A	0.02	0.11	10.62	67.42	2.18	12.95	5.13
pSZ1462克隆B	0.03	0.11	8.88	66.83	2.30	15.32	5.16
pSZ1462克隆C	0.03	0.11	9.28	66.65	2.27	15.19	5.14
pSZ1462克隆D	0.02	0.09	8.30	66.36	2.27	16.52	5.01

[0505] 结果显示，与野生型比较，具有任一种萼距花C16:0偏好型硫酯酶构建体的转化子具有显著受影响的C16:0脂肪酸水平以及在较小程度上受影响的C14:0脂肪酸水平、连同在C18:1脂肪酸水平上的急剧下降。与pSZ1417转化子比较，在这两种构建体中的转运肽的差异可以解释在pSZ1462转化子中的C16:0脂肪酸水平增加。

[0506] 将来自油棕(Elaeis guiniensis)(非洲油棕)的两种C16:0偏好型硫酯酶引入到

Prototheca moriformis UTEX 1435遗传背景中，这两种C16:0偏好型硫酯酶相应于Genbank登录号AAD422220.2(SEQ ID NO:152)和ABD83939(SEQ ID NO:162)中的氨基酸序列，分别称作油棕棕榈酰基-ACP硫酯酶和油棕棕榈酰基-ACP硫酯酶PATE。在如以上说明的序列表中列出这些密码子优化的cDNA序列和氨基酸序列。这两种硫酯酶共享显著水平的氨基酸一致性(超过94％)，但是它们在非洲油棕植物中的对应作用仍不清楚。编码油棕棕榈酰基-ACP硫酯酶的构建体被称作pSZ1437，而编码油棕棕榈酰基-ACP硫酯酶PATE的构建体被称作pSZ1436。两种表达构建体均含有针对用于整合至核基因组中的6S基因组区域的5’(SEQ ID NO:82)和3’(SEQ ID NO:84)同源重组靶向序列(在构建体侧翼)和在莱茵衣藻β-微管蛋白启动子/5’UTR和普通小球藻硝酸盐还原酶3’UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。在两种构建体中，油棕硫酯酶编码区在Prototheca moriformis Amt03启动子/5’UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3’UTR的控制下。将这两种构建体转化到Prototheca moriformis遗传背景中，并且在含有蔗糖作为唯一碳源的平板上筛选阳性克隆。选择阳性克隆的亚组并且使其在脂质产生条件下生长，并且使用如以上所述的直接酯交换法确定脂质谱。将所选择的克隆的脂质谱汇总在以下表20中。

[0507] 表20.油棕C16:0偏好型硫酯酶转基因无绿藻细胞的脂质谱。

[0508]

株	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2
野生型	0.01	0.03	1.17	25.86	2.84	58.33	9.16
pSZ1437 克隆 A	0.02	0.07	3.82	55.86	3.12	24.47	10.48
pSZ1437 克隆 B	0.02	0.05	3.01	53.23	3.47	28.70	9.26
pSZ1437 克隆 C	0.02	0.06	3.33	53.20	3.30	26.64	11.19
pSZ1437 克隆 D	0.02	0.05	3.08	52.88	3.60	27.94	10.16
pSA1437 克隆 E	0.02	0.05	3.01	52.84	3.48	28.46	9.87
pSZ1436 克隆 A	0.01	0.04	1.48	29.54	3.33	52.26	10.58
pSZ1436 克隆 B	0.01	0.04	1.48	29.43	3.33	52.11	10.76
pSZ1436 克隆 C	0.01	0.04	1.50	29.25	3.38	52.07	10.89
pSZ1436 克隆 D	0.01	0.04	1.51	29.18	3.41	51.80	11.17
pSZ1436 克隆 E	0.01	0.04	1.54	29.14	3.56	51.43	11.42

[0509] 当与野生型比较时，由pSZ1437编码的油棕C16:0偏好型硫酯酶对于C16:0脂肪酸水平具有显著影响并且对于C14:0脂肪酸水平具有较小程度的影响，并且在C18:1脂肪酸水平上急剧下降。出人意料地，由pSZ1436编码的油棕C16:0偏好型硫酯酶PATE尽管与由pSZ1436编码的硫酯酶具有显著水平的氨基酸一致性，但对于C16:0或C14:0脂肪酸水平具

有相对很低的活性。

[0510] 将来自甘蓝型油菜的C18:0偏好型硫酯酶引入到Prototheca moriformis UTEX 1435遗传背景中，并且在如以上说明的序列表中列出密码子优化的cDNA序列和氨基酸序列。该表达构建体含有针对用于整合至核基因组中的6S基因组区域的5'(SEQ ID NO:82)和3'(SEQ ID NO:84)同源重组靶向序列(在构建体侧翼)和在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。该甘蓝型油菜编码区在Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR的控制下。将整个甘蓝型油菜表达盒称作pSZ1358并且转化到Prototheca moriformis遗传背景中。在含有蔗糖作为唯一碳源的平板上筛选阳性克隆。选择阳性克隆的亚组并且使其在脂质产生条件下生长，并且使用如以上所述的直接酯交换法确定脂质谱。将所选择的克隆的脂质谱汇总在以下表21中。

[0511] 表21.甘蓝型油菜C18:0偏好型硫酯酶转基因无绿藻细胞的脂质谱。

[0512]

株	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2
野生型	0.00	0.04	1.18	25.44	3.42	57.97	6.98
pSZ1358 克隆A	0.07	0.31	1.51	33.27	27.26	27.37	7.50
pSZ1358 克隆B	0.07	0.33	1.60	34.73	26.71	26.52	7.32

[0513] 结果显示，与野生型比较，具有甘蓝型油菜C18:0偏好型硫酯酶转基因的转化子具有显著受影响的C18:0脂肪酸水平和在较小程度上受影响的C16:0脂肪酸水平、以及在C18:1脂肪酸水平上的急剧下降。

[0514] 将来自蓖麻的脂肪酰基-ACP硫酯酶引入到Prototheca moriformis UTEX 1435遗传背景中，并且在如以上说明的序列表中列出密码子优化的cDNA序列和氨基酸序列。该表达构建体含有针对用于整合至核基因组中的6S基因组区域的5'(SEQ ID NO:82)和3'(SEQ ID NO:84)同源重组靶向序列(在构建体侧翼)和在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。该蓖麻编码区在Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR的控制下。将蓖麻天然转运肽也替换为来自原始小球藻硬脂酰去饱和酶(SEQ ID NO:49)的转运肽，并且将带有替换的转运肽的硫酯酶的cDNA序列作为SEQ ID NO:155列出。将整个蓖麻表达盒称作pSZ1375并且转化到Prototheca moriformis遗传背景中。在含有蔗糖作为唯一碳源的平板上筛选阳性克隆。选择阳性克隆的亚组并且使其在脂质产生条件下生长，并且使用如以上所述的直接酯交换法确定脂质谱。将所选择的克隆的脂质谱汇总在以下表22中。

[0515] 表22.蓖麻ACP-硫酯酶转基因无绿藻细胞的脂质谱。

[0516]

株	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2
野生型	0.01	0.03	0.98	24.65	3.68	62.48	6.26
pSZ1375 克隆 A	0.01	0.03	0.91	18.34	2.55	67.93	8.35
pSZ1375 克隆 B	0.01	0.03	0.97	18.51	2.47	67.83	8.25
pSZ1375 克隆 C	0.01	0.03	0.93	18.65	2.84	67.58	7.90
pSZ1375 克隆 D	0.01	0.03	0.92	18.90	2.30	67.48	8.37

[0517] 结果显示,带有蓖麻硫酯酶转基因的转化子具有受影响的C16:0脂肪酸水平以及在较小程度上受影响的C18:0脂肪酸水平。同样,当与野生型水平相比时,存在着C18:1脂肪酸水平的同时增加。

[0518] 实例4:用多个外源异源硫酯酶基因转化无绿藻属

[0519] 使用以上披露的方法转化微藻株Prototheca moriformis(UTEX 1435)以便在单个克隆中表达多种硫酯酶。在单个克隆中表达多种硫酯酶允许微藻产生具有多种脂肪酸谱的油,这些脂肪酸谱与单独表达任何单一硫酯酶时产生的那些完全不同(如在前述实例中所展示的)。首先利用同源重组用樟树硫酯酶(一种C14偏好型硫酯酶)连同蔗糖转化酶基因、来自酿酒酵母的suc2(选择是在蔗糖上生长的能力)转化Prototheca moriformis(UTEX 1435)。用于这种同源重组构建体的DNA来自Prototheca moriformis基因组DNA的KE858区域,如在以上第III节所描述的。这种构建体的相关部分在序列表中作为SEQ ID NO:51列出。在含有蔗糖的平板上筛选阳性克隆。然后用三个盒的其中之一对阳性克隆进行再转化,这三个盒每一个编码对抗生素G418的抗性以及一种另外的硫酯酶:(1)来自蓖距花的硫酯酶基因(C8-10偏好型),SEQ ID NO:52;(2)来自加州月桂的硫酯酶基因(C12偏好型),SEQ ID NO:53;或来自美洲榆的硫酯酶基因(广谱;C10-C16偏好型),SEQ ID NO:54。包括在序列表中的是每一构建体的相关部分的序列。在含有蔗糖的、具有50μg/ml G418的培养基上筛选表达两种硫酯酶基因的克隆。选择阳性克隆并且对生长与脂质谱进行测定。表23汇总了代表性阳性克隆的脂质谱(以面积%表示)。

[0520] 表23.用多个硫酯酶转化的Prototheca moriformis的脂质谱。

[0521]

脂肪酸	UTEX 1435	UTEX 1435 + *樟树* TE	UTEX 1435 + *樟树* TE 遗传背景		
			+ *蓖距花* TE	+ *加州月桂* TE	+ *美洲榆* TE
C8:0	0	0	0.19	0	0.06

[0522]

C10:0	0.02	0.02	2.16	0.07	1.87
C12:0	0.05	0.66	0.53	13.55	1.61
C14:0	1.65	10.52	7.64	8.0	14.58
C16:0	28.0	22.56	22.31	19.98	29.53
C18:0	2.9	6.67	3.23	2.24	2.93
C18:1	53.8	47.78	48.54	42.55	37.3
C18:2	10.95	12.3	11.76	10.13	8.9
C18:3 α	0.8	0.93	0.91	0.91	0.76
总饱和物（面积%）	32.62	40.43	36.06	43.84	50.58

[0523] 另外，在22℃使具有樟树与加州月桂硫酯酶的双硫酯酶克隆在含有2％蔗糖的、具有50mg/L G418的培养基中生长。在这些培养条件下从这个株获得的脂肪酸谱是：C8:0(0)、C10:0(0.10)、C12:0(31.03)、C14:0(7.47)、C16:0(15.20)、C18:0(0.90)、C18:1(30.60)、C18:2(12.44)、以及C18:3α(1.38)，具有54.7的总饱和物。

[0524] 产生了具有两个含有樟树硫酯酶的同源重组构建体(一个靶向6S区域并且另一个靶向KE858区域)的双硫酯酶克隆。代表性阳性克隆具有以下脂肪酸谱：0％C8:0、0.06％C10:0、5.91％C12:0、43.27％C14:0、19.63％C16:0、0.87％C18:0、13.96％C18:1、以及13.78％C18:2，具有69.74％的总饱和物。这个克隆具有超过49％的C12-C14水平，超过野生型细胞中C12-C14水平的37倍。

[0525] 以上数据显示多种硫酯酶可以成功地在微藻中共表达。多种硫酯酶的共表达导致改变的脂肪酸谱，这些脂肪酸谱不仅与野生型株系显著不同，而且与通过表达这些单独的硫酯酶的任何一种而获得的脂肪酸谱也显著不同。具有重叠的链长特异性的多种硫酯酶的表达可以导致那些特异性脂肪酸的累积增加。

[0526] 在Prototheca moriformis中表达异源硫酯酶(单独地或组合地)不仅改变宿主株中的脂肪酸/脂质谱，而且当与目前从多种植物作物可获得的油(表5)相比时，这些谱是未发现于其他目前可获得的系统中的真正独特的油的谱。这些转基因株系不仅显示出与未转化的野生型株系的显著差异，而且它们具有与在表5中所显示的任何商业用油显著不同的谱。作为实例，椰子油和棕榈仁油都具有范围从5.5％-17％的C8-C10脂肪酸水平。表达湿地蓖距花C8偏好型硫酯酶或蓖距花C10偏好型硫酯酶的转基因株系在任何地方分别积累从3.66％至8.65％。这些C8-C10脂肪酸水平类似于椰子油和棕榈仁，然而这些转基因藻株缺乏显著更高的C12:0脂肪酸，并且它们具有极高的C16:0(在转基因株系中为23％，而在椰子油或棕榈仁油中分别为11％-16％)和/或18:1(在转基因株系中为50％-57％，而在椰子油或棕榈仁油中分别为8％-19％)。

[0527] 通过表达Cuphea wrightii硫酯酶在UTEX1435株中产生富含月桂酸酯和肉豆蔻酸酯的油：Cuphea wrightii的种子已经显示出积累含有超过25％的C10:0和超过65％的C12:0脂肪酸的油。两种FatB硫酯酶，CwFatB1(Gen Bank登录号U56103)和CwFatB2(Gen Bank登录号U56104)，已经从Cuphea wrightii克隆(如在Leonard等人，Plant Mol.Biol.(植物分

子生物学)34(4):669-79(1997)中所描述的)并且在拟南芥中表达(如在Leonard等人,Plant J.(植物杂志)13(5):621-8(1998)中所描述的)。表达CwFatB1和CwFatB2的拟南芥转基因株系的脂肪酸谱显示出增加的高达16%至25%的C12:0脂肪酸种类(Leonard等人,1998,上文)。在此我们证明了通过在UTEX1435株中表达Cuphea wrightii硫酯酶CwFatB1和CwFatB2而产生富含月桂酸酯和肉豆蔻酸酯的油的能力。在此处描述的实例中,表达CwFatB1和CwFatB2的转基因株系是使用之前描述的转化方法产生的。

[0528] CwFatB1和CwFatB2的氨基酸序列显示于下文,其中预测的叶绿体靶向序列加下划线。使用这些一级氨基酸序列来合成用于转化构建体的相应基因。利用如前描述的UTEX 1435株的优选密码子使用对这两种基因的核苷酸序列针对在UTEX 1435株中表达进行了优化。

[0529] CwFatB1(U56103):

[0530] MVAAAASSAFFSVPTPGTSPKPGKFGNWPSSLSVPFKPDNGGFVKANASAHPKANGSAVNLKSGSLETPPRSFINQLPDLSMLLSKITTVFGAAEKQWKRPGMLVEPFGVDRIFQDGVFFRQSFSIRSYEIGVDRTASIETLMNIFQETSLNHCKSIGLLNDGFGRTPEMCKRDLIWVVTKIQVEVNRYPTWGDTIEVNTWVSESGKNGMGRDWLISDCRTGEILIRATSVWAMMNQNTRRLSKFPYEVRQEIAPHFVDSAPVIEDDRKLHKLDVKTGDSIRDGLTPRWNDLDVNQHVNNVKYIGWILKSVPIEVFETQELCGVTLEYRRECGRDSVLESVTTMDPAKEGDRCVYQHLLRLEDGADITIGRTEWRPKNAGANGAISSGKTSNGNSVS(SEQ ID NO:186)

[0531] CwFatB2(U56104):

[0532] MVVAAAASSAFFPVPAPRPTPKPGKFGNWPSSLSQPFKPKSNPNGRFQVKANVSPHPKANGSAVSLKSGSLNTLEDPPSSPPPRTFLNQLPDWSRLRTAITTVFVAAEKQFTRLDRKSKRPDMLVDWFGSETIVQDGLVFRERFSIRSYEIGADRTASIETLMNHLQDTSLNHCKSVGLLNDGFGRTPEMCTRDLIWVLTKMQIVVNRYPTWGDTVEINSWFSQSGKIGMGREWLISDCNTGEILVRATSAWAMMNQKTRRFSKLPCEVRQEIAPHFVDAPPVIEDNDRKLHKFDVKTGDSICKGLTPGWNDFDVNQHVSNVKYIGWILESMPTEVLETQELCSLTLEYRRECGRESVVESVTSMNPSKVGDRSQYQHLLRLEDGADIMKGRTEWRPKNAGTNRAIST(SEQ ID NO:187)

[0533] 用C.wrightii硫酯酶转化UTEX 1435:在这个实例中,使用UTEX 1435株作为向其中引入表达C.wrightii FatB1和FatB2硫酯酶的盒的受体株。这些转化构建体含有允许在蔗糖上选择的盒(酿酒酵母suc2基因)连同这些硫酯酶。使用生物射弹如前所述对细胞进行转化。在含有2%蔗糖的培养基上直接转化细胞。产生转化构建体使得这些硫酯酶的表达受莱因衣藻β-微管蛋白启动子或受内源UTEX 1435Amt3启动子驱动。

[0534] 产生了这些硫酯酶盒的另外的变体,其中天然高等植物转运肽被藻类转运肽替代。在这些构建体中使用的转运肽指定如下:TP1编码一种针对衍生自UTEX250的硬脂酰ACP去饱和酶的转运肽;TP2编码一种针对衍生自UTEX 1435的硬脂酰ACP去饱和酶的转运肽;TP3编码一种针对衍生自UTEX 1435的Δ12脂肪酸去饱和酶的转运肽;并且TP4编码一种针对衍生自UTEX 1435的异戊烯二磷酸合成酶的转运肽。在这个实例中使用的构建体列于以下表24中。

[0535] 表24.TE构建体。

[0536]

构建体	说明
构建体1	6S-CrbTub_suc2_nr::CrbTub_CwFatB1_nr-6S

构建体2	6S-CrbTub_suc2_nr::CrbTub_CwFatB2_nr-6S
构建体3	6S-CrbTub_suc2_nr::Amt3_CwFatB1_nr-6S
构建体4	6S-CrbTub_suc2_nr::Amt3_CwFatB2_nr-6S
构建体5	6S-CrbTub_suC2_nr::CrbTub_TP1-CwFatB1_nr-6S
构建体6	6S-CrbTub_suc2_nr::CrbTub_TP2-CwFatB1_nr-6S
构建体7	6S-CrbTub_suc2_nr::CrbTub_TP3-CwFatB1_nr-6S
构建体8	6S-CrbTub_suc2_nr::CrbTub_TP4-CwFatB1_nr-6S
构建体9	6S-CrbTub_suc2_nr::CrbTub_TP1-CwFatB2_nr-6S
构建体10	6S-CrbTub_suc2_nr::CrbTub_TP2-CwFatB2_nr-6S
构建体11	6S-CrbTub_suc2_nr::CrbTub_TP3-CwFatB2_nr-6S
构建体12	6S-CrbTub_suc2_nr::CrbTub_TP4-CwFatB2_nr-6S

[0537] 表达suc2和C.wrightii FatB1硫酯酶的转化DNA(构建体1):在以下给出了该转化构建体的序列,6S-CrbTub_suc2_nr::CrbTub_CwFatB1_nr-6S,指定为构建体1。相关的限制性位点以小写字母、粗体以及下划线指示并且分别是:5'-3'SapI、KpnI、AscI、MfeI、BamHI、EcoRI、SpeI、AscI、XhoI、SacI以及SapI。SapI位点界定转化DNA的5′和3′末端。该构建体的在5'和3'侧翼的下划线序列代表来自UTEX1435的基因组DNA,其允许经由同源重组的该转化DNA的靶向整合(6S区域)。在5′至3′方向继续向前,通过小写字母、加框文本指示驱动酿酒酵母suc2基因(编码蔗糖水解活性,由此允许该株在蔗糖上生长)表达的莱茵衣藻β-微管蛋白启动子。suc2的起始子ATG和终止子TGA由大写字母、粗斜体指示,同时编码区以小写字母斜体指示。普通小球藻硝酸盐还原酶3′UTR由小写字母粗体文本指示,后接间隔区。驱动C.wrightii TE(CwFatB1)表达的莱茵衣藻β-微管蛋白启动子由加框文本指示。该硫酯酶(CwFatB1)的起始子ATG和终止子TGA由大写字母、粗斜体文本指示,同时编码区的其余部分以小写字母斜体指示。该硫酯酶的预测的质体靶向序列位于在序列中的起始子ATG与AscI位点之间。普通小球藻硝酸盐还原酶3′UTR由粗体文本指示。

[0538] 构建体1:

[0539]

gctcttcgccgccgccactcctgctcgagcgcgcccgcgcgtgcgccgccagcgccttggccttttcgccgcgctcgtgcgcgtc
gctgatgtccatcaccaggtccatgaggtctgccttgcgccggctgagccactgcttcgtccgggcggccaagaggagcatgagg
gaggactcctggtccagggtcctgacgtggtcgcggctctgggagcgggccagcatcatctggctctgccgcaccgaggccgcct
ccaactggtcctccagcagccgcagtcgccgccgaccctggcagaggaagacaggtgaggggggtatgaattgtacagaacaac
cacgagccttgtctaggcagaatccctaccagtcatggctttacctggatgacggcctgcgaacagctgtccagcgaccctcgctgc

[0540]

cgccgcttctcccgcacgcttctttccagcaccgtgatggcgcgagccagcgccgcacgctggcgctgcgcttcgccgatctgagg
acagtcggggaactctgatcagtctaaaccccccttgcgcgttagtgttgccatcctttgcagaccggtgagagccgacttgttgtgcg
ccacccccccacaccacctcctcccagaccaattctgtcacctttttggcgaaggcatcggcctcggcctgcagagaggacagcagt
gcccagccgctgggggttggcggatgcacgctca**ggtacc**ctttcttgcgctatgacacttccagcaaaaggtagggcgggctgc
gagacggcttcccggcgctgcatgcaacaccgatgatgcttcgaccccccgaagctccttcggggctgcatgggcgctccgatgc
cgctccagggcgagcgctgtttaaatagccaggccccgattgcaaagacattatagcgagctaccaaagcatattcaaacaccta
gatcactaccacttctacacaggccactcgagcttgtgatcgcactccgctaaggggcgcctcttcccttcgtttcagtcacaaccc
gcaaac**ggcgcgcc**atatca***ATG***ctgctgcaggccttcctgttcctgctggccggcttcgccgccaagatcagcgcctccatga
cgaacgagacgtccgaccgccccctggtgcacttcacccccaacaagggctggatgaacgaccccaacggcctgtggtacga
cgagaaggacgccaagtggcacctgtacttccagtacaacccgaacgacaccgtctgggggacgcccttgttctggggccac
gccacgtccgacgacctgaccaactgggaggaccagcccatcgccatcgccccgaagcgcaacgactccggcgccttctccg
gctccatggtggtggactacaacaacacctccggcttcttcaacgacaccatcgacccgcgccagcgctgcgtggccatctgga
cctacaacaccccggagtccgaggagcagtacatctcctacagcctggacggcggctacaccttcaccgagtaccagaagaa
cccgtgctggccgccaactccacccagttccgcgacccgaaggtcttctggtacgagccctcccagaagtggatcatgaccgc
ggccaagtcccaggactacaagatcgagatctactcctccgacgacctgaagtcctggaagctggagtccgcgttcgccaacg
agggcttcctcggctaccagtacgagtgccccggcctgatcgaggtccccaccgagcaggaccccagcaagtcctactgggtg
atgttcatctccatcaaccccggcgccccggccggcggctccttcaaccagtacttcgtcggcagcttcaacggcacccacttcg
aggccttcgacaaccagtcccgcgtggtggacttcggcaaggactactacgccctgcagaccttcttcaacaccgacccgacct
acgggagcgccctgggcatcgcgtgggcctccaactgggagtactccgccttcgtgcccaccaaccccctggcgctcctccatgt
ccctcgtgcgcaagttctccctcaacaccgagtaccaggccaacccggagacggagctgatcaacctgaaggccgagccgat
cctgaacatcagcaacgccggccccctggagccggttcgccaccaacaccacgttgacgaaggccaacagctacaacgtcga
cctgtccaacagcaccggcaccctggagttcgagctggtgtacgccgtcaacaccacccagacgatctccaagtccgtgttcgc
ggacctctccctctggttcaagggcctggaggaccccgaggagtacctccgcatgggcttcgaggtgtccgcgtcctccttcttcc
tggaccgcgggaacagcaaggtgaagttcgtgaaggagaaccctacttcaccaaccgcatgagcgtgaacaaccagccctt
caagagcgagaacgacctgtcctactacaaggtgtacggcttgctggaccagaacatcctggagctgtacttcaacgacggcg
acgtcgtgtccaccaacacctacttcatgaccaccgggaacgccctgggctccgtgaacatgacgacgggggtggacaacctg
ttctacatcgacaagttccaggtgcgcgaggtcaag***TGA***caattggcagcagcagctcggatagtatcgacacactctgga
cgctggtcgtgtgatggactgttgccgccacacttgctgccttgacctgtgaatatccctgccgcttttatcaaacagcctcagt
gtgtttgatcttgtgtgtacgcgcttttgcgagttgctagctgcttgtgctatttgcgaataccacccccagcatcccccttccctcg
tttcatatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgcccct
cgcacagccttggtttgggctccgcctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaa
gtagtgggatgggaacacaaatggaggatccgcgtctcgaacagagcgcgcagaggaacgctgaaggtctcgcctctgtcg
cacctcagcgcggcatacaccacaataaccacctgacgaatgcgcttggttcttcgtccattagcgaagcgtccggttcacacacgt
gccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggtcgaaacgttcacag**cctagggatatcgaattc**ctttct
tgcgctatgacacttccagcaaaaggtagggcgggctgcgagacggcttcccggcgctgcatgcaacaccgatgatgcttcgacc
ccccgaagctccttcggggctgcatgggcgctccgatgccgctccagggcgagcgctgtttaaatagccaggccccgattgcaa
agacattatagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatcgcactc
cgctaaggggcgcctcttcctcttcgtttcagtcacaacccgcaaac**actagt*ATG***gtggccgccgccgccagcagcgccttc
ttcagcgtgcccacccccggcaccagccccaagcccggcaagttcggcaactggcccagcagcctgagcgtgcccttcaagc
ccgacaacggcggcttccacgtgaaggccaacgccagcgcccacg**ggcgcgcc**ccccaaggccaacggcagcgccgtgaa
cctgaagtccggcagcctggagacccccccccgcagcttcatcaaccagctgcccgacctgagcatgctgctgagcaagatca
ccaccgtgttcggcgccgccgagaagcagtggaagcgccccggcatgctggtggagcccttcggcgtggaccgcatcttccag
gacggcgtgttcttccgccagagcttcagcatccgcagctacgagatcggcgtggaccgcaccgccagcatcgagaccctgat
gaacatcttccaggagaccagcctgaaccactgcaagagcatcggcctgctgaacgacggcttcggccgcaccccgagatg
tgcaagcgcgacctgatctgggtggtgaccaagatccaggtggaggtgaaccgctaccccacctgggggcgacaccatcgagg
tgaacacctgggtgagcgagagcggcaagaacggcatgggccgcgactggctgatcagcgactgccgcaccggcgagatc
ctgatccgcgccaccagcgtgtgggccatgatgaaccagaacacccgccgcctgagcaagttcccctacgaggtgcgccagg

[0541]

agatcgcccccccacttcgtggacagcgcccccgtgatcgaggacgaccgcaagctgcacaagctggacgtgaagaccggcg
acagcatccgcgacggcctgacccccccgctggaacgacctggacgtgaaccagcacgtgaacaacgtgaagtacatcggct
ggattctgaagtccgtgcccatcgaggtgttcgagacccaggagctgtgcggcgtgaccctggagtaccgccgcgagtgcggc
cgcgacagcgtgctggagagcgtgaccaccatggaccccgccaaggagggcgaccgctgcgtgtaccagcacctgctgcgc
ctggaggacggcgccgacatcaccatcggccgcaccgagtggcgccccaagaacgccggcgccaacggcgccatcagca
*gcggcaagaccagcaacggcaacagcgtgagc**TGA**ttaattaa**ctcgaggcagcagcagctcggatagtatcgacacact***
ctggacgctggtcgtgtgatggactgttgccgccacacttgctgccttgacctgtgaatatccctgccgcttttatcaaacagcc
tcagtgtgtttgatcttgtgtgtacgcgcttttgcgagttgctagctgcttgtgctatttgcgaataccacccccagcatccccttc
cctcgtttcatatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactg
cccctcgcacagccttggtttgggctccgcctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacg
ggaagtagtgggatgggaacacaaatggaaagcttgagctcttgttttccagaaggagttgctccttgagcctttcattctcagcct*
cgataacctccaaagccgctctaattgtggagggggttcgaatttaaaagcttggaatgttggttcgtgcgtctggaacaagcccaga
cttgttgctcactgggaaaaggaccatcagctccaaaaaacttgccgctcaaaccgcgtacctctgctttcgcgcaatctgccctgttg
aaatcgccaccacattcatattgtgacgcttgagcagtctgtaattgcctcagaatgtggaatcatctgccccctgtgcgagcccatgc
caggcatgtcgcgggcgaggacacccgccactcgtacagcagaccattatgctacctcacaatagttcataacagtgaccatatttct
cgaagctccccaacgagcacctccatgctctgagtggccacccccggccctggtgcttgcggagggcaggtcaaccggcatgg
ggctaccgaaatccccgaccggatcccaccacccccgcgatgggaagaatctctccccgggatgtgggcccaccaccagcacaa
cctgctggcccaggcgagcgtcaaaccataccacacaaatatccttggcatcggccctgaattccttctgccgctctgctacccggtg
*cttctgtccgaagcaggggttgctagggatcgctccgagtccgcaaacccttgtcgcgtggcggggcttgttcgagctt**gaagagc***
(SEQ ID NO: 188)

[0542] 表达suc2和C.wrightii FatB2硫酯酶的转化DNA(构建体2):该转化构建体(6S-CrbTub_suc2_nr::CrbTub_CwFatB2_nr-6S,指定为构建体2)是通过利用SpeI和AscI限制性位点(以小写字母、以粗体和下划线指示)用密码子优化的CwFatB2基因替换来自构建体1的CwFatB1基因而产生的。该硫酯酶(CwFatB2)的起始子ATG和终止子TGA以大写字母、粗斜体文本指示,而编码区的其余部分以小写字母斜体指示。该硫酯酶的预测的质体靶向序列位于在序列中的起始子ATG与AscI位点之间。

[0543] 构建体2(部分):

[0544]

actagt*ATG*gtggtggccgccgccgccagcagcgccttcttccccgtgcccgccccccgccccacccccaagcccggcaag*
ttcggcaactggcccagcagcctgagccagcccttcaagcccaagagcaaccccaacggccgcttccaggtgaaggccaac
*gtgagcccccacg**ggcgcgcc**cccaaggccaacggcagcgccgtgagcctgaagtccggcagcctgaacaccctggagga*
cccccccagcagcccccccccccgcaccttcctgaaccagctgcccgactggagccgcctgcgcaccgccatcaccaccgtgt
tcgtggccgccgagaagcagttcacccgcctggaccgcaagagcaagcgccccgacatgctggtggactggttcggcagcga
gaccatcgtgcaggacggcctggtgttccgcgagcgcttcagcatccgcagctacgagatcggcgccgaccgcaccgccagc
atcgagaccctgatgaaccacctgcaggacaccagcctgaaccactgcaagagcgtgggcctgctgaacgacggcttcggcc
gcacccccgagatgtgcacccgcgacctgatctgggtgctgaccaagatgcagatcgtggtgaaccgctaccccacctggggc
gacaccgtggagatcaacagctggttcagccagagcggcaagatcggcatgggccgcgagtggctgatcagcgactgcaac
accggcgagatcctggtgcgcgccaccagcgcctgggccatgatgaaccagaagacccgccgcttcagcaagctgccctgcg
aggtgcgccaggagatcgcccccccacttcgtggacgcccccccccgtgatcgaggacaacgaccgcaagctgcacaagttcga
cgtgaagaccggcgacagcatctgcaagggcctgacccccggctggaacgacttcgacgtgaaccagcacgtgagcaacgt
gaagtacatcggctggattctggagagcatgcccaccgaggtgctggagacccaggagctgtgcagcctgaccctggagtac
cgccgcgagtgcggccgcgagagcgtggtggagagcgtgaccagcatgaaccccagcaaggtgggcgaccgcagccagt
accagcacctgctgcgcctggaggacggcgccgacatcatgaagggccgcaccgagtggcgccccaagaacgccggcacc
*aaccgcgccatcagcacc**TGA**ttaattaa**ctcgag*** (SEQ ID NO: 189)

[0545] 表达suc2和由amt3启动子驱动的C.wrightii FatB1和FatB2硫酯酶的转化DNA(构

建体3和4)：这些转化构建体(6S-CrbTub_suc2_nr::Amt3_CwFatB1_nr-6S，指定为构建体3，以及6S-CrbTub_suc2_nr::Amt3_CwFatB2_nr-6S，指定为构建体4)是通过用衍生自UTEX1435的、作为EcoRI和SpeI限制性片段的Amt3启动子替换来自构建体1和构建体2的、驱动这些硫酯酶的CrbTub启动子(在下文以小写字母、粗体和下划线指示)而产生的。Amt3启动子区域通过小写字母加框文本指示。

[0546] 构建体3和4(部分)：

[0547]

gaattcggcgctatgacacttccagcaaaaggtagggcgggctgcgagacggcttcccggcgctgcatgcaacaccgatgatgct
tcgaccccccgaagctccttcggggctgcatgggcgctccgatgccgctccagggcgagcgctgtttaaatagccaggcccccga
ttgcaaagacattatagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatcg
cactccgctaagggggcgcctcttcctcttcgtttcagtcggccgacaggacgcgcgtcaaaggtgctggtcgtgtatgccctggcc
ggcaggtcgttgctgctgctggttagtgattccgcaaccctgattttggcgtcttattttggcgtggcaaacgctggcgcccgcgagcc
gggccggcggcgatgcggtgccccacggctgccggaatccaagggaggcaagagcgcccgggtcagttgaagggctttacgc
gcaaggtacagccgctcctgcaaggctgcgtggtggaattggacgtgcaggtcctgctgaagttcctccaccgcctcaccagcgg
acaaagcaccggtgtatcaggtccgtgtcatccactctaaagagctcgactacgacctactgatggccctagattcttcatcaaaaac
gcctgagacacttgcccaggattgaaactccctgaagggaccaccaggggccctgagttgttccttccccccgtggcgagctgcca
gccaggctgtacctgtgatcgaggctggcgggaaaataggcttcgtgtgctcaggtcatgggaggtgcaggacagctcatgaaac
gccaacaatcgcacaattcatgtcaagctaatcagctatttcctcttcacgagctgtaattgtcccaaaattctggtctaccgggggtga
tccttcgtgtacgggcccttccctcaaccctaggtatgcgcgcatgcggtcgccgcgcaactcgcgcgagggccgagggtttggga
cgggccgtcccgaaatgcagttgcacccggatgcgtggcacctttttgcgataatttatgcaatggactgctctgcaaaattctggct
ctgtcgccaaccctaggatcagcggcgtaggatttcgtaatcattcgtcctgatggggagctaccgactaccctaatatcagcccgac
tgcctgacgccagcgtccacttttgtgcacacattccattcgtgcccaagacatttcattgtggtgcgaagcgtccccagttacgctca
cctgtttcccgacctccttactgttctgtcgacagagcgggcccacaggccggtcgcagcc**actagt** (SEQ ID NO: 190)

[0548] 在藻类转运肽控制下表达C.wrightii FatB1硫酯酶的转化DNA：这些转化构建体(6S-CrbTub_suc2_nr::CrbTub_TP1-CwFatB1_nr-6S，指定为构建体5；构建体6S-CrbTub_suc2_nr::CrbTub_TP2-CwFatB1_nr-6S，指定为构建体6；构建体6S-CrbTub_suc2_nr::CrbTub_TP3-CwFatB1_nr-6S，指定为构建体7以及构建体6S-CrbTub_suc2_nr::CrbTub_TP4-CwFatB1_nr-6S，指定为构建体8)是通过用在以下显示为SpeI和AscI限制性片段(用小写字母、粗体和下划线指示)的相应的藻类转运肽替换来自构建体1的CwFatB1的天然转运肽而产生的。所产生的藻类转运肽序列位于在以下序列中的起始子ATG与AscI位点之间。

[0549] 构建体5-12(部分)：

[0550] TP1(UTEX250硬脂酰ACP去饱和酶转运肽序列)

[0551]

actagt*ATG*gccaccgcatccactttctcggcgttcaatgcccgctgcggcgacctgcgtcgctcggcgggctccgggccccg
gcgcccagcgaggcccctccccgtgcgc**gggcgcgcc** (SEQ ID NO: 191)

[0552] TP2(UTEX1435硬脂酰ACP去饱和酶转运肽序列)

[0553]

actagt*ATG*gcttccgcggcattcaccatgtcggcgtgccccgcgatgactggcagggcccctggggcacgtcgctccggac

[0554] ggccagtcgccacccgcctgaggg**ggcgcgcc** (SEQ ID NO: 192)

[0555] TP3(UTEX1435Δ12脂肪酸去饱和酶转运肽序列)

[0556]

actagt*ATG*gctatcaagacgaacaggcagcctgtggagaagcctccgttcacgatcgggacgctgcgcaaggccatcccc gcgcactgtttcgagcgctcggcgcttcgtg**ggcgcgcc** (SEQ ID NO: 193)

[0557] TP4(UTEX1435异戊烯二磷酸合成酶转运肽序列)

[0558]

actagt*ATG*acgttcggggtcgccctcccggccatgggccgcggtgtctcccttccccggcccagggtcgcggtgcgcgccc agtcggcgagtcaggttttggagagcg**ggcgcgcc** (SEQ ID NO: 194)

[0559] 在藻类转运肽控制下表达C.wrightii FatB2硫酯酶的转化DNA:这些转化构建体(6S-CrbTub_suc2_nr::CrbTub_TP1-CwFatB2_nr-6S,指定为构建体9;构建体6S-CrbTub_suc2_nr::CrbTub_TP2-CwFatB2_nr-6S,指定为构建体10;构建体6S-CrbTub_suc2_nr::CrbTub_TP3-CwFatB2_nr-6S,指定为构建体11以及构建体6S-CrbTub_suc2_nr::CrbTub_TP4-CwFatB2_nr-6S,指定为构建体12)是通过用在以上显示为SpeI和AscI限制性片段(用小写字母、粗体和下划线指示)的相应的藻类转运肽替换来自构建体2的CwFatB2的天然转运肽而产生的。该藻类转运肽序列位于在以上序列中的起始子ATG与AscI位点之间。

[0560] 从表达Cuphea wrightii硫酯酶的株系产生的脂肪酸谱:将用上述构建体转化的株系在允许如前所述的油产生的条件下进行培养。使野生型UTEX 1435在葡萄糖上生长,同时使所有由UTEX 1435的转化而产生的转基因系在蔗糖上生长。针对各测试的构建体,分析了四个转化子对脂肪酸谱的影响。转基因株系的脂肪酸谱显示于以下表25至28中。

[0561] 表达CwFatB1和CwFatB2的拟南芥转基因株系(表25)显示出对C16:0脂肪酸积累连同C14:0和C12:0脂肪酸积累显著的影响(来自Leonard等人,1998,上文)。

[0562] 如从表26中可见,表达CwFatB1(构建体1和构建体3)的、具有天然高等植物转运肽的转基因UTEX 1435系显示出主要对C14:0脂肪酸积累的影响以及在较小程度上对C12:0脂肪酸积累的影响。表达CwFatB2(构建体2和构建体4)的、具有天然高等植物转运肽的转基因UTEX 1435系显示出对C12:0和C14:0脂肪酸积累的显著影响,其中对C12:0的影响比对C14:0的影响更显著。在这两种启动子(CrbTub(构建体1和构建体2)与Amt3(构建体3和构建体4))之间的比较证明:当用Amt3启动子进行驱动时,表达CwFatB1和CwFatB2的转基因系显示出对C10:0、C12:0、C14:0以及C16:0脂肪酸更显著的影响。

[0563] 对其中CwFatB1硫酯酶的表达是由四种不同的藻类叶绿体靶向序列(构建体5、6、7以及8)驱动的转基因系的分析显示:任何这些藻类转运肽都比天然高等植物转运肽更高效地将该硫酯酶靶向该质体(将构建体5-8中的C12:0和C14:0水平(表27)与构建体1中的那些(表26)进行比较)。这些转基因系的进一步分析揭示:在这四种藻类转运肽中,TP2(UTEX1435硬脂酰ACP去饱和酶叶绿体靶向序列)和TP3(UTEX 1435Δ12脂肪酸去饱和酶叶绿体靶向序列)显示出对C12:0与C14:0积累更大的影响。在将该CwFatB1靶向的UTEX 1435中的质体方面这两种转运肽(TP1和TP2)显得更好。

[0564] 对其中CwFatB2硫酯酶的表达是由四种不同的藻类叶绿体靶向序列(构建体9、10、11以及12)驱动的转基因系的分析显示:这些藻类转运肽中仅有一种产生比天然高等植物转运肽优越的性能,即TP-1,如所见的更显著的对C12:0和C14:0脂肪酸的影响(将构建体9与构建体2进行比较)。

[0565] 这些C.wrightii硫酯酶在UTEX 1435中表达时的影响与在拟南芥属中表达时显著

不同。表达CwFatB1和CwFatB2的拟南芥转基因系(表25)显示出对C16:0脂肪酸积累连同C14:0和C12:0脂肪酸积累的显著影响。然而,在UTEX 1435中表达的CwFatB1和CwFatB2并不显示相同水平的对C16:0脂肪酸的影响。在所有UTEX 1435转基因系中(表达CwFatB1,并且一些表达CwFatB2(构建体10、11、12))的C12:C14的比率与表达这种硫酯酶的拟南芥转基因系类似(表29)。然而,与这些拟南芥转基因系相比较,这些UTEX 1435系显示出显著更低的C14:C16比率(表29)。用构建体2、4以及9产生的UTEX 1435转基因系中的C12:C14和C14:C16比率与表达相同硫酯酶的拟南芥转基因系显著不同(表29)。因此,CwFatB1和CwFatB2在UTEX 1435中的表达产生一种油谱,该油谱与在表达相同硫酯酶的拟南芥属转基因系中产生的油谱显著不同。在这些UTEX 1435转基因系中的油谱还与在野生型UTEX 1435中的那些截然不同。

[0566] 最后,由在这个实例中描述的转基因系所产生的修饰的油还与在表达来自加州月桂的C12:0特异性硫酯酶的转基因UTEX 1435系中所产生的富含月桂酸酯的油(先前描述)显著不同。

[0567] 总之,这些数据显示:(1)Cuphea wrightii硫酯酶在UTEX 1435中的表达对脂肪酸谱具有显著影响并且产生独特的油;(2)CwFatB1硫酯酶在UTEX 1435株中的表达导致富含肉豆蔻酸酯的油的产生;(3)CwFatB2硫酯酶在UTEX 1435中的表达导致富含月桂酸酯和肉豆蔻酸酯两者的油的产生;以及(4)就所产生的中链脂肪酸的绝对水平和它们彼此的相对比率两者而言,CwFatB1和FatB2在藻类中的表达所产生的谱与在模式高等植物系统中所产生的那些截然不同。

[0568] 表25.在UTEX 1435、拟南芥野生型(Ath)以及表达CwFatB1(CwFatB1-Ath)和CwFatB2(CwFatB2-Ath)的拟南芥转基因系中的脂肪酸谱(表示为面积%)。

[0569]

样品ID	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3	C20:0	C20:1
UTEX1435	0.01	0.03	0.93	23.83	3.27	61.85	8.08	0.53	0.31	0.08
Ath	0.00	0.00	0.00	8.40	3.80	13.00	29.20	20.10	2.40	19.30
Cw FatB1-Ath	0.00	7.10	24.40	22.80	3.30	4.50	14.10	12.90	3.00	6.00
Cw FatB2-Ath	4.40	16.40	15.30	18.10	3.90	4.90	13.90	13.60	2.80	5.70

[0570] 表26.在表达构建体1、构建体2、构建体3以及构建体4的UTEX 1435转基因系中的脂肪酸谱(表示为面积%)。

[0571]

构建体	样品ID	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3	C20:0	C20:1
构建体1	1A	0.01	0.43	3.17	19.84	1.66	60.34	12.38	0.45	0.23	0.02
	1B	0.01	0.52	3.45	19.81	2.03	60.65	11.42	0.42	0.26	0.02
	1C	0.01	0.59	3.62	20.53	2.24	59.64	11.29	0.42	0.26	0.02
	1D	0.01	0.67	3.92	21.97	1.96	58.62	10.82	0.44	0.24	0.02
构建体2	2A	0.63	7.47	5.64	18.74	2.36	52.11	10.98	0.49	0.28	0.02
	2B	0.82	7.77	5.83	19.84	2.62	51.98	9.24	0.50	0.24	0.02
	2C	0.82	9.57	6.31	18.64	1.66	50.89	10.42	0.43	0.21	0.01
	2D	0.90	10.04	7.11	17.99	2.34	49.03	10.63	0.47	0.26	0.02
构建体3	3A	0.04	2.85	12.09	28.04	2.69	39.02	12.35	1.05	0.25	0.05
	3B	0.03	2.90	13.39	28.01	2.02	41.47	10.01	0.71	0.21	0.05
	3C	0.04	3.30	14.10	27.91	2.09	40.50	9.92	0.71	0.21	0.04
	3D	0.04	3.71	15.10	27.88	2.01	39.56	9.62	0.68	0.21	0.06
构建体4	4A	1.43	11.78	8.87	19.67	2.04	43.80	10.28	0.77	0.21	0.06
	4B	1.39	12.26	9.29	16.88	1.48	44.20	12.18	0.86	0.19	0.08
	4C	1.73	13.42	9.55	18.93	2.05	41.92	10.30	0.81	0.21	0.04
	4D	1.82	14.18	10.07	18.56	1.93	41.22	10.21	0.78	0.19	0.05

[0572] 表27.在表达构建体5、构建体6、构建体7以及构建体8的UTEX 1435转基因系中的脂肪酸谱(表示为面积%)。

[0573]

构建体	样品 ID	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3	C20:0	C20:1
构建体5	5A	0.02	0.62	4.33	23.30	2.00	57.43	10.00	0.59	0.26	0.09
	5B	0.02	0.78	5.11	25.06	2.45	55.65	8.76	0.58	0.26	0.07
	5C	0.02	1.20	7.41	24.48	1.87	52.65	10.18	0.60	0.24	0.08
	5D	0.02	1.33	7.56	24.55	1.87	52.59	9.90	0.54	0.26	0.09
构建体6	6A	0.02	0.56	4.01	24.23	2.50	57.04	9.32	0.57	0.28	0.08
	6B	0.02	0.69	4.97	23.41	1.92	55.59	10.98	0.61	0.25	0.09
	6C	0.02	1.14	7.07	25.05	2.11	53.23	9.23	0.60	0.23	0.07
	6D	0.05	5.10	19.88	21.43	1.29	40.40	9.89	0.63	0.20	0.08
构建体7	7A	0.02	1.39	8.36	25.39	2.00	51.37	9.37	0.55	0.22	0.07
	7B	0.02	1.42	7.59	24.77	2.12	53.17	8.87	0.47	0.23	0.08
	7C	0.02	1.49	7.82	24.87	2.10	52.45	9.14	0.53	0.24	0.08
	7D	0.03	2.15	11.01	25.64	1.85	47.35	9.92	0.51	0.23	0.08
构建体8	8A	0.02	0.81	5.28	23.38	2.03	56.10	10.14	0.54	0.26	0.09
	8B	0.02	0.88	5.77	23.58	1.91	54.91	10.58	0.57	0.24	0.09
	8C	0.02	1.27	7.57	24.28	1.93	52.95	9.88	0.54	0.24	0.08
	8D	0.02	1.43	5.02	21.52	2.63	58.32	9.14	0.52	0.28	0.09

[0574] 表28.在表达构建体9、构建体10、构建体11以及构建体12的UTEX 1435转基因系中的脂肪酸谱(表示为面积%)。

[0575]

构建体	样品 ID	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2	C18:3	C20:0	C20:1
构建体9	9A	0.90	8.96	6.59	19.24	2.20	51.17	8.91	0.56	0.24	0.08
	9B	1.00	9.07	6.31	19.55	2.21	51.43	8.47	0.53	0.25	0.07
	9C	1.06	12.08	8.79	17.57	1.70	46.54	10.29	0.58	0.21	0.07
	9D	1.27	13.05	8.67	17.70	1.78	46.28	9.38	0.57	0.22	0.07
构建体10	10A	0.52	5.55	5.00	20.60	2.03	53.35	10.66	0.71	0.24	0.07
	10B	0.53	5.76	5.16	20.63	1.92	52.80	10.90	0.67	0.23	0.06
	10C	0.47	5.86	5.20	19.54	1.89	53.34	11.41	0.62	0.26	0.06
	10D	0.87	8.59	6.85	19.65	1.98	49.65	10.21	0.69	0.23	0.08
构建体11	11A	0.21	2.48	2.85	20.80	1.99	57.69	11.53	0.70	0.27	0.11
	11B	0.22	2.80	3.01	21.30	1.89	57.28	11.07	0.66	0.27	0.11
	11C	0.29	3.22	3.38	21.33	2.09	56.92	10.42	0.71	0.25	0.05
	11D	0.28	4.01	4.01	18.79	1.69	56.08	12.73	0.64	0.26	0.09
构建体12	12A	0.65	6.01	5.43	21.50	2.10	52.70	9.28	0.75	0.24	0.10
	12B	0.52	6.58	5.62	18.71	1.78	52.59	11.84	0.66	0.26	0.11
	12C	0.78	8.60	6.88	19.01	1.69	50.40	10.45	0.69	0.23	0.08
	12D	0.72	8.75	7.07	17.74	1.54	50.57	11.29	0.68	0.22	0.10

[0576] 表29.表达CwFatB1(构建体1、3、5、6、7、8)和CwFatB2(构建体2、4、9、10、11、12)的UTEX 1435转基因系中连同表达CwFatB1(CwFatB1-Ath)和CwFatB2(CwFatB2-Ath)的拟南芥转基因系中的C12:C14和C14:C16的比率。

[0577]

样品ID	平均12:14比率	平均14:16比率
CwFatB1-Ath	0.291	1.070
构建体1	0.155	0.172
构建体3	0.233	0.489
构建体5	0.250	0.250
构建体6	0.174	0.397

构建体7	0.185	0.345
构建体8	0.190	0.254
CwFatB2-Ath	1.072	0.845
构建体2	1.396	0.332
构建体4	1.365	0.512
构建体9	1.419	0.414
构建体10	1.152	0.277
构建体11	0.938	0.163
构建体12	1.191	0.328

[0578] 实例5:内源的氮依赖性元绿藻属启动子的鉴定

[0579] A.内源的氮依赖性启动子的鉴定和以及表征。

[0580] 使用标准技术从Prototheca moriformis(UTEX 1435)产生cDNA文库。使这些Prototheca moriformis细胞在氮充满条件下生长48小时。然后将5%接种物(v/v)转移至低氮条件并且每24小时收获这些细胞,持续七天。在培养24小时之后,培养基中的氮供应完全耗尽。使用干冰和异丙醇将收集的样品立即冷冻。接着从这些冷冻的细胞团块样品中分离总RNA并且将各样品的一部分保存以用于RT-PCR研究。使从这些样品收获的总RNA的剩余部分经受polyA选择。然后将来自各条件的等摩尔量的经polyA选择的RNA汇集并且用来在载体pcDNA 3.0(Invitrogen公司)中产生cDNA文库。从得到的汇集的cDNA文库随机挑选大致1200个克隆并且使其经受双链测序。从这1200个序列中选择大约68个不同的cDNA并且将其用于设计cDNA特异性引物以供在实时RT-PCR研究中使用。

[0581] 将从保存的细胞团块样品中分离的RNA作为实时RT-PCR研究中的底物,该研究使用以上产生的cDNA特异性引物集。将这种保存的RNA转化成cDNA并且用作用于针对68个基因特异性引物集的每一个的RT-PCR的底物。阈值循环或C_T数目被用来指示所收集的每一RNA样品之中的68个cDNA的每一个的在整个时程的相对转录丰度。将在氮充满与氮耗尽之间的条件下显示出显著增加(大于三倍)的cDNA标记为其表达受氮耗尽上调的潜在基因。如在说明书中所讨论的,氮耗尽/限制在产油微生物中是一个已知的脂肪生成诱导因素。

[0582] 为了从在氮耗尽/限制过程中其表达上调的cDNA中鉴定假定启动子/5'UTR序列,从在氮充满条件下生长的Prototheca moriformis(UTEX 1435)分离总DNA并且然后使用454测序技术(罗氏公司)使其经受测序。使用BLAST将以上根据RT-PCR结果标记为上调的cDNA与由454基因组测序读数产生的装配重叠群进行比较。将cDNA的5'末端映射到特定的重叠群,并且在可能的情况下,使用大于500bp的5'侧翼DNA来推定地鉴定启动子/UTR。接着证实启动子/5'UTR的存在并且使用基因组DNA的PCR扩增将其进行克隆。使用单独的cDNA 5'末端来设计3'引物并且使用454重叠群装配(contig assemblies)的5'末端来设计5'基因特异性引物。

[0583] 作为第一次筛选,将一个假定启动子(从Aat2(铵转运蛋白,SEQ ID NO:63)分离的5'UTR/启动子)克隆到具有原始小球藻硬脂酰ACP去饱和酶转运肽的樟树C14硫酯酶构建体中,从而替换C.sorokinana谷氨酸脱氢酶启动子。这一构建体列作SEQ ID NO:81。为了测试该假定启动子,使用以上描述的方法,将该硫酯酶构建体转化到Prototheca moriformis细胞中,以便在低/无氮条件下通过针对C14/C12脂肪酸增加的筛选来证实实际启动子活性。

可以使用相同方法来完成对从该cDNA/基因组筛选中分离的假定的氮调节的启动子的类似测试。

[0584] 其他的从该cDNA/基因组筛选中分离的假定的氮调节的启动子5'UTR是：

启动子/5'UTR	SEQ ID NO.	增加倍数
FatB/A 启动子/5'UTR	SEQ ID NO: 55	n/a
NRAMP 金属转运蛋白启动子/5'UTR	SEQ ID NO: 56	9.65
Flap 鞭毛-相关蛋白启动子/5'UTR	SEQ ID NO: 57	4.92
SulfRed 亚硫酸盐还原酶启动子/5'UTR	SEQ ID NO: 58	10.91
SugT 糖转运蛋白启动子/5'UTR	SEQ ID NO: 59	17.35
Amt03-铵转运蛋白 03 启动子/5'UTR	SEQ ID NO: 60	10.1
Amt02-铵转运蛋白 02 启动子/5'UTR	SEQ ID NO: 61	10.76
Aat01-氨基酸转运蛋白 01 启动子/5'UTR	SEQ ID NO: 62	6.21
Aat02-氨基酸转运蛋白 02 启动子/5'UTR	SEQ ID NO: 63	6.5
Aat03-氨基酸转运蛋白 03 启动子/5'UTR	SEQ ID NO: 64	7.87
Aat04-氨基酸转运蛋白 04 启动子/5'UTR	SEQ ID NO: 65	10.95
Aat05-氨基酸转运蛋白 05 启动子/5'UTR	SEQ ID NO: 66	6.71

[0585]

[0586] 增加倍数是指在低氮培养基中在培养24小时之后cDNA丰度的增加倍数。

[0587] 为了进一步认识这些假定启动子/5'UTR的潜在调节，选择了八个序列用于进一步测试：(1)FatB/A；(2)SulfRed亚硫酸盐还原酶；(3)SugT糖转运蛋白；(4)Amt02-铵转运蛋白02；(5)Aat01-氨基酸转运蛋白01；(6)Aat03-氨基酸转运蛋白03；(7)Aat04-氨基酸转运蛋白04；以及(8)Aat05-氨基酸转运蛋白05。利用Illumina测序阅读对在以下不同时间点从Prototheca moriformis细胞分离的RNA进行了更高分辨率的转录组分析：T0(种子)；从种子接种后20小时；32小时；48小时；62小时；以及114小时。T0(种子)处的培养基是氮充满的，而在20小时和更长的时间点处，培养基含有很少或不含氮。然后将从每一时间点分离的RNA产生的装配的转录重叠群独立地与先前鉴定的八个转录本中的每一个进行序列比对(blasted)。结果汇总在以下表30中。

[0588] 表30.八个假定启动子/5'UTR的转录组表达谱。

[0589]

cDNA	TS	T20	T32	T48	T62	T114

[0590]

氨基酸转运蛋白_01	绝对的	98	96	321	745	927	1300
	相对的	1	0.98	3.28	7.61	9.47	13.28
氨基酸转运蛋白_03	绝对的	7	21	51	137	102	109
	相对的	1	2.95	7.2	19.42	14.47	15.45
氨基酸转运蛋白_04	绝对的	1	6	25	90	131	160
	相对的	1	5.16	21.29	74.97	109.35	133.31
氨基酸转运蛋白_05	绝对的	109	88	123	210	214	273
	相对的	1	0.81	1.13	1.93	1.97	2.51
铵转运蛋白_02	绝对的	683	173	402	991	1413	1397
	相对的	1	0.25	0.59	1.45	2.07	2.04
fatA/B-1_cDNA	绝对的	13	36	654	617	544	749
	相对的	1	2.8	51.57	48.65	42.9	59.1
糖转运蛋白_01	绝对的	25	25	106	261	266	251
	相对的	1	1	4.22	10.4	10.63	10
亚硫酸盐还原酶_01	绝对的	634	238	138	145	163	155
	相对的	1	0.38	0.22	0.22	0.26	0.24

[0591] 根据以上总结的结果，这些转录本中的一些显示出随时间增加的积累，虽然有趣的是亚硫酸盐还原酶mRNA显示出随时间明显减少的mRNA积累。

[0592] 将这八个假定启动子/5'UTR区域克隆在樟树硫酯酶编码区的上游，该樟树硫酯酶编码区的天然转运肽被去掉并且用来自原始小球藻(UTEX 250)硬脂酰ACP去饱和酶的转运肽取代。使用来自6S基因组序列的DNA将每一个假定启动子/5'UTR区域构建体经由同源重组引入到Prototheca moriformis UTEX 1435中。同样包含在该构建体中的是来自酿酒酵母的suc2蔗糖转化酶基因，该蔗糖转化酶基因用于在含有蔗糖的培养基/平板上进行阳性克隆选择。针对Aat01的构建体的相关部分的cDNA序列作为SEQ ID NO:67在序列表中列出。对于其他构建体而言，使用了相同的骨架，仅有的变化是该假定启动子/5'UTR序列。产生了另外的控制型转基因株系，其中使用莱茵衣藻β-微管蛋白启动子来驱动樟树硫酯酶基因的表达。这个启动子已经显示出驱动感兴趣的基因的组成型表达，并且因此提供了一种有用的对照，当受到所测试的不同的假定N-调节的启动子/5'UTR驱动时，相对于该对照来测量相同的硫酯酶的信息。

[0593] 一旦产生这些转基因克隆，则进行三个单独的实验。前两个实验通过经由RT-PCR、脂肪酸谱以及培养物上清液中的氨水平而测量稳态硫酯酶mRNA水平来评定所有八个假定启动子的潜在氮调节能力(nitrogen regulatability)。最初使克隆在28℃在富氮种子培养基(1g/L硝酸铵—15mM作为氨的氮，4g/L酵母提取物)中在搅拌(200rpm)下生长24至48小时，此时将200D单位(A_{750})用于接种50ml的低氮培养基(0.2g/L硫酸铵—3mM作为氨的氮，0.2g/L酵母提取物)。每24小时对细胞进行取样，持续6天，并且恰好在转换至低氮条件之前也采集样品。然后使用Trizol试剂(根据生产商建议的方法)将来自每一样品的细胞的一部分用于总RNA提取。氨测定揭示，在低氮培养基的情况下在24小时之后在上清液中的氨水平

落到检测限(约100μM)以下。

[0594] 对于实时RT-PCR而言,针对每一时间点将所有RNA水平标准化到在Prototheca moriformis(UTEX 1435)中表达的内部对照RNA的水平。该内部对照RNA(称作cd189)是ARG9基因的产物,该基因编码N-乙酰鸟氨酸转氨酶。在这些实验中用于实时RT-PCR的引物集是:

[0595]

基因特异性	引物序列5'-3'	SEQ ID NO:
樟树TE正向	TACCCCGCCTGGGGCGACAC	SEQ ID NO: 68
樟树TE反向	CTTGCTCAGGCGGCGGGTGC	SEQ ID NO: 69
cd189正向	CCGGATCTCGGCCAGGGCTA	SEQ ID NO: 70
cd189反向	TCGATGTCGTGCACCGTCGC	SEQ ID NO: 71

[0596] 还产生了来自每一时间点的每一转化子的脂质谱并且将其与RT-PCR结果进行比较。基于氨水平、RT-PCR结果以及C12-C14脂肪酸水平的变化,可以得出结论:氨基酸转运蛋白01(Aat-01)、氨基酸转运蛋白04(Aat-04)、以及氨基酸转运蛋白02(Amt-02)序列确实含有功能性氮可调节启动子/5'UTR。

[0597] 根据RT-PCR结果,Aat-01显示出驱动稳态樟树硫酯酶mRNA水平高达比对照(莱茵衣藻β-微管蛋白启动子)高四倍的能力。mRNA水平还与氮限制和C12-C14脂肪酸水平的显著增加相关。这些结果证明,与Aat-01启动子相关联的5'UTR在脂质生物合成下驱动蛋白质合成方面可能比对照莱茵衣藻启动子更有效。类似于Aat-01启动子,Aat-04启动子能够驱动mRNA积累高达比莱茵衣藻对照启动子驱动的mRNA积累高五倍。然而,Aat-04启动子构建体仅仅产生了适度的影响C12-C14脂肪酸水平的能力。这些数据证明,Aat-04启动子显然可受氮耗尽调节,但是与该启动子关联的UTR作为转录增强子可能功能不良。最后,Amt-02启动子与Aat-01启动子类似,因为它能够驱动mRNA积累高达比对照启动子驱动的mRNA积累高三倍。mRNA水平还与氮限制和C12-C14脂肪酸水平的显著增加相关。总之,证明了这些启动子中的所有三个是受氮调节的。

[0598] B.铵转运蛋白3(amt03)启动子的进一步表征以及不同硫酯酶的表达。

[0599] 如上所述,鉴定了称作铵转运蛋白02和03(amt02和amt03)的部分cDNA。连同这两部分cDNA,还鉴定了称作铵转运蛋白01(amt01)的一个第三部分cDNA。将该部分cDNA与假定的翻译的氨基酸序列的比对进行了比较。结果显示在这三个序列中amt01关系较远,而amt02与amt03仅有一个单一氨基酸不同。

[0600] 启动子/5'UTR最初是通过将这些部分cDNA序列与上述Roche 454基因组DNA装配和Illumina转录组装配进行序列比对用计算机模拟而产生的。对显示出与编码amt01、amt02、以及amt03的cDNA一致性的转录重叠群进行了鉴定,然而,这些转录重叠群不能在这三个mRNA之间区分,因为这些重叠群包含所有这三者共享的序列。Roche 454基因组DNA装配给出了与amt02和amt03cDNA序列的匹配(hits)并且包含N-端蛋白质序列。进行PCR来克隆5'侧翼区。用来确认克隆amt02和amt03启动子/UTR的PCR引物是:

[0601] Amt03正向:5'-GGAGGAATTCGGCCGACAGGACGCGCGTCA-3'(SEQ ID NO:85)

[0602] Amt03反向:5'-GGAGACTAGTGGCTGCGACCGGCCTGTG-3'(SEQ ID NO:86)

[0603] Amt02正向:5'-GGAGGAATTCTCACCAGCGGACAAAGCACCG-3'(SEQ ID NO:87)

[0604] Amt02反向:5'-GGAGACTAGTGGCTGCGACCGGCCTCTGG-3'(SEQ ID NO:88)

[0605] 在两情形中，这些5’与3’引物包含对于预期地克隆进入表达载体有用的限制性酶切位点以验证这些启动子/5’UTR区域的功能。

[0606] 在通过这一组合的计算机模拟以及基于PCR的方法而克隆的DNA与编码amt02(SEQ ID NO:61)以及amt03(SEQ ID NO:60)的原始cDNA进行了逐对比对。这些比对的结果显示了这些原始cDNA与这些克隆的基因组序列之间的显著不同，这说明铵转运蛋白似乎代表一个多样的基因家族。另外，amt03的基于该组合方法的启动子/5’UTR克隆与原始amt03序列不同，而amt02序列则是一致的。进行了进一步的实验来表征该amt03启动子/UTR序列(SEQ ID NO:89)并且描述于下。

[0607] 通过克隆这个假定启动子/UTR序列测试了以上鉴定的amt03启动子/UTR序列(SEQ ID NO:89)，以便驱动四种不同硫酯酶的表达。该表达盒包含针对基因组的6S基因座的上游和下游同源重组序列(分别是SEQ ID NOs:82和84)。该盒还包含酿酒酵母SUC2蔗糖转化酶cDNA以便能够在含蔗糖培养基上进行阳性克隆选择。该蔗糖转化酶表达受莱茵衣藻β-微管蛋白启动子的驱动并且还包含普通小球藻硝酸盐还原酶3’UTR。然后将该amt03启动子/UTR序列克隆在该蔗糖转化酶盒的下游，继之为来自以下四种硫酯酶基因之一的框内硫酯酶cDNA序列：(1)来自樟树的C14硫酯酶；(2)来自加州月桂的C12硫酯酶；(3)来自美洲榆的C10-C16硫酯酶；或(4)来自蓖距花的C10硫酯酶并且还包含普通小球藻硝酸盐还原酶3’UTR。该C14樟树硫酯酶、C12加州月桂硫酯酶、C10-C16美洲榆硫酯酶都包含来自原始小球藻硬脂酰ACP去饱和酶的转运肽。该C10蓖距花硫酯酶包含来自Prototheca moriformis Δ12脂肪酸去饱和酶(FAD)的转运肽。在所有情况下，对这些序列针对在Prototheca moriformis中表达进行了密码子优化。针对前述的硫酯酶构建体的序列描述于序列表中：

[0608]

amt03 *启动子/UTR::樟树硫酯酶构建体*	SEQ ID NO: 90
樟树硫酯酶构建体	SEQ ID NO: 91
加州月桂硫酯酶构建体	SEQ ID NO: 92
美洲榆硫酯酶构建体	SEQ ID NO: 93
蓖距花硫酯酶构建体	SEQ ID NO: 94

[0609] 通过如在实例2中描述的进入野生型Prototheca moriformis细胞的生物射弹转化方法产生了转基因系并且在含有蔗糖的平板/培养基上进行选择。然后针对其脂肪酸谱被改变的程度来筛选阳性系。然后另外地分析了用上述四个构建体的每一个进行转化而得到的四个系。系76表达樟树C14硫酯酶，系37表达加州月桂C12硫酯酶，系60表达美洲榆C10-C16硫酯酶，并且系56表达蓖距花C10硫酯酶。使每个系在含有蔗糖作为唯一碳源的培养基中生长48小时并且在14、24、36以及48小时取出细胞样品(种子培养)用于通过直接酯交换到脂肪酸甲酯和随后的借助于GC-FID的分析(以上所述)来确定脂肪酸谱并且用于分离总RNA。在48小时结束时，将这些细胞用来接种不具有氮或具有低水平氮(含有蔗糖作为唯一碳源)的、维持在pH 5.0(柠檬酸盐缓冲的，0.05M终浓度)或pH 7.0(HEPES缓冲的，0.1M终浓度)的培养基。在12、24、72以及108小时取出培养物样品(脂质产生)用于脂肪酸谱分析和总RNA分离。这些培养物的氨测定揭示，在低氮培养基的情况下在24小时之后氨水平落到检测界限(大约100μM)以下。

[0610] 对从在以上每一时间点收集的总RNA进行硫酯酶的mRNA水平实时RT-PCR测定并且

将所有mRNA水平相对于内部对照RNA(cd189)的水平进行标准化。在实时PCR中使用的引物集显示于下表31：

[0611] 表31.用于实时PCR的引物集。

[0612]

基因特异性	引物序列 5'-3'	SEQ ID NO:
樟树 TE 正向	TACCCCGCCTGGGGCGACAC	SEQ ID NO: 68
樟树 TE 反向	CTTGCTCAGGCGGCGGGTGC	SEQ ID NO: 69
加州月桂 TE 正向	CTGGGCGACGGCTTCGGCAC	SEQ ID NO: 95
加州月桂 TE 反向	AAGTCGCGGCGCATGCCGTT	SEQ ID NO: 96
美洲榆 TE 正向	CCCAGCTGCTCACCTGCACC	SEQ ID NO: 97
美洲榆 TE 反向	CACCCAAGGCCAACGGCAGCGCCGTG	SEQ ID NO: 98
萼距花 TE 正向	TACCCCGCCTGGGGCGACAC	SEQ ID NO: 99
萼距花 TE 反向	AGCTTGGACAGGCGGCGGGT	SEQ ID NO: 100
cd189 反向	TCGATGTCGTGCACCGTCGC	SEQ ID NO: 71
cd189 正向	CCGGATCTCGGCCAGGGCTA	SEQ ID NO: 70

[0613] 来自种子培养阶段的每一时间点的脂肪酸谱的结果显示出极小的来自硫酯酶的影响。随着脂质产生阶段的开始，这些脂肪酸谱受到显著影响，与维持在pH 5.0的培养物比较，对于维持在pH 7.0的培养物而言这些增加显著得多。虽然pH在7.0与pH 5.0目标脂肪酸积累之间的差异的大小随着每一测试的硫酯酶而变化，但是总体效果是相同的：在pH 5.0生长的细胞显示出显著较低水平的目标脂肪酸积累，但是与对照野生型细胞相比较多。

[0614] 从这些相同的样品分离的RNA的分析与脂肪酸谱数据非常好地相关联，因为培养pH对每一硫酯酶的稳态mRNA水平具有明显影响。将脂肪酸积累数据与mRNA数据一起考虑，由amt03启动子/UTR驱动的硫酯酶基因表达的pH调节在转录水平、mRNA稳定性或两者上被明显介导。另外，观察到加州月桂mRNA的稳态水平与萼距花mRNA的稳态水平相比低四个对数级。这一观察与这样的假设一致：单独的mRNA序列可能在控制表达中起作用。这些数据暗示通过amt03家族转运蛋白的在Prototheca moriformis中的铵摄入与pH直接偶联。

[0615] 对于从用驱动美洲榆C10-C16硫酯酶表达的构建体amt03启动子/UTR转化Prototheca moriformis细胞而产生的十二个系进行了另外的脂肪酸谱分析。系60，如上所述，是以下分析的一部分。以下表32显示了与野生型对照一起分析的十二个系中的三个的脂质谱。

[0616] 表32.包含由amt03启动子/UTR驱动的美洲榆TE的转化子的脂肪酸谱。

[0617]

面积%	C8:0	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2	总饱和物
野生型	0.00	0.01	0.04	1.27	27.20	3.85	58.70	7.18	32.36
系40	2.38	20.61	3.41	28.41	29.92	1.91	8.57	3.74	86.64
系44	1.50	20.16	4.44	31.88	26.66	1.88	6.95	5.42	86.50
系60	0.98	14.56	3.15	27.49	31.76	2.14	12.23	6.36	80.06

[0618] 如在上表中展示，在株系40的情形中与野生型相比，总饱和物的水平比野生型的

总饱和物水平大幅增加了2.6倍多(所分析的来自这十二个株系的总饱和物的范围是从大约63％至超过86％)。另外,当在这些水平表达时,美洲榆硫酯酶显著降低了不饱和物的水平,尤其是C18:1以及C18:2(参见系40和44),在系44的情况下,C18:1水平与野生型相比降低了8倍以上。同样,美洲榆硫酯酶(由amt03启动子驱动)大大增加了中链脂肪酸的水平。系44显示出大于56％的C10:0-C14:0水平(比在野生型株系中所见到的水平大约高42倍)以及大于57％的C8:0-C14:0水平。用Amt03启动子驱动美洲榆硫酯酶表达的构建体转化的另外的株具有代表性的脂质谱:0.23％C8:0、9.64％C10:0、2.62％C12:0、31.52％C14:0、37.63％C16:0、5.34％C18:0、7.05％C18:1、以及5.03％C18:2,具有在86.98％的总饱和物百分比。

[0619] 从用驱动蓖距花C10硫酯酶(SEQ ID NO:94)表达的构建体amt03启动子/UTR(SEQ ID NO:89)转化Prototheca moriformis细胞而产生了另外的脂质谱。选择表达这种构建体的阳性克隆并且使其在pH 7.0条件下生长。来自一个阳性克隆的代表性脂质谱是:9.87％C8:0、23.97％C10:0、0.46％C12:0、1.24％C14:0、10.24％C16:0、2.45％C18:0、42.81％C18:1、以及7.32％C18:2。这个克隆具有33.84的C8-C10百分比。

[0620] 总之,这些数据说明,amt03启动子/UTR、以及其他与其类似的启动子可以用作一种受紧密调节的启动子,这对于表达一种具有潜在毒性的化合物可能是特别有用的并且需要基因表达的严格增强。Prototheca moriformis在宽范围的pH方案(至少pH 5.0至7.0)下生长的能力使这种微生物与调节元件如amt03启动子/UTR结合特别有用。另外,以上的脂质谱数据证明amt03启动子/UTR驱动基因表达的令人印象深刻的能力。

[0621] 实例6:改变微藻Prototheca moriformis中的饱和脂肪酸的水平

[0622] A.通过基因敲除途径减少硬脂酰ACP去饱和酶和Δ12脂肪酸去饱和酶表达

[0623] 作为使用基于生物信息学的方法基于cDNA的基因组筛选(来自Prototheca moriformis(UTEX 1435)的基因组DNA的Illumia转录组和Roche 454测序)的部分,鉴定到涉及脂肪酸去饱和的特定两组基因:硬脂酰ACP去饱和酶(SAD)和Δ12脂肪酸去饱和酶(Δ12FAD)。硬脂酰ACP去饱和酶酶是脂质合成途径的部分,并且它们发挥作用以将双键导入脂肪酰基链中,例如,从C18:0脂肪酸合成C18:1脂肪酸。Δ12脂肪酸去饱和酶也是脂质合成途径的部分,并且它们发挥作用以将双键导入已经不饱和的脂肪酸中,例如,从C18:1脂肪酸合成C18:2脂肪酸。使用基于在生物信息学工作期间已经鉴定的两类脂肪酸去饱和酶基因的探针,Southern印迹分析表明每一类去饱和酶基因可能由多个家族成员组成。另外,编码硬脂酰ACP去饱和酶的基因落入两个不同的家族。基于这些结果,设计了三个基因破坏构建体以通过靶向每个去饱和酶酶家族内的更高度保守的编码区潜在地破坏多个基因家族成员。

[0624] 使用以下序列,设计3个同源重组靶向构建体:(1)Δ12脂肪酸去饱和酶(d12FAD)家族成员的编码序列的高度保守部分和(2)靶向两个不同SAD家族的每个成员的两个构建体,每个构建体具有来自每个家族的编码序列保守区域。这种策略将选择标记基因(赋予水解蔗糖的能力的来自酿酒酵母的suc2蔗糖转化酶盒)嵌入这些高度保守的编码区(靶向多个家族成员)内,而不是经典的基因替换策略,其中同源重组将靶向该靶向基因的侧翼区。

[0625] 通过生物射弹转化,使用上文描述的方法,将所有构建体导入细胞中,并且将构建体在射击到细胞中之前线性化。在含有蔗糖的平板/培养基上选择转化子,并且使用上述方

法测定脂质谱的变化。下文列出了来自三个靶向构建体中的每一个的相关序列。

[0626]

描述	SEQ ID NO:
来自源于靶向构建体的 d12FAD 的编码区的 5'序列	SEQ ID NO: 72
来自源于靶向构建体的 d12FAD 的编码区的 3'序列	SEQ ID NO: 73
d12FAD 靶向构建体 cDNA 序列	SEQ ID NO: 74
来自 SAD2A 的编码区的 5'序列	SEQ ID NO: 75
来自 SAD2A 的编码区的 3'序列	SEQ ID NO: 76
SAD2A 靶向构建体 cDNA 序列	SEQ ID NO: 77
来自 SAD2B 的编码区的 5'序列	SEQ ID NO: 78
来自 SAD2B 的编码区的 3'序列	SEQ ID NO: 79
SAD2B 靶向构建体 cDNA 序列	SEQ ID NO: 80

[0627] 挑出来自用每种构建体转化的代表性阳性克隆，然后确定这些克隆的脂质谱(以面积％表示)并且将其汇总在以下表33中。

[0628] 表33.去饱和酶敲除的脂质谱。

[0629]

脂肪酸	d12FAD KO	SAD2A KO	SAD2B KO	wt UTEX 1435
C8:0	0	0	0	0
C10:0	0.01	0.01	0.01	0.01

[0630]

C12:0	0.03	0.03	0.03	0.03
C14:0	1.08	0.985	0.795	1.46
C16:0	24.42	25.335	23.66	29.87
C18:0	6.85	12.89	19.555	3.345
C18:1	58.35	47.865	43.115	54.09
C18:2	7.33	10.27	9.83	9.1
C18:3 α	0.83	0.86	1	0.89
C20:0	0.48	0.86	1.175	0.325

[0631] 每种构建体对所希望类别的脂肪酸产生可度量的影响，并且在所有3种情况下，C18:0水平明显增加，尤其随着两个SAD敲除的情况下。进一步比较来自SAD敲除的多个克隆表明，SAD2B敲除系在C18:1脂肪酸方面比随SAD2A敲除系所观察到的C18:1脂肪酸水平具有明显更大的减少。

[0632] 使用上文描述的方法在Prototheca moriformis背景中产生另外的Δ12脂肪酸去饱和酶(FAD)敲除。为了鉴定Δ12FAD的潜在同源物，使用了以下引物，以便扩增编码假定FAD的基因组区域：

[0633] 引物1 5'-TCACTTCATGCCGGCGGTCC-3' SEQ ID NO:101

[0634] 引物2 5'-GCGCTCCTGCTTGGCTCGAA-3' SEQ ID NO:102

[0635] 从使用以上引物的Prototheca moriformis基因组DNA的基因组扩增产生的序列是高度相似的，但是表明多个基因或Δ12FAD等位基因存在于无绿藻属中。

[0636] 基于这种结果，设计了两个基因破坏构建体以寻求失活一个或多个Δ12FAD基因。该策略会将蔗糖转化酶(来自酿酒酵母的suc2)盒(因此赋予水解蔗糖能力作为选择标记)嵌入高度保守的编码区中，而不是使用经典的基因替换策略。第一构建体，称作pSZ1124，含有在莱茵衣藻β-微管蛋白启动子和普通小球藻硝酸盐还原酶3'UTR侧翼的5'和3'基因组靶向序列(酿酒酵母suc2盒)，其中所述莱茵衣藻β-微管蛋白启动子驱动酿酒酵母suc2基因的表达。第二构建体，称作pSZ1125，含有在莱茵衣藻β-微管蛋白启动子和普通小球藻硝酸盐还原酶3'UTR侧翼的5'和3'基因组靶向序列，其中所述莱茵衣藻β-微管蛋白启动子驱动酿酒酵母suc2基因的表达。构建体的相关序列列出在序列表中：

[0637]

pSZ1124 (FAD2B) 5'基因组靶向序列	SEQ ID NO:103
pSZ1124 (FAD2B) 3'基因组靶向序列	SEQ ID NO:104
酿酒酵母 suc2 盒	SEQ ID NO:105
pSZ1125 (FAD2C) 5'基因组靶向序列	SEQ ID NO:106
pSZ1125 (FAD2C) 3'基因组靶向序列	SEQ ID NO:107

[0638] 将pSZ1124和pSZ1125分别导入Prototheca moriformis背景，并且基于水解蔗糖的能力选择阳性克隆。表34汇总了在使用pSZ1124和pSZ1125打靶载体的两个转基因系中所获得的脂质谱(以面积%表示，使用上文描述的方法产生)。

[0639] 表34.Δ12FAD敲除的脂质谱

[0640]

	C10:0	C12:0	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3 α

[0641]

亲本	0.01	0.03	1.15	26.13	1.32	4.39	57.20	8.13	0.61
FAD2B	0.02	0.03	0.80	12.84	1.92	0.86	74.74	7.08	0.33
FAD2C	0.02	0.04	1.42	25.85	1.65	2.44	66.11	1.39	0.22

[0642] 含有FAD2B(pSZ1124)的转基因构建体在脂质谱方面产生十分有趣和出乎意料的结果，在于预期下降的C18:2水平仅下降大约一个面积%。然而，C18:1脂肪酸水平显著地增加，几乎完全以显著下降的C16:0水平为代价。含有FAD2C(pSZ1125)的转基因构建体也产生脂质谱变化：C18:2的水平显著降低，同时C18:1水平相应增加。

[0643] 牛脂模拟

[0644] 将从如以上所述的SAD2B敲除实验产生的一个阳性克隆选择为用作进一步引入C14偏好型脂肪酰基-ACP硫酯酶基因的背景。引入樟树C14偏好型硫酯酶的构建体包含针对6S基因组区域的靶向序列(允许转化DNA通过同源重组进行靶向整合)并且该表达构建体包含驱动具有普通小球藻硝酸盐还原酶3'UTR的neoR基因表达的莱茵衣藻β-微管蛋白启动子，随后是驱动具有原始小球藻硬脂酰ACP去饱和酶转运肽的、具有一个第二普通小球藻硝酸盐还原酶3'UTR的密码子优化的樟树硫酯酶的表达的一个第二莱茵衣藻β-微管蛋白启动子。该5'6S基因组供体序列在SEQ ID NO:82中列出；该3'6S基因组供体序列在SEQ ID NO:

84中列出,并且该樟树硫酯酶的相关表达构建体在SEQ ID NO:83中列出。

[0645] 使用上述的生物射弹方法进行转化并且允许这些细胞在含有2%蔗糖的平板上恢复24小时。在此时间之后,将这些细胞再悬浮并且重新铺板在含有2%蔗糖和50μg/ml G418的用于选择的平板上。从所产生的阳性克隆中选出九个克隆用于脂质生产和脂质谱分析。将这九个转基因克隆(具有SAD2B KO并且表达樟树C14偏好型硫酯酶)如上述进行培养并且针对脂质谱进行分析。这些结果汇总在以下表35中。动物脂的脂质谱也包括在下表35中(国家研究委员会1976:动物产品的脂肪含量和组成(National Research Council 1976:Fat Content and Composition of Animal Product))。

[0646] 表35.硫酯酶转化的克隆的脂质谱。

[0647]

	C10:0	C12:0	C14:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3	C20
SAD2BKO 樟树 TE 克隆 1	0.01	0.33	6.13	24.24	0.19	11.08	42.03	13.45	0.98	0.73
SAD2BKO 樟树 TE 克隆 2	0.01	0.16	3.42	23.80	0.40	9.40	50.62	10.2	0.62	0.70
SAD2BKO 樟树 TE 克隆 3	0.01	0.20	4.21	25.69	0.40	7.79	50.51	9.37	0.66	0.63
SAD2BKO 樟树 TE 克隆 4	0.01	0.21	4.29	23.57	0.31	9.44	50.07	10.07	0.70	0.70

[0648]

SAD2BKO 樟树 TE 克隆 5	0.01	0.18	3.87	24.42	0.32	9.24	49.75	10.17	0.71	0.71
SAD2BKO 樟树 TE 克隆 6	0.01	0.28	5.34	23.78	0.33	9.12	49.12	10.00	0.68	0.70
SAD2BKO 樟树 TE 克隆 7	0.01	0.15	3.09	23.07	0.32	10.08	51.21	10.00	0.66	0.74
SAD2BKO 樟树 TE 克隆 8	0.01	0.29	5.33	24.62	0.37	7.02	49.67	10.74	0.69	0.70
SAD2BKO 樟树 TE 克隆 9	0.01	0.12	2.74	25.13	0.30	10.17	50.18	9.42	0.71	0.71
wt UTEX 1435	0.01	0.02	0.96	23.06	0.79	3.14	61.82	9.06	0.46	0.27
SAD2BKO	0.01	0.03	0.80	23.66	0.13	19.56	43.12	9.83	1.00	1.18
动物脂	0.00	0.00	4.00	26.00	3.00	14.00	41.00	3.00	1.00	0.00

[0649] 如在表35中所见,这些转基因系的脂质谱与动物脂的脂质谱十分类似。总的来看,这些数据证明了将特定转基因背景(在这种情况下是SAD2B敲除)与C14偏好型硫酯酶(来自樟树)结合来产生可以产生类似于动物脂的脂质谱的转基因藻株的效用。

[0650] B.下调无绿藻细胞中Δ12去饱和酶基因(FADc)的RNAi途径

[0651] 将通过RNAi下调FADc(Δ12去饱和酶基因)基因表达的载体导入到Prototheca moriformis UTEX 1435遗传背景中。使用酿酒酵母suc2蔗糖转化酶基因作为选择标记,从而赋予阳性克隆在作为唯一碳源的蔗糖上生长的能力。第一个类型的构建体利用FADc编码区第一外显子的一部分,这个部分以顺式方式连接于其第一内含子,后接处于反方向的第一外显子重复单元。当表达为mRNA时,这种类型的构建体理论上导致发夹RNA的形成。产生了这种第一类型的两个构建体,一个构建体由Prototheca moriformis Amt03启动子(SEQ ID NO:89)驱动,称作pSZ1468,并且第二构建体由莱茵衣藻β-微管蛋白启动子(SEQ ID NO:114)驱动,称作pSZ1469。第二个类型的构建体利用处于反义方向的大FADc外显子2,由Prototheca moriformis Amt03启动子(SEQ ID NO:89)驱动,称作pSZ1470,或者由莱茵衣藻β-微管蛋白启动子(SEQ ID NO:114)驱动,称作pSZ1471。所有4个构建体具有酿酒酵母suc2蔗糖转化酶盒(SEQ ID NO:159)和针对用于整合至核基因组中的6S基因组区域的5'(SEQ ID NO:82)和3'(SEQ ID NO:84)同源重组靶向序列(在构建体的侧翼)。每种RNAi构建体的FADc部分的序列连同每种构建体的相关部分在序列表中列出为:

[0652]

描述	SEQ ID NO:
pSZ1468 FADc RNAi 发夹盒	SEQ ID NO: 163
pSZ1468 构建体的相关部分	SEQ ID NO: 164
pSZ1469 FADc RNAi 发夹盒	SEQ ID NO: 165
pSZ1469 构建体的相关部分	SEQ ID NO: 166
pSZ1470 FADc 外显子 2 RNAi 盒	SEQ ID NO: 167
pSZ1470 构建体的相关部分	SEQ ID NO: 168
pSZ1471 FADc 外显子 2 RNAi 盒	SEQ ID NO: 169
pSZ1471 构建体的相关部分	SEQ ID NO: 170

[0654] 将这4个构建体的每一个转化到Prototheca moriformis背景中,并且使用含有蔗糖作为唯一碳源的平板筛选阳性克隆。从每个转化中挑取阳性克隆,并且选择亚组以确定pSZ1468、pSZ1469、pSZ1470和pSZ1471中所含的发夹和反义盒对脂肪酸谱的影响。在产生脂质的条件下使从每个转化中选择的克隆生长,并且使用如上文所述的直接酯交换法确定脂质谱。来自每个转化的代表性脂质谱汇总在以下表36中。野生型1和2细胞是未转化的Prototheca moriformis细胞,它们作为阴性对照与每个转化子进行操作。

[0655] 表36.含有RNAi构建体以下调Δ12去饱和酶基因(FADc)表达的Prototheca moriformis细胞的脂质谱。

[0656]

株	C10:0	C12:0	C14:0	C16:0	C18:0	C18:1	C18:2
野生型 1	0.01	0.03	1.20	27.08	4.01	57.58	7.81
pSZ1468 克隆 A	0.01	0.04	1.33	25.95	3.68	65.60	1.25
pSZ1468 克隆 B	0.01	0.03	1.18	23.43	2.84	65.32	4.91
pSZ1468 克隆 C	0.01	0.04	1.34	23.18	4.27	63.65	5.17
pSZ1468 克隆 D	0.01	0.03	1.24	23.00	3.85	61.92	7.62
pSZ1470 克隆 A	0.01	0.03	1.23	24.79	4.33	58.43	8.92
pSZ1470 克隆 B	0.01	0.03	1.26	24.91	4.14	57.59	9.64
pSZ1470 克隆 C	0.01	0.03	1.21	23.35	4.75	58.52	9.70
野生型 2	0.01	0.03	0.98	24.65	3.68	62.48	6.26
pSZ1469 克隆 A	0.01	0.03	1.05	21.74	2.71	71.33	1.22
pSZ1469 克隆 B	0.01	0.03	1.01	22.60	2.98	70.19	1.27
pSZ1469 克隆 C	0.01	0.03	1.03	19.82	2.38	72.95	1.82
pSZ1469 克隆 D	0.01	0.03	1.03	20.54	2.66	70.96	2.71
pSZ1471 克隆 A	0.01	0.03	1.03	18.42	2.63	66.94	8.55
pSZ1471 克隆 B	0.01	0.03	0.94	18.61	2.58	67.13	8.66
pSZ1471 克隆 C	0.01	0.03	1.00	18.31	2.46	67.41	8.71
pSZ1471 克隆 D	0.01	0.03	0.93	18.82	2.54	66.84	8.77

[0657] 以上总结的结果显示,发夹构建体pSZ1468和pSZ1469显示出特定的预期表型,即,分别与野生型1和野生型2相比,C18:2脂肪酸水平降低和C18:1脂肪酸水平增加。反义构建体pSZ1470和pSZ1471不导致C18:2脂肪酸水平下降,但是当分别与野生型1和野生型2相比时,反而显示略微增加,和C16:0脂肪酸水平略微下降。

[0658] C.外源硬脂酰-ACP去饱和酶的表达

[0659] 将油橄榄硬脂酰-ACP去饱和酶(GenBank登录号AAB67840.1)导入到Prototheca moriformis UTEX 1435遗传背景中。表达构建体含有针对用于整合至核基因组中的6S基因组区域的5'(SEQ ID NO:82)和3'(SEQ ID NO:84)同源重组靶向序列(在构建体侧翼)和在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。油橄榄硬脂酰-ACP去饱和酶编码区在Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR的控制下,并且将天然转运肽替换为原始小球藻硬脂酰-ACP去饱和酶转运肽(SEQ ID NO:49)。密码子优化的cDNA序列和氨基酸序列(具有替换的转运肽)在序列表中分别作为SEQ ID NO:171和SEQ ID NO:172列出。将整个油橄榄SAD表达盒称作pSZ1377并且转化到Prototheca moriformis遗传背景中。在含有蔗糖作为唯一碳源的平板上筛选阳性克隆。选择阳性克隆的亚组并且使其在脂质产生条件下生长,并且使用如上文所述的直接酯交换法确定脂质谱。所选择的克隆的脂质谱汇总在以下表37中。

[0660] 表37.油橄榄硬脂酰-ACP去饱和酶转基因Prototheca moriformis细胞的脂质谱。

[0661]

株	C14:0	C16:0	C18:0	C18:1	C18:2
野生型	0.88	22.82	3.78	64.43	6.54
pSZ1377 克隆 A	0.94	18.60	1.50	69.45	7.67
pSZ1377 克隆 B	0.93	18.98	1.35	69.12	7.67
pSZ1377 克隆 C	0.93	19.01	2.31	68.56	7.43

[0662] 以上总结的结果表明,导入异源去饱和酶(在这种情况下来自油橄榄的硬脂酰-ACP去饱和酶)可以导致更高水平的C18:1脂肪酸和C18:0和C16:0脂肪酸水平的同时下降。

[0663] 实例7:工程化无绿藻属以生产羟化脂肪酸

[0664] 将蓖麻油酸12-羟化酶(Rc12hydro)(GenBank登录号AAC49010.1)导入到Prototheca moriformis UTEX 1435遗传背景中。该表达构建体含有针对用于整合至核基因组中的6S基因组区域的5'(SEQ ID NO:82)和3'(SEQ ID NO:84)同源重组靶向序列(在构建体侧翼)和在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。该蓖麻油酸12-羟化酶编码区在Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR的控制下。在序列表中分别作为SEQ ID NO:173和SEQ ID NO:174列出密码子优化的cDNA序列和氨基酸序列。将整个Rc12hydro表达盒称作pSZ1419并且转化到Prototheca moriformis遗传背景中。在含有蔗糖作为唯一碳源的平板上筛选阳性克隆。选择阳性克隆的亚组并且使其在脂质产生条件下生长,并且使用GC/MS方法针对蓖麻油酸12-羟化酶的产物(即,蓖麻油酸)进行筛选。

[0665] 用来检测任何在阳性Rc12hydro转基因中的蓖麻油酸的GC/MS方法如下。将阳性克隆和野生型对照的样品进行干燥并且然后悬浮于2.2mL的在10:1(v/v)甲醇-甲苯中的4.5H_2SO_4中。然后将该混合物在间断声处理和涡旋下在70℃-75℃加热3.5小时。在冷却至室温之后,添加2mL的6%K_2CO_3(水性)和2mL庚烷,然后将该混合物进行剧烈涡旋并且通过在900rpm离心2分钟来加速各层的分离。将上层移出并且用氮流浓缩至干燥。向得到的油添加500μL干吡啶和500μL BSTFA/1%TMCS(赛默飞世尔公司(Thermo Scientific))。将得到的溶液在70℃-75℃加热1.5小时并且然后用氮流浓缩至干燥。然后将残余物再悬浮于2mL的庚烷中并且进行再浓缩。将样品再悬浮于2mL的庚烷中并且通过GC/MS在Thermo Trace GC Ultra/DSQII系统上以EI模式使用任一蓖麻油酸甲酯TMS的基峰的选择性粒子监测(m/z 187)进行分析。将脂肪酸甲酯在Restek Rxi-5Sil MS柱上进行分离(0.25mm ID,30m长,0.5μm膜厚度),使用氦作为载气,流速为1mL/min。将该柱的初始温度保持在130℃持续4分钟,接着是4℃/min的倾斜升温至终温度240℃。通过将保留时间和全扫描质谱与按上述处理的蓖麻油酸真实试样进行比较来证实样品中蓖麻油酸的存在。图2显示了与蓖麻油酸标准品和野生型对照相比较的代表性阳性转基因克隆的GC保留时间。该阳性转基因克隆在RT:33.22/33.23处具有一个可衍生的峰,这与蓖麻油酸标准品中的一个类似峰相应,从而指示可衍生蓖麻油酸存在于该转基因克隆和该阳性对照两者中。在野生型对照样品中完全缺乏这个峰。

[0666] 实例8:用可替代的选择标记将无绿藻属工程化

[0667] A.可分泌的α-半乳糖苷酶在Prototheca moriformis中的表达

[0668] 先前已经在PCT申请号PCT/US2009/66142中描述了在无绿藻种类中表达异源蔗糖转化酶基因的方法和效果，所述专利特此通过引用进行结合。在这个实例中检测了其他异源多糖降解酶的表达。测试了具有编码α-半乳糖苷酶的以下外源基因之一的Prototheca moriformis UTEX 1435在蜜二糖(α-D-gal-glu)上生长的能力：来自卡尔酵母(Saccharomyces carlbergensis)的MEL1基因(相应于NCBI登录号P04824的氨基酸序列)、来自黑曲霉的AglC(相应于NCBI登录号Q9UUZ4的氨基酸序列)、以及来自高等植物Cyamopsis tetragobobola(瓜儿豆)的α-半乳糖苷酶(相应于NCBI登录号P14749的氨基酸序列)。将以上登录号和相应的氨基酸序列特此通过引用进行结合。在所有情况下，根据Prototheca moriformis中的优选密码子使用对基因进行优化。该表达盒的相关部分连同序列表编号列于下文。所有表达盒使用用于稳定基因组整合的5'和3'Clp同源重组靶向序列、莱茵衣藻TUB2启动子/5'UTR、以及普通小球藻硝酸盐还原酶3'UTR。

[0669]

*卡尔酵母MEL1*氨基酸序列	SEQ ID NO: 108
*卡尔酵母MEL1*氨基酸序列信号肽	SEQ ID NO: 109
*卡尔酵母MEL1*转化盒	SEQ ID NO: 110
*卡尔酵母MEL1*序列（密码子优化）	SEQ ID NO: 111
5' Clp同源重组靶向序列	SEQ ID NO: 112
3' Clp同源重组靶向序列	SEQ ID NO: 113
*莱茵衣藻TUB2*启动子/5'UTR	SEQ ID NO: 114
*普通小球藻*硝酸盐还原酶3'UTR	SEQ ID NO: 115

[0670]

*黑曲霉AlgC*氨基酸序列	SEQ ID NO: 116
*黑曲霉AlgC*氨基酸序列信号肽	SEQ ID NO: 117
黑曲霉*AlgC*序列（密码子优化）	SEQ ID NO: 118
*黑曲霉AlgC*转化盒	SEQ ID NO: 119
*瓜儿豆*α-半乳糖苷酶氨基酸序列	SEQ ID NO: 120
*瓜儿豆*α-半乳糖苷酶序列（密码子优化）	SEQ ID NO: 121
*瓜儿豆*α-半乳糖苷酶转化盒	SEQ ID NO: 122

[0671] 利用如在以上实例2中所述的生物射弹转化方法，用含有卡尔酵母MEL1、黑曲霉AlgC、或瓜儿豆α-半乳糖苷酶基因的三个表达盒中的每一个对Prototheca moriformis细胞进行转化。使用含有2％蜜二糖作为唯一碳源的平板来筛选阳性克隆。对于瓜儿豆表达盒转化子而言，平板上没有出现集落。从含有卡尔酵母MEL1转化子和黑曲霉AlgC转化子的平板上挑选阳性克隆。使用靶向黑曲霉AlgC 3'UTR和3'Clp同源重组靶向序列的一部分的引物进行PCR来证实转化DNA的整合。

[0672] 5'引物普通小球藻3'UTR：下游Clp序列(SEQ ID NO:123)

[0673] ACTGCAATGCTGATGCACGGGA

[0674] 3'引物普通小球藻3'UTR：下游Clp序列(SEQ ID NO:124)

[0675] TCCAGGTCCTTTTCGCACT

[0676] 作为阴性对照,还将来自未转化的Prototheca moriformis细胞的基因组DNA用该引物集进行扩增。从来自这些野生型细胞的基因组DNA未扩增出产物。

[0677] 对于来自卡尔酵母MEL1转化子和黑曲霉AlgC转化子(如通过PCR所证实的)中的每一个的一些阳性克隆,测试了它们在液体培养基中在作为唯一碳源的密二糖上生长的能力。在以上实例1中所述的条件和基础培养基中、以密二糖作为唯一碳源使这些选择的克隆生长3天。在这一时间期间,所有含有α-半乳糖苷酶编码基因的克隆都健壮生长,而未转化的野生型株系和表达酿酒酵母SUC2蔗糖转化酶的Prototheca moriformis在该密二糖培养基上都生长不良。这些结果说明α-半乳糖苷酶编码基因可以用作一个用于转化的选择标记。同样,这些数据说明在卡尔酵母MEL1(SEQ ID NO:109)或黑曲霉AlgC(SEQ ID NO:117)中存在的天然信号肽有用于将蛋白质靶向到Prototheca moriformis细胞的周质。

[0678] B.无绿藻属中硫胺素营养缺陷的THIC基因补偿

[0679] 使用外源THIC基因的表达来检测Prototheca moriformis细胞中的硫胺素营养缺陷。植物和藻类中的硫胺素生物合成典型地是在质体中进行的,因此多数参与其产生的核编码蛋白将需要有效地靶向到质体。Prototheca moriformis细胞的DNA测序和转录组测序揭示:除了THIC之外,所有编码硫胺素生物合成酶的基因都存在于基因组中。为了在生物化学水平剖析负责硫胺素营养缺陷的损害,在五种不同方案下对Prototheca moriformis细胞的生长进行了检测。(1)在2μM盐酸硫胺素的存在下;(2)无硫胺素;(3)无硫胺素,但具有2μM羟乙基噻唑(THZ);(4)无硫胺素,但具有2μM的2-甲基-4-氨基-5-(氨甲基)嘧啶(PYR);以及(5)无硫胺素,但具有2μM的THZ以及2μM的PYR。在这5种不同条件下的生长实验的结果表明:Prototheca moriformis细胞能够进行从头合成,但如果提供PYR前体则仅产生焦磷酸硫胺素(TPP)。这个结果与以下假设一致:Prototheca moriformis的硫胺素营养缺陷是由于不能从氨基咪唑核糖核苷酸合成羟甲基嘧啶磷酸酯(HMP-P),而这是由THIC酶来转化催化的。

[0680] 使用如在以上实例2中所述的生物射弹转化方法对Prototheca moriformis细胞进行转化,表达胶球藻属C-169 THIC(相应于JGI蛋白ID 30481的氨基酸序列,并且特此通过引用而结合)和作为选择性标记的酿酒酵母SUC2蔗糖转化酶。这种表达构建体包含来自胶球藻属C-169 THIC的天然转运肽、到基因组DNA 6S区域的上游和下游同源重组靶向序列、莱茵衣藻TUB2启动子/5'UTR区域(SEQ ID NO:104)、以及普通小球藻硝酸盐还原酶3'UTR(SEQ ID NO:115)。酿酒酵母SUC2表达也由莱茵衣藻TUB2启动子/5'UTR区域(SEQ ID NO:114)驱动并且包含普通小球藻硝酸盐还原酶3'UTR(SEQ ID NO:115)。根据Prototheca moriformis中的优选密码子使用来优化基因。相关表达盒序列在序列表中列出并且详述于下文:

	胶球藻属 C-169 THIC 氨基酸序列	SEQ ID NO: 125
	胶球藻属 C-169 THIC 氨基酸序列天然转运肽	SEQ ID NO: 126
	胶球藻属 C-169 THIC 转化盒	SEQ ID NO: 127
[0681]	*胶球藻属* C-169 *THIC* 序列（密码子优化）	SEQ ID NO: 128
	酿酒酵母 SUC2 序列（密码子优化）	SEQ ID NO: 129
	5' 6S 同源重组靶向序列	SEQ ID NO: 82
	3' 6S 同源重组靶向序列	SEQ ID NO: 84

[0682] 在不含硫胺素并且含有蔗糖作为唯一碳源的平板上进行阳性克隆的筛选。利用在胶球藻属C-169 THIC基因之内结合的5'引物以及在6S基因座中的转化DNA的下游退火的3'引物通过PCR来证实阳性克隆。还使用Southern印迹测定来确认经PCR证实的阳性克隆。

[0683] 为了观察野生型Prototheca moriformis细胞的硫胺素营养缺陷，需要首先耗尽细胞的内部硫胺素储备。为了测试在无硫胺素的培养基中的生长，首先使细胞在含有2μM硫胺素的培养基中生长至静止期并且然后将这些细胞在无硫胺素的培养基中稀释至在750nm(OD750)处大约0.05的光密度。然后使这些稀释的细胞在无硫胺素的培养基中再一次生长至静止期(大约2-3天)。这些硫胺素耗尽的细胞被用来在无硫胺素的培养基中接种用于生长研究的培养物。使野生型细胞在具有葡萄糖作为碳源的培养基(具有或不具有硫胺素)中生长并且使具有天然转运肽胶球藻属C-169 THIC构建体的阳性克隆具有蔗糖作为唯一碳源的培养基中生长。通过监测在750nm处的吸光度来量度生长。这些生长实验的结果表明：与无硫胺素的培养基中的野生型细胞相比较，表达转基因的株系在无硫胺素的培养基中的生长实质性地更为良好。然而，这些转化子未能达到野生型细胞在含有硫胺素的培养基中的生长速率和细胞密度。在无硫胺素的培养基中这些转化子克隆的生长的量与整合的胶球藻属酶的拷贝数之间存在着强的关联(即，该转基因的拷贝数越多，则这些细胞在无硫胺素的培养基中生长越好)。

[0684] 使用包含胶球藻属THIC、拟南芥THIC基因、以及集胞藻种类PCC 6803thiC基因的表达构建体产生了另外的转化子。在胶球藻属和拟南芥THIC基因的情况下，将天然转运肽序列替换为来自原始小球藻硬脂酰-ACP去饱和酶(SAD)基因的转运肽序列。集胞藻种类是一种蓝藻并且该thiC蛋白不包含天然转运肽序列。在集胞藻种类thiC构建体中，来自原始小球藻SAD基因的转运肽序列与集胞藻种类thiC的N端融合。在所有情况下，对这些序列针对在Prototheca moriformis中表达进行了密码子优化。所有前述三种构建体都包含针对基因组的6S区域的上游和下游同源重组靶向序列(SEQ ID NOs:82和84)、原始小球藻肌动蛋白启动子/5'UTR、以及原始小球藻EF1A基因3'UTR。所有三种构建体都包含由莱茵衣藻TUB2启动子/5'UTR(SEQ ID NO:114)驱动的neoR基因并且包含普通小球藻3'UTR(SEQ ID NO:115)，从而赋予通过G418的选择性。拟南芥THIC的氨基酸序列相应于NCBI登录号NP_180524并且集胞藻种类thiC的氨基酸序列相应于NCBI登录号NP_442586，特此将两个序列通过引用而结合。相关表达盒序列在序列表中列出并且详述于下文：

[0685]

具有*原始小球藻*转运肽的胶球藻属 *THIC* 表达构建体	SEQ ID NO: 130
具有*原始小球藻*转运肽的胶球藻属 *THIC*	SEQ ID NO: 131
*原始小球藻*肌动蛋白启动子/5' UTR	SEQ ID NO: 132
原始小球藻 EF1A 3'UTR	SEQ ID NO: 133
拟南芥 THIC 表达构建体	SEQ ID NO: 134
具有*原始小球藻*转运肽的*拟南芥 THIC*	SEQ ID NO: 135
具有天然转运肽的*拟南芥*THIC 氨基酸序列	SEQ ID NO: 136
集胞藻种类 thiC 表达构建体	SEQ ID NO: 137
具有*原始小球藻*转运肽的集胞藻种类 *thiC*	SEQ ID NO: 138
*集胞藻种类*thiC 氨基酸序列	SEQ ID NO: 139
neoR 基因	SEQ ID NO: 140

[0686] 在含有G418的平板上筛选阳性克隆并且从每一转化挑选一些克隆用于通过PCR进行验证。使用以下引物通过PCR来确认包含胶球藻属C-169(具有原始小球藻转运肽)、拟南芥以及集胞藻种类PCC 6803THIC基因的转化DNA分别在基因组6S基因座中的整合：

[0687] 5'THIC胶球藻属确认引物序列(SEQ ID NO:141)

[0688] ACGTCGCGACCCATGCTTCC

[0689] 3'THIC确认引物序列(SEQ ID NO:142)

[0690] GGGTGATCGCCTACAAGA

[0691] 5'THIC拟南芥确认引物序列(SEQ ID NO:143)

[0692] GCGTCATCGCCTACAAGA

[0693] 5'thiC集胞藻种类确认引物序列(SEQ ID NO:144)

[0694] CGATGCTGTGCTACGTGA

[0695] 在含有G418的培养基中,使用来自每一个这些不同构建体的转化子的经选择的证实的阳性克隆对硫胺素耗尽细胞(如上述)进行了生长实验。所有转化子都能够在无硫胺素的培养基中生长(健壮程度不同)。在无硫胺素培养基中的这些转化子的生长与在含有硫胺素的培养基上的野生型细胞的比较显示出以下有关它们在无硫胺素培养基中支持生长的能力的等级评定：(1)拟南芥转化子；(2)胶球藻属C-169(具有原始小球藻转运肽)转化子；以及(3)集胞藻种类转化子。这些结果说明：虽然拟南芥THIC的单一拷贝能够补偿Prototheca moriformis细胞中的硫胺素营养缺陷,但是需要胶球藻属C-169的多个拷贝(具有天然转运肽序列或来自原始小球藻的转运肽序列)以及集胞藻种类THIC以允许在没有硫胺素的情况下快速生长。考虑到来自不同来源的不同THIC的结果的变异性,任何特定的THIC基因完全补偿无绿藻种类中存在的损害的能力不是可预测的。

[0696] 对这三个THIC氨基酸序列进行了比对。虽然与来自胶球藻属和拟南芥的THIC相比较来自集胞藻种类的thiC之间存在显著的序列保守(在氨基酸水平上41％的一致性),蓝藻蛋白在N端缺少一个结构域,该结构域在藻类和植物蛋白质中是非常保守的。尽管缺少该结构域(并且推测上导致结构差异),表达集胞藻种类thiC的构建体能够至少部分地恢复Prototheca moriformis细胞中的硫胺素原养型(prototrophic)。

[0697] 实例9:燃料生产

[0698] A.使用螺旋榨机和压榨辅料(press aid)从微藻提取油

[0699] 利用转鼓式干燥器使含有38％油(以DCW计算)的微藻生物质干燥,导致获得5％-5.5％的水分含量。将这种生物质传送至L250弗氏压碎器(French L250press)中。将30.4kg(67lbs.)生物质传送通过该压碎器,并且没有回收油。传送相同的与不同百分比的作为压榨辅料的柳枝稷组合的干燥微生物生物质通过该压碎器。干燥微生物生物质和20％w/w柳枝稷的组合产生最好的总体油回收百分比。然后使榨饼经受己烷提取并且20％柳枝稷条件下的最终产率是61.6％的总可获得油(以重量计算)。具有超过50％油干细胞重量的生物质不需要使用压榨辅料(如柳枝稷)来释放油。在PCT申请号PCT/US2010/31108中描述了使用螺旋榨机从微藻中提取油的其他方法,所述专利通过引用结合在此。

[0700] B.从无绿藻油生产生物柴油

[0701] 使来自Prototheca moriformis(UTEX 1435)的脱胶油(根据上述方法产生)经受酯交换以产生脂肪酸甲酯。结果显示于以下表38中。

[0702] 该油的脂质谱是:

[0703]

C10:0	0.02
C12:0	0.06
C14:0	1.81
C14.1	0.07

[0704]

C16:0	24.53
C16:1	1.22
C18:0	2.34
C18:1	59.21
C18:2	8.91
C18:3	0.28
C20:0	0.23
C20:1	0.10
C20:1	0.08
C21:0	0.02
C22:0	0.06
C24:0	0.10

[0705] 表38.来自Prototheca moriformis甘油三酯油的生物柴油谱。

[0706]

方法	测试		结果	单位
ASTM D6751 A1	生物柴油掺混燃料的冷浸过滤性	过滤时间	120	秒
		过滤体积	300	ml
ASTM D93	彭斯基-马顿（Pensky-Martens）闭杯闪点	使用的程序	A	
		修正的闪点	165.0	°C
ASTM D2709	中间馏分燃料中的水和沉积物（离心法）	沉积物和水	0.000	体积%
EN 14538	通过 ICP OES 确定 Ca 和 Mg 的含量	（Ca 和 Mg）的和	< 1	mg/kg
EN 14538	通过 ICP OES 确定 Ca 和 Mg 的含量	（Na 和 K）的和	< 1	mg/kg
ASTM D445	运动/动态粘度	运动粘度 @104 °F/40°C	4.873	mm^2/s
ASTM D874	来自润滑油和添加剂的硫酸盐灰分	硫酸盐灰分	< 0.005	Wt %
ASTM D5453	通过紫外荧光确定在轻烃、火花点火式发动机燃料、柴油发动机燃料以及发动机油中的总硫	硫，mg/kg	1.7	mg/kg
ASTM D130	腐蚀铜带	生物柴油-Cu 腐蚀 50°C（122°F）/3 hr	1a	
ASTM D2500	浊点	浊点	6	°C
ASTM D4530	微量碳残余物	平均微量法碳残余物	< 0.10	Wt %
ASTM D664	通过电位滴定的石油产品的酸值	使用的程序	A	
		酸值	0.20	mg KOH/g
ASTM D6584	通过气相色谱确定 B-100 生物柴油甲基酯中的游离甘油和总甘油	游离甘油	< 0.005	Wt %
		总甘油	0.123	Wt %
ASTM D4951	通过 ICP-AES 的润滑油中的添加元素	磷	0.000200	Wt %
ASTM D1160	在减压下的石油产品的蒸馏	IBP	248	°C
		AET @ 5%回收	336	°C
		AET @ 10%回收	338	°C
		AET @ 20%回收	339	°C
		AET @ 30%回收	340	°C
		AET @ 40%回收	342	°C
		AET @ 50%回收	344	°C

[0707]

		AET @ 60%回收	345	°C
		AET @ 70%回收	347	°C
		AET @ 80%回收	349	°C
		AET @ 90%回收	351	°C
		AET @ 95%回收	353	°C
		FBP	362	°C
		%回收	98.5	%
		%损失	1.5	%
		%残留	0.0	%
		冷阱体积	0.0	ml
		IBP	248	°C
EN 14112	氧化稳定性的确定（加速氧化试验）	氧化稳定性	> 12	hr
		操作温度（通常是110 deg C）	110	°C
ASTM D4052	通过数字式密度仪的液体密度	API 重力@ 60°F	29.5	°API
ASTM D 6890	点火延迟（ID）以及衍生十六烷值（DCN）的确定	衍生十六烷值（DCN）	> 61.0	

[0708] 该生物柴油的脂质谱与该原料油的脂质谱高度类似。可以使通过这些方法提供的其他油以及本发明的组合物经受酯交换以产生具有多种脂质谱的生物柴油，这些脂质谱包括(a)至少4%C8-C14；(b)至少0.3%C8；(c)至少2%C10；(d)至少2%C12；以及(3)至少30%C8-C14。

[0709] 根据ASTM D6751 A1方法，生产的生物柴油的冷浸过滤性是300ml体积120秒。这种测试包括过滤300ml的B100，冷却至40°F持续16小时，允许加温至室温，并且在真空下使用0.7微米的具有不锈钢支撑的玻璃纤维过滤器进行过滤。本发明的油可以进行酯交换以产生具有少于120秒、少于100秒、以及少于90秒的冷浸时间的生物柴油。

[0710] C.可再生柴油的生产

[0711] 使来自Prototheca moriformis(UTEX 1435)的脱胶油(根据上述方法生产并且具有与在以上这个实例中用来制造生物柴油的油相同的脂质谱)经受酯交换以产生可再生柴油。

[0712] 首先将该油氢化处理以去除氧和甘油骨架，从而产出n-石蜡。然后使这些n-石蜡经受裂解和异构化。该材料的色谱图显示于图1中。然后使该材料经受冷过滤，这去除了大约5％的C18材料。该冷过滤之后，将总体积材料转到(cut to)闪点并且针对闪点、ASTM D-86蒸馏分布、浊点以及粘度进行评估。闪点是63℃；粘度是2.86cSt(厘沱)；浊点是4℃。ASTM D-86蒸馏值显示于表39中：

[0713] 表39.ASTM D-86蒸馏值。

[0714] 以℃的读数：

[0715]

体积	温度
IBP	173
5	217.4

[0716]

10	242.1
15	255.8
20	265.6
30	277.3
40	283.5
50	286.6
60	289.4
70	290.9
80	294.3
90	300
95	307.7
FBP	331.5

[0717] 所产生的材料的T10-T90是57.9℃。可以采用在此披露的油的氢化处理、异构化以及其他的共价修饰方法、以及在此披露的蒸馏和分提方法(如冷过滤),使用根据在此披露的方法生产的甘油三酯油来产生具有其他T10-T90范围(如20℃、25℃、30℃、35℃、40℃、45℃、50℃、60℃以及65℃)的可再生柴油组合物。

[0718] 所产生的材料的T10是242.1℃。可以采用在此披露的油的氢化处理、异构化以及其他的共价修饰方法、以及在此披露的蒸馏和分提方法(如冷过滤)来产生具有其他T10值(如在180与295之间、在190与270之间、在210与250之间、在225与245之间、以及至少290的T10)的可再生柴油组合物。

[0719] 所产材料的T90是300℃。可以采用在此披露的油的氢化处理、异构化以及其他的共价修饰方法、以及在此披露的蒸馏和分提方法(如冷过滤)来产生具有其他T90值(如在280与380之间、在290与360之间、在300与350之间、在310与340之间、以及至少290的T90)的可再生柴油组合物。

[0720] 所产生的材料的FBP是300℃。可以采用在此披露的油的氢化处理、异构化以及其他的共价修饰方法、以及在此披露的蒸馏和分提方法(如冷过滤)来产生具有其他FBP值(如在290与400之间、在300与385之间、在310与370之间、在315与360之间、以及至少300的FBP)的可再生柴油组合物。

[0721] 可以使通过本发明的方法和组合物提供的其他油经受氢化处理、异构化以及其他共价修饰的组合,这些油包括具有多种脂质谱的油,这些脂质谱包括(a)至少4%C8-C14;(b)至少0.3%C8;(c)至少2%C10;(d)至少2%C12;以及(3)至少30%C8-C14。

[0722] 虽然已经结合其具体实施例描述了本发明,应当理解能够进一步修改。本申请旨在涵盖通常遵循本发明原理的本发明的任何变型、用途或适应性改变,并且包括这样的本披露的偏离,如落入本发明所属领域内已知或惯常的实践范围内或和如可能适用于在上文中所述的本质特征的偏离。

[0723] 实例10:将微生物工程化以生产C18:2和C18:3甘油脂

[0724] 在藻类和植物中的脂质合成开始于葡萄糖或其他碳源通过质体丙酮酸脱氢酶复合体转化为乙酰CoA。接下来,乙酰CoA羧化酶(ACCase)利用碳酸氢盐作为底物产生3-C化合物,丙二酰CoA。然后β-酮脂酰-ACP(酰基载体蛋白质)合成酶III(KAS III)催化丙二酰CoA与乙酰CoA之间的第一缩合反应以产生4-C化合物。由KAS I催化连续的直到C16:0的2-C加成。由KAS II催化最后的到C18:0的2-C延伸。硫酯酶(TE)终止酰基-ACP的延长。可溶性酶硬

脂酰ACP去饱和酶(SAD)对于C18:0-ACP底物具有活性并且在Δ9位置形成双键,从而产生油酸-ACP。将产生的C18:1脂肪酸通过油酸酯或广谱特异性TE的作用从ACP上释放。

[0725] 所有脂肪酸,一旦在质体中从ACP上释放,则被运输至ER,在此发生脂质(TAG)生物合成。广而言之,在高等植物的ER中有两条脂质生物合成途径,然而这两条途径无疑地在一定程度上共享底物。脂肪酰基CoA独立途径利用展示出非常具有选择性的底物特异性的酰基溶血卵磷脂酰基转移酶(acyllysophosphatidylcholine acyl transferase)在磷脂酰胆碱(P-胆碱)部分之间转移脂肪酰基基团,最终将它们转移至二酰甘油(DAG)。二酰基甘油酰基转移酶(DGAT)进行最后的转移,将来自酰基CoA底物的脂肪酰基基团转移至DAG,从而产生最终的三酰甘油。另一方面,脂肪酰基CoA依赖途径使用脂肪酰基CoA作为底物通过甘油磷酸酰基转移酶(GPAT)、溶血磷脂酸酰基转移酶(LPAAT)以及DGAT的作用将脂肪酰基基团分别地转移到甘油-3-磷酸、溶血磷脂酸(LPA)以及DAG上。

[0726] 在将微生物工程化以合成包含C18:2和C18:3的TAG中有用的酶包括β-酮脂酰-ACP合成酶II(KAS II)、硬脂酰ACP去饱和酶(SAD)、硫酯酶(包括油酸特异性硫酯酶)、脂肪酸去饱和酶(FAD)以及甘油脂去饱和酶(如ω-6脂肪酸去饱和酶、ω-3脂肪酸去饱和酶、或ω-6油酸去饱和酶)。可以将这些不同的酶单独地或组合地在微生物中过表达以增加C18:2或C18:3或TAG产量。增加KAS II酶活性的表达将碳积累从棕榈酸脂(C16:0)推向硬脂酸脂(C18:0)及以后。一些候选KAS II酶的氨基酸序列显示于以下表40中。披露的KAS II序列来自高等植物种类,这些种类特异性地产生水平提高的油酸、亚油酸或亚麻酸脂肪酸。技术人员将能够鉴定针对KAS II的其他基因,包括但不限于麻疯树(GenBank登录号ABJ90469.2)、大豆(GenBank登录号AAW88762.1)、美洲油棕榈(GenBank登录号ACQ41833.1)、拟南芥(GenBank登录号AAL91174)、葡萄(GenBank登录号CBI27767)、以及陆地棉(GenBank登录号ADK23940.1)。

[0727] 表40.示例性的KAS II酶。

[0728]

KAS II酶	SEQ ID NO
蓖麻	SEQ ID NO:175
向日葵	SEQ ID NO:176
甘蓝型油菜	SEQ ID NO:177
大豆	SEQ ID NO:178
P.moriformis	SEQ ID NO:179

[0729] 通过一种或多种脂质途径酶的微生物过表达而实现将水平增加的硬脂酸酯转变为油酸(C18:1),以便用于水平提高的亚油酸和亚麻酸脂肪酸的生产。在使不饱和物水平提高中具有效用的两种另外的酶活性是硬脂酰ACP去饱和酶(SAD)和油酸特异性硫酯酶。首先完成了将水平增加的硬脂酸酯(C18:0)通过这些酶中的一种或两种的作用转化为油酸,用于亚油酸和亚麻酸脂肪酸的形成。

[0730] 对于用于使油酸水平提高的过表达有用的示例性的SAD酶的氨基酸序列引用于表41中。另外,来自Prototheca moriformis的内源SAD(SEQ ID NO:180)对于增加C18:1水平也是有效的。披露的SAD序列来自高等植物种类,这些种类特异性地产生水平提高的油酸、亚油酸或亚麻酸脂肪酸。技术人员将能够鉴定针对SAD的其他基因。

[0731] 表41.示例性的SAD酶。

[0732]

SAD酶	SEQ ID NO	GenBank ID No.
蓖麻	SEQ ID NO:196	ACG59946.1
向日葵	SEQ ID NO:197	AAB65145.1
芥菜	SEQ ID NO:198	AAD40245.1
大豆	SEQ ID NO:199	ACJ39209.1
油橄榄	SEQ ID NO:200	AAB67840.1
油桐	SEQ ID NO:201	ADC32803.1

[0733] SAD酶对于C18:0-ACP底物具有活性并且在Δ9位置形成碳-碳双键,从而产生油酸-ACP。使产生的C18:1脂肪酸通过油酸酯或广谱特异性TE的作用从ACP上释放。我们在此处的这些实例中显示了油橄榄硬脂酰-ACP去饱和酶(登录号:AAB67840.1;SEQ ID NO:172)或红花(Carthamus tinctorius)ACP硫酯酶(登录号:AAA33019.1;SEQ ID NO:195)的过表达导致C18:1脂肪酸的积累增加。对于使油酸脂肪酸水平增加有用的示例性油酸硫酯酶的氨基酸序列引用于表42中。普通技术人员将能够鉴定针对油酸硫酯酶的其他基因。

[0734] 表42.用于提高的油酸脂肪酸生产的示例性硫酯酶。

[0735]

硫酯酶	SEQ ID NO	GenBank ID No.
向日葵	SEQ ID NO:202	AAL79361.1
白菜型油菜	SEQ ID NO:203	AAC49002.1
麻疯树	SEQ ID NO:204	ABX82799.3
玉米	SEQ ID NO:205	ACG40089.1
玉米	SEQ ID NO:206	ACG42559.1

[0736] 脂肪酸去饱和酶是另外的在增加微生物中的亚油酸和亚麻酸脂肪酸方面具有效用的酶。具体而言,两种酶活性,FAD 2和FAD 3,在微生物中提供了增加的亚油酸和亚麻酸脂肪酸积累。示例性的FAD 2和FAD 3酶的氨基酸序列显示于以下表43中。

[0737] 表43.示例性的FAD 2和FAD 3酶。

[0738]

FAD 2 和 FAD 3 酶	SEQ ID NO
亚麻（Linus usitatissimum ）12 去饱和	SEQ ID NO: 181

[0739]

酶	
亚麻 15 去饱和酶	SEQ ID NO: 182
亚麻 15 去饱和酶	SEQ ID NO: 183
红花 12 去饱和酶	SEQ ID NO: 184
向日葵 12 去饱和酶	SEQ ID NO: 185

[0740] 除了在表43中列出的那些酶之外,示例性的Δ12FAD酶的氨基酸序列在表44中列出。适合用于在微生物中过表达的其他Δ12FAD酶引用于表45中。示例性的Δ15FAD以及对

于增加不饱和脂肪酸和TAG水平有用的其他酶的氨基酸序列在表46中列出。另外的甘油脂去饱和酶提供于表47中。

[0741] 表44.用于增加亚油脂肪酸生产的示例性Δ12FAD酶。

[0742]

Δ12 FAD 酶	SEQ ID NO	GenBank ID No.
红花	SEQ ID NO: 207	ADM48790.1
陆地棉	SEQ ID NO: 208	CAA71199.1
大豆	SEQ ID NO: 209	BAD89862.1
玉米	SEQ ID NO: 210	ABF50053.1
Prototheca moriformis 等位基因 *1*	SEQ ID NO: 211	
Prototheca moriformis 等位基因 *2*	SEQ ID NO: 212	

[0743] 表45.另外的适合用于在微生物中过表达以增加亚油酸或亚麻酸的Δ12FAD酶。

Δ12 FAD 酶	GenBank 登录号
斑鸠菊（Vernonia galamensis）	AAF04094.1
斑鸠菊（Vernonia galamensis）	AAF04093.1
倒吊蜡烛（Wrightia tinctoria）	ADK47520.1

[0744]

油橄榄	AAW63041.1
油桐	AAN87573
拟南芥	AAA32782.1
亚麻荠	ADU18247.1
亚麻荠	ADU18248.1
亚麻荠	ADU18249.1
红花	ADK94440.1
大豆	BAD89862.1

大豆	DQ532371.1
陆地棉	AAL37484.1
亚麻	ACF49507.1
亚麻	ACF49508.1
[0745] *月见草（Oenothera biennis）*	ACB47482
酿酒酵母	NP_011460.1
玉米	ACG37433.1
白菜型油菜	CAD30827.1

[0746] 表46.用于增加亚麻酸脂肪酸生产的示例性Δ15FAD酶。

[0747]

Δ15 FAD 酶	SEQ ID NO	GenBank ID No.
甘蓝型油菜（Brassica napa）	SEQ ID NO: 213	AAA32994.1
亚麻荠	SEQ ID NO: 214	
亚麻荠	SEQ ID NO: 215	
大豆	SEQ ID NO: 216	ACF19424.1
油桐	SEQ ID NO: 217	AAF12821.1
蓖麻	SEQ ID NO: 218	EEF36775.1
亚麻	SEQ ID NO: 219	ADV92272.1
Prototheca moriformis 等位基因 1	SEQ ID NO: 220	
Prototheca moriformis 等位基因 2	SEQ ID NO: 221	

[0748] 表47.适合用于在微生物中过表达以增加亚麻酸或增加不饱和TAG水平的甘油脂去饱和酶。

[0749]

甘油脂去饱和酶	GenBank登录号
大豆(Glycine max)	ACD69577.1
Glycine max cultivar volania	ACS15381.1
野大豆(Glycine soja)	P48621.1
拟南芥	NP_180559.1
大花亚麻(Linum grandiflorum)	BAG70949.1
玉米	BAA22440.1
油橄榄	ABG88130.2
麻疯树	ABX82798.1
油桐	CAB45155.1

油桐	AAD13527.1

[0750] 利用在此披露的方法将在此披露的氨基酸序列表达于在微生物中。可以将编码序列针对在微生物中表达进行优化。例如，为了在Prototheca moriformis中表达，使用了如在此于表2中披露的优选密码子使用。

[0751] 实例11：将微生物工程化用于增加亚麻酸不饱和脂肪酸和甘油酯的产生

[0752] 如在实例10中所述，Δ15去饱和酶在C18:2(亚油酸)脂肪酸或脂肪酰基分子的15位置处催化双键的形成，由此产生C18:3(亚麻酸)脂肪酸或脂肪酰基分子。某些产生富含亚麻酸不饱和脂肪酸的油的高等植物种类(包括甘蓝型油菜(Bn)、亚麻荠(Cs)以及亚麻)是编码Δ15去饱和酶的基因的来源，这些基因可以在微生物中表达以影响脂肪酸谱。这个实例描述了使用编码Δ15去饱和酶的多核苷酸将微生物工程化，其中该转化的微生物的脂肪酸谱已经富含亚麻酸。

[0753] 根据在实例2中详述的生物射弹转化方法，将Protheca moriformis UTEX 1435的以经典方式诱变(为了更高的油产量)的衍生物(A株)用在表49中列出的每一个质粒构建体单独地转化。每个构建体含有针对用于整合至核基因组中的6S基因组区域的5'(SEQ ID NO:82)和3'(SEQ ID NO:84)同源重组靶向序列(在构建体侧翼)和在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且该蔗糖转化酶基因用作选择标记。根据表2，所有蛋白质编码区经密码子优化以反映在Prototheca moriformis UTEX 1435核基因中固有的密码子偏倚。来自甘蓝型油菜的去饱和酶基因(Bn FAD3，AAA32994)、来自亚麻荠FAD-7的去饱和酶基因、以及来自亚麻的去饱和酶基因(Lu FAD3A和Lu FAD3B，GenBank登录号ABA02172和ABA02173)的编码区各自在Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR的控制下。一个FLAG®表位序列被编码在这些重组去饱和酶基因序列的N端细胞质环中。

[0754] 表49.用来转化Protheca moriformis(UTEX 1435)A株的质粒构建体。

[0755]

质粒构建体	相关序列元件	SEQ ID NO:
pSZ2124	6S::Crβtub:ScSuc2:Cvnr::PmAmt03:3xFlag-BnFad3:Cvnr::6S	SEQ ID NO: 222
pSZ2125	6S::Crβtub:ScSuc2:Cvnr::PmAmt03:3xFlag-CsFad7-1:Cvnr::6S	SEQ ID NO: 223
pSZ2126	6S::Crβtub:ScSuc2:Cvnr::PmAmt03:3xFlag-LuFad3A:Cvnr::6S	SEQ ID NO: 224
pSZ2127	6S::Crβtub:ScSuc2:Cvnr::PmAmt03:3xFlag-LuFad3B:Cvnr::6S	SEQ ID NO: 225

[0756] 将构建体pSZ2124、pSZ2125、pSZ2126、以及pSZ2127中的每一个单独地转化到A株中。在含有蔗糖作为唯一碳源的平板上选择初级转化子。将各个转化子以克隆方式纯化并且在适合于脂质产生的条件下在pH 7.0生长，与在实例1中披露的那些类似。从来自每种转化子的干燥生物质制备脂质样品。将来自每种转化子的20mg-40mg的干燥生物质再悬浮于2mL的在MeOH中的5％H_2SO_4中，并且添加含有适当量的适合的内部标准品(C19:0)的200μl甲

苯。将该混合物进行短暂的声处理以分散该生物质，然后在70℃-75℃加热3.5小时。添加2mL的庚烷以提取脂肪酸甲酯，接着添加2mL的6％K_2CO_3(水性)以中和酸。将该混合物剧烈搅拌，并且将一部分上层转移至含有Na_2SO_4(无水)的小瓶中用于使用标准FAME GC/FID(脂肪酸甲酯气相色谱火焰电离检测)方法进行气相色谱分析。使用标准脂肪酸甲酯气相色谱火焰电离(FAME GC/FID)检测方法分析脂肪酸谱。从一组代表性克隆得到的脂肪酸谱(表示为总脂肪酸的面积％)显示于表50中，这些代表性克隆由pSZ2124、pSZ2125、pSZ2126以及pSZ2127的A株转化产生。为了比较，将从未转化A株对照细胞获得的脂质的脂肪酸谱另外地呈现在表50中。

[0757] 表50.经工程化而表达高等植物的外源去饱和酶的Prototheca moriformis细胞的不饱和C18:1、C18:2、以及C18:3脂肪酸谱。

[0758]

		总脂肪酸的%		
株	样品	C18:1	C18:2	C18:3
未转化的A株	1	55.68	7.99	0.63
	2	55.31	8.16	0.7
pSZ2124 甘蓝型油菜FAD3	转化子1	60.42	1.62	7.69
	转化子2	59.7	0.6	8.49
	转化子3	60.56	1.19	8.15
	转化子4	59.85	0.8	7.9
pSZ2125 亚麻荠FAD7	转化子1	57.11	9.45	1.4
	转化子2	57.5	8.56	1.39
	转化子3	56.27	8.78	1.39
	转化子4	52.57	9.39	1.7
pSZ2126 亚麻FAD3A	转化子1	58.97	0.84	9.67
	转化子2	57.93	1.36	11.92
	转化子3	59.37	0.58	10.33
	转化子4	59.05	0.49	10.24
pSZ2127 亚麻FAD3B	转化子1	59.6	1.67	8.67
	转化子2	59.73	1.02	8.6
	转化子3	60.04	1.6	8.74
	转化子4	58.57	0.89	9.05

[0759] 未转化的Prototheca moriformis的(UTEX 1435)株A株展示出包含少于1％C18:3脂肪酸的脂肪酸谱。相反，表达高等植物脂肪酸去饱和酶的A株的脂肪酸谱显示出增加的C18:3脂肪酸组成，增加范围从大约2倍至17倍。工程化的表达亚麻的FAD3A或FAD3B或者甘蓝型油菜的FAD3基因产物的株系显示出最大程度的C18:3增加(表50)。18:3与总C18不饱和物的比率在未转化株中是大约1％，而在转化株中范围是从大约2％至17％。18:2与总C18:0的比率在未转化株中是大约12％-13％，而在转化株中范围是从大约1％至15％，在FAD3A转化子中水平最低。这些数据证明了多核苷酸的效用和效果，这些多核苷酸允许Δ15去饱和酶脂肪酸去饱和酶的外源表达而改变工程化微生物的脂肪酸谱，并且尤其是增加微生物细胞中18:3脂肪酸的浓度。

[0760] 实例12:通过发夹RNA途径将微生物工程化用于增加硬脂酸和硬脂酸酯的产量

[0761] 硬脂酸ACP去饱和酶(SAD)是脂质合成途径中的一部分。它们起作用以将双键引入到脂肪酰基链中。例如,SAD酶催化从C18:0脂肪酸(硬脂酸)合成C18:1脂肪酸。如在实例6中所显示的,通过靶向基因破坏而中断Prototheca moriformis的SAD2等位基因,导致工程化微生物的脂肪酸谱中的C18:0脂肪酸水平的可测量的增加。这个实例描述了使用编码下调SAD2表达的发夹RNA的多核苷酸将微生物工程化,其中该转化的微生物的脂肪酸谱已经富含饱和C18:0脂肪酸。

[0762] 设计了在表51中列出的四种构建体(pSZ2139-pSZ2142)以减弱Prototheca moriformis SAD2基因产物的表达。每种构建体包含编码靶向Prototheca moriformis SAD2 mRNA转录本的发夹RNA的不同核酸序列(具有大小范围从180碱基对至240碱基对的茎长)、以及针对用于整合至核基因组中的6S基因组区域的5'(SEQ ID NO:82)和3'(SEQ ID NO:84)同源重组靶向序列(在构建体侧翼)和在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。编码每种构建体的SAD2RNA发夹的多核苷酸在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制之下。

[0763] 表51.用来转化Prototheca moriformis(UTEX 1435)A株的质粒构建体。

[0764]

质粒构建体	相关序列元件	SEQ ID NO:
SZ2139 发夹 A	6S::Crβtub:ScSuc2:Cvnr:Crβtub:PmSAD2-hpA:Cvnr::6S	SEQ ID NO: 226
SZ2140 发夹 B	6S::Crβtub:ScSuc2:Cvnr:Crβtub:PmSAD2-hpB:Cvnr::6S	SEQ ID NO: 227
SZ2141 发夹 C	6S::Crβtub:ScSuc2:Cvnr:Crβtub:PmSAD2-hpC:Cvnr::6S	SEQ ID NO: 228
SZ2142 发夹 D	6S::Crβtub:ScSuc2:Cvnr:Crβtub:PmSAD2-hpD:Cvnr::6S	SEQ ID NO: 229

[0765] 根据在实例2中详述的生物射弹转化方法,将Prototheca moriformis UTEX 1435的以经典方式诱变(为了更高的油产量)的衍生物(A株)用在表51中列出的质粒构建体单独地转化。将初级转化子在含有蔗糖作为唯一碳源的琼脂平板上进行筛选,以克隆方式纯化,并且使其在标准脂质产生条件下生长。使用如在实例11中所述的直接酯交换方法来确定脂肪酸谱。从一组代表性克隆得到的脂肪酸谱(表示为总脂肪酸的面积%)显示于以下表52中,这些代表性克隆由用pSZ2139、pSZ2140、pSZ214以及pSZ2142转化A株而产生。为了比较,将从未转化A株对照细胞获得的脂质的脂肪酸谱另外地呈现在表52中。

[0766] 表52.经工程化而表达靶向硬脂酰ACP去饱和酶基因/基因产物的发夹RNA构建体的Prototheca moriformis细胞的不饱和C18:1、C18:2、以及C18:2脂肪酸谱。

[0767]

		总脂肪酸的%			
株/质粒构建体	转化子	C18:0	C18:1	C18:2	% 比率 C18 饱和物:C18 不饱和物
A 株	未转化	2.77	60.74	7.27	4
A 株/pSZ2139 发夹 A	转化子 1	6.39	51.69	9.06	11
	转化子 2	5.49	52.89	9.25	9
	转化子 3	3.39	56.12	8.85	5
	转化子 4	3.24	54.55	8.62	5
A 株/pSZ2140 发夹 B	转化子 1	22.14	36.14	8.13	50
	转化子 2	17.19	41.17	8.31	35
	转化子 3	9.45	49.81	8.79	16
	转化子 4	5.61	53.8	9.02	9
A 株/pSZ2141 发夹 C	转化子 1	20.7	40.96	6.45	44
	转化子 2	16.33	45.57	7.31	31
	转化子 3	13.43	44.79	9.04	25
	转化子 4	12.7	46.25	9.98	23
	转化子 5	8.47	50.65	9.12	14
A 株/pSZ2142 发夹 D	转化子 1	26.99	30.93	8.31	69
	转化子 2	10.96	47.27	9.9	19
	转化子 3	8.64	50.77	11.7	14
	转化子 4	7.67	49.76	9.39	13

[0768] 在表52中呈现的数据表明了SAD2发夹RNA构建体的表达对宿主生物的C18:0和C18:1脂肪酸谱的明显影响。包含SAD2发夹RNA构建体的A株转化子的脂肪酸谱显示出饱和C18:0脂肪酸百分比的增加，伴随不饱和C18:1脂肪酸的同时减少。未转化株的脂肪酸谱包含大约3％C18:0。转化株的脂肪酸谱包含大于3％至大概27％的C18:0。C18:0与总C18不饱和物的比率在未转化株中是大约4％，而在转化株中范围是从大约5％至69％。这些数据表明了多核苷酸SAD RNA发夹构建体在Prototheca moriformis中的成功表达和用途，以改变工程化宿主微生物中的饱和脂肪酸的百分比，并且尤其是增加微生物细胞中的C18:0脂肪酸的浓度和减少C18:1脂肪酸。

[0769] 实例13：通过β-酮脂酰-ACP合成酶II基因的过表达来改变微藻的脂肪酸谱。

[0770] β-酮脂酰-ACP合成酶II在脂肪酸生物合成过程中催化C16:0-ACP的2-碳延伸至C18:0-ACP。产生质粒构建体以评定宿主细胞的脂肪酸谱是否可以由于KASII基因的表达而受影响。KASII基因序列的来源选自Prototheca moriformis UTEX 1435或高等植物(大豆，GenBank登录号AAW88763；向日葵，GenBank登录号ABI18155；以及蓖麻，GenBank登录号AAA33872)。

[0771] 使用实例2的方法，将Prototheca moriformis UTEX 1435的以经典方式诱变的(为了更高的油产量)衍生物(A株)用在表53中的以下质粒构建体之一单独地转化。每种构建体含有针对用于整合至核基因组中的6S基因组区域的5’(SEQ ID NO:82)和3’(SEQ ID NO:84)同源重组靶向序列(在构建体侧翼)和在莱茵衣藻β-微管蛋白启动子/5’UTR和普通小球藻硝酸盐还原酶3’UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母

suc2表达盒作为SEQ ID NO:29列出并且用作选择标记。对于每种构建体,KASII编码区在Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:37)和普通小球藻硝酸盐还原酶3'UTR的控制下。将每种KASII酶的天然转运肽替换为原始小球藻硬脂酰-ACP去饱和酶转运肽(SEQ ID NO:54)。根据表2,所有蛋白质编码区经密码子优化以反映Prototheca moriformis UTEX 1435核基因中固有的密码子偏倚。

[0772] 表53.用来转化Prototheca moriformis(UTEX 1435)A株的质粒构建体。

[0773]

质粒构建体	KASII酶的来源	序列元件	SEQ ID. NO:
SZ1747	*大豆(Glm)*	6S::β-tub:suc2:nr::Amt03:S106SAD:GlmKASII:nr::6S	SEQ ID NO: 230
SZ1750	*向日葵(Ha)*	6S::β-tub:suc2:nr::Amt03:S106SAD:HaKASII:nr::6S	SEQ ID NO: 231
SZ1754	*蓖麻(Rc)*	6S::β-tub:suc2:nr::Amt03:S106SAD:RcKASII:nr::6S	SEQ ID NO: 232
SZ2041	*Prototheca moriformis(Pm)*	6S::β-tub:suc2:nr::Amt03:S106SAD:PmKASII:nr::6S	SEQ ID NO: 233

[0774] 构建体6S::β-Tub:suc2:nr::Amt03:S106SAD:PmKASII:nr::6S中的相关限制性位点以小写字母、粗体以及下划线指示并且分别是:5'-3'BspQ I、Kpn I、Xba I、Mfe I、BamH I、EcoR I、Spe I、Asc I、Cla I、Sac I、BspQ I。BspQI位点界定转化DNA的5'和3'末端。粗体、小写字母序列代表来自A株的允许通过同源重组在6S基因座处靶向整合的基因组DNA。在5'至3'方向继续向前,通过加框文本指示驱动酵母蔗糖转化酶基因(赋予A株代谢蔗糖的能力)的表达的莱茵衣藻β-微管蛋白启动子。转化酶的起始子ATG和终止子TGA由大写字母、粗斜体指示,同时编码区以小写字母斜体指示。普通小球藻硝酸盐还原酶3'UTR由小写字母加下划线文本指示,后接由加框斜体文本指示的Prototheca moriformis内源amt03启动子。PmKASII的起始子ATG和终止子TGA密码子由大写字母、粗斜体指示,同时该基因的其余部分以粗斜体指示。原始小球藻S106硬脂酰-ACP去饱和酶转运肽位于起始子ATG与Asc I位点之间。普通小球藻硝酸盐还原酶3'UTR再次由小写字母加下划线文本指示,后接由小写字母粗体文本指示的A株6S基因组区域。构建体6S::β-tub:suc2:nr::Amt03:S106SAD:PmKASII:nr::6S的相关核苷酸序列作为SEQ ID.NO:234提供于序列表中。包含原始小球藻S106硬脂酰-ACP去饱和酶转运肽的、密码子优化的PmKASII序列作为SEQ ID.NO:235提供于序列表中。SEQ ID NO:236提供了SEQ ID NO.235的蛋白质翻译。

[0775]

gctcttcgccgccgccactcctgctcgagcgcgcccgcgcgtgcgccgccagcgccttggccttttcgccgcgctcgtgcgcg
tcgctgatgtccatcaccaggtccatgaggtctgccttgcgccggctgagccactgcttcgtccgggcggccaagaggagca
tgagggaggactcctggtccagggtcctgacgtggtcgcggctctgggagcgggccagcatcatctggctctgccgcaccg
aggccgcctccaactggtcctccagcagccgcagtcgccgccgaccctggcagaggaagacaggtgaggggggtatgaat
tgtacagaacaaccacgagccttgtctaggcagaatccctaccagtcatggctttacctggatgacggcctgcgaacagctg
tccagcgaccctcgctgccgccgcttctcccgcacgcttctttccagcaccgtgatggcgcgagccagcgccgcacgctggc
gctgcgcttcgccgatctgaggacagtcggggaactctgatcagtctaaaccccccttgcgcgttagtgttgccatcctttgcag
accggtgagagccgacttgttgtgcgccacccccccacaccacctcctcccagaccaattctgtcacctttttggcgaaggcat
cggcctcggcctgcagagaggacagcagtgcccagccgctgggggttggcggatgcacgctcaggtaccctttcttgcgctat
gacacttccagcaaaaggtagggcgggctgcgagacggcttcccggcgctgcatgcaacaccgatgatgcttcgacccccccgaa
gctccttcggggctgcatgggcgctccgatgccgctccagggcgagcgctgtttaaatagccaggcccccgattgcaaagacatta
tagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatcgcactccgctaagg
gggcgcctcttcctcttcgtttcagtcacaacccgcaaactctagaatatca***ATG***ctgctgcaggccttcctgttcctgctggccggc
ttcgccgccaagatcagcgcctccatgacgaacgagacgtccgaccgccccctggtgcacttcaccccccaacaagggctggat
gaacgaccccaacggcctgtggtacgacgagaaggacgccaagtggcacctgtacttccagtacaacccgaacgacaccgt
ctgggggacgcccttgttctggggccacgccacgtccgacgacctgaccaactgggaggaccagcccatcgccatcgccccg
aagcgcaacgactccggcgccttctccggctccatggtggtggactacaacaacacctccggcttcttcaacgacaccatcgac
ccgcgccagcgctgcgtggccatctggacctacaacacccccggagtccgaggagcagtacatctcctacagcctggacggcg
gctacaccttcaccgagtaccagaagaaccccgtgctggccgccaactccacccagttccgcgacccgaaggtcttctggtacg
agccctcccagaagtggatcatgaccgcggccaagtcccaggactacaagatcgagatctactcctccgacgacctgaagtcc
tggaagctggagtccgcgttcgccaacgagggcttcctcggctaccagtacgagtgccccggcctgatcgaggtccccaccga
gcaggaccccagcaagtcctactgggtgatgttcatctccatcaaccccggcgccccggccggcggctccttcaaccagtacttc
gtcggcagcttcaacggcacccacttcgaggccttcgacaaccagtcccgcgtggtggacttcggcaaggactactacgccctg
cagaccttcttcaacaccgacccgacctacgggagcgccctgggcatcgcgtgggcctccaactgggagtactccgccttcgtg
cccaccaaccccctggcgctcctccatgtccctcgtgcgcaagttctccctcaacaccgagtaccaggccaacccggagacgga
gctgatcaacctgaaggccgagccgatcctgaacatcagcaacgccggcccctggagccggttcgccaccaacaccacgttg
acgaaggccaacagctacaacgtcgacctgtccaacagcaccggcaccctggagttcgagctggtgtacgccgtcaacacca
cccagacgatctccaagtccgtgttcgcggacctctccctctggttcaagggcctggaggaccccgaggagtacctccgcatgg
gcttcgaggtgtccgcgtcctccttcttcctggaccgcgggaacagcaaggtgaagttcgtgaaggagaacccctacttcaccaa
ccgcatgagcgtgaacaaccagcccttcaagagcgagaacgacctgtcctactacaaggtgtacggcttgctggaccagaac
atcctggagctgtacttcaacgacggcgacgtcgtgtccaccaacacctacttcatgaccaccgggaacgccctgggctccgtg
*aacatgacgacgggggtggacaacctgttctacatcgacaagttccaggtgcgcgaggtcaag**TGA**caattg*gcagcagca
gctcggatagtatcgacacactctggacgctggtcgtgtgatggactgttgccgccacacttgctgccttgacctgtgaatatccctgc
cgcttttatcaaacagcctcagtgtgtttgatcttgtgtgtacgcgcttttgcgagttgctagctgcttgtgctatttgcgaataccacccccc
agcatccccttccctcgtttcatatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctg
ctcactgcccctcgcacagccttggtttgggctccgcctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgca
cgggaagtagtgggatgggaacacaaatgga**ggatcc**cgcgtctcgaacagagcgcgcagaggaacgctgaaggtctcgcctc
tgtcgcacctcagcgcggcatacaccacaataaccacctgacgaatgcgcttggttcttcgtccattagcgaagcgtccggttcacac
acgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggtcgaaacgttcacagcctagggatatc**gaattc***g*
gccgacaggacgcgcgtcaaaggtgctggtcgtgtatgccctggccggcaggtcgttgctgctgctggttagtgattccgcaacc
ctgattttggcgtcttatttggcgtggcaaacgctggcgcccgcgagccgggccggcggcgatgcggtgccccacggctgccg
gaatccaagggaggcaagagcgcccgggtcagttgaagggctttacgcgcaaggtacagccgctcctgcaaggctgcgtggt
ggaattggacgtgcaggtcctgctgaagttcctccaccgcctcaccagcggacaaagcaccggtgtatcaggtccgtgtcatcc
actctaaagaactcgactacgacctactgatggccctagattcttcatcaaaaacgcctgagacacttgcccaggattgaaactc
cctgaagggaccaccaggggccctgagttgttccttccccccgtggcgagctgccagccaggctgtacctgtgatcgaggctgg

[0776]

cgggaaaataggcttcgtgtgctcaggtcatgggaggtgcaggacagctcatgaaacgccaacaatcgcacaattcatgtcaa
gctaatcagctatttcctcttcacgagctgtaattgtcccaaaattctggtctaccgggggtgatccttcgtgtacgggcccttccctc
aaccctaggtatgcgcgcatgcggtcgccgcgcaactcgcgcgagggccgagggtttgggacgggccgtcccgaaatgcagt
tgcacccggatgcgtggcacctttttgcgataatttatgcaatggactgctctgcaaaattctggctctgtcgccaaccctaggatc
agcggcgtaggatttcgtaatcattcgtcctgatggggagctaccgactaccctaatatcagcccgactgcctgacgccagcgtc
cacttttgtgcacacattccattcgtgcccaagacatttcattgtggtgcgaagcgtccccagttacgctcacctgtttcccgacctcc
*ttactgttctgtcgacagagcgggcccacaggccggtcgcagcc***actagt*ATGgccaccgcatccactttctcggcgttcaat**
gcccgctgcggcgacctgcgtcgctcggcgggctccgggccccggcgcccagcgaggcccctccccgtgcgcgggcgcgcc
gccgccgccgccgacgccaaccccgcccgccccgagcgccgcgtggtgatcaccggccagggcgtggtgacctccctgggc
cagaccatcgagcagttctactcctccctgctggagggcgtgtccggcatctcccagatccagaagttcgacaccaccggctac
accaccaccatcgccggcgagatcaagtccctgcagctggacccctacgtgcccaagcgctgggccaagcgcgtggacgac
gtgatcaagtacgtgtacatcgccggcaagcaggccctggagtccgccggcctgcccatcgaggccgccggcctggccggcg
ccggcctggaccccgccctgtgcggcgtgctgatcggcaccgccatggccggcatgacctccttcgccgccggcgtggaggcc
ctgacccgcggcggcgtgcgcaagatgaaccccttctgcatcccctctccatctccaacatgggcggcgccatgctggccatg
gacatcggcttcatgggccccaactactccatctccaccgcctgcgccaccggcaactactgcatcctgggcgccgccgaccac
atccgccgcggcgacgccaacgtgatgctggccggcggcgccgacgccgccatcatcccctccggcatcggcggcttcatcgc
ctgcaaggccctgtccaagcgcaacgacgagcccgagcgcgcctcccgccccctgggacgccgaccgcgacggcttcgtgat
gggcgagggcgccggcgtgctggtgctggaggagctggagcacgccaagcgccgcggcgccaccatcctggccgagctggt
gggcggcgccgccacctccgacgcccaccacatgaccgagcccgacccccagggccgcggcgtgcgcctgtgcctggagc
gcgccctggagcgcgcccgcctggcccccgagcgcgtgggctacgtgaacgcccacggcacctccacccccgccggcgacg
tggccgagtaccgcgccatccgcgccgtgatcccccaggactccctgcgcatcaactccaccaagtccatgatcggccacctgc
tgggcggcgccggcgccgtggaggccgtggccgccatccaggccctgcgcaccggctggctgcaccccaacctgaacctgg
agaaccccgcccccggcgtggaccccgtggtgctggtgggcccccgcaaggagcgcgccgaggacctggacgtggtgctgt
ccaactccttcggcttcggcggccacaactcctgcgtgatcttccgcaagtacgacgagatggactacaaggaccacgacggc
gactacaaggaccacgacatcgactacaaggacgacgacgacaag*TGA*atcgatagatctcttaaggcagcagcagctcg
gatagtatcgacacactctggacgctggtcgtgtgatggactgttgccgccacacttgctgccttgacctgtgaatatccctgccgcttt
tatcaaacagcctcagtgtgtttgatcttgtgtgtacgcgcttttgcgagttgctagctgcttgtgctatttgcgaataccacccccagcat
ccccttccctcgtttcatatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcact
gcccctcgcacagccttggtttgggctccgcctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacggga
agtagtgggatgggaacacaaatggaaagcttaattaa**gagctcttgttttccagaaggagttgctccttgagcctttcattctcag**
cctcgataacctccaaagccgctctaattgtggagggggttcgaatttaaaagcttggaatgttggttcgtgcgtctggaaca
agcccagacttgttgctcactgggaaaaggaccatcagctccaaaaaacttgccgctcaaaccgcgtacctctgctttcgcg
caatctgccctgttgaaatcgccaccacattcatattgtgacgcttgagcagtctgtaattgcctcagaatgtggaatcatctg
ccccctgtgcgagcccatgccaggcatgtcgcgggcgaggacacccgccactcgtacagcagaccattatgctacctcaca
atagttcataacagtgaccatatttctcgaagctccccaacgagcacctccatgctctgagtggccacccccggccctggtg
cttgcggagggcaggtcaaccggcatggggctaccgaaatccccgaccggatcccaccacccccgcgatgggaagaatct
ctccccgggatgtgggcccaccaccagcacaacctgctggcccaggcgagcgtcaaaccataccacacaaatatccttggc
atcggccctgaattccttctgccgctctgctacccggtgcttctgtccgaagcaggggttgctagggatcgctccgagtccgca
aacccttgtcgcgtggcggggcttgttcgagcttgaagagc

[0777] 在将每种质粒构建体单独地转化进入A株后,在包含蔗糖作为唯一碳源的琼脂平板上选择阳性克隆。如在先前的实例中,将初级转化子以克隆方式纯化并且在标准脂质产生条件下在pH 7生长并且从每种转化子的干燥生物质制备脂质样品。使用如在实例11中所述的直接酯交换方法来确定脂肪酸谱。与未转化A株对照的那些脂肪酸谱比较,针对每种质粒构建体,从一组代表性克隆得到的脂肪酸谱(表示为总脂肪酸的面积%)汇总在以下表54中,这些代表性克隆由一些阳性转化子的A株脂肪酸谱(表示为面积%)的转化而产生。

[0778] 表54.经工程化而过表达KAS II基因的Prototheca moriformis细胞的脂肪酸谱。

[0779]

质粒构建体	KASII 来源	转化子	%C14:0	%C16:0	%C18:0	%C18:1	%C18:2
无	无过表达	1	1.36	28.69	2.92	56.36	8.16
		2	1.35	28.13	3.57	55.63	8.79
		3	1.22	25.74	2.82	60.6	7.31
		4	1.22	25.74	2.82	60.6	7.31
pSZ1747	大豆	1	2.23	25.34	2.69	57.35	9.53
		2	2.18	25.46	2.74	57.35	9.46
		3	2.18	25.33	2.89	57.34	9.5
		4	2.2	25.69	2.66	57.28	9.43
		5	2.17	25.38	3.03	56.99	9.72
pSZ1750	向日葵	1	2.43	26.82	2.72	55.17	9.87
		2	2.44	27.14	2.62	54.89	9.81
		3	2.61	26.9	2.67	54.43	10.25
		4	1.96	30.32	2.87	53.87	8.26
		5	2.55	27.64	2.98	53.82	10.07
pSZ1754	蓖麻	1	1.84	24.41	2.89	59.26	9.08
		2	1.3	25.04	2.81	58.75	9.65
		3	1.27	25.98	2.76	58.33	9.22
		4	1.95	25.34	2.77	58.15	9.22
		5	1.3	26.53	2.75	57.87	9.09
pSZ2041	*P.moriformis*	1	1.63	11.93	3.62	70.95	9.64
		2	1.85	11.63	3.34	69.88	10.93
		3	1.84	12.01	3.81	69.56	10.45
		4	1.63	14.22	3.72	68.86	9.6
		5	1.67	15.04	3.05	68.63	9.24

[0780] 使用pSZ2041利用在表2中指示的密码子频率进行了进一步密码子优化的Prototheca moriformis(UTEX 1435)KASII基因的过表达后,观察到C16:0链长的明显减少,伴随C18:1长度脂肪酸的同时增加。在表达由β-微管蛋白启动子驱动的Prototheca moriformis(UTEX 1435)KASII基因的构建体的转化后,观察到类似的脂肪酸谱变化。

[0781] 这些结果显示,密码子优化的无绿藻属脂质生物合成基因的过表达可以改变基因工程化微藻的脂肪酸谱。具体地说,KASII基因的过表达可以将C18脂肪酸的百分比从未转化细胞中的大约68%增加至大约84%。

[0782] 实例14:通过KASI等位基因的靶向敲除改变工程化无绿藻属中的中链脂肪酸水平

[0783] β-酮脂酰-ACP合成酶I(KASI)在脂肪酸生物合成过程中催化C4:0、C6:0、C8:0、C10:0、C12:0、以及C14:0脂肪酰基-ACP分子的2-碳延伸。在这个实例中,产生了一种敲除质粒构建体pSZ2014来评定在KASI基因位点靶向破坏后对宿主细胞的脂肪酸谱的影响。

[0784] 使用在实例2中所述的生物射弹转化方法,用pSZ2014构建体转化Prototheca moriformis UTEX 1435的以经典方式诱变的(为了更高的油产量)衍生物(A株)。pSZ2014含有在莱茵衣藻β-微管蛋白启动子和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母

suc2转化酶表达盒，其借助于KASI基因特异性同源区域而位于任一侧的侧翼，以便使该构建体靶向整合进入Prototheca moriformis基因组的KASI基因座。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。用于核基因组整合的KASI基因座的靶向区域的相关序列显示于下文并且列于SEQ ID NO:238和SEQ ID NO:239中。pSZ2014中的相关限制性位点(以小写字母、粗体和下划线指示)分别是5'-3'BspQ 1、Kpn I、AscI、Xho I、Sac I、BspQ I，显示于以下序列中。BspQI位点界定转化DNA的5'和3'末端。粗体、小写字母序列代表来自A株的允许通过同源重组在KASI基因座处靶向整合的基因组DNA。在5'至3'方向继续向前，通过加框文本指示驱动密码子优化的酵母蔗糖转化酶基因(赋予A株代谢蔗糖的能力)的表达的莱茵衣藻β-微管蛋白启动子。suc2转化酶的起始子ATG密码子和终止子TGA密码子由大写字母、粗斜体指示，同时编码区以小写字母斜体指示。普通小球藻硝酸盐还原酶3'UTR由小写字母加下划线文本指示。pSZ2014的转化序列显示于下文并且作为SEQ ID NO:237列出。

[0785]

gctcttcgctcaccgcgtgaattgctgtcccaaacgtaagcatcatcgtggctcggtcacgcgatcctggatccggggatcct
agaccgctggtggagagcgctgccgtcggattggtggcaagtaagattgcgcaggttggcgaagggagagaccaaaacc
ggaggctggaagcgggcacaacatcgtattattgcgtatagtagagcagtggcagtcgcatttcgaggtccgcaacggatct
cgcaagctcgctacgctcacagtaggagaaaggggaccactgcccctgccagaatggtcgcgaccctctccctcgccggcc
ccgcctgcaacacgcagtgcgtatccggcaagcgggctgtcgccttcaaccgcccccatgttggcgtccgggctcgatcagg
tgcgctgaggggggtttggtgtgcccgcgcctctgggcccgtgtcggccgtgcggacgtggggccctgggcagtggatcagc
agggtttgcgtgcaaatgcctataccggcgattgaatagcgatgaacgggatacggttgcgctcactccatgcccatgcgac
cccgtttctgtccgccagccgtggtcgcccgggctgcgaagcgggaccccacccagcgcattgtgatcaccggaatgggcgt
ggggtaccctttcttgcgctatgacacttccagcaaaaggtagggcgggctgcgagacggcttcccggcgctgcatgcaacaccg
atgatgcttcgaccccccgaagctccttcggggctgcatgggcgctccgatgccgctccagggcgagcgctgtttaaatagccagg
ccccccgattgcaaagacattatagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagct
tgtgatcgcactccgctaagggggcgcctcttcctcttcgtttcagtcacaacccgcaaac**ggcgcgcc*ATG***ctgctgcaggcctt
cctgttcctgctggccggcttcgccgccaagatcagcgcctccatgacgaacgagacgtccgaccgcccccctggtgcacttcac
ccccaacaagggctggatgaacgaccccaacggcctgtggtacgacgagaaggacgccaagtggcacctgtacttccagta
caacccgaacgacaccgtctgggggacgcccttgttctggggccacgccacgtccgacgacctgaccaactgggaggacca
gcccatcgccatcgccccgaagcgcaacgactccggcgccttctccggctccatggtggtggactacaacaacacctccggctt
cttcaacgacaccatcgacccgcgccagcgctgcgtggccatctggacctacaacacccccggagtccgaggagcagtacatct
cctacagcctggacggcggctacaccttcaccgagtaccagaagaaccccgtgctggccgccaactccacccagttccgcga
cccgaaggtcttctggtacgagccctcccagaagtggatcatgaccgcggccaagtcccaggactacaagatcgagatctact
cctccgacgacctgaagtcctggaagctggagtccgcgttcgccaacgagggcttcctcggctaccagtacgagtgccccggc
ctgatcgaggtccccaccgagcaggaccccagcaagtcctactgggtgatgttcatctccatcaaccccggcgccccggccgg
cggctccttcaaccagtacttcgtcggcagcttcaacggcacccacttcgaggccttcgacaaccagtcccgcgtggtggacttc
ggcaaggactactacgccctgcagaccttcttcaacaccgacccgacctacgggagcgccctgggcatcgcgtgggcctccaa
ctgggagtactccgccttcgtgcccaccaaccccctggcgctcctccatgtccctcgtgcgcaagttctccctcaacaccgagtacc
aggccaacccggagacggagctgatcaacctgaaggccgagccgatcctgaacatcagcaacgccggcccctggagccgg
ttcgccaccaacaccacgttgacgaaggccaacagctacaacgtcgacctgtccaacagcaccggcaccctggagttcgagc
tggtgtacgccgtcaacaccacccagacgatctccaagtccgtgttcgcggacctctccctctggttcaagggcctggaggaccc
cgaggagtacctccgcatgggcttcgaggtgtccgcgtcctccttcttcctggaccgcgggaacagcaaggtgaagttcgtgaa
ggagaaccctacttcaccaaccgcatgagcgtgaacaaccagcccttcaagagcgagaacgacctgtcctactacaaggtg

[0786]

*tacggcttgctggaccagaacatcctggagctgtacttcaacgacggcgacgtcgtgtccaccaacacctacttcatgaccacc gggaacgccctgggctccgtgaacatgacgacgggggtggacaacctgttctacatcgacaagttccaggtgcgcgaggtca ag**TGA**caattg*gcagcagcagctcggatagtatcgacacactctggacgctggtcgtgtgatggactgttgccgccacacttgct gccttgacctgtgaatatccctgccgcttttatcaaacagcctcagtgtgtttgatcttgtgtgtacgcgcttttgcgagttgctagctgctt gtgctatttgcgaataccacccccagcatcccccttccctcgtttcatatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatc cctcagcgctgctcctgctcctgctcactgcccctcgcacagccttggtttgggctccgcctgtattctcctggtactgcaacctgtaaa ccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaggatcgta**gagctccacctgcatccgcctgg cgctcgaggacgccggcgtctcgcccgacgaggtcaactacgtcaacgcgcacgccacctccaccctggtgggcgacaag gccgaggtgcgcgcggtcaagtcggtctttggcgacatgaagggcatcaagatgaacgccaccaagtccatgatcgggca ctgcctgggcgccgccggcggcatggaggccgtcgccacgctcatggccatccgcaccggctgggtgcaccccaccatcaa ccacgacaaccccatcgccgaggtcgacggcctggacgtcgtcgccaacgccaaggcccagcacaaaatcaacgtcgcca tctccaactccttcggcttcggcgggcacaactccgtcgtcgcctttgcgcccttccgcgagtaggcggagcgagcgcgcttg gctgaggagggaggcggggtgcgagcccttggctgcgcgcgatactctccccgcacgagcagactccacgcgcctgaatc tacttgtcaacgagcaaccgtgtgttttgtccgtggccattcttattatttctccgactgtggccgtactctgtttggctgtgcaag caccgaagagc***

[0787] 将质粒构建体pSZ2014转化进入A株后，在含有蔗糖作为唯一碳源的平板上选择阳性克隆。将初级转化子以克隆方式纯化并且使其在标准脂质产生条件下生长。从每种转化子的干燥生物质制备脂质样品。使用如在实例11中所述的直接酯交换方法来确定脂肪酸谱。与未转化A株对照的那些脂肪酸谱比较，一些阳性转化子的脂肪酸谱(表示为总脂肪酸的面积%)汇总在以下表55中。

[0788] 表55.包含选择标记以破坏内源KASI等位基因的工程化Prototheca moriformis细胞的脂肪酸谱。

[0789]

	转化子	% C14:0	% C16:0	% C18:0	% C18:1	% C18:2
A 株	未转化对照	1.22	25.61	2.82	60.76	7.44
A 株 pSZ2014	转化子 1	1.65	32.55	2.17	53.99	7.43
	转化子 2	2.25	30.04	2.57	55.86	7.12
	转化子 3	3.51	31.22	1.90	53.85	7.00
	转化子 4	4.09	31.51	2.57	53.14	6.21
	转化子 5	4.68	34.47	1.94	49.75	6.49
	转化子 6	5.68	37.98	1.83	44.76	6.75
	转化子 7	5.82	37.82	1.93	44.84	6.44

[0790] 如在上表55中所显示的，KASI等位基因的靶向中断影响了转化微生物的脂肪酸谱。包含pSZ2014转化载体的A株的脂肪酸谱显示出C14:0和C16:0脂肪酸的组成增加，伴随C18:1脂肪酸的同时减少。在所有转化子中，C18:0脂肪酸减少。在一些转化中，KASI等位基因的中断进一步导致了一种脂肪酸谱，该脂肪酸谱包括相对于未转化A株生物的脂肪酸谱而言降低的C18:2脂肪酸百分比。

[0791] 因此，我们通过内源KASI的破坏使总C14脂肪酸的百分比增加了大约35％至400％并且使C16脂肪酸的百分比增加了大约30％至50％。

[0792] 这些数据证明了内源KASI等位基因的靶基因中断改变宿主微生物的脂肪酸谱的效用。

[0793] 实例15:结合基因修饰方法以改变无绿藻属中的脂肪酸谱

[0794] 在这个实例中,显示了将敲除KASII等位基因与同时过表达展示出中链脂肪酸水解偏好性的外源硫酯酶的基因修饰在微生物中的结合来改变该宿主生物的脂肪酸谱。

[0795] 根据在实例2中详述的生物射弹转化方法,用质粒构建体pSZ1283初始转化Prototheca moriformis(UTEX 1435)的以经典方式诱变(为了更高的油产量)的衍生物(C株)。pSZ1283(SEQ ID NO:258)(先前描述于PCT申请号PCT/US2011/038463和PCT/US2011/038463中)含有Cuphea wrightii FATB2(CwTE2)硫酯酶的编码序列、针对用于整合至核基因组中的6S基因组区域的5'(SEQ ID NO:82)和3'(SEQ ID NO:84)同源重组靶向序列(在构建体侧翼)、以及在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。CwTE2编码区在Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR(SEQ ID NO:32)的控制下。根据表2,CwTE2和suc2的蛋白编码区经密码子优化以反映Prototheca moriformis UTEX 1435核基因中固有的密码子偏倚。

[0796] 将pSZ1283转化进入C株后,在含有蔗糖作为唯一碳源的琼脂平板上选择阳性克隆。然后将初级转化子以克隆方式纯化并且选择了一个单个转化子(B株),用于进一步基因修饰。将这个基因工程化株系用质粒构建体pSZ2110(SEQ ID NO:240)转化,以便中断KASII等位基因1基因座。pSZ2110(写为KASII‘5::CrbTub:NeoR:nr::KASII-‘3)包含在莱茵衣藻β-微管蛋白启动子和普通小球藻硝酸盐还原酶3'UTR的控制下的新霉素抗性(NeoR)表达盒(赋予对G418的抗性),其借助于KASII基因特异性同源区域而位于任一侧的侧翼,以便使该构建体靶向整合进入Prototheca moriformis基因组的KASII基因座。在pSZ2110构建体中的相关限制性位点从5'-3'是BspQ 1、KpnI、XbaI、MfeI、BamHI、EcoRI、SpeI、XhoI、Sac I、以及BspQI,以小写字母、粗体、以及下划线格式指示。BspQI位点界定转化DNA的5'和3'末端。该转化构建体的5'和3'末端处的粗体、小写字母序列代表来自UTEX 1435的、经由同源重组靶向整合至KASII等位基因1基因座的基因组DNA。莱茵衣藻β-微管蛋白由小写字母加框文本指示。NeoR的起始子ATG和终止子TGA由大写字母斜体指示,同时编码区以小写字母斜体指示。

[0797] 3'UTR由小写字母加下划线文本显示。

[0798]

gctcttcgctggcctgttgccggacgatccgtgtcgtcgagactgcattttgttttgggtgtggggctggggtactggatggctt gagggcatgacttttctgatggagaagattgcaatgagatcatttgggtcgtctatttgtttgctgtgcaagagggtttactgg tatctggcaccagcttttggcccgtgcccgtttgatggacgcgtgacaggcaggcgtcctggaaagcacagacaccgtacgt acgaccttgacctccccccttctccacacggcaggtgcgaggctgcccacggcgtcgaggcgggcggtgcgccgggcatg gtcccgcatcgcgcgcgcggcggccgcggccgacgcaaaccccgcccgccctgagcgccgcgtggtcatcacgggccagg gcgtggtgaccagcctgggccagacgatcgagcagttttacagcagcctgctggagggcgtgagcggcatctcgcagatac agaagttcgacaccacgggctacacgacgacgatcgcgggcgagatcaagtcgctgcagctggacccgtacgtgcccaag

[0799]

cgctgggcgaagcgcgtggacgacgtgataaagtacgtctacatcgcgggcaagcaggcgctggagagcgccggcctgcc
gatcgaggcggcggggctggcgggcgcggggctggacccggcgctgtgcggcgtgctcatcggcaccgccatggcgggc
atgacgtctttcgcggcgggcgtggaggcgctgacgcgcggtaccctttcttgcgctatgacacttccagcaaaaggtagggcg
ggctgcgagacggcttcccggcgctgcatgcaacaccgatgatgcttcgacccccgaagctccttcggggctgcatgggcgctc
cgatgccgctccagggcgagcgctgtttaaatagccaggcccccgattgcaaagacattatagcgagctaccaaagccatattcaa
acacctagatcactaccacttctacacaggccactcgagcttgtgatcgcactccgctaaggggggcgcctcttcctcttcgtttcagtc
acaacccgcaaac**tctaga**atatca*ATGatcgagcaggacggcctccacgccggctcccccgccgcctgggtggagcgcct*
gttcggctacgactgggcccagcagaccatcggctgctccgacgccgccgtgttccgcctgtccgcccagggccgccccgtgct
gttcgtgaagaccgacctgtccggcgccctgaacgagctgcaggacgaggccgcccgcctgtcctggctggccaccaccggc
gtgccctgcgccgccgtgctggacgtggtgaccgaggccggccgcgactggctgctgctgggcgaggtgcccggccaggacc
tgctgtcctcccacctggcccccgccgagaaggtgtccatcatggccgacgccatgcgccgcctgcacaccctggaccccgcc
acctgccccttcgaccaccaggccaagcaccgcatcgagcgcgcccgcacccgcatggaggccggcctggtggaccaggac
gacctggacgaggagcaccagggcctggcccccgccgagctgttcgcccgcctgaaggcccgcatgcccgacggcgagga
cctggtggtgacccacggcgacgcctgcctgcccaacatcatggtggagaacggccgcttctccggcttcatcgactgcggccg
cctgggcgtggccgaccgctaccaggacatcgccctggccacccgcgacatcgccgaggagctgggcggcgagtgggccga
*ccgcttcctggtgctgtacggcatcgccgcccccgactcccagcgcatcgccttctaccgcctgctggacgagttcttcTGA***caa**
ttggcagcagcagctcggatagtatcgacacactctggacgctggtcgtgtgatggactgttgccgccacacttgctgccttgacctg
tgaatatccctgccgcttttatcaaacagcctcagtgtgtttgatcttgtgtgtacgcgcttttgcgagttgctagctgcttgtgctatttgcg
aataccacccccagcatccccttccctcgtttcatatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatccctcagcgctg
ctcctgctcctgctcactgcccctcgcacagccttggtttgggctccgcctgtattctcctggtactgcaacctgtaaaccagcactgca
atgctgatgcacgggaagtagtgggatgggaacacaaatggaggatccactagttctagagcggccgccaccgcggtg**gagctc**
ggcggcgtgcgcaagatgaaccccttttgcatcccctctccatctccaacatgggcggcgcgatgctggcgatggacatcg
gcttcatgggccccaactactccatctccacggcctgcgcgacgggcaactactgcatcctgggcgcggcggaccacatccg
gcgcggcgacgcaaacgtgatgctggccggcggcgcggacgcggccatcatcccctcgggcatcggcggcttcatcgcgtg
caaggcgctgagcaagcgcaacgacgagcccgagcgcgcgagccggccctgggacgccgaccgcgacggcttcgtcatg
ggcgagggcgccggcgtgctggtgctggaggagctggagcacgccaagcgccgcggcgcgaccattttggctgaattagtt
ggcggcgcggccacctcggacgcgcaccacatgaccgagcccgacccgcagggccgcggcgtgcgcctctgcctcgagcg
cgcgctcgagcgcgcgcgcctcgcgcccgagcgcgtcggctacgtcaacgcgcacggcaccagcacgcccgcgggcgac
gtggccgagtaccgcgccatccgcgccgtcatcccgcaggactcactacgcatcaactccacaaagtccatgatcgggcac
ctgctcggcggcgccggcgcggtcgaggccgtggccgccatccaggccctgcgcaccggctggctccaccccaacttgaac
ctcgagaaccccgcgcctggcgtcgaccccgtcgtgctcgtgggctcttcc

[0800] 用pSZ2110转化B株后,在含有G418的选择性琼脂平板上选择阳性克隆。然后将初级转化子以克隆方式纯化并且使其以蔗糖作为碳源、在标准脂质产生条件下、在pH 5.0和pH 7.0两者下生长。如在实例11中所述从每种转化子的干燥生物质制备脂质样品。在以下表56中呈现了5种阳性转化子(T1-T5)的脂肪酸谱(表示为总脂肪酸的面积%)、在作为唯一碳源的蔗糖上生长的B株的谱(U1)、以及在以作为唯一碳源的葡萄糖上生长的未转化UTEX 1435的谱(U1)。

[0801] 表56.经多重工程化以消融内源KASII基因产物并且表达Cuphea wrightii硫酯酶的Prototheca moriformis(UTEX 1435)的脂肪酸谱。

[0802]

		株						
脂肪酸	pH	UTEX 1435	B 株	B 株 pSZ2011				
		U1	U1	T1	T2	T3	T4	T5
% C10:0	pH 5.0	0.01	*	0.02	0.03	0.03	0.03	0.07

[0803]

	pH 7.0	0.01	5.35	4.94	4.79	4.70	4.10	4.12
% C12:0	pH 5.0	0.04	*	0.09	0.39	0.41	0.41	0.86
	pH 7.0	0.04	27.06	25.31	25.03	25.02	23.54	23.47
% C14:0	pH 5.0	1.30	*	0.92	1.13	1.14	1.15	1.50
	pH 7.0	1.56	15.20	14.49	14.22	14.50	15.84	15.85
% C16:0	pH 5.0	25.89	*	36.07	35.05	35.35	35.23	35.05
	pH 7.0	29.80	13.89	14.96	14.88	15.11	22.62	22.94
% C18:0	pH 5.0	2.84	*	1.76	1.79	1.84	1.82	1.80
	pH 7.0	3.00	1.44	1.57	1.71	1.49	1.37	1.33
% C18:1	pH 5.0	60.34	*	49.82	50.21	49.97	50.01	49.13
	pH 7.0	54.96	28.57	29.84	30.45	30.27	24.18	23.96
% C18:2	pH 5.0	7.40	*	8.39	8.59	8.47	8.52	8.75
	pH 7.0	8.15	6.85	7.19	7.16	7.15	6.51	6.65

[0804] *未测试

[0805] 如在表56中所显示的,CwTE2在Prototheca moriformis(UTEX 1435)B株中的表达在转化的微生物的脂肪酸谱方面的影响显著改变。表达CwTE2并且在pH 7.0进行培养以从Amt03启动子促进CwTE2表达的株(B株)的脂肪酸谱显示:相对于未转化UTEX 1435的脂肪酸谱,其C10:0、C12:0、以及C14:0脂肪酸组成增加,伴随C16:0和C18:1脂肪酸组成的同时减少。随后修饰B株以中断KASII等位基因(编码一种催化C16:0 2-碳延伸至C18:0脂肪酸的酶),导致在pH 5.0培养时存在于新近工程化的株系的脂质谱中的C16:0脂肪酸增加,伴随C18:1脂肪酸的同时减少。在pH 5.0的转化子增殖表明除了造成脂肪酸谱改变的原因的硫酯酶之外的KASII等位基因敲除的影响,因为这种培养基的pH对于Amt03启动子的活性不是最优的。在pH7.0处产生脂质后(因此表达CwTE2),pSZ2011转化子展示出一个脂肪酸谱,相对于UTEX 1435株的谱,该脂肪酸谱在C10:0、C12:0、以及C14:0脂肪酸的组成方面增加,伴随C16:0和C18:1脂肪酸的同时减少。当在pH 7.0培养时,一些pSZ2011转化子展示出一个脂肪酸谱,相对于这些转化子的亲本株(B株)在pH 7.0培养时的脂肪酸谱,该脂肪酸谱富含C16:0脂肪酸并且仍然伴随C18:1脂肪酸组成的进一步减少。

[0806] 这些数据证明了多重基因修饰影响宿主生物的脂肪酸谱的效用。另外,这个实例展示了针对内源KASII等位基因的靶基因中断以改变宿主微生物的脂肪酸谱的重组多核苷酸的用途。

[0807] 实例16:结合基因修饰方法以改变无绿藻属中的棕榈酸组成

[0808] 在这个实例中,显示了将敲除KASII等位基因与同时过表达展示出对于C14和C16脂肪酸水解的优先特异性的外源硫酯酶的基因修饰在微生物中的结合来改变该宿主生物的脂肪酸谱。

[0809] 根据在实例2中详述的生物射弹转化方法,用质粒构建体pSZ2004转化Prototheca moriformis(UTEX 1435)的以经典方式诱变(为了更高的油产量)的衍生物(A株)。pSZ2004(写为KASII_5'_btub-SUC2-nr_2X_Amt03-Ch16TE2-nr_KASII_3')包含萼距花脂肪酰基-ACP硫酯酶(Ch16TE2,GenBank#Q39513)、用于在核基因组的KASII基因座处靶向整合的5'和3'同源重组靶向序列(在构建体侧翼)、以及在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。Ch16TE2是一种显示出针对C14和C16脂肪酸的优先特异性的硫酯酶。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。Ch16TE编码区在串联重复的Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR的控制下。根据表2,Ch16TE和suc2的蛋白编码区经密码子优化以反映在Prototheca moriformis UTEX 1435核基因中固有的密码子偏倚。pSZ2004作为SEQ ID NO:250存在于序列表中。

[0810] 用pSZ2004转化后,在含有蔗糖作为唯一碳源的平板上选择初级转化子。将各个转化子以克隆方式进行纯化并且使其在标准脂质产生条件下在pH 7.0生长,与在实例1中披露的条件类似。使用标准脂肪酸甲酯气相色谱火焰电离(FAME GC/FID)检测方法分析脂肪酸谱,如在实例11中所述。从由转化载体pSZ2004的转化产生的代表性克隆得到的脂肪酸谱(表示为总脂肪酸的面积%)显示于表57中。从在包含葡萄糖作为唯一碳源的脂质产生条件下生长的未转化株获得的脂质的脂肪酸谱另外地呈现在表57中。

[0811] 表57.经多重工程化以消融内源KASII基因产物并且表达萼距花硫酯酶的Prototheca moriformis(UTEX 1435)的脂肪酸谱。

[0812]

脂肪酸	UTEX 1435	pSZ2004
%C10:0	0.01	0.00
%C12:0	0.04	0.09
%C14:0	1.27	6.42
%C16:0	27.20	69.97
%C18:0	3.85	1.84
%C18:1	58.70	13.69
%C18:2	7.18	7.15

[0813] 如在以上表57中所显示的,具有用于选择标记表达的表达盒和及C14/C16偏好型硫酯酶的KASII等位基因的靶向中断影响了转化的微生物的脂肪酸谱。包含pSZ2004转化载体的株的脂肪酸谱显示出C14:0和C16:0脂肪酸的组成增加,伴随C18:0和C18:1脂肪酸的同时减少。未转化的Prototheca moriformis的(UTEX 1435)株展示出包含大约27%C16脂肪酸和大约58%C18:1脂肪酸的脂肪酸谱。相反,在KASII基因座被一个盒(使萼距花脂肪酰基-ACP硫酯酶和选择标记能够表达)破坏的株的脂肪酸谱包括大约70%C16脂肪酸和大约14%脂肪酸。C16:0的水平增加了超过2.5倍。这些数据显示,可以将外源基因过表达的基因修饰和内源基因消融进行结合以改变宿主生物中的脂肪酸谱。

[0814] 为了比较,在KASII基因座被一个盒(使蔗糖转化酶基因能够表达)破坏的株的脂肪酸谱提供了一个具有大约35%C16脂肪酸和大约50%C18:1脂肪酸的株。

[0815] 这些数据证明了多核苷酸的效用和效果,这些多核苷酸允许硫酯酶的外源表达而

改变工程化微生物的脂肪酸谱，并且尤其是增加C14和C16脂肪酸的浓度并且伴随地通过用所述多核苷酸的KASII等位基因的靶向破坏而引起微生物细胞中C18:0和C18:1脂肪酸减少。

[0816] 实例17：将微生物工程化以生产亚油酸不饱和脂肪酸和甘油脂

[0817] 某些Δ12脂肪酸去饱和酶可以催化C18:1脂肪酸或脂肪酰基分子中的双键形成，由此产生C18:2(亚油酸)脂肪酸或脂肪酰基分子。某些产生富含亚油酸不饱和脂肪酸的油的植物种类(包括陆地棉、红花、大豆、向日葵、以及玉米)是编码Δ12去饱和酶的基因的来源，这些基因可以在微生物中表达以影响脂肪酸谱。这个实例描述了编码Δ12去饱和酶的多核苷酸将微生物工程化的用途，其中该转化的微生物的脂肪酸谱已经富含亚油酸。

[0818] 根据在实例2中详述的生物射弹转化方法，用在表58中列出的以下质粒构建体之一转化Prototheca moriformis UTEX 1435的以经典方式诱变(为了更高的油产量)的衍生物(A株)。每种构建体含有针对用于整合至核基因组中的6S基因组区域的5'(SEQ ID NO:82)和3'(SEQ ID NO:84)同源重组靶向序列(在构建体侧翼)以及在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。根据表2，所有蛋白编码区经密码子优化以反映Prototheca moriformisUTEX 1435中固有的密码子偏倚。来自陆地棉的去饱和酶基因(Gh，GenBank登录号CAA71199)、来自红花的去饱和酶基因(Ct GenBank登录号ADM48789)、来自大豆的去饱和酶基因(Gm，GenBank登录号BAD89862)、来自向日葵的去饱和酶基因(Ha，GenBank登录号AAL68983)以及来自玉米的去饱和酶基因(Zm，GenBank登录号ABF50053)的编码区各自在Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR的控制下。

[0819] 表58.用来转化Prototheca moriformis(UTEX 1435)A株的质粒构建体。

[0820]

质粒构建体	相关序列元件	SEQ ID NO:
pSZ2150	6S::Crβtub:ScSuc2:Cvnr::PmAmt03:CtFad2-2:Cvnr::6S	SEQ ID NO: 241
pSZ2151	6S::Crβtub:ScSuc2:Cvnr::PmAmt03:GlmFad2-2:Cvnr::6S	SEQ ID NO: 242
pSZ2152	6S::Crβtub:ScSuc2:Cvnr::PmAmt03:HaFad2:Cvnr::6S	SEQ ID NO: 243
pSZ2153	6S::Crβtub:ScSuc2:Cvnr::PmAmt03:ZmFad2:Cvnr::6S	SEQ ID NO: 244
pSZ2172	6S::Crβtub:ScSuc2:Cvnr::PmAmt03:GhFad2:Cvnr::6S	SEQ ID NO: 245

[0821] 将在表58中列出的构建体的每一个单独地转化进入A株。在含有蔗糖作为唯一碳源的平板上选择初级转化子。将各个转化子以克隆方式进行纯化并且使其在标准脂质产生条件下在pH 7.0生长，与如在实例1中披露的条件类似。使用在实例11中所述的标准脂肪酸甲酯气相色谱火焰电离(FAME GC/FID)检测方法分析脂肪酸谱。从由表58的相应A株转化而产生的一组代表性克隆得到的脂肪酸谱显示于表59中。为了比较，将从未转化的A株对照细胞获得的脂质的脂肪酸谱另外地呈现在表59中。

[0822] 表59.经工程化而表达外源FAD去饱和酶的Prototheca moriformis细胞的C18:1、

C18:2、以及C18:3脂肪酸谱。

[0823]

		总脂肪酸的%				
株	样品	C18:1	C18:2	C18:3	总C18不饱和物（总脂肪酸的%）	C18多不饱和物的%/总C18不饱和物
A株	未转化	55.37	8.18	0.7	64.25	13.82
A株 pSZ2150 *Ct* FAD2-2	转化子1	58.29	11.96	0.75	71	17.90
	转化子2	59.00	10.29	0.84	70.13	15.87
	转化子3	52.41	10.36	1.16	63.93	18.02
	转化子4	56.77	10.17	0.8	67.74	16.19
	转化子5	57.19	10.15	0.79	68.13	16.06
A株 pSZ2151 *Glm* FAD2-2	转化子1	58.09	10.32	0.86	69.27	16.14
	转化子2	59.6	11.87	0.86	72.33	17.60
	转化子3	58.93	11.54	0.83	71.3	17.35
	转化子4	58.4	12.29	0.9	71.59	18.42
	转化子5	58.27	10.8	0.83	69.9	16.64
	转化子6	58.85	10.48	0.82	70.15	16.11
A株 pSZ2152 *Ha* FAD2	转化子1	59.30	10.02	0.82	70.14	15.45
	转化子2	58.45	9.87	0.81	69.13	15.45
	转化子3	59.38	9.89	0.79	70.06	15.24
	转化子4	59.54	9.79	0.81	70.14	15.11
	转化子5	59.07	9.92	0.82	69.81	15.38
	转化子6	59.57	10.02	0.55	70.14	15.07
A株 pSZ2153 *Zm* FAD2	转化子1	64.30	11.18	0.89	76.37	15.80
	转化子2	58.54	10.49	0.88	69.91	16.26
	转化子3	58.80	9.95	0.81	69.56	15.47
	转化子4	56.18	10.81	1.16	68.15	17.56
	转化子5	58.82	10.02	0.83	69.67	15.57
	转化子6	58.72	10.06	0.86	69.64	15.68
A株 pSZ2172 *Gh*-FLAG-FAD2	转化子1	55.71	10.85	0.84	67.4	17.34
	转化子2	56.12	10.29	0.75	67.16	16.44
	转化子3	54.14	12	0.96	67.1	19.31
	转化子4	55.72	11.68	0.75	68.15	18.24

[0824] 未转化的Prototheca moriformis(UTEX 1435)株展示出包含低于8.5%C18:2脂肪酸的脂肪酸谱。如在表59中所显示的，表达高等植物脂肪酸去饱和酶的A株的脂质谱显示出增加的C18:2脂肪酸。总C18不饱和脂肪酸从大约64%增加至大约67%-72%。同样，总C18多不饱和脂肪酸(C18:2和C18:3)与总的合并的C18不饱和脂肪酸(C18:1、C18:2以及C18:3)的比率从少于14%增加至超过19%。这些数据证明了多核苷酸的效用和效果，这些多核苷酸允许Δ12去饱和酶脂肪酸去饱和酶的外源表达而改变工程化微生物的脂肪酸谱。

[0825] 实例18:通过脂肪酸去饱和酶的多重等位基因破坏来改变工程化微生物的脂肪酸水平

[0826] 这个实例描述了使用一种转化载体用包含选择标记和编码外源SAD酶的序列的转

化盒破坏Prototheca moriformis FADc基因座，从而将微生物工程化，其中该转化的微生物的脂肪酸谱已经被改变。

[0827] 根据实例2中详述的生物射弹转化方法，用转化构建体pSZ1499(SEQ ID NO:246)转化Prototheca moriformis(UTEX 1435)的以经典方式诱变(为了更高的油产量)的衍生物(C株)。pSZ1499包含油橄榄硬脂酰-ACP去饱和酶基因的核苷酸序列，对该序列针对在Prototheca moriformis UTEX 1435中表达进行了密码子优化。pSZ1499表达构建体含有针对用于整合至核基因组中的FADc基因组区域的5'(SEQ ID NO:247)和3'(SEQ ID NO:248)同源重组靶向序列(在构建体侧翼)以及在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。油橄榄硬脂酰-ACP去饱和酶编码区在Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR的控制下，并且将天然转运肽替换为原始小球藻硬脂酰-ACP去饱和酶转运肽(SEQ ID NO:49)。将整个油橄榄SAD表达盒称作pSZ1499并且可以写为FADc5'_btub-Suc2-nr_amt03-S106SAD-OeSAD-nr-FADc3'。

[0828] 在含有蔗糖作为唯一碳源的平板上选择初级转化子。将各个转化子以克隆方式进行纯化并且使其在标准脂质产生条件下在pH 7.0生长，与如在实例1中披露的条件类似。使用如在实例11中所述的标准脂肪酸甲酯气相色谱火焰电离(FAME GC/FID)检测方法分析脂肪酸谱。从由该转化载体的转化而产生的一组代表性克隆得到的脂肪酸谱显示于表60中。从在包含葡萄糖作为唯一碳源的脂质产生条件下(pH 5.0)生长的未转化株(C株)获得的脂质的脂肪酸谱另外地呈现在表60中。

[0829] 表60.经多重工程化以敲除内源FADc等位基因并且表达油橄榄硬脂酰-ACP去饱和酶的Prototheca moriformis(UTEX 1435)的脂肪酸谱。

[0830]

株	转化子	% C16:0	% C18:0	% C18:1	% C18:2
C 株	未转化	28.50	3.72	57.70	7.04
C 株	未转化	28.57	3.69	57.61	7.07
C 株 pSZ1499	转化子 1	20.37	1.13	74.38	0.01
	转化子 2	19.98	1.16	74.60	0.00
	转化子 3	20.10	1.16	74.70	0.00
	转化子 4	21.13	1.21	73.86	0.00

[0831]

	转化子 5	19.95	1.11	74.58	0.00
	转化子 6	20.20	1.14	74.61	0.00
	转化子 7	20.72	1.15	74.15	0.00
	转化子 8	20.06	1.11	74.44	0.00
	转化子 9	19.86	1.18	74.88	0.00

[0832] 如在表60中所显示的，用pSZ1499对C株的转化影响了转化的微生物的脂肪酸谱。未转化的Prototheca moriformis的(UTEX 1435)株(C株)展示出包含少于60％C18:1脂肪

酸和多于7％C18:2脂肪酸的脂肪酸谱。相反，用pSZ1499转化的C株展示出具有C18:1脂肪酸的组成增加以及C18:0和C18:2脂肪酸的同时减少的脂肪酸谱。在用pSZ1499转化的C株的脂肪酸谱中未检测到C18:2脂肪酸。在p SZ1499转化子中不存在可检测到的C18:2脂肪酸表明：使用含有用于整合进入多个FADc基因组座位的同源重组靶向序列的pSZ1499进行的转化已经完全破坏了FAD活性。

[0833] 进行Southern印迹分析来验证多个FADc等位基因被pSZ1499转化载体中断。使用标准的分子生物学方法从C株和pSZ1499转化子中提取基因组DNA。将来自每个样品的DNA在用限制性酶PstI消化之后在0.8％的琼脂糖凝胶上进行电泳。将来自这种凝胶的DNA转移到Nylon+膜(安玛西亚公司(Amersham))上，然后使其与相应于FADc 3’区域的、P32标记的多核苷酸探针杂交。图3显示了pSZ1499转化盒、Prototheca moriformis的(UTEX 1435)的两个经测序的FADc等位基因、以及被pSZ1499转化载体破坏的等位基因的预测大小的图谱。FADc等位基因1包含PstI限制性位点，而FADc等位基因2则没有。无论哪个等位基因被破坏，SAD盒的整合会将PstI限制性位点引入被破坏的FADc等位基因，导致一个约6kb的片段在Southern上分辨。图4显示了Southern印迹分析的结果。在两转化子中都在约6kb处检测到杂交带。没有检测到指示未中断的等位基因的更小的杂交带。这些结果表明两个FADc等位基因都被pSZ1499破坏。

[0834] 用SAD表达盒对FADc脂肪酸去饱和酶的两个等位基因的消融导致包含大约74％C18:1的脂肪酸谱。总之，这些数据证明了多核苷酸的效用和效果，这些多核苷酸允许FAD等位基因的敲除以及伴随的硬脂酰-ACP去饱和酶的外源表达而改变工程化微生物的脂肪酸谱。

[0835] 实例19：从工程化微生物中产生的加工油的特征

[0836] 用转化载体pSZ1500(SEQ ID NO:251)转化Prototheca moriformis的(UTEX 1435)的方法和效果先前已经描述于PCT申请号PCT/US2011/038463和PCT/US2011/038463中。

[0837] 根据在实例2中详述的生物射弹转化方法，用转化构建体pSZ1500转化Prototheca moriformis(UTEX 1435)的以经典方式诱变(为了更高的油产量)的衍生物(C株)。将初级转化子在含有蔗糖作为唯一碳源的琼脂平板上进行选择，以克隆方式纯化，并且选择一个单个的工程化株系(D株)用于分析。如在此描述使D株生长。然后使用标准方法从所产生的生物质进行油的己烷提取，并且确定所得到的甘油三酯油不含残留己烷。在PCT申请号PCT/US2010/31108中描述了使用螺旋榨机从微藻中提取油的其他方法，特此通过引用而结合。

[0838] 然后使用熟知的植物油加工方法将从D株的生物质提取的油进行精炼、漂白、以及除臭。这些工序产生一种油样品RBD469，根据通过美国油化学协会(American Oil Chemists’Society)、美国检测与材料学会(American Society for Testing and Materials)以及国际标准化组织(International Organization for Standardization)所定义的方法，使其经受一些分析测试方案。在以下表60中汇总了这些分析的结果。

[0839] 表60.油样品RBD469的分析结果。

[0840]

方法编号	测试说明	结果	单位
AOCS Ca 3a-46	不溶性杂质	< 0.01	%
AOCS Ca 5a-40	脂肪酸（油酸）	0.02	%
AOCS Ca 5a-40	酸值	0.04	mg KOH/g
AOCS CA 9f-57	中性油	98.9	%
D97	浊点	-15	deg C
D97	倾点	-18	deg C
	卡尔费休水分	0.01	%
AOCS Cc 13d-55（修改的）	叶绿素	< 0.01	ppm
	碘值	78.3	g I_2/100g
AOCS Cd 8b-90	过氧化物值	0.31	meq/kg
ISO 6885	对甲氧基苯胺值	0.65	
AOCS Cc 18-80	滴熔点（Dropping Melting point）（Mettler）	6.2	deg C
AOCS Cd 11d-96	甘油三酯（Tricylglicerides）	98.6	%
AOCS Cd 11d-96	甘油单酯	< 0.01	%
AOCS Cd 11d-96	甘油二酯（Diglicerides）	0.68	%
AOCS Cd 20-91	总极性化合物	2.62	%
IUPAC, 2.507 and 2.508	氧化的和聚合的甘油三酯（Tricylglicerides）	17.62	%
AOCS Cc 9b-55	闪点	244	deg C
AOCS Cc 9a-48	烟点	232	deg C
AOCS Cd 12b-92	氧化稳定指数 Rancimat（110°C）	31.6	小时
AOCS Ca 6a-40	未皂化的物质	2.28	%

[0841] 根据AOCS方法对相同批的Prototheca moriformis D株RBD469油针对痕量元素含量、固体脂肪含量、以及洛维邦得色(Lovibond color)进行分析。这些分析的结果呈现在以下表61、62、以及63中。

[0842] 表61.RBD469油的ICP元素分析。

[0843]

方法编号	测试说明	结果（ppm）
AOCS Ca 20-99 和	磷	1.09

[0844]

AOCS Ca 17-01（修改的）	钙	0.1
	镁	0.04
	铁	< 0.02
	硫	28.8
	铜	< 0.05
	钾	< 0.50
	钠	< 0.50
	硅	0.51
	硼	0.06
	铝	< 0.20
	铅	< 0.20
	锂	< 0.02
	镍	< 0.20
	钒	< 0.05
	锌	< 0.02
	砷	< 0.20
	汞	< 0.20
	镉	< 0.03
	铬	< 0.02
	锰	< 0.05
	银	< 0.05
	钛	< 0.05
	硒	< 0.50
UOP779	有机氯化物	< 1
UOP779	无机氯化物	7.24
AOCS Ba 4e-93	氮	6.7

[0845] 表62.RBD469油的固体脂肪含量分析

[0846]

方法编号	固体脂肪含量	结果
AOCS Cd 12b-93	在 10°C 的固体脂肪含量	0.13%
AOCS Cd 12b-93	在 15°C 的固体脂肪含量	0.13%
AOCS Cd 12b-93	在 20°C 的固体脂肪含量	0.28%
AOCS Cd 12b-93	在 25°C 的固体脂肪含量	0.14%
AOCS Cd 12b-93	在 30°C 的固体脂肪含量	0.08%
AOCS Cd 12b-93	在 35°C 的固体脂肪含量	0.25%

[0847] 表63.RBD469油的洛维邦得色分析

[0848]

方法编号	颜色	结果	单位
AOCS Cc 13j-97	红	2	单位
AOCS Cc 13j-97	黄	27	单位

[0849] 使RBD469油经受酯交换以生产脂肪酸甲酯(FAME)。RBD469的所得到的脂肪酸甲酯谱显示于表64中：

[0850] 表64.RBD469油的脂肪酸甲酯谱

[0851]

脂肪酸	面积%
C10	0.01
C12:0	0.04
C14:0	0.64
C15:0	0.08
C16:0	8.17
C16:1异构体	0.39
C16:1	0.77
C17:0	0.08
C18:0	1.93
C18:1	85.88
C18:1	0.05
C18:2	0.05
C20:0	0.3
C20:1	0.06
C20:1	0.44
C22:0	0.11
C23:0	0.03
C24:0	0.1
鉴定的总FAME	99.13

[0852] 实例20:具有改变的脂肪酸谱的工程化微藻

[0853] 如上文所述,整合异源基因以敲除或敲低无绿藻种类中的特定内源脂质途径酶可以改变工程化微生物的脂肪酸谱。在这个实例中,产生质粒构建体来评定宿主细胞的脂质谱是否可以因敲除或敲低内源脂肪酰基-ACP硫酯酶基因FATA1而受影响。

[0854] A.通过敲除内源Prototheca moriformis硫酯酶基因改变脂质谱

[0855] 使用实例2的方法,用在表65中的以下质粒构建体之一转化Prototheca moriformis UTEX 1435的以经典方式诱变(为了更高的油产量)的衍生物(A株)。每种构建体含有用于整合到核基因组以中断内源FATA1基因的区域和在莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的控制下的酿酒酵母suc2蔗糖转化酶编码区。这种酿酒酵母suc2表达盒作为SEQ ID NO:159列出并且用作选择标记。所有蛋白编码区经密码子优化以反映根据表2的Prototheca moriformis UTEX 1435核基因中固有的密码子偏倚。下文显示用于核基因组整合的FATA1基因的靶向区域的相关序列。

[0856] 描述 SEQ ID NO:

[0857] 用于整合进入FATA1基因座的5'序列 SEQ ID NO:253

[0858] 用于整合进入FATA1基因座的3'序列 SEQ ID NO:254

[0859] 表65.用来转化Prototheca moriformis(UTEX 1435)A株的质粒构建体

[0860] 质粒构建体 序列元件

[0861] 1 FATA1-CrbTub_yInv_nr-FATA1

[0862] 2 FATA1-CrbTub_yInv_nr::amt03_CwTE2_nr-FATA1

[0863] 在构建体FATA1-CrbTub_yInv_nr-FATA1(SEQ ID NO:255)中的相关限制性位点以小写字母、粗体和加下划线指示并且分别是:5'-3'BspQ 1、Kpn I、Asc I、Mfe I、Sac I、BspQ I。BspQI位点界定转化DNA的5'和3'末端。粗体、小写字母序列代表来自A株的允许通过同源重组在FATA1基因座处靶向整合的基因组DNA。在5'至3'方向继续向前,通过加框文本指示驱动酵母蔗糖转化酶基因(赋予A株代谢蔗糖的能力)表达的莱茵衣藻β-微管蛋白启动子。转化酶的起始子ATG和终止子TGA由大写字母、粗斜体指示,同时编码区以小写字母斜体指示。普通小球藻硝酸盐还原酶3'UTR由小写字母加下划线文本指示,后接由小写字母粗体文本指示的A株FATA1基因组区域:

[0864]

gctcttcggagtcactgtgccactgagttcgactggtagctgaatggagtcgctgctccactaaacgaattgtcagcaccgcc
agccggccgaggacccgagtcatagcgagggtagtagcgcgccatggcaccgaccagcctgcttgccagtactggcgtctc
ttccgcttctctgtggtcctctgcgcgctccagcgcgtgcgcttttccggtggatcatgcggtccgtggcgcaccgcagcggcc
gctgcccatgcagcgccgctgcttccgaacagtggcggtcagggccgcacccgcggtagccgtccgtccggaacccgccca
agagttttgggagcagcttgagccctgcaagatggcggaggacaagcgcatcttcctggaggagcaccggtgcgtggaggt
ccggggctgaccggccgtcgcattcaacgtaatcaatcgcatgatgatcagaggacacgaagtcttggtggcggtggccag
aaacactgtccattgcaagggcatagggatgcgttccttcacctctcatttctcatttctgaatccctccctgctcactctttctcc
tcctccttcccgttcacgcagcattcggggtaccctttcttgcgctatgacacttccagcaaaaggtagggcgggctgcgagacgg
cttcccggcgctgcatgcaacaccgatgatgcttcgaccccccgaagctccttcggggctgcatgggcgctccgatgccgctccag
ggcgagcgctgtttaaatagccaggcccccgattgcaaagacattatagcgagctaccaaagccatattcaaacacctagatcacta
ccacttctacacaggccactcgagcttgtgatcgcactccgctaaggggggcgcctcttcctcttcgtttcagtcacaacccgcaaac**g**
gcgcgcc*ATG*ctgctgcaggccttcctgttcctgctggccggcttcgccgccaagatcagcgcctccatgacgaacgagacgt
ccgaccgccccctggtgcacttcacccccaacaagggctggatgaacgaccccaacggcctgtggtacgacgagaaggacg
ccaagtggcacctgtacttccagtacaacccgaacgacaccgtctgggggacgcccttgttctggggccacgccacgtccgac
gacctgaccaactgggaggaccagcccatcgccatcgccccgaagcgcaacgactccggcgccttctccggctccatggtggt
ggactacaacaacacctccggcttcttcaacgacaccatcgacccgcgccagcgctgcgtggccatctggacctacaacaccc
cggagtccgaggagcagtacatctcctacagcctggacggcggctacaccttcaccgagtaccagaagaaccccgtgctggc
cgccaactccacccagttccgcgacccgaaggtcttctggtacgagccctcccagaagtggatcatgaccgcggccaagtccc
aggactacaagatcgagatctactcctccgacgacctgaagtcctggaagctggagtccgcgttcgccaacgagggcttcctcg
gctaccagtacgagtgccccggcctgatcgaggtccccaccgagcaggaccccagcaagtcctactgggtgatgttcatctcca
tcaaccccggcgccccggccggcggctccttcaaccagtacttcgtcggcagcttcaacggcacccacttcgaggccttcgaca
accagtcccgcgtggtggacttcggcaaggactactacgccctgcagaccttcttcaacaccgacccgacctacgggagcgcc
ctgggcatcgcgtgggcctccaactgggagtactccgccttcgtgcccaccaacccctggcgctcctccatgtccctcgtgcgca
agttctccctcaacaccgagtaccaggccaacccggagacggagctgatcaacctgaaggccgagccgatcctgaacatcag
caacgccggcccctggagccggttcgccaccaacaccacgttgacgaaggccaacagctacaacgtcgacctgtccaacag
caccggcaccctggagttcgagctggtgtacgccgtcaacaccacccagacgatctccaagtccgtgttcgcggacctctccctc

[0865]

tggttcaagggcctggaggaccccgaggagtacctccgcatgggcttcgaggtgtccgcgtcctccttcttcctggaccgcggga
acagcaaggtgaagttcgtgaaggagaaccccctacttcaccaaccgcatgagcgtgaacaaccagcccttcaagagcgaga
acgacctgtcctactacaaggtgtacggcttgctggaccagaacatcctggagctgtacttcaacgacggcgacgtcgtgtccac
caacacctacttcatgaccaccgggaacgccctgggctccgtgaacatgacgacgggggtggacaacctgttctacatcgaca
*agttccaggtgcgcgaggtcaag**TGA***caattggcagcagcagctcggatagtatcgacacactctggacgctggtcgtgtgat
ggactgttgccgccacacttgctgccttgacctgtgaatatccctgccgcttttatcaaacagcctcagtgtgtttgatcttgtgtgtacgc
gcttttgcgagttgctagctgcttgtgctatttgcgaataccacccccagcatcccctccctcgtttcatatcgcttgcatcccaaccgc
aacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgcccctcgcacagccttggtttgggctccgcctgtat
tctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaggatcgtag
agctcactagtatcgatttcgaa**gacagggtggttggctggatggggaaacgctggtcgcgggattcgatcctgctgcttatatc**
ctccctggaagcacacccacgactctgaagaagaaaacgtgcacacacacaacccaaccggccgaatatttgcttccttatc
ccgggtccaagagagactgcgatgcccccctcaatcagcatcctcctccctgccgcttcaatcttccctgcttgcctgcgcccg
cggtgcgccgtctgcccgcccagtcagtcactcctgcacaggcccccttgtgcgcagtgctcctgtaccctttaccgctccttcc
attctgcgaggcccccctattgaatgtattcgttgcctgtgtggccaagcgggctgctgggcgcgccgccgtcgggcagtgctc
ggcgactttggcggaagccgattgttcttctgtaagccacgcgcttgctgctttgggaagagaaggggggggtactgaatg
gatgaggaggagaaggaggggtattggtattatctgagttgggtgaagagc

[0866] 为了将Cuphea wrightii ACP-硫酯酶2(CwTE2)基因(登录号:U56104)在FATA1基因座处导入A株,产生了一种构建体,它在Prototheca moriformis Amt03启动子/5'UTR(SEQ ID NO:89)和普通小球藻硝酸盐还原酶3'UTR的控制下表达CwTE2基因的蛋白编码区。

已经在A株中表达的构建体可以被写为FATA1-CrbTub_yInv_nr::amt03_CwTE2_nr-FATA1(SEQ ID NO:256)。

[0867] 在构建体FATA1-CrbTub_yInv_nr::amt03_CwTE2_nr-FATA1中的相关限制性位点以小写字母、粗体和加下划线指示并且分别是:5'-3'BspQ 1、Kpn I、Asc I、Mfe I、BamH I、EcoR I、Spe I、Asc I、Pac I、Sac I、BspQ I。BspQI位点界定转化DNA的5'和3'末端。粗体、小写字母序列代表来自A株的允许通过同源重组在FATA1基因座处靶向整合的基因组DNA。在5'至3'方向继续向前,通过加框文本指示驱动酵母蔗糖转化酶基因(赋予A株代谢蔗糖的能力)表达的莱茵衣藻β-微管蛋白启动子。转化酶的起始子ATG和终止子TGA由大写、粗斜体指示,同时编码区以小写字母斜体指示。普通小球藻硝酸盐还原酶3'UTR由小写体加下划线文本指示,后接由加框斜体文本指示的Prototheca moriformis的内源Amt03启动子。C.wrightii ACP-硫酯酶的起始子ATG和终止子TGA密码子由大写字母、粗斜体指示,同时ACP-硫酯酶编码区的其余部分以粗斜体指示。普通小球藻硝酸盐还原酶3'UTR同样由小写字母加下划线文本显示,后接由小写字母粗体文本显示的A株FATA1基因组区域。

[0868]

gctcttcggagtcactgtgccactgagttcgactggtagctgaatggagtcgctgctccactaaacgaattgtcagcaccgcc
agccggccgaggacccgagtcatagcgagggtagtagcgcgccatggcaccgaccagcctgcttgccagtactggcgtctc
ttccgcttctctgtggtcctctgcgcgctccagcgcgtgcgcttttccggtggatcatgcggtccgtggcgcaccgcagcggcc
gctgcccatgcagcgccgctgcttccgaacagtggcggtcagggccgcacccgcggtagccgtccgtccggaacccgccca
agagttttgggagcagcttgagccctgcaagatggcggaggacaagcgcatcttcctggaggagcaccggtgcgtggaggt
ccggggctgaccggccgtcgcattcaacgtaatcaatcgcatgatgatcagaggacacgaagtcttggtggcggtggccag
aaacactgtccattgcaagggcatagggatgcgttccttcacctctcatttctcatttctgaatccctccctgctcactctttctcc
tcctccttcccgttcacgcagcattcggggtaccctttcttgcgctatgacacttccagcaaaaggtagggcgggctgcgagacgg
cttcccggcgctgcatgcaacaccgatgatgcttcgaccccccgaagctccttcggggctgcatgggcgctccgatgccgctccag
ggcgagcgctgtttaaatagccaggcccccgattgcaaagacattatagcgagctaccaaagccatattcaaacacctagatcacta

[0869]

ccacttctacacaggccactcgagcttgtgatcgcactccgctaaggggcgcctcttcctcttcgtttcagtcacaacccgcaaacg
gcgcgcc*ATG*ctgctgcaggccttcctgttcctgctggccggcttcgccgccaagatcagcgcctccatgacgaacgagacgt
ccgaccgccccctggtgcacttcacccccaacaagggctggatgaacgaccccaacggcctgtggtacgacgagaaggacg
ccaagtggcacctgtacttccagtacaacccgaacgacaccgtctgggggacgccttgttctggggccacgccacgtccgac
gacctgaccaactgggaggaccagcccatcgccatcgccccgaagcgcaacgactccggcgccttctccggctccatggtggt
ggactacaacaacacctccggcttcttcaacgacaccatcgacccgcgccagcgctgcgtggccatctggacctacaacaccc
cggagtccgaggagcagtacatctcctacagcctggacggcggctacaccttcaccgagtaccagaagaaccccgtgctggc
cgccaactccacccagttccgcgacccgaaggtcttctggtacgagccctcccagaagtggatcatgaccgcggccaagtccc
aggactacaagatcgagatctactcctccgacgacctgaagtcctggaagctggagtccgcgttcgccaacgagggcttcctcg
gctaccagtacgagtgccccggcctgatcgaggtccccaccgagcaggaccccagcaagtcctactgggtgatgttcatctcca
tcaaccccggcgccccggccggcggctccttcaaccagtacttcgtcggcagcttcaacggcacccacttcgaggccttcgaca
accagtcccgcgtggtggacttcggcaaggactactacgccctgcagaccttcttcaacaccgacccgacctacgggagcgcc
ctgggcatcgcgtgggcctccaactgggagtactccgccttcgtgcccaccaaccccctggcgctcctccatgtccctcgtgcgca
agttctccctcaacaccgagtaccaggccaacccggagacggagctgatcaacctgaaggccgagccgatcctgaacatcag
caacgccggcccctggagccggttcgccaccaacaccacgttgacgaaggccaacagctacaacgtcgacctgtccaacag
caccggcaccctggagttcgagctggtgtacgccgtcaacaccacccagacgatctccaagtccgtgttcgcggacctctccctc
tggttcaagggcctggaggaccccgaggagtacctccgcatgggcttcgaggtgtccgcgtcctccttcttcctggaccgcggga
acagcaaggtgaagttcgtgaaggagaacccctacttcaccaaccgcatgagcgtgaacaaccagcccttcaagagcgaga
acgacctgtcctactacaaggtgtacggcttgctggaccagaacatcctggagctgtacttcaacgacggcgacgtcgtgtccac
caacacctacttcatgaccaccgggaacgccctgggctccgtgaacatgacgacgggggtggacaacctgttctacatcgaca
agttccaggtgcgcgaggtcaag*TGA*caattggcagcagcagctcggatagtatcgacacactctggacgctggtcgtgtgat
ggactgttgccgccacacttgctgccttgacctgtgaatatccctgccgcttttatcaaacagcctcagtgtgtttgatcttgtgtgtacgc
gcttttgcgagttgctagctgcttgtgctatttgcgaataccacccccagcatcccccttccctcgtttcatatcgcttgcatcccaaccgc
aacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgcccctcgcacagccttggtttgggctccgcctgtat
tctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatgga**ggatcc**cgc
gtctcgaacagagcgcgcagaggaacgctgaaggtctcgcctctgtcgcacctcagcgcggcatacaccacaataaccacctgac
gaatgcgcttggttcttcgtccattagcgaagcgtccggttcacacacgtgccacgttggcgaggtggcaggtgacaatgatcggtg
gagctgatggtcgaaacgttcacagcctagggatatc**gaattc***ggccgacaggacgcgcgtcaaaggtgctggtcgtgtatgcc*
ctggccggcaggtcgttgctgctgctggttagtgattccgcaaccctgattttggcgtcttatttttggcgtggcaaacgctggcgccc
gcgagccgggccggcggcgatgcggtgccccacggctgccggaatccaagggaggcaagagcgcccgggtcagttgaag
ggctttacgcgcaaggtacagccgctcctgcaaggctgcgtggtggaattggacgtgcaggtcctgctgaagttcctccaccgc
ctcaccagcggacaaagcaccggtgtatcaggtccgtgtcatccactctaaagagctcgactacgacctactgatggccctaga
ttcttcatcaaaaacgcctgagacacttgcccaggattgaaactccctgaagggaccaccaggggccctgagttgttccttcccc
ccgtggcgagctgccagccaggctgtacctgtgatcgaggctggcgggaaaataggcttcgtgtgctcaggtcatgggaggtg
caggacagctcatgaaacgccaacaatcgcacaattcatgtcaagctaatcagctatttcctcttcacgagctgtaattgtcccaa
aattctggtctaccgggggtgatccttcgtgtacgggcccttccctcaaccctaggtatgcgcgcatgcggtcgccgcgcaactcg
cgcgagggccgagggtttgggacgggccgtcccgaaatgcagttgcacccggatgcgtggcaccttttttgcgataatttatgca
atggactgctctgcaaaattctggctctgtcgccaaccctaggatcagcggcgtaggatttcgtaatcattcgtcctgatggggag
ctaccgactaccctaatatcagcccgactgcctgacgccagcgtccacttttgtgcacacattccattcgtgcccaagacatttcat
tgtggtgcgaagcgtccccagttacgctcacctgtttcccgacctccttactgttctgtcgacagagcgggcccacaggccggtcg
*cagcc***actagt***atggtggtggccgccgccgccagcagcgccttcttccccgtgcccgcccccccgccccacccccaagcccggc*
aagttcggcaactggcccagcagcctgagccagcccttcaagcccaagagcaaccccaacggccgcttccaggtgaaggcc
aacgtgagccccccacgggcgcgcccccaaggccaacggcagcgccgtgagcctgaagtccggcagcctgaaacaccctgga
ggaccccccccagcagccccccccccccgcaccttcctgaaccagctgcccgactggagccgcctgcgcaccgccatcaccacc
gtgttcgtggccgccgagaagcagttcacccgcctggaccgcaagagcaagcgccccgacatgctggtggactggttcggca
gcgagaccatcgtgcaggacggcctggtgttccgcgagcgcttcagcatccgcagctacgagatcggcgccgaccgcaccgc
cagcatcgagaccctgatgaaccacctgcaggacaccagcctgaaccactgcaagagcgtgggcctgctgaacgacggcttc

[0870]

ggccgcacccccgagatgtgcacccgcgacctgatctgggtgctgaccaagatgcagatcgtggtgaaccgctaccccacctg gggcgacaccgtggagatcaacagctggttcagccagagcggcaagatcggcatgggccgcgagtggctgatcagcgactg caacaccggcgagatcctggtgcgcgccaccagcgcctgggccatgatgaaccagaagacccgccgcttcagcaagctgcc ctgcgaggtgcgccaggagatcgccccccacttcgtggacgcccccccccgtgatcgaggacaacgaccgcaagctgcacaa gttcgacgtgaagaccggcgacagcatctgcaagggcctgacccccggctggaacgacttcgacgtgaaccagcacgtgag caacgtgaagtacatcggctggattctggagagcatgcccaccgaggtgctggagacccaggagctgtgcagcctgaccctg gagtaccgccgcgagtgcggccgcgagagcgtggtggagagcgtgaccagcatgaaccccagcaaggtgggcgaccgca gccagtaccagcacctgctgcgcctggaggacggcgccgacatcatgaagggccgcaccgagtggcgccccaagaacgcc ggcaccaaccgcgccatcagcaccTGAttaattaactcgaggcagcagcagctcggatagtatcgacacactctggacgctg gtcgtgtgatggactgttgccgccacacttgctgccttgacctgtgaatatccctgccgcttttatcaaacagcctcagtgtgtttgatctt gtgtgtacgcgcttttgcgagttgctagctgcttgtgctatttgcgaataccacccccagcatcccccttccctcgtttcatatcgcttgcat cccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcctgctcactgcccctcgcacagccttggtttgggct ccgcctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatgga aagctt**gagctc**ttgttttccagaaggagttgctccttgagcctttcattctcagcctcgataacctccaaagccgctctaattgtggagg gggttcgaa**gacagggtggttggctggatggggaaacgctggtcgcgggattcgatcctgctgcttatatcctccctggaagc acacccacgactctgaagaagaaaacgtgcacacacacaacccaaccggccgaatatttgcttccttatcccgggtccaag agagactgcgatgccccctcaatcagcatcctcctccctgccgcttcaatcttccctgcttgcctgcgcccgcggtgcgccgt ctgcccgcccagtcagtcactcctgcacaggccccttgtgcgcagtgctcctgtaccctttaccgctccttccattctgcgaggc cccctattgaatgtattcgttgcctgtgtggccaagcgggctgctgggcgcgccgccgtcgggcagtgctcggcgactttggc ggaagccgattgttcttctgtaagccacgcgcttgctgctttgggaagagaaggggggggtactgaatggatgaggagga gaaggaggggtattggtattatctgagttgggtgaagagc**

[0871] 将质粒构建体1或2单独地转化进入A株后，在含有蔗糖作为唯一碳源的琼脂平板上选择阳性克隆。如在先前的实例中，将初级转化子以克隆方式纯化并且使其在pH 7在标准脂质产生条件下生长并且从每种转化子的干生物质制备脂质样品。使用如在实例11中所述的直接酯交换方法来确定脂肪酸谱。在表66中显示了与未转化的A株对照的脂肪酸谱比较，从一组代表性克隆得到的脂肪酸谱(表示为总脂肪酸的面积%)，这些代表性克隆通过用构建体1转化而产生。在表67中显示了与未转化的A株对照的脂肪酸谱相比较，从一组代表性克隆得到的脂肪酸谱(表示为总脂肪酸的面积%)，这些代表性克隆通过用构建体2转化而产生。

[0872] 表66.含有破坏内源FATA1等位基因的选择标记的Prototheca moriformis细胞的脂肪酸谱。

[0873]

转化	% C14:0	% C16:0	% C18:0	% C18:1	% C18:2
野生型	1.23	25.68	2.83	60.54	7.52
转化子 1	0.86	16.95	1.75	68.44	9.78
转化子 2	0.85	17.33	1.71	68.57	9.31
转化子 3	0.82	17.40	1.78	68.55	9.22
转化子 4	0.84	17.43	1.78	68.25	9.53
转化子 5	0.75	17.64	2.02	69.02	8.61

[0874] 呈现在表66中的结果显示，消融宿主的内源FATA1等位基因改变了工程化微藻的脂肪酸谱。将选择标记靶向到内源FATA1等位基因对转化的微生物的脂肪酸谱的影响是C16:0脂肪酸的明显减少，伴随C18:1脂肪酸的同时增加。

[0875] 表67.含有破坏内源FATA1等位基因的选择标记和外源硫酯酶的Prototheca

moriformis细胞的脂肪酸谱。

[0876]

	转化子	碳源	% C10:0	% C12:0	% C14:0	% C16:0	% C18:0	% C18:1	% C18:2
A株	野生型	葡萄糖	0.01	0.04	1.38	28.83	3.00	56.05	8.21
	野生型	葡萄糖	0.01	0.04	1.50	29.38	3.00	55.29	8.23
	野生型	葡萄糖/果糖	0.01	0.05	1.48	28.58	3.20	57.14	7.27
	野生型	葡萄糖/果糖	0.01	0.04	1.54	29.05	3.23	56.47	7.32
>2拷贝	1	葡萄糖/果糖	4.29	19.98	9.17	20.68	3.47	34.38	6.37
	2	葡萄糖/果糖	3.11	16.17	9.91	15.97	1.57	45.72	5.81
	3	蔗糖	4.84	24.22	11.56	19.48	2.67	29.56	6.02
	4	蔗糖	3.24	16.67	10.39	16.34	1.43	44.41	6.00
1-2拷贝	1	葡萄糖/果糖	0.18	1.64	1.85	14.43	2.12	70.30	7.63
	2	葡萄糖/果糖	0.18	1.56	1.74	13.56	2.25	71.04	7.72
	3	蔗糖	0.19	1.69	1.89	13.79	3.15	69.97	7.68
	4	蔗糖	0.15	1.26	1.49	13.44	2.73	71.46	7.77

[0877] 与将选择标记单独靶向到宿主的FATA1等位基因相一致，选择标记与外源硫酯酶的整合导致工程化微藻的脂肪酸谱的改变。如在以上表67中所显示的，将外源硫酯酶基因靶向以中断FATA1等位基因导致C16:0脂肪酸产生的明显减少。CwTE2硫酯酶在FATA1基因座处的表达还影响中链脂肪酸和C18:1脂肪酸产生至这样的程度，这种程度取决于所分析的转化子中存在的外源硫酯酶活性水平。在扩增的转基因在靶整合位点处的拷贝数与硫酯酶水平之间存在良好的一致性，正如对脂肪酸谱或重组蛋白积累(通过蛋白质印迹法评定)的影响所揭示的。

[0878] 其中CwTE2基因已经经历扩增的转基因系显示C10:0-C14:0脂肪酸的明显增加和C18:1脂肪酸的同时减少。相反，其中CwTE2经历很少或不经历扩增的那些转化子与较低的外源硫酯酶表达一致，导致中链脂肪酸略微增加和对C18:1脂肪酸增加的大得多的影响。

[0879] 总之，这些数据显示，宿主的内源FATA1等位基因的靶向破坏改变了工程化微藻的脂质谱。这些数据证明了多核苷酸的效用和效果，这些多核苷酸允许FATA等位基因的靶向

破坏以改变工程化微生物细胞的脂肪酸谱，尤其是减少C16脂肪酸的浓度和增加C18:1脂肪酸的浓度。这些数据另外地证明了多核苷酸的效用和效果，这些多核苷酸允许FATA等位基因的靶向破坏同时伴随地表达外源硫酯酶以改变工程化微生物细胞的脂肪酸谱，尤其是减少C16脂肪酸的浓度。

[0880] B.通过敲低内源Prototheca moriformis硫酯酶基因改变脂质谱

[0881] 将一种构建体导入Prototheca moriformis UTEX 1435A株遗传背景，该构建体通过RNAi下调Prototheca moriformis FATA1基因表达。使用酿酒酵母suc2蔗糖转化酶基因作为选择标记，从而赋予在作为唯一碳源的蔗糖上生长的能力。该构建体利用FatA1编码区的第一外显子，其后接内源内含子，和处于反方向的第一外显子的重复单元。包括针对6S基因组区域的5’和3’同源重组靶向序列(在构建体侧翼)，分别作为SEQ ID NO:82和84列出，用于将发夹构建体整合至核基因组中。这种构建体命名为6S::β-Tub:suc2:nr::β-tub:发夹FatA:nr::6S。

[0882] 在6S::β-Tub:suc2:nr::β-tub:发夹FatA:nr::6S中的相关限制性酶切位点以小写证明、粗体和加下划线显示并且分别是：5’-3’BspQ 1、Kpn I、Mfe I、BamH I、EcoR I、Spe I、Xho I、Sac I、BspQ I。BspQI位点界定转化DNA的5’和3’末端。粗体、小写字母序列代表来自A株的允许通过同源重组在6s基因座处靶向整合的基因组DNA。在5’至3’方向继续向前，通过加框文本指示驱动酵母蔗糖转化酶基因(赋予A株代谢蔗糖的能力)表达的莱茵衣藻β-微管蛋白启动子。转化酶的起始子ATG和终止子TGA由大写字母、粗斜体指示，同时编码区以小写字母斜体指示。普通小球藻硝酸盐还原酶3’UTR由小写字母加下划线文本指示，后接由加框斜体文本指示的驱动发夹FatA1表达的第二莱茵衣藻β-微管蛋白启动子。FatA1的起始ATG密码子由大写字母、粗斜体指示，同时FatA1编码区的第一外显子的其余部分以大写字母指示。FatA基因的内含子由加下划线的大写字母指示，并且连接区域以加下划线的大写字母指示，在FatA1内含子/反向第一外显子交界处产生粗斜体，以便有助于这些载体中的RNA剪接。FatA1的倒位的第一外显子由大写字母显示指示。普通小球藻硝酸盐还原酶3’UTR同样由小写字母加下划线文本指示，后接由小写字母粗体文本指示的A株6S基因组区域。这种RNAi构建体的FATA部分的序列作为SEQ ID NO:257列出。

[0883]

gctcttcgccgccgccactcctgctcgagcgcgcccgcgcgtgcgccgccagcgccttggccttttcgccgcgctcgtgcgcg
tcgctgatgtccatcaccaggtccatgaggtctgccttgcgccggctgagccactgcttcgtccgggcggccaagaggagca
tgagggaggactcctggtccagggtcctgacgtggtcgcggctctgggagcgggccagcatcatctggctctgccgcaccg
aggccgcctccaactggtcctccagcagccgcagtcgccgccgaccctggcagaggaagacaggtgaggggggtatgaat
tgtacagaacaaccacgagccttgtctaggcagaatccctaccagtcatggctttacctggatgacggcctgcgaacagctg
tccagcgacccctcgctgccgccgcttctcccgcacgcttctttccagcaccgtgatggcgcgagccagcgccgcacgctggc
gctgcgcttcgccgatctgaggacagtcggggaactctgatcagtctaaaccccccttgcgcgttagtgttgccatcctttgcag
accggtgagagccgacttgttgtgcgccacccccccacaccacctcctcccagaccaattctgtcacctttttggcgaaggcat
cggcctcggcctgcagagaggacagcagtgcccagccgctgggggttggcggatgcacgctcaggtaccctttcttgcgctat
gacacttccagcaaaaggtagggcgggctgcgagacggcttcccggcgctgcatgcaacaccgatgatgcttcgaccccccgaa
gctccttcgggggctgcatgggcgctccgatgccgctccagggcgagcgctgtttaaatagccaggcccccgattgcaaagacatta

[0884]

tagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagcttgtgatcgcactccgctaagg
gggcgcctcttcctcttcgtttcagtcacaacccgcaaac**tctaga**atatca***ATG***ctgctgcaggccttcctgttcctgctggccgg
cttcgccgccaagatcagcgcctccatgacgaacgagacgtccgaccgccccctggtgcacttcacccccaacaagggctgg
atgaacgaccccaacggcctgtggtacgacgagaaggacgccaagtggcacctgtacttccagtacaacccgaacgacacc
gtctgggggacgcccttgttctggggccacgccacgtccgacgacctgaccaactgggaggaccagcccatcgccatcgcccc
gaagcgcaacgactccggcgccttctccggctccatggtggtggactacaacaacacctccggcttcttcaacgacaccatcga
cccgcgccagcgctgcgtggccatctggacctacaacaccccggagtccgaggagcagtacatctcctacagcctggacggc
ggctacaccttcaccgagtaccagaagaaccccgtgctggccgccaactccacccagttccgcgacccgaaggtcttctggtac
gagccctcccagaagtggatcatgaccgcggccaagtcccaggactacaagatcgagatctactcctccgacgacctgaagtc
ctggaagctggagtccgcgttcgccaacgagggcttcctcggctaccagtacgagtgccccggcctgatcgaggtccccaccg
agcaggaccccagcaagtcctactgggtgatgttcatctccatcaaccccggcgccccggccggcggctccttcaaccagtact
tcgtcggcagcttcaacggcacccacttcgaggccttcgacaaccagtcccgcgtggtggacttcggcaaggactactacgccc
tgcagaccttcttcaacaccgacccgacctacgggagcgccctgggcatcgcgtgggcctccaactgggagtactccgccttcg
tgcccaccaaccctggcgctcctccatgtccctcgtgcgcaagttctccctcaacaccgagtaccaggccaacccggagacgg
agctgatcaacctgaaggccgagccgatcctgaacatcagcaacgccggcccctggagccggttcgccaccaacaccacgtt
gacgaaggccaacagctacaacgtcgacctgtccaacagcaccggcaccctggagttcgagctggtgtacgccgtcaacacc
acccagacgatctccaagtccgtgttcgcggacctctccctctggttcaagggcctggaggaccccgaggagtacctccgcatg
ggcttcgaggtgtccgcgtcctccttcttcctggaccgcgggaacagcaaggtgaagttcgtgaaggagaacccctacttcacca
accgcatgagcgtgaacaaccagcccttcaagagcgagaacgacctgtcctactacaaggtgtacggcttgctggaccagaa
catcctggagctgtacttcaacgacggcgacgtcgtgtccaccaacacctacttcatgaccaccgggaacgccctgggctccgt
gaacatgacgacgggggtggacaacctgttctacatcgacaagttccaggtgcgcgaggtcaag***TGA***caattggcagcagc
agctcggatagtatcgacacactctggacgctggtcgtgtgatggactgttgccgccacacttgctgccttgacctgtgaatatccctg
ccgcttttatcaaacagcctcagtgtgtttgatcttgtgtgtacgcgcttttgcgagttgctagctgcttgtgctatttgcgaataccacccc
cagcatccccttccctcgtttcatatcgcttgcatcccaaccgcaacttatctacgctgtcctgctatccctcagcgctgctcctgctcct
gctcactgcccctcgcacagccttggtttgggctccgcctgtattctcctggtactgcaacctgtaaaccagcactgcaatgctgatgc
acgggaagtagtgggatgggaacacaaatgga**ggatcc**cgcgtctcgaacagagcgcgcagaggaacgctgaaggtctcgcct
ctgtcgcacctcagcgcggcatacaccacaataaccacctgacgaatgcgcttggttcttcgtccattagcgaagcgtccggttcaca
cacgtgccacgttggcgaggtggcaggtgacaatgatcggtggagctgatggtcgaaacgttcacagcctagggatatc**gaattc**
*ctttcttgcgctatgacacttccagcaaaaggtagggcgggctgcgagacggcttcccggcgctgcatgcaacaccgatgatgc
ttcgacccccgaagctccttcggggctgcatgggcgctccgatgccgctccagggcgagcgctgtttaaatagccaggccccc
gattgcaaagacattatagcgagctaccaaagccatattcaaacacctagatcactaccacttctacacaggccactcgagctt
gtgatcgcactccgctaaggggggcgcctcttcctcttcgtttcagtcacaacccgcaaac***ATG***GCACCGACC
AGCCTGCTTGCCAGTACTGGCGTCTCTTCCGCTTCTCTGTGGTCCTCTGCGCGCT
CCAGCGCGTGCGCTTTTCCGGTGGATCATGCGGTCCGTGGCGCACCGCAGCGGC
CGCTGCCCATGCAGCGCCGCTGCTTCCGAACAGTGGCGGTCAGGGCCGCACCCG
CGGTAGCCGTCCGTCCGGAACCCGCCCAAGAGTTTTGGGAGCAGCTTGAGCCCT
GCAAGATGGCGGAGGACAAGCGCATCTTCCTGGAGGAGCACCGGTGCGTGGAG
GTCCGGGGCTGACCGGCCGTCGCATTCAACGTAATCAATCGCATGATGATCAGA
GGACACGAAGTCTTGGTGGCGGTGGCCAGAAACACTGTCCATTGCAAGGGCAT
AGGGATGCGTTCCTTCACCTCTCATTTCTCATTTCTGAATCCCTCCCTGCTCACTC
TTTCTCCTCCTCCTTCCCGTTCACGCAG***CATTCGGGGCAACGAGGTGGGCCC***GTG
CTCCTCCAGGAAGATGCGCTTGTCCTCCGCCATCTTGCAGGGCTCAAGCTGCTC
CCAAAACTCTTGGGCGGGTTCCGGACGGACGGCTACCGCGGGTGCGGCCCTGA
CCGCCACTGTTCGGAAGCAGCGGCGCTGCATGGGCAGCGGCCGCTGCGGTGCG
CCACGGACCGCATGATCCACCGGAAAAGCGCACGCGCTGGAGCGCGCAGAGGA
CCACAGAGAAGCGGAAGAGACGCCAGTACTGGCAAGCAGGCTGGTCGGTGCCA
T**atcgat**agatctcttaaggca*gcagcagctcggatagtatcgacacactctggacgctggtcgtgtgatggactgttgccgccaca*
cttgctgccttgacctgtgaatatccctgccgcttttatcaaacagcctcagtgtgtttgatcttgtgtgtacgcgcttttgcgagttgctag

[0885]

ctgcttgtgctatttgcgaataccacccccagcatcccctttccctcgtttcatatcgcttgcatcccaaccgcaacttatctacgctgtcct
gctatccctcagcgctgctcctgctcctgctcactgcccctcgcacagccttggtttgggctccgcctgtattctcctggtactgcaacc
tgtaaaccagcactgcaatgctgatgcacgggaagtagtgggatgggaacacaaatggaaagcttaattaa**gagctcttgttttcca**
gaaggagttgctccttgagcctttcattctcagcctcgataacctccaaagccgctctaattgtggagggggttcgaatttaaa
agcttggaatgttggttcgtgcgtctggaacaagcccagacttgttgctcactgggaaaaggaccatcagctccaaaaaact
tgccgctcaaaccgcgtacctctgctttcgcgcaatctgccctgttgaaatcgccaccacattcatattgtgacgcttgagcag
tctgtaattgcctcagaatgtggaatcatctgccccctgtgcgagcccatgccaggcatgtcgcgggcgaggacacccgcca
ctcgtacagcagaccattatgctacctcacaatagttcataacagtgaccatatttctcgaagctccccaacgagcacctcca
tgctctgagtggccacccccggccctggtgcttgcggagggcaggtcaaccggcatggggctaccgaaatccccgaccgg
atcccaccacccccgcgatgggaagaatctctccccgggatgtgggcccaccaccagcacaacctgctggcccaggcgag
cgtcaaaccataccacacaaatatccttggcatcggccctgaattccttctgccgctctgctacccggtgcttctgtccgaagc
aggggttgctagggatcgctccgagtccgcaaacccttgtcgcgtggcggggcttgttcgagcttgaagagc

[0886] 6S::β-Tub:suc2:nr::β-tub:发夹FatA:nr::6S的表达导致沉默靶FatA基因产物的发夹RNA形成。将构建体6S::β-Tub:suc2:nr::β-tub:hairpin FatA:nr::6S转化进入A株之后,在含有蔗糖作为唯一碳源的琼脂平板上选择阳性克隆。将初级转化子以克隆方式纯化并且使其在标准脂质产生条件下在pH 5.0生长,然后从每种转化子的干燥生物质制备脂质样品。使用如在实例11中所述的直接酯交换方法来确定脂肪酸谱。在表68中呈现了与未转化的A株对照的脂肪酸谱比较,从由转化产生的一组代表性克隆得到的脂肪酸谱(表示为总脂肪酸的面积%)。

[0887] 表68.含有下调FATA表达的RNA发夹构建体的Prototheca moriformis细胞的脂肪酸谱。

[0888]

转化子	% C10:0	% C12:0	% C14:0	% C16:0	% C16:1	% C18:0	% C18:1	% C18:2
未转化	0.01	0.03	1.23	25.68	0.96	2.83	60.54	7.52
转化子 1	0.01	0.03	0.71	15.10	1.05	1.67	72.08	8.27
转化子 2	0.01	0.03	0.81	15.66	1.16	1.56	70.03	9.61
转化子 3	0.01	0.03	1.09	22.67	1.05	2.12	63.18	8.66
转化子 4	0.01	0.04	1.14	23.31	1.01	2.23	62.83	8.26

[0889] 在表68中呈现的数据显示了FATA发夹RNA构建体的表达对宿主生物的C16和C18:1脂肪酸谱的明显影响。包含FATA发夹RNA构建体的A株转化子的脂肪酸谱显示出C18:1脂肪酸百分比的增加,伴随C16脂肪酸的同时减少。这些数据表明多核苷酸FATA RNA发夹构建体在Prototheca moriformis中的成功表达和用途,以改变工程化宿主微生物的脂肪酸谱,并且尤其是增加微生物细胞中C18:1脂肪酸的浓度和减少C16脂肪酸。

[0890] 实例21:将Chlorella sorokinian工程化

[0891] 可以通过修改如在此讨论的Dawson等人传授的方法和载体来完成根据本发明的重组基因在Chlorella sorokinian中的表达。简言之,Dawson等人,《当代微生物学》(Current Microbiology)第35卷(1997)第356-362页报道了用质粒DNA对Chlorella sorokinian的稳定核转化。使用微粒轰击转化方法,Dawson将编码完全的普通小球藻硝酸盐还原酶基因(NR,GenBank登录号U39931)的质粒pSV72-NRg引入到突变体Chlorella

sorokinian(NR-突变体)中。在不使用硝酸盐作为氮源的情况下,这些NR-突变体不能够生长。硝酸盐还原酶催化硝酸盐转化为亚硝酸盐。在转化之前,Chlorella sorokinian NR-突变体不能够在包含硝酸盐(NO_3^-)作为唯一氮源的培养基上生长超过小集落阶段。将普通小球藻NR基因产物在NR-突变体Chlorella sorokinian中的表达用作为一种选择标记以挽救硝酸盐代谢缺陷。用pSV72-NRg质粒转化后,获得了稳定表达普通小球藻NR基因产物的NR-突变体Chlorella sorokinian,其能够在含有硝酸盐作为唯一碳源的琼脂平板上生长超过小集落阶段。通过Southern分析进行稳定转化子的DNA的评估并且通过RNA酶保护进行稳定转化子的RNA的评估。在琼脂平板(pH 7.4)上进行转化的Chlorella sorokinian(NR突变体)的选择和维持,这些琼脂平板包含0.2g/L $MgSO_4$、0.67g/L KH_2PO_4、3.5g/L K_2HPO_4、1.0g/L $Na_3C_6H_5O_7 \cdot H_2O$以及16.0g/L琼脂、适当的氮源(例如,NO_3^-)、微量营养素、以及碳源。Dawson还报道了Chlorella sorokinian以及Chlorella sorokinian NR突变体在液体培养基中的增殖。Dawson报道质粒pSV72-NRg以及普通小球藻硝酸盐还原酶基因的启动子和3'UTR/终止子适合于使异源基因能够在Chlorella sorokinian NR-突变体中表达。Dawson还报道普通小球藻硝酸盐还原酶基因产物的表达适合用作Chlorella sorokinian NR-突变体中的选择标记。

[0892] 在本发明的一个实施例中,构建了载体pSV72-NRg(包含编码用作选择标记的普通小球藻硝酸盐还原酶(CvNR)基因产物的核苷酸序列)并且将其修饰以进一步包含脂质生物合成途径表达盒序列,由此产生一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在Chlorella sorokinian中表达进行了密码子优化以反映在根据表69A-D的Chlorella sorokinian的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的CvNR启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的CvNR 3'UTR/终止子。该转化构建体可以另外地包含与Chlorella sorokinian基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括微粒轰击的转化技术或其他已知方法实现用该转化载体稳定转化Chlorella sorokinian。CvNR基因产物的活性可以用作为一种选择标记以挽救Chlorella sorokinian NR突变株的氮同化缺陷以及选择稳定表达该转化载体的Chlorella sorokinian NR突变体。适合于Chlorella sorokinian脂质生产的生长培养基包括但不限于0.5g/L KH_2PO_4、0.5g/L K_2HPO_4、0.25g/L $MgSO_4$-7H2O、连同补充的微量营养素以及适当的氮源和碳源(Patterson,《脂质》(Lipids)Vol.5:7(1970),pp.597-600)。可以通过在此描述的标准脂质提取和分析方法来评定Chlorella sorokinian脂质的脂肪酸谱的评估。

[0893] 实例22-44:前言与表格

[0894] 以下实例22-24描述了根据本发明的各种微生物的工程化。为了改变微生物的脂肪酸谱,可以将微生物进行基因修饰,其中将内源或外源脂质生物合成途径酶进行表达、过表达、或减弱。对一种微生物进行基因工程化以改变其有关脂肪酸不饱和度的脂肪酸谱或者减少或增加脂肪酸链长的步骤包括:转化载体(例如,质粒)的设计和构建,用一种或多种载体转化微生物,转化的微生物(转化子)的选择,转化的微生物的生长、以及由该工程化微生物产生的脂质的脂肪酸谱的分析。

[0895] 改变宿主生物的脂肪酸谱的转基因可以在众多真核微生物中表达。转基因在真核微生物中表达的实例可以在科学文献中找到,这些真核微生物包括莱茵衣藻、椭圆小球藻、Chlorella saccarophila、普通小球藻、凯氏小球藻、Chlorella sorokinian、雨生红球藻、胸状盘藻、强壮团藻、杜氏盐藻、绿色杜氏藻、杜氏盐藻、极锐新月藻-瘦狭新月藻-沿岸新月藻复合体、微拟球藻种类、假微型海链藻、三角褐指藻、腐生舟形藻、梭形筒柱藻)、隐秘小环藻、Symbiodinium microadriacticum、前沟藻种类、角毛藻种类、高山被孢霉(Mortierella alpina)、以及解脂耶罗威亚酵母(Yarrowia lipolytica)。这些表达技术可以与本发明的教导结合以产生具有改变的脂肪酸谱的工程化微生物。

[0896] 改变宿主生物的脂肪酸谱的转基因还可以在众多原核微生物中表达。转基因在包括混浊红球菌(Rhodococcus opacus)的产油微生物中的表达的实例可以在文献中找到。这些表达技术可以与本发明的教导结合以产生具有改变的脂肪酸谱的工程化微生物。

[0897] 表69A-D.密码子偏好列表。

[0898]

氨基酸	密码子	*Chlorella sorokinian*	*普通小球藻*	*椭圆小球藻*	*凯氏小球藻*	*杜氏盐藻*	*强壮团藻*	*雨生红球藻*
Ala	GCG	0.20	0.25	0.15	0.14	0.09	0.25	0.21
Ala	GCA	0.05	0.24	0.32	0.10	0.17	0.13	0.27
Ala	GCT	0.12	0.16	0.26	0.18	0.31	0.26	0.17
Ala	GCC	0.63	0.35	0.27	0.58	0.43	0.36	0.35
Arg	AGG	0.03	0.09	0.10	0.09	0.26	0.08	0.14
Arg	AGA	0.04	0.05	0.14	0.01	0.09	0.03	0.05
Arg	CGG	0.06	0.19	0.09	0.06	0.06	0.17	0.15
Arg	CGA	0.00	0.10	0.08	0.00	0.08	0.08	0.10
Arg	CGT	0.06	0.09	0.37	0.14	0.12	0.22	0.13
Arg	CGC	0.81	0.48	0.22	0.71	0.40	0.43	0.42
Asn	AAT	0.04	0.16	0.43	0.06	0.27	0.23	0.21
Asn	AAC	0.96	0.84	0.57	0.94	0.73	0.77	0.79
Asp	GAT	0.13	0.25	0.47	0.12	0.40	0.35	0.27
Asp	GAC	0.87	0.75	0.53	0.88	0.60	0.65	0.73
Cys	TGT	0.06	0.13	0.43	0.09	0.20	0.17	0.27
Cys	TGC	0.94	0.87	0.57	0.91	0.80	0.83	0.64
End	TGA	0.00	0.72	0.14	0.14	0.36	0.24	0.70
End	TAG	0.33	0.11	0.29	0.00	0.00	0.18	0.22
End	TAA	0.67	0.17	4.00	0.86	0.64	0.59	0.09
Gln	CAG	0.42	0.40	0.15	0.40	0.27	0.29	0.33
Gln	CAA	0.04	0.04	0.21	0.40	0.27	0.07	0.10
Glu	GAG	0.53	0.50	0.33	0.40	0.27	0.53	0.49

[0899]

Glu	GAA	0.02	0.06	0.31	0.40	0.27	0.11	0.07
Gly	GGG	0.04	0.16	0.19	0.08	0.10	0.12	0.22
Gly	GGA	0.02	0.11	0.13	0.07	0.13	0.12	0.11
Gly	GGT	0.03	0.12	0.39	0.24	0.25	0.23	0.15
Gly	GGC	0.91	0.61	0.29	0.96	0.51	0.53	0.52
His	CAT	0.14	0.16	0.30	0.08	0.25	0.35	0.27
His	CAC	0.86	0.84	0.70	0.93	0.75	0.65	0.73
Ile	ATA	0.00	0.04	0.07	0.01	0.04	0.08	0.09
Ile	ATT	0.15	0.30	0.63	0.29	0.31	0.35	0.29
Ile	ATC	0.85	0.66	0.65	0.69	0.65	0.57	0.62
Leu	TTG	0.03	0.07	0.03	0.05	0.14	0.14	0.16
Leu	TTA	0.00	0.01	0.32	0.00	0.02	0.03	0.02
Leu	CTG	0.72	0.61	0.34	0.61	0.60	0.45	0.53
Leu	CTA	0.01	0.03	0.03	0.04	0.04	0.07	0.07
Leu	CTT	0.04	0.08	0.16	0.06	0.06	0.14	0.09
Leu	CTC	0.20	0.20	0.12	0.24	0.14	0.17	0.13
Lys	AAG	0.98	0.94	0.54	0.98	0.90	0.90	0.84
Lys	AAA	0.02	0.06	0.46	0.02	0.10	0.10	0.16
Met	ATG	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Phe	TTT	0.28	0.32	0.42	0.31	0.24	0.27	0.35
Phe	TTC	0.72	0.68	0.58	0.69	0.76	0.73	0.65
Pro	CCG	0.18	0.31	0.09	0.07	0.04	0.34	0.15
Pro	CCA	0.06	0.17	0.36	0.07	0.04	0.20	0.24
Pro	CCT	0.10	0.14	0.25	0.17	0.04	0.19	0.29
Pro	CCC	0.66	0.38	0.29	0.69	0.04	0.27	0.32
Ser	AGT	0.03	0.04	0.14	0.02	0.08	0.08	0.07
Ser	AGC	0.27	0.38	0.18	0.18	0.31	0.27	0.31
Ser	TCG	0.12	0.14	0.08	0.10	0.02	0.19	0.10
Ser	TCA	0.03	0.08	0.14	0.08	0.09	0.09	0.14
Ser	TCT	0.09	0.11	0.26	0.18	0.19	0.14	0.13
Ser	TCC	0.47	0.24	0.20	0.44	0.30	0.24	0.24
Thr	ACG	0.11	0.20	0.13	0.05	0.12	0.27	0.19
Thr	ACA	0.01	0.20	0.32	0.07	0.20	0.12	0.23
Thr	ACT	0.12	0.13	0.29	0.12	0.24	0.20	0.18
Thr	ACC	0.76	0.47	0.26	0.76	0.44	0.41	0.40
Trp	TGG	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Tyr	TAT	0.07	0.15	0.43	0.27	0.28	0.24	0.19
Tyr	TAC	0.93	0.85	0.57	0.73	0.72	0.76	0.81
Val	GTG	0.71	0.54	0.37	0.60	0.54	0.46	0.62
Val	GTA	0.00	0.05	0.25	0.03	0.09	0.07	0.09
Val	GTT	0.11	0.14	0.24	0.09	0.14	0.17	0.09
Val	GTC	0.18	0.27	0.14	0.28	0.23	0.30	0.21

[0900]

氨基酸	密码子	极锐新月藻-瘦狭新月藻-沿岸新月藻复合体	绿色杜氏藻	杜氏盐藻	胸状盘藻	三角褐指藻	扁形角毛藻（Chaetoceros compressum）
Ala	GCG	0.48	0.13	0.15	0.43	0.15	0.08

[0901]

Ala	GCA	0.10	0.27	0.20	0.09	0.10	0.37
Ala	GCT	0.15	0.25	0.27	0.08	0.23	0.36
Ala	GCC	0.26	0.35	0.39	0.41	0.52	0.18
Arg	AGG	0.04	0.25	0.22	0.13	0.02	0.14
Arg	AGA	0.00	0.06	0.05	0.00	0.04	0.29
Arg	CGG	0.18	0.08	0.12	0.40	0.10	0.00
Arg	CGA	0.00	0.06	0.06	0.05	0.12	0.19
Arg	CGT	0.13	0.15	0.13	0.08	0.41	0.38
Arg	CGC	0.64	0.39	0.43	0.35	0.31	0.00
Asn	AAT	0.04	0.17	0.23	0.07	0.30	0.58
Asn	AAC	0.96	0.83	0.77	0.93	0.65	0.42
Asp	GAT	0.30	0.38	0.40	0.11	0.41	0.53
Asp	GAC	0.70	0.62	0.60	0.89	0.59	0.47
Cys	TGT	0.06	0.24	0.17	0.20	0.39	0.44
Cys	TGC	0.94	0.76	0.83	0.90	0.61	0.56
终止	TGA	0.75	0.31	0.37	0.50	0.06	0.50
终止	TAG	0.00	0.15	0.14	0.00	0.13	0.00
终止	TAA	0.25	0.54	0.49	0.50	0.81	0.50
Gln	CAG	0.53	0.36	0.32	0.31	0.23	0.16
Gln	CAA	0.09	0.12	0.08	0.07	0.14	0.19
Glu	GAG	0.31	0.44	0.51	0.56	0.21	0.28
Glu	GAA	0.06	0.09	0.09	0.07	0.42	0.37
Gly	GGG	0.31	0.14	0.10	0.18	0.08	0.12
Gly	GGA	0.06	0.11	0.12	0.09	0.34	0.33
Gly	GGT	0.09	0.22	0.22	0.07	0.30	0.39
Gly	GGC	0.53	0.54	0.56	0.65	0.28	0.16
His	CAT	0.33	0.25	0.25	0.43	0.28	0.84
His	CAC	0.67	0.75	0.75	0.57	0.72	0.16
Ile	ATA	0.03	0.03	0.03	0.07	0.03	0.12
Ile	ATT	0.23	0.25	0.31	0.33	0.51	0.65
Ile	ATC	0.74	0.72	0.66	0.59	0.46	0.23
Leu	TTG	0.04	0.11	0.12	0.04	0.26	0.11
Leu	TTA	0.00	0.01	0.01	0.00	0.02	0.14
Leu	CTG	0.31	0.60	0.61	0.64	0.15	0.05
Leu	CTA	0.01	0.05	0.04	0.01	0.05	0.08
Leu	CTT	0.04	0.07	0.08	0.05	0.18	0.51
Leu	CTC	0.60	0.16	0.14	0.26	0.34	0.11
Lys	AAG	0.86	0.87	0.89	0.93	0.75	0.52
Lys	AAA	0.14	0.13	0.11	0.07	0.25	0.48
Met	ATG	1.00	1.00	1.00	1.00	1.00	1.00
Phe	TTT	0.09	0.25	0.29	0.10	0.44	0.65
Phe	TTC	0.91	0.75	0.71	0.90	0.56	0.35
Pro	CCG	0.28	0.10	0.08	0.53	0.29	0.05
Pro	CCA	0.15	0.10	0.17	0.09	0.12	0.45
Pro	CCT	0.12	0.10	0.30	0.04	0.20	0.33
Pro	CCC	0.44	0.10	0.45	0.34	0.40	0.17

[0902]

Ser	AGT	0.04	0.09	0.06	0.02	0.12	0.14
Ser	AGC	0.05	0.31	0.32	0.20	0.12	0.07
Ser	TCG	0.22	0.04	0.06	0.42	0.19	0.08
Ser	TCA	0.16	0.08	0.10	0.09	0.06	0.31
Ser	TCT	0.05	0.17	0.15	0.07	0.15	0.23
Ser	TCC	0.47	0.31	0.30	0.20	0.35	0.18
Thr	ACG	0.30	0.16	0.13	0.42	0.23	0.10
Thr	ACA	0.06	0.21	0.18	0.03	0.13	0.38
Thr	ACT	0.22	0.18	0.23	0.08	0.19	0.27
Thr	ACC	0.42	0.46	0.46	0.47	0.45	0.25
Trp	TGG	1.00	1.00	1.00	1.00	1.00	1.00
Tyr	TAT	0.07	0.16	0.21	0.12	0.18	0.67
Tyr	TAC	0.93	0.84	0.79	0.88	0.82	0.33
Val	GTG	0.50	0.64	0.62	0.57	0.22	0.30
Val	GTA	0.02	0.03	0.05	0.04	0.09	0.27
Val	GTT	0.06	0.11	0.11	0.04	0.22	0.10
Val	GTC	0.42	0.22	0.23	0.35	0.47	0.33

[0903]

氨基酸	密码子	梭形筒柱藻	强壮前沟藻(Amphidinium carterae)	Symbiodinium micro-adriacticum	微绿球藻种类	隐秘小环藻	Navicula pellicul osa	假微型海链藻	莱茵衣藻
Ala	GCG	0.07	0.17	0.22	0.24	0.11	0.00	0.11	0.35
Ala	GCA	0.14	0.33	0.26	0.10	0.16	0.13	0.25	0.08
Ala	GCT	0.35	0.29	0.20	0.17	0.45	0.44	0.33	0.13
Ala	GCC	0.43	0.20	0.32	0.48	0.27	0.44	0.30	0.43
Arg	AGG	0.09	0.15	0.27	0.00	0.09	0.05	0.18	0.05
Arg	AGA	0.14	0.03	0.27	0.00	0.05	0.10	0.17	0.01
Arg	CGG	0.06	0.08	0.09	0.00	0.04	0.05	0.06	0.20
Arg	CGA	0.16	0.18	0.09	0.29	0.08	0.35	0.11	0.04
Arg	CGT	0.34	0.18	0.09	0.14	0.47	0.20	0.34	0.09
Arg	CGC	0.22	0.40	0.18	0.57	0.28	0.25	0.15	0.62
Asn	AAT	0.42	0.37	0.21	0.00	0.25	0.47	0.43	0.09
Asn	AAC	0.58	0.63	0.79	1.00	0.75	0.53	0.57	0.91
Asp	GAT	0.54	0.54	0.50	0.20	0.52	0.20	0.56	0.14
Asp	GAC	0.46	0.46	0.50	0.80	0.48	0.80	0.44	0.86
Cys	TGT	0.44	0.75	0.50	0.00	0.29	0.10	0.54	0.10
Cys	TGC	0.56	0.25	0.50	1.00	0.71	0.90	0.46	0.90
终止	TGA	0.13	0.50	1.00	0.00	0.10	0.00	0.31	0.27
终止	TAG	0.10	0.00	0.00	0.00	0.00	0.00	0.38	0.22
终止	TAA	0.77	0.50	0.00	1.00	0.90	1.00	0.31	0.52
Gln	CAG	0.12	0.33	0.28	0.41	0.19	0.21	0.16	0.38
Gln	CAA	0.25	0.15	0.17	0.00	0.17	0.28	0.19	0.04
Glu	GAG	0.23	0.41	0.50	0.59	0.38	0.17	0.40	0.55
Glu	GAA	0.39	0.10	0.06	0.00	0.26	0.34	0.26	0.03

[0904]

Gly	GGG	0.06	0.19	0.32	0.10	0.10	0.03	0.12	0.11
Gly	GGA	0.47	0.10	0.12	0.05	0.45	0.28	0.51	0.06
Gly	GGT	0.35	0.34	0.16	0.25	0.22	0.13	0.23	0.11
Gly	GGC	0.12	0.37	0.40	0.60	0.24	0.56	0.14	0.72
His	CAT	0.39	0.12	0.40	0.00	0.42	1.00	0.50	0.11
His	CAC	0.61	0.88	0.60	1.00	0.58	0.00	0.50	0.89
Ile	ATA	0.06	0.05	0.00	0.00	0.04	0.00	0.08	0.03
Ile	ATT	0.42	0.53	0.38	0.14	0.53	0.73	0.38	0.22
Ile	ATC	0.52	0.42	0.63	0.86	0.42	0.27	0.54	0.75
Leu	TTG	0.26	0.35	0.39	0.22	0.20	0.16	0.29	0.04
Leu	TTA	0.09	0.01	0.00	0.00	0.03	0.00	0.05	0.01
Leu	CTG	0.09	0.22	0.39	0.09	0.06	0.12	0.08	0.73
Leu	CTA	0.05	0.00	0.04	0.00	0.03	0.04	0.06	0.03
Leu	CTT	0.37	0.31	0.13	0.04	0.39	0.36	0.20	0.05
Leu	CTC	0.13	0.12	0.04	0.65	0.29	0.32	0.32	0.15
Lys	AAG	0.60	0.93	0.85	1.00	0.70	0.83	0.76	0.95
Lys	AAA	0.40	0.07	0.15	0.00	0.30	0.17	0.24	0.05
Met	ATG	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Phe	TTT	0.37	0.21	0.25	0.20	0.31	0.78	0.38	0.16
Phe	TTC	0.63	0.79	0.75	0.80	0.69	0.22	0.62	0.84
Pro	CCG	0.11	0.14	0.18	0.08	0.10	0.21	0.16	0.33
Pro	CCA	0.33	0.42	0.09	0.08	0.16	0.29	0.31	0.08
Pro	CCT	0.32	0.22	0.41	0.25	0.35	0.21	0.31	0.13
Pro	CCC	0.24	0.22	0.32	0.58	0.39	0.29	0.23	0.47
Ser	AGT	0.12	0.13	0.09	0.00	0.09	0.13	0.18	0.04
Ser	AGC	0.09	0.24	0.14	0.13	0.08	0.28	0.11	0.35
Ser	TCG	0.13	0.03	0.05	0.00	0.15	0.25	0.17	0.25
Ser	TCA	0.12	0.25	0.05	0.00	0.12	0.08	0.12	0.05
Ser	TCT	0.30	0.16	0.23	0.13	0.39	0.25	0.23	0.07
Ser	TCC	0.24	0.19	0.45	0.75	0.18	0.03	0.19	0.25
Thr	ACG	0.09	0.14	0.10	0.28	0.10	0.18	0.21	0.30
Thr	ACA	0.15	0.28	0.10	0.00	0.15	0.09	0.19	0.08
Thr	ACT	0.39	0.12	0.10	0.17	0.33	0.41	0.28	0.10
Thr	ACC	0.37	0.47	0.70	0.56	0.43	0.32	0.32	0.52
Trp	TGG	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Tyr	TAT	0.38	0.32	0.20	0.00	0.38	0.20	0.39	0.10
Tyr	TAC	0.62	0.68	0.80	1.00	0.62	0.80	0.61	0.90
Val	GTG	0.11	0.65	0.67	0.31	0.16	0.18	0.29	0.67
Val	GTA	0.06	0.05	0.00	0.00	0.09	0.09	0.16	0.03
Val	GTT	0.38	0.08	0.11	0.15	0.42	0.09	0.28	0.07
Val	GTC	0.46	0.21	0.22	0.54	0.33	0.64	0.27	0.22

[0905]

氨基酸	密码子	解脂耶罗威亚酵母	高山被孢霉	混浊红球菌
Ala	GCG	0.08	0.14	0.35

[0906]

Ala	GCA	0.11	0.12	0.14
Ala	GCT	0.35	0.29	0.09
Ala	GCC	0.46	0.45	0.43
Arg	AGG	0.05	0.05	0.05
Arg	AGA	0.13	0.06	0.02
Arg	CGG	0.12	0.06	0.26
Arg	CGA	0.52	0.09	0.12
Arg	CGT	0.11	0.32	0.11
Arg	CGC	0.07	0.42	0.44
Asn	AAT	0.17	0.15	0.21
Asn	AAC	0.83	0.85	0.79
Asp	GAT	0.35	0.42	0.24
Asp	GAC	0.65	0.58	0.76
Cys	TGT	0.46	0.13	0.26
Cys	TGC	0.54	0.87	0.74
终止	TGA	0.16	0.05	0.72
终止	TAG	0.38	0.25	0.17
终止	TAA	0.46	0.70	0.11
Gln	CAG	0.33	0.36	0.28
Gln	CAA	0.08	0.06	0.06
Glu	GAG	0.44	0.49	0.45
Glu	GAA	0.14	0.09	0.22
Gly	GGG	0.05	0.03	0.18
Gly	GGA	0.28	0.29	0.15
Gly	GGT	0.32	0.32	0.20
Gly	GGC	0.34	0.36	0.48
His	CAT	0.34	0.27	0.20
His	CAC	0.66	0.73	0.80
Ile	ATA	0.03	0.01	0.05
Ile	ATT	0.44	0.33	0.14
Ile	ATC	0.53	0.66	0.81
Leu	TTG	0.09	0.27	0.09
Leu	TTA	0.02	0.00	0.01
Leu	CTG	0.37	0.26	0.41
Leu	CTA	0.05	0.02	0.03
Leu	CTT	0.18	0.12	0.06
Leu	CTC	0.29	0.32	0.40
Lys	AAG	0.84	0.91	0.80
Lys	AAA	0.16	0.09	0.20
Met	ATG	1.00	1.00	1.00
Phe	TTT	0.38	0.39	0.09
Phe	TTC	0.62	0.61	0.91
Pro	CCG	0.10	0.07	0.52
Pro	CCA	0.10	0.08	0.09
Pro	CCT	0.32	0.36	0.07
Pro	CCC	0.47	0.49	0.32

[0907]

Ser	AGT	0.07	0.05	0.08
Ser	AGC	0.11	0.14	0.23
Ser	TCG	0.16	0.32	0.33
Ser	TCA	0.08	0.08	0.07
Ser	TCT	0.28	0.12	0.05
Ser	TCC	0.30	0.29	0.24
Thr	ACG	0.11	0.17	0.28
Thr	ACA	0.14	0.10	0.11
Thr	ACT	0.26	0.23	0.07
Thr	ACC	0.49	0.49	0.53
Trp	TGG	1.00	1.00	1.00
Tyr	TAT	0.18	0.20	0.18
Tyr	TAC	0.82	0.80	0.82
Val	GTG	0.33	0.22	0.37
Val	GTA	0.05	0.02	0.05
Val	GTT	0.26	0.27	0.10
Val	GTC	0.36	0.49	0.49

[0908] 表70.脂质生物合成途径蛋白。

[0909]

3-酮脂酰 ACP 合成酶

荨距花 3-酮脂酰-ACP 合成酶（GenBank 登录号 AAC68861.1）、*Cuphea wrightii β-酮脂酰-ACP 合成酶 II*（GenBank 登录号 AAB37271.1）、*披针叶荨距花* β-酮脂酰-ACP 合成酶 IV（GenBank 登录号 CAC59946.1）、*Cuphea wrightii β-酮脂酰-ACP 合成酶 II*（GenBank 登录号 AAB37270.1）、*蓖麻酮脂酰-ACP 合成酶*（GenBank 登录号 XP_002516228）、陆地棉*酮脂酰-ACP 合成酶*（GenBank 登录号 ADK23940.1）、*大豆质体 3-酮脂酰-ACP 合成酶 II-A*（GenBank 登录号 AAW88763.1）、*油棕* β-酮脂酰-ACP 合成酶 II（GenBank 登录号 AAF26738.2）、*向日葵*质体 3-酮脂酰-ACP 合成酶 I（GenkBank 登录号 ABM53471.1）、*大豆* 3-酮脂酰-ACP 合成酶 I（GenkBank 登录号 NP_001238610.1）、*向日葵*质体 3-酮脂酰-ACP 合成酶 II（GenBank 登录号 ABI18155.1）、*甘蓝型油菜 β-酮脂酰-ACP 合成酶 2*（GenBank 登录号 *AAF61739.1*）、*紫苏 (Perilla frutescens)* β-酮脂酰-ACP 合成酶 II（GenBank 登录号 AAC04692.1）、*向日葵* β-酮脂酰-ACP 合成酶 II（GenBank 登录号 ABI18155）、*蓖麻* β-酮脂酰-ACP 合成酶 II（GenBank 登录号 AAA33872）、*雨生红球藻* β-酮脂酰酰基载体蛋白合成酶（GenBank 登录号 HM560033.1）、*麻疯树 β* 酮脂酰-ACP 合成酶 I（GenBank 登录号 ABJ90468.1）、*毛果杨 (Populus trichocarpa)* β-酮脂酰-ACP 合成酶 I（GenBank 登录号 XP_002303661.1）、*芫荽 (Coriandrum sativum)* β-酮脂酰-ACP 合成酶 I（GenBank 登录号 AAK58535.1）、*拟南芥* 3-氧酰基-[酰基载体蛋白]合成酶 I（GenBank 登录号 NP_001190479.1）、*葡萄* 3-氧酰基-[酰基载体蛋白]合成酶 I（GenBank 登录号 XP_002272874.2）

脂肪酰基-ACP 硫酯酶

[0910]

*加州月桂*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAC49001）、*樟树*脂肪酰基-ACP 硫酯酶（GenBank 登录号 Q39473）、*加州月桂*脂肪酰基-ACP 硫酯酶（GenBank 登录号 Q41635）、*肉豆蔻*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAB71729）、*肉豆蔻*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAB71730）、*油棕*脂肪酰基-ACP 硫酯酶（GenBank 登录号 ABD83939）、*油棕*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAD42220）、*毛白杨*脂肪酰基-ACP 硫酯酶（GenBank 登录号 ABC47311）、*拟南芥*脂肪酰基-ACP 硫酯酶（GenBank 登录号 NP_172327）、*拟南芥*脂肪酰基-ACP 硫酯酶（GenBank 登录号 CAA85387）、*拟南芥*脂肪酰基-ACP 硫酯酶（GenBank 登录号 CAA85388）、*陆地棉*脂肪酰基-ACP 硫酯酶（GenBank 登录号 Q9SQI3）、*披针叶萼距花*脂肪酰基-ACP 硫酯酶（GenBank 登录号 CAA54060）、*萼距花*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAC72882）、*Cuphea calophylla subsp. mesostemon* 脂肪酰基-ACP 硫酯酶（GenBank 登录号 ABB71581）、*披针叶萼距花*脂肪酰基-ACP 硫酯酶（GenBank 登录号 CAC19933）、*油棕*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAL15645）、*萼距花*脂肪酰基-ACP 硫酯酶（GenBank 登录号 Q39513）、*陆地棉*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAD01982）、*葡萄*脂肪酰基-ACP 硫酯酶（GenBank 登录号 CAN81819）、*山竹*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAB51525）、*长叶马府油树（Madhuca longifolia）*脂肪酰基-ACP 硫酯酶（GenBank 登录号 ABI18986）、*长叶马府油树*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAX51637）、*甘蓝型油菜*脂肪酰基-ACP 硫酯酶（GenBank 登录号 ABH11710)、*甘蓝型油菜*脂肪酰基-ACP 硫酯酶（GenBank 登录号 CAA52070.1）、*水稻*（籼型栽培品种组）脂肪酰基-ACP 硫酯酶（GenBank 登录号 EAY86877）、*水稻*（粳型栽培品种组）脂肪酰基-ACP 硫酯酶（GenBank 登录号 NP_001068400）、*水稻*（籼型栽培品种组）脂肪酰基-ACP 硫酯酶（GenBank 登录号 EAY99617）、*萼距花*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAC49269）、*美洲榆脂肪酰基*-ACP 硫酯酶（GenBank 登录号 AAB71731）、*披针叶萼距花*脂肪酰基-ACP 硫酯酶（GenBank 登录号 CAB60830）、*湿地萼距花*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAC49180）、*德国鸢尾*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAG43858）、*德国鸢尾*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAG43858.1）、*湿地萼距花*脂肪酰基-ACP 硫酯酶（GenBank 登录号 AAC49179）、*肉豆蔻脂肪酰基*-ACP 硫酯酶（GenBank 登录号 AAB71729）、*肉豆蔻脂肪酰基*-ACP 硫酯酶（GenBank 登录号 AAB717291.1）、*萼距花*脂肪酰基-ACP 硫酯酶 GenBank 登录号 U39834）、*加州月桂*脂肪酰基-ACP 硫酯酶（GenBank 登录号 M94159）、*樟树脂肪酰基*-ACP 硫酯酶（GenBank 登录号 U31813）、*蓖麻脂肪酰基*-ACP 硫酯酶（GenBank 登录号 ABS30422.1）、*向日葵*酰基-ACP 硫酯酶（GenBank 登录号 AAL79361.1）、*麻疯树*酰基-ACP 硫酯酶（GenBank 登录号 ABX82799.3）、*玉米*油酰基-酰基载体蛋白硫酯酶,（GenBank 登录号 ACG40089.1）、*雨生红球藻*脂肪酰基-ACP 硫酯酶（GenBank 登录号 HM560034.1）

[0911]

去饱和酶

*亚麻*脂肪酸去饱和酶 3C（GenBank 登录号 ADV92272.1）、*蓖麻* ω-3 脂肪酸去饱和酶（内质网，假定的）（GenBank 登录号 EEF36775.1）、*油桐* ω-3 脂肪酸去饱和酶（GenBank 登录号 AAF12821）、*大豆*叶绿体 ω3 脂肪酸去饱和酶亚型 2（GenBank 登录号 ACF19424.1）、*Prototheca moriformis* FAD-Dω3 去饱和酶（SEQ ID NO: 221）、*Prototheca moriformis* 亚油酸去饱和酶（SEQ ID NO: 220）、*红花* Δ12 去饱和酶（GenBank 登录号 ADM48790.1）、*陆地棉* ω-6 去饱和酶（GenBank 登录号 CAA71199.1）、*大豆*微粒体去饱和酶（GenBank 登录号 BAD89862.1）、*玉米*脂肪酸去饱和酶（GenBank 登录号 ABF50053.1）、*甘蓝型油菜*亚油酸去饱和酶（GenBank 登录号 AAA32994.1）、*亚麻荠* ω-3 去饱和酶（SEQ ID NO: 214）、*Prototheca moriformis* Δ12 去饱和酶等位基因 2（SEQ ID NO: 212）、*亚麻荠* ω-3 FAD7-1（SEQ ID NO: 215）、*向日葵*硬脂酰-ACP 去饱和酶（GenBank 登录号 AAB65145.1）、*蓖麻*硬脂酰-ACP 去饱和酶（GenBank 登录号 AACG59946.1）、*芥菜 (Brassica juncea)* 质体 Δ-9-硬脂酰-ACP 去饱和酶（GenBank 登录号 AAD40245.1）、*大豆*硬脂酰-ACP 去饱和酶（GenBank 登录号 ACJ39209.1）、*油橄榄*硬脂酰-ACP 去饱和酶（GenBank 登录号 AAB67840.1）、*油桐*硬脂酰-酰基-载体蛋白去饱和酶（GenBank 登录号 ADC32803.1）、*播娘蒿 (Descurainia sophia)* Δ-12 脂肪酸去饱和酶（GenBank 登录号 ABS86964.2）、*Euphorbia lagascae* Δ12-油酸去饱和酶（GenBank 登录号 AAS57577.1）、*普通小球藻 Δ12 脂肪酸去饱和酶 (GenBank 登录号* ACF98528）、*普通小球藻 ω-3 脂肪酸去饱和酶 (GenBank 登录号* BAB78717）、*雨生红球藻* ω-3 脂肪酸去饱和酶（GenBank 登录号 HM560035.1）、*雨生红球藻*硬脂酰-ACP-去饱和酶 GenBank 登录号 EF586860.1、*雨生红球藻*硬脂酰-ACP-去饱和酶 GenBank 登录号 EF523479.1

油酸 12-羟化酶

*蓖麻*油酸 12-羟化酶（GenBank 登录号 AAC49010.1）、*Physaria lindheimeri* 油酸 12-羟化酶（GenBank 登录号 ABQ01458.1）、*Physaria lindheimeri* 突变体双功能油酸 12-羟化酶:去饱和酶（GenBank 登录号 ACF17571.1）、*Physaria lindheimeri* 双功能油酸 12-羟化酶:去饱和酶（GenBank 登录号 ACQ42234.1）、*Physaria lindheimeri* 双功能油酸 12-羟化酶:去饱和酶（GenBank 登录号 AAC32755.1）、*琴叶拟南芥 (Arabidopsis lyrata subsp. Lyrata)*（GenBank 登录号 XP_002884883.1）

[0912] 实例22:将普通小球藻工程化

[0913] 可以通过修改如在此讨论的由Chow和Tung等人传授的方法和载体来完成根据本发明的重组基因在普通小球藻中的表达。简言之,Chow和Tung等人,Plant Cell Reports(《植物细胞报道》),第18卷(1999),第778-780页报道了用质粒DNA对普通小球藻的稳定核转化。使用电穿孔转化方法,Chow和Tung将质粒pIG121-Hm(GenBank登录号AB489142)引入

到普通小球藻中。pIG121-Hm的核苷酸序列包含编码β-葡糖醛酸酶(GUS)报告基因产物的序列,该序列可操作地连接至该GUS蛋白编码序列上游的CaMV 35S启动子并且进一步可操作地连接至该GUS蛋白编码序列下游的胭脂碱合成酶(nos)基因的3'UTR/终止子。质粒pIG121-Hm的序列进一步包含潮霉素B抗生素抗性盒。这种潮霉素B抗生素抗性盒包含可操作地连接至编码潮霉素磷酸转移酶(hpt,GenBank登录号BAH24259)基因产物的序列的CaMV 35S启动子。在转化之前,普通小球藻不能在包含50μg/ml潮霉素B的培养基中增殖。用pIG121-Hm质粒转化后,获得普通小球藻的转化子,这些转化子能在包含50μg/ml潮霉素B的培养基中增殖。Hpt基因产物在普通小球藻中的表达使得转化的普通小球藻能够在50μg/ml潮霉素B存在的情况下增殖,由此建立了潮霉素B抗性盒作为供在普通小球藻中使用的选择标记的效用。GUS报告基因的可检测活性表明,CaMV 35S启动子和nos 3'UTR适合于使异源基因能够在普通小球藻中表达。通过Southern分析进行稳定转化子的基因组DNA的评估。在含有YA培养基(琼脂和4g/L酵母提取物)的琼脂平板上进行转化的普通小球藻的选择和维持。如Chow和Tung所讨论,进行普通小球藻在液体培养基中的增殖。已经描述了普通小球藻在除了YA培养基以外的培养基中的增殖(例如,参见Chader等人,Revue des Energies Renouvelabes,第14卷(2011),第21-26页以及Illman等人,《酶与微生物技术》(Enzyme and Microbial Technology),第27卷(2000),第631-635页)。Chow和Tung报道,质粒pIG121-Hm、CaMV 35S启动子、以及根癌土壤杆菌胭脂碱合成酶基因3'UTR/终止子适合于使异源基因能够在普通小球藻中表达。另外,Chow和Tung报道,潮霉素B抗性盒适合用作普通小球藻中的选择标记。适合于使异源基因能在普通小球藻中表达的另外的质粒、启动子、3'UTR/终止子、以及选择标记已经讨论于Chader等人,Revue des Energies Renouvelabes,第14卷(2011),第21-26页中。

[0914] 在本发明的一个实施例中,构建了pIG121-Hm(包含编码用作选择标记的潮霉素B基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在普通小球藻中表达进行了密码子优化以反映根据表69A-D的普通小球藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的CaMV 35S启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的根癌土壤杆菌胭脂碱合成酶基因3'UTR/终止子。该转化构建体可以另外地包含与普通小球藻基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括电穿孔的转化技术或其他已知方法实现用该转化载体稳定转化普通小球藻。潮霉素B抗性基因产物的活性可以用作用于在(但不限于)包含潮霉素的琼脂培养基上选择用该转化载体转化的普通小球藻的标记。适合于普通小球藻脂质产生的生长培养基包括但不限于补充有痕量金属以及任选地1.5g/L NaNO3的BG11培养基(0.04g/L KH_2PO_4、0.075g/L $CaCl_2$、0.036g/L柠檬酸、0.006g/L柠檬酸铁铵、1mg/L EDTA、以及0.02g/L Na_2CO_3)。适合于培养普通小球藻用于脂质产生的另外的培养基包括,例如Watanabe培养基(包含1.5g/L KNO_3、1.25g/L KH_2PO_4、1.25g $l^{-1}MgSO_4 \cdot 7H_2O$、20mg $l^{-1}FeSO_4 \cdot 7H_2O$连同微量营养素)以及低氮培养基(包含203mg/l $(NH_4)_2HPO_4$、2.236g/l KCl、2.465g/l $MgSO_4$、1.361g/l KH_2PO_4以及10mg/l $FeSO_4$),如由Illman等人,《酶与微生物

技术》(Enzyme and Microbial Technology),第27卷(2000),第631-635页所报道。可以通过在此描述的标准脂质提取和分析方法来评定普通小球藻脂质的脂肪酸谱的评估。

[0915] 实例23:将椭圆小球藻工程化

[0916] 可以通过修改如在此讨论的由Chen等人传授的方法和载体来完成根据本发明的重组基因在椭圆小球藻中的表达。简言之,Chen等人,Current Genetics(《当代遗传学》),第39:5卷(2001),第365-370页报道了用质粒DNA对椭圆小球藻的稳定转化。使用电穿孔转化方法,Chow将质粒pBinUΩNP-1引入到椭圆小球藻中。pBinUΩNP-1的核苷酸序列包含编码嗜中性粒细胞肽-1(NP-1)兔基因产物的序列,该序列可操作地连接至该NP-1蛋白编码区的上游的玉米泛素(ubi1)基因启动子并且可操作地连接至该NP-1蛋白编码区的下游的胭脂碱合成酶(nos)基因的3'UTR/终止子。质粒pBinUΩNP-1的序列进一步包含G418抗生素抗性盒。这种G418抗生素抗性盒包含编码氨基糖苷3'-磷酸转移酶(aph 3')基因产物的序列。该aph 3'基因产物赋予对抗生素G418的抗性。在转化之前,椭圆小球藻不能在包含30μg/ml G418的培养基中增殖。用pBinUΩNP-1质粒转化后,获得椭圆小球藻的转化子,这些转化子能在包含30μg/ml G418的选择性培养基中增殖。aph 3'基因产物在椭圆小球藻中的表达使得转化的椭圆小球藻能够在30μg/ml G418存在的情况下增殖,由此建立了G418抗生素抗性盒作为供在椭圆小球藻中使用的选择标记的效用。NP-1基因产物的可检测活性表明,ubi1启动子和nos 3'UTR适合于使异源基因能在椭圆小球藻中表达。通过Southern分析进行稳定转化子的基因组DNA的评估。在具有15μg/mL G418(用于液体培养)或具有30μg/mL G418(用于包含1.8%琼脂的固体培养)的Knop培养基(包含0.2g/L K_2HPO_4、0.2g/L $MgSO_4 \cdot 7H_2O$、0.12g/L KCl、以及10mg/L FeCl3,pH 6.0-8.0,补充有0.1%的酵母提取物和0.2%的葡萄糖)上进行转化的椭圆小球藻的选择和维持。已经描述了椭圆小球藻在除了Knop培养基以外的培养基中的增殖(例如,参见Cho等人,《水产科学》(Fisheries Science),第73:5卷(2007),第1050-1056页,Jarvis和Brown,《当代遗传学》(Current Genetics),第19卷(1991),第317-321页以及Kim等人,《海洋生物技术》(Marine Biotechnology),第4卷(2002),第63-73页)。已经报道了适合于使异源基因能够在椭圆小球藻中表达的另外的质粒、启动子、3'UTR/终止子、以及选择标记(参见Jarvis与Brown以及Kim等人,《海洋生物技术》,第4卷(2002),第63-73页)。Chen报道了质粒pBinUΩNP-1、ubi1启动子、以及根癌土壤杆菌胭脂碱合成酶基因3'UTR/终止子适合于使外源基因能够在椭圆小球藻中表达。另外,Chen报道了在pBinUΩNP-1上编码的G418抗性盒适合用作椭圆小球藻中的选择标记。

[0917] 在本发明的一个实施例中,构建了载体pBinUΩNP-1(包含编码用作选择标记的aph 3'基因产物(赋予对G418的抗性)的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在椭圆小球藻中表达进行了密码子优化以反映根据表69A-D的椭圆小球藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以独地可操作地连接至该蛋白编码序列上游的玉米ubi1启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的根癌土壤杆菌胭脂碱合成酶基因3'UTR/终止子。该转化构建体可以另外地包含与椭圆小球藻基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括电穿孔的转化技

术或其他已知方法实现用该转化载体稳定转化椭圆小球藻。aph 3’基因产物的活性可以用作用于在(但不限于)包含G418的Knop琼脂培养基上选择用该转化载体转化的椭圆小球藻的标记。适合于椭圆小球藻脂质产生的生长培养基包括但不限于Knop培养基以及由Jarvis和Brown以及Kim等人报道的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定椭圆小球藻脂质的脂肪酸谱的评估。

[0918] 实例24:将凯氏小球藻工程化

[0919] 可以通过修改如在此讨论的由El-Sheekh等人传授的方法和载体来完成根据本发明的重组基因在凯氏小球藻中的表达。简言之,El-Sheekh等人,Biologia Plantarium,第42:2卷(1999),第209-216页报道了用质粒DNA对凯氏小球藻的稳定转化。使用微粒轰击转化方法,El-Sheekh将质粒pBI121(GenBank登录号AF485783)引入到凯氏小球藻中。质粒pBI121包含卡那霉素/新霉素抗生素抗性盒。这种卡那霉素/新霉素抗生素抗性盒包含根癌土壤杆菌胭脂碱合成酶(nos)基因启动子、编码针对卡那霉素和G418抗性的新霉素磷酸转移酶II(nptII)基因产物的序列(GenBank登录号AAL92039)、以及根癌土壤杆菌胭脂碱合成酶(nos)基因的3’UTR/终止子。pBI121进一步包含编码β-葡糖醛酸酶(GUS)报告基因产物的序列,该序列可操作地连接至CaMV 35S启动子并且可操作地连接至nos基因的3’UTR/终止子。在转化之前,凯氏小球藻不能在包含15μg/ml卡那霉素的培养基中增殖。用pBI121质粒转化后,获得凯氏小球藻的转化子,这些转化子能在包含15mg/L卡那霉素的选择性培养基中增殖。nptII基因产物在凯氏小球藻中的表达使得能够在15mg/L卡那霉素存在的情况下增殖,由此建立了卡那霉素/新霉素抗生素抗性盒作为用于在凯氏小球藻中使用的选择标记的效用。GUS基因产物的可检测活性表明,CaMV 35S启动子和nos 3’UTR适合于使异源基因能够在凯氏小球藻中表达。通过Southern分析来进行稳定转化子的基因组DNA的评估。如由El-Sheekh报道,在含有YEG培养基(1%酵母提取物,1%葡萄糖)和15mg/L卡那霉素的半固体琼脂平板上进行转化的凯氏小球藻的选择和维持。El-Sheekh还报道了凯氏小球藻在YEG液体培养基中的增殖。适合于培养凯氏小球藻用于脂质产生的另外的培养基披露于Sato等人,《脂质的BBA分子与细胞生物学》(BBA Molecular and Cell Biology of Lipids),第1633卷(2003),第27-34页中。El-Sheekh报道了质粒pBI121、CaMV启动子、以及胭脂碱合成酶基因3’UTR/终止子适合于使异源基因能够在凯氏小球藻中表达。另外,El-Sheekh报道了在pBI121上编码的卡那霉素/新霉素抗性盒适合用作凯氏小球藻中的选择标记。

[0920] 在本发明的一个实施例中,构建了载体pBI121(包含编码用作选择标记的卡那霉素/新霉素抗性基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在凯氏小球藻中表达进行了密码子优化以反映根据表69A-D的凯氏小球藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的CaMV 35S启动子并且可操作地连接至该蛋白编码序列的在3’区域或下游的根癌土壤杆菌胭脂碱合成酶基因3’UTR/终止子。该转化构建体可以另外地包含与凯氏小球藻基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括微粒轰击的转化技术或其

他已知方法实现用该转化载体稳定转化凯氏小球藻。nptII基因产物的活性可以用作用于在(但不限于)包含卡那霉素或新霉素的YEG琼脂培养基上选择用该转化载体转化的凯氏小球藻的标记。适合于凯氏小球藻脂质产生的生长培养基包括但不限于YEG培养基以及由Sato等人报道的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定凯氏小球藻脂质的脂肪酸谱的评估。

[0921] 实例25:将杜氏盐藻工程化

[0922] 可以通过修改如在此讨论的由Walker等人传授的方法和载体来完成根据本发明的重组基因在杜氏盐藻中的表达。简言之,Walker等人,《应用藻类学杂志》(Journal of Applied Phycology),第17卷(2005),第363-368页报道了用质粒DNA对杜氏盐藻的稳定核转化。使用电穿孔转化方法,Walker将质粒pDbleFLAG1.2引入到杜氏盐藻中。pDbleFLAG1.2包含编码博来霉素抗生素抗性盒的序列,该序列包含编码针对抗生素抗腐草霉素抗性的印度斯坦链异壁菌(Streptoalloteichus hindustanus)博来霉素结合蛋白(ble)的序列并且可操作地连接至杜氏盐藻核酮糖-1,5-二磷酸羧化酶/加氧酶小亚基基因(rbcS1,GenBank登录号AY530155)的启动子和3'UTR。在转化之前,杜氏盐藻不能在包含1mg/L腐草霉素的培养基中增殖。用pDbleFLAG1.2质粒转化后,获得杜氏盐藻的转化子,这些转化子能在包含1mg/L腐草霉素的选择性培养基中增殖。ble基因产物在杜氏盐藻中的表达使得能够在1mg/L腐草霉素存在的情况下增殖,由此建立了博来霉素抗生素抗性盒作为用于在杜氏盐藻中使用的选择标记的效用。通过Southern分析来进行稳定转化子的基因组DNA的评估。如由Walker报道,在进一步含有4.5g/L NaCl和1mg/L腐草霉素的杜氏藻培养基(DM,如Provasoli等人,Archiv fur Mikrobiologie,第25卷(1957),第392-428页所描述)上进行转化的杜氏盐藻的选择和维持。适合于培养杜氏盐藻用于脂质产生的另外的培养基讨论于Takagi等人,《生物科学与生物工程杂志》(Journal of Bioscience and Bioengineering),第101:3卷(2006),第223-226页以及Massart和Hanston,Proceedings Venice 2010,Third International Symposium on Energy from Biomass and Waste(第三届国际生物质与废弃物能源大会,会议记录,威尼斯2010)中。Walker报道了质粒pDbleFLAG1.2以及杜氏盐藻核酮糖-1,5-二磷酸羧化酶/加氧酶小亚基基因的启动子和3'UTR适合于使在杜氏盐藻中异源表达能够实现。另外,Walker报道了在pDbleFLAG1.2上编码的博来霉素抗性盒适合用作杜氏盐藻中的选择标记。

[0923] 在本发明的一个实施例中,构建了载体pDbleFLAG1.2(包含编码用作选择标记的ble基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在杜氏盐藻中表达进行了密码子优化以反映根据表69A-D的杜氏盐藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的rbcS1启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的rbcS1 3'UTR/终止子。该转化构建体可以另外地包含与杜氏盐藻基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括电穿孔的转化技术或其他已知方法实现用该转化载体稳定转化杜氏盐藻。ble基因产物的活性可以用作用于在(但不限于)包含腐草霉素的DM培养基上选择用该

转化载体转化的杜氏盐藻的标记。适合于杜氏盐藻脂质产生的生长培养基包括但不限于DM培养基以及由Takagi等人以及Massart和Hanston所描述的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定杜氏盐藻脂质的脂肪酸谱的评估。

[0924] 实例26:将强壮团藻工程化

[0925] 可以通过修改如在此讨论的由Hallman和Rappel等人传授的方法和载体来完成根据本发明的重组基因在强壮团藻中的表达。简言之,Hallman以及Rappel等人,The Plant Journal(《植物杂志》),第17卷(1999),第99-109页报道了用质粒DNA对强壮团藻的稳定核转化。使用微粒轰击转化方法,Hallman和Rappel将pzeoE质粒引入到强壮团藻中。pzeoE质粒包含编码博来霉素抗生素抗性盒的序列,该序列包含编码针对抗生素吉欧霉素(zeocin)抗性的印度斯坦链异壁菌(Streptoalloteichus hindustanus)博来霉素结合蛋白(ble)的序列(可操作地连接至强壮团藻β-微管蛋白基因(GenBank登录号L24547)的启动子和3'UTR)。在转化之前,强壮团藻不能在包含1.5μg/ml吉欧霉素的培养基中增殖。用pzeoE质粒转化后,获得强壮团藻的转化子,这些转化子能在包含大于20μg/ml吉欧霉素的选择性培养基中增殖。ble基因产物在强壮团藻中的表达使得能够在20μg/ml吉欧霉素存在的情况下增殖,由此建立了博来霉素抗生素抗性盒作为用于在强壮团藻中使用的选择标记的效用。通过Southern分析来进行稳定转化子的基因组DNA的评估。如Hallman和Rappel报道,在含有1mg/L腐草霉素的团藻培养基(VM,如由Provasoli和Pintner,《藻类生态学特别版2》(The Ecology of Algae,Special Publication No.2)(1959),Tyron,C.A.与Hartman,R.T.,eds.,Pittsburgh:Univeristy of Pittsburgh(匹兹堡:匹兹堡大学),第88-96页所描述)上进行转化的强壮团藻的选择和维持。适合于培养强壮团藻用于脂质产生的培养基还由Starr在Starr R,C,.Dev Biol Suppl.(《发育生物学增刊》),第4卷(1970),第59-100页中讨论。Hallman和Rappel报道了质粒pzeoE以及强壮团藻β-微管蛋白基因的启动子和3'UTR适合于允许在强壮团藻中异源表达。另外,Hallman和Rappel报道了在pzeoE上编码的博来霉素抗性盒适合用作强壮团藻中的选择标记。已经报道了适合于使异源基因能够在强壮团藻中表达以及适合用作强壮团藻中的选择标记的另外的质粒、启动子、3'UTR/终止子、以及选择标记(例如参见Hallamann和Sumper,Proceedings of the National Academy of Sciences(《美国科学院院报》),第91卷(1994),第11562-11566页以及Hallman和Wodniok,Plant Cell Reports(《植物细胞报道》),第25卷(2006),第582-581页)。

[0926] 在本发明的一个实施例中,构建了载体pzeoE(包含编码用作选择标记的ble基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在强壮团藻中表达进行了密码子优化以反映根据表69A-D的强壮团藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的强壮团藻β-微管蛋白启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的强壮团藻β-微管蛋白3'UTR/终止子。该转化构建体可以另外地包含与强壮团藻基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。本领域的技术人员可以鉴定这样的在强壮团藻基因组序列内的同源区域(引用于Prochnik等人的出版物《科学》(Science),第329:5988卷(2010),第223-226页

中)。通过熟知的包括微粒轰击的转化技术或其他已知方法实现用该转化载体稳定转化强壮团藻。ble基因产物的活性可以用作用于在(但不限于)包含吉欧霉素的VM培养基上选择用该转化载体转化的强壮团藻的标记。适合于强壮团藻脂质产生的生长培养基包括但不限于VM培养基以及由Starr讨论的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定强壮团藻脂质的脂肪酸谱的评估。

[0927] 实例27:将雨生红球藻工程化

[0928] 可以通过修改如在此讨论的由Steinbrenner和Sandmann等人传授的方法和载体来完成根据本发明的重组基因在雨生红球藻中的表达。简言之,Steinbrenner和Sandmann等人,Applied and Environmental Microbiology(《应用与环境微生物学》),第72:12卷(2006),第7477-7484页报道了用质粒DNA对雨生红球藻的稳定核转化。使用微粒轰击转化方法,Steinbrenner将质粒pPlat-pds-L504R引入到雨生红球藻中。质粒pPlat-pds-L504R包含达草灭抗性盒,该达草灭抗性盒包含启动子、蛋白编码序列、以及雨生红球藻八氢番茄红素去饱和酶基因(Pds,GenBank登录号AY781170)的3'UTR,其中Pds的蛋白编码序列在位置504被修饰(由此将亮氨酸改变为精氨酸)以编码一个赋予对除草剂达草灭的抗性的基因产物(Pds-L504R)。在用pPlat-pds-L504R转化之前,雨生红球藻不能在包含5μM达草灭的培养基上增殖。用pPlat-pds-L504R质粒转化后,获得雨生红球藻的转化子,这些转化子能在包含5μM达草灭的选择性培养基中增殖。Pds-L504R基因产物在雨生红球藻中的表达使得能够在5uM达草灭存在的情况下增殖,由此建立了达草灭除草剂抗性盒作为用于在雨生红球藻中使用的选择标记的效用。通过Southern分析来进行稳定转化子的基因组DNA的评估。如由Steinbrenner报道,在补充有2.42g/L Tris-乙酸盐以及5mM达草灭的、包含OHA培养基(OHM,0.41g/L KNO_3、0.03g/L Na_2HPO_4、0.246g/L $MgSO_4 \cdot 7H_2O$、0.11g/L $CaCl_2 \cdot 2H_2O$、2.62mg/L柠檬酸$Fe_{(III)}$x H_2O、0.011mg/L $CoCl_2 \cdot 6H_2O$、0.012mg/L $CuSO_4 \cdot 5H_2O$、0.075mg/L Cr_2O_3、0.98mg/L $MnCl_2 \cdot 4H_2O$、0.12mg/L Na_2MoO_4x $2H_2O$、0.005mg/L SeO_2以及25mg/L生物素、17.5mg/L硫胺素、以及15mg/L维生素B12)的琼脂平板上进行转化的雨生红球藻的选择和维持。Steinbrenner和Sandmann使用基础培养基(如由Kobayashi等人,Applied and Environmental Microbiology(应用与环境微生物学),第59卷(1993),第867-873页所描述的基础培养基)进行了雨生红球藻在液体培养中的增殖。Steinbrenner和Sandmann报道了pPlat-pds-L504R质粒以及雨生红球藻八氢番茄红素去饱和酶基因的启动子和3'UTR适合于允许在雨生红球藻中异源表达。另外,Steinbrenner和Sandmann报道了在pPlat-pds-L504R上编码的达草灭抗性盒适合用作雨生红球藻中的选择标记。已经报道了适合于使异源基因能够在雨生红球藻中表达的另外的质粒、启动子、3'UTR/终止子、以及选择标记(参见Kathiresan等人,Journal of Phycology(《藻类学杂志》),第45卷(2009),第642-649页)。

[0929] 在本发明的一个实施例中,构建了载体pPlat-pds-L504R(包含编码用作选择标记的Pds-L504R基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在雨生红球藻中表达进行了密码子优化以反映根据表69A-D的雨生红球藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序

列上游的雨生红球藻pds基因启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的雨生红球藻pds基因3'UTR/终止子。该转化构建体可以另外地包含与雨生红球藻基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括微粒轰击的转化技术或其他已知方法实现用该转化载体稳定转化雨生红球藻。Pds-L504R基因产物的活性可以用作用于在(但不限于)包含达草灭的OHA培养基上选择用该转化载体转化的雨生红球藻的标记。适合于雨生红球藻脂质产生的生长培养基包括但不限于基础培养基以及由Kobayashi等人,Kathiresan等人,以及Gong和Chen,《应用藻类学杂志》(Journal of Applied Phycology),第9:5卷(1997),第437-444页所描述的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定雨生红球藻脂质的脂肪酸谱的评估。

[0930] 实例28:将极锐新月藻-瘦狭新月藻-沿岸新月藻复合体工程化

[0931] 可以通过修改如在此讨论的由Abe等人传授的方法和载体来完成根据本发明的重组基因在极锐新月藻-瘦狭新月藻-沿岸新月藻复合体中的表达。简言之,Abe等人,Plant Cell Physiology(《植物细胞生理学》),第52:9卷(2011),第1676-1685页报道了用质粒DNA对极锐新月藻-瘦狭新月藻-沿岸新月藻复合体的稳定核转化。使用微粒轰击转化方法,Abe将质粒pSA106引入到极锐新月藻-瘦狭新月藻-沿岸新月藻复合体中。质粒pSA106包含编码博来霉素抗性盒,该抗性盒包含编码印度斯坦链异壁菌博来霉素结合蛋白基因(ble,GenBank登录号CAA37050)的序列,该基因可操作地连接至极锐新月藻-瘦狭新月藻-沿岸新月藻复合体叶绿素a/b-结合蛋白基因(CAB,GenBank登录号AB363403)的启动子和3'UTR。在用pSA106转化之前,极锐新月藻-瘦狭新月藻-沿岸新月藻复合体不能在包含3μg/ml腐草霉素的培养基上增殖。用pSA106转化后,获得极锐新月藻-瘦狭新月藻-沿岸新月藻复合体的转化子,这些转化子能在包含3μg/ml腐草霉素的选择性培养基中增殖。ble基因产物在极锐新月藻-瘦狭新月藻-沿岸新月藻复合体中的表达使能够在3μg/ml腐草霉素存在的情况下增殖,由此建立了博来霉素抗生素抗性盒作为用于在极锐新月藻-瘦狭新月藻-沿岸新月藻复合体中使用的选择标记的效用。通过Southern分析来进行稳定转化子的基因组DNA的评估。如由Abe报道,首先在具有C培养基(0.1g/L KNO_3、0.015g/L $Ca(NO_3)_2 \cdot 4H2O$、0.05g/L甘油磷酸酯-Na2、0.04g/L $MgSO_4 \cdot 7H_2O$、0.5g/L三(羟甲基)氨基甲烷、痕量矿物质、生物素、维生素B_1和B_{12})的顶层琼脂上进行转化的极锐新月藻-瘦狭新月藻-沿岸新月藻复合体的选择和维持,并且然后将其分离至包含补充有腐草霉素的C培养基的琼脂平板上。如Abe报道,进行了在液体培养物中的极锐新月藻-瘦狭新月藻-沿岸新月藻复合体在C培养基中的增殖。适合于极锐新月藻-瘦狭新月藻-沿岸新月藻复合体的增殖的另外的液体培养基由Sekimoto等人,《DNA研究》(DNA Research),10:4(2003),第147-153页讨论。Abe报道了质粒pSA106以及极锐新月藻-瘦狭新月藻-沿岸新月藻复合体CAB基因的启动子和3'UTR适合于使异源基因能够在极锐新月藻-瘦狭新月藻-沿岸新月藻复合体中表达。另外,Abe报道了在pSA106上编码的博来霉素抗性盒适合用作极锐新月藻-瘦狭新月藻-沿岸新月藻复合体中的选择标记。已经报道了适合于使异源基因能够在极锐新月藻-瘦狭新月藻-沿岸新月藻复合体中表达的另外的质粒、启动子、3'UTR/终止子、以及选择标记(参见Abe等人,Plant Cell Physiology(《植物细胞生理学》),第49卷(2008),第625-632页)。

[0932] 在本发明的一个实施例中,构建了载体pSA106(包含编码用作选择标记的ble基因

产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在极锐新月藻-瘦狭新月藻-沿岸新月藻复合体中表达进行了密码子优化以反映根据表69A-D的极锐新月藻-瘦狭新月藻-沿岸新月藻复合体的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的极锐新月藻-瘦狭新月藻-沿岸新月藻复合体CAB基因启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的极锐新月藻-瘦狭新月藻-沿岸新月藻复合体CAB基因3'UTR/终止子。该转化构建体可以另外地包含与极锐新月藻-瘦狭新月藻-沿岸新月藻复合体基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括微粒轰击的转化技术或其他已知方法实现用该转化载体稳定转化极锐新月藻-瘦狭新月藻-沿岸新月藻复合体。ble基因产物的活性可以用作用于在(但不限于)包含腐草霉素的C培养基上选择用该转化载体转化的极锐新月藻-瘦狭新月藻-沿岸新月藻复合体的标记。适合于极锐新月藻-瘦狭新月藻-沿岸新月藻复合体脂质产生的生长培养基包括但不限于C培养基以及由Abe等人和Sekimoto等人报道的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定极锐新月藻-瘦狭新月藻-沿岸新月藻复合体脂质的脂肪酸谱的评估。

[0933] 实例29:将绿色杜氏藻工程化

[0934] 可以通过修改如在此讨论的由Sun等人传授的方法和载体来完成根据本发明的重组基因在绿色杜氏藻中的表达。简言之,Sun等人,《基因》(Gene),第377卷(2006),第140-149页报道了用质粒DNA对绿色杜氏藻的稳定核转化。使用电穿孔转化方法,Sun将编码完全的绿色杜氏藻硝酸盐还原酶基因的质粒pDVNR引入到突变体绿色杜氏藻(绿色杜氏藻NR-突变体)中。在不使用硝酸盐作为氮源的情况下,这些NR-突变体不能够生长。硝酸盐还原酶催化硝酸盐转化为亚硝酸盐。在转化之前,绿色杜氏藻NR-突变体不能在包含硝酸盐(NO_3^-)作为唯一氮源的培养基中增殖。将绿色杜氏藻NR基因产物在NR-突变体绿色杜氏藻中的表达用作为一种选择标记以挽救硝酸盐代谢缺陷。用pDVNR质粒转化后,获得了稳定表达绿色杜氏藻NR基因产物的NR-突变体绿色杜氏藻,其能够在包含硝酸盐作为唯一氮源的琼脂平板上生长。通过Southern分析来进行稳定转化子的DNA的评估。在包含5mM KNO_3的琼脂平板上进行转化的绿色杜氏藻(NR突变体)的选择和维持。Sun还报道了绿色杜氏藻和绿色杜氏藻NR突变体在液体培养基中的增殖。适合于绿色杜氏藻的增殖的另外的培养基由Gordillo等人,《应用藻类学杂志》(Journal of Applied Phycology),第10:2卷(1998),第135-144页以及由Moulton和Burford,《水生生物学》(Hydrobiologia),第204-205:1卷(1990),第401-408页报道。Sun报道了质粒pDVNR以及绿色杜氏藻硝酸盐还原酶基因的启动子和3'UTR/终止子允许在绿色杜氏藻NR-突变体中的异源表达。Sun还报道了绿色杜氏藻硝酸盐还原酶基因产物的表达适合用作绿色杜氏藻NR-突变体中的选择标记。

[0935] 在本发明的一个实施例中,构建了载体pDVNR(包含编码用作选择标记的绿色杜氏藻硝酸盐还原酶(DvNR)基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在绿色杜氏藻中表达进

行了密码子优化以反映根据表69A-D的绿色杜氏藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白，密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的DvNR启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的DvNR 3'UTR/终止子。该转化构建体可以另外地包含与绿色杜氏藻基因组同源的区域，用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括电穿孔的转化技术或其他已知方法实现用该转化载体稳定转化绿色杜氏藻。DvNR基因产物的活性可以用作为一种选择标记以挽救绿色杜氏藻NR突变株的氮同化缺陷以及选择稳定表达该转化载体的绿色杜氏藻NR-突变体。适合于绿色杜氏藻脂质产生的生长培养基包括但不限于由Sun等人、Moulton和Burford以及Gordillo等人讨论的那些。可以通过在此描述的标准脂质提取和分析方法来评定绿色杜氏藻脂质的脂肪酸谱的评估。

[0936] 实例30:将盐生杜氏藻(Dunaliella salina)工程化

[0937] 可以通过修改如在此讨论的由Geng等人传授的方法和载体来完成根据本发明的重组基因在盐生杜氏藻中的表达。简言之，Geng等人，《应用藻类学杂志》(Journal of Applied Phycology)，第15卷(2003)，第451-456页报道了用质粒DNA对盐生杜氏藻的稳定转化。使用电穿孔转化方法，Geng将质粒pUΩHBsAg-CAT引入到盐生杜氏藻中。pUΩHBsAg-CAT包含乙型肝炎表面抗原(HBsAG)表达盒，该表达盒包含编码乙型肝炎表面抗原的序列，该序列可操作地连接至该HBsAG蛋白编码区的上游的玉米ubi1启动子并且可操作地连接至该HBsAG蛋白编码区的下游的根癌土壤杆菌胭脂碱合成酶基因(nos)的3'UTR/终止子。pUΩHBsAg-CAT进一步包含氯霉素抗性盒，该抗性盒包含编码赋予对抗生素氯霉素的抗性的氯霉素乙酰转移酶(CAT)基因产物的序列，该序列可操作地连接至猿猴病毒40的启动子和增强子。在用pUΩHBsAg-CAT转化之前，盐生杜氏藻不能在包含60mg/L的培养基上增殖。用pUΩHBsAg-CAT质粒转化后，获得盐生杜氏藻的转化子，这些转化子能在包含60mg/L氯霉素的选择性培养基中增殖。CAT基因产物在盐生杜氏藻中的表达使得能够在60mg/L氯霉素存在的情况下增殖，由此建立了氯霉素抗性盒作为用于在盐生杜氏藻中使用的选择标记的效用。HBsAg基因产物的可检测活性表明ubi1启动子和nos3'UTR/终止子适合于允许盐生杜氏藻中的基因表达。通过Southern分析来进行稳定转化子的基因组DNA的评估。Geng报道了在包含具有60mg/L氯霉素的约翰逊培养基(Johnson's medium)(J1，由Borowitzka和Borowitzka(eds)，Micro-algal Biotechnology.Cambridge University Press，Cambridge(《微藻生物技术》，剑桥大学出版社，剑桥)，第460-461描述)的琼脂平板上进行的转化的盐生杜氏藻的选择和维持。Geng在具有60mg/L氯霉素的J1培养基上进行了盐生杜氏藻的液体增殖。已经讨论了盐生杜氏藻在除了J1培养基以外的培养基中的增殖(参见Feng等人，《分子生物学报道》(Mol.Bio.Reports)，第36卷(2009)，第1433-1439页以及Borowitzka等人，《水生生物学》(Hydrobiologia)，第116-117:1卷(1984)，第115-121页)。已经由Feng等人报道了适合于使异源基因能够在盐生杜氏藻中表达的另外的质粒、启动子、3'UTR/终止子、以及选择标记。Geng报道了质粒pUΩHBsAg-CAT、ubi1启动子、以及根癌土壤杆菌胭脂碱合成酶基因3'UTR/终止子适合于使外源基因能够在盐生杜氏藻中表达。另外，Geng报道了在pUΩHBsAg-CAT上编码的CAT抗性盒适合用作盐生杜氏藻中的选择标记。

[0938] 在本发明的一个实施例中，构建了载体pUΩHBsAg-CAT(包含编码用作选择标记的

CAT基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在盐生杜氏藻中表达进行了密码子优化以反映根据表69A-D的盐生杜氏藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的ubi1启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的根癌土壤杆菌胭脂碱合成酶基因3'UTR/终止子。该转化构建体可以另外地包含与盐生杜氏藻基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括电穿孔的转化技术或其他已知方法实现用该转化载体稳定转化盐生杜氏藻。CAT基因产物的活性可以用作用于在(但不限于)包含氯霉素的J1培养基中选择用该转化载体转化的盐生杜氏藻的选择标记。适合于盐生杜氏藻脂质产生的生长培养基包括但不限于J1培养基以及由Feng等人和Borowitzka等人描述的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定盐生杜氏藻脂质的脂肪酸谱的评估。

[0939] 实例31:将胸状盘藻工程化

[0940] 可以通过修饰如在此讨论的由Lerche和Hallman等人传授的方法和载体来完成根据本发明的重组基因在胸状盘藻中的表达。简言之,Lerche和Hallman等人,《BMC生物技术》(BMC Biotechnology),第9:64卷,2009报道了用质粒DNA对胸状盘藻的稳定核转化。使用微粒轰击转化方法,Lerche将质粒pPmr3引入到胸状盘藻中。质粒pPmr3包含巴龙霉素抗性盒,该抗性盒包含编码针对抗生素巴龙霉素的抗性的龟裂链霉菌(Streptomyces rimosus)的氨基糖苷3'-磷酸转移酶(aphVIII)基因产物(GenBank登录号AAB03856)的序列,该序列可操作地连接至该aphVIII蛋白编码区的上游的强壮团藻hsp70A-rbcS3杂合启动子并且可操作地连接至该aphVIII蛋白编码区的下游的强壮团藻rbcS3基因的3'UTR/终止子。在用pPmr3转化之前,胸状盘藻不能在包含0.06μg/ml巴龙霉素的培养基上增殖。用pPmr3转化后,获得胸状盘藻的转化子,这些转化子能在包含0.75μg/ml以及更高的μg/ml的巴龙霉素的选择性培养基中增殖。aphVIII基因产物在胸状盘藻中的表达使得能够在0.75以及更高的μg/ml的巴龙霉素存在的情况下增殖,由此建立了巴龙霉素抗生素抗性盒作为用于在胸状盘藻中使用的选择标记的效用。通过Southern分析来进行稳定转化子的基因组DNA的评估。Lerche和Hallman报道了在具有1.0μg/ml巴龙霉素的液体贾沃斯基培养基(Jaworski's medium)(20mg/L $Ca(NO_3)_2 \cdot 4H_2O$、12.4mg/L KH_2PO_4、50mg/L $MgSO_4 \cdot 7H_2O$、15.9mg/L $NaHCO_3$、2.25mg/L EDTA-FeNa、2.25mg/L EDTA Na_2、2.48g/L H_3BO_3、1.39g/L $MnCl_2.4H_2O$、1mg/L $(NH_4)_6MO_7O_24.4H_2O$、0.04mg/L维生素B12、0.04mg/L硫胺素-HCl、0.04mg/L生物素、80mg/L $NaNO_3$、36mg/L $Na_4HPO_4.12H_2O$)中进行转化的胸状盘藻的选择和维持。由Lerche和Hallman进一步讨论了适合于使异源基因能够在胸状盘藻中表达的另外的质粒、启动子、3'UTR/终止子、以及选择标记。Lerche和Hallman报道了质粒pPmr3、强壮团藻hsp70A-rbcS3杂合启动子、以及强壮团藻rbcS3基因的3'UTR/终止子适合于使外源基因能够在胸状盘藻中表达。另外,Lerche和Hallman报道了在pPmr3上编码的巴龙霉素抗性盒适合用作胸状盘藻中的选择标记。

[0941] 在本发明的一个实施例中,构建了载体pPmr3(包含编码用作选择标记的aphVIII

基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列，由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白，对每种蛋白编码序列针对在胸状盘藻中表达进行了密码子优化以反映根据表69A-D的胸状盘藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白，密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的强壮团藻hsp70A-rbcS3杂合启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的强壮团藻rbcS3基因3'UTR/终止子。该转化构建体可以另外地包含与胸状盘藻基因组同源的区域，用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括微粒轰击的转化技术或其他已知方法可以实现用该转化载体稳定转化胸状盘藻。aphVIII基因产物的活性可以用作用于在(但不限于)包含巴龙霉素的贾沃斯基培养基中选择用该转化载体转化的胸状盘藻的选择标记。适合于胸状盘藻脂质产生的生长培养基包括贾沃斯基培养基以及由Stein,《美国植物学杂志》(American Journal of Botany),第45:9卷(1958),第664-672页报道的培养基。可以通过在此描述的标准脂质提取和分析方法来评定胸状盘藻脂质的脂肪酸谱的评估。

[0942] 实例32:将三角褐指藻工程化

[0943] 可以通过修改如在此讨论的由Apt等人传授的方法和载体来完成根据本发明的重组基因在三角褐指藻中的表达。简言之,Apt等人,《分子遗传学和普通遗传学》(Molecular and General Genetics),第252卷(1996)第572-579页报道了用载体DNA对三角褐指藻的稳定核转化。使用微粒轰击转化技术,Apt将质粒pfcpA引入到三角褐指藻中。质粒pfcpA包含博来霉素抗性盒,该抗性盒包含编码针对抗生素腐草霉素和吉欧霉素的抗性的印度斯坦链异壁菌博来霉素结合蛋白(ble)的序列,该序列可操作地连接至该ble蛋白编码区的上游的三角褐指藻藻褐素叶绿素a结合蛋白基因(fcpA)的启动子并且可操作地连接至该ble蛋白编码区的在3'区域或下游的三角褐指藻fcpA基因的3'UTR/终止子。在用pfcpA转化之前,三角褐指藻不能在包含50ug/ml吉欧霉素的培养基上增殖。用pfcpA转化后,获得三角褐指藻的转化子,这些转化子能在包含50μg/ml吉欧霉素的选择性培养基中增殖。ble基因产物在三角褐指藻中的表达使得能够在50μg/ml吉欧霉素存在的情况下增殖,由此建立了博来霉素抗生素抗性盒作为用于在三角褐指藻中使用的选择标记的效用。通过Southern分析来进行稳定转化子的基因组DNA的评估。Apt报道了在包含具有50mg/L吉欧霉素的LDM培养基(如由Starr和Zeikus,《藻类学杂志》(Journal of Phycology),第29卷,增刊,(1993)报道)的琼脂平板上进行的转化的三角褐指藻的选择和维持。Apt报道了在含有50mg/L吉欧霉素的LDM培养基中的三角褐指藻转化子的液体增殖。已经讨论了三角褐指藻在除了LDM培养基以外的培养基中的增殖(由Zaslavskaia等人,《科学》(Science),第292卷(2001),第2073-2075卷,以及由Radokovits等人,《代谢工程》(Metabolic Engineering),第13卷(2011),第89-95页)。已经由Apt等人、由Zaslavskaia等人、以及由Radokovits等人在相同的报道中报道了适合于使异源基因能够在三角褐指藻中表达的另外的质粒、启动子、3'UTR/终止子、以及选择标记。Apt报道了质粒pfcp、三角褐指藻fcpA启动子和3'UTR/终止子适合于使外源基因能够在三角褐指藻中表达。另外,Apt报道了在pfcpA上编码的博来霉素抗性盒适合用作三角褐指藻中的选择标记。

[0944] 在本发明的一个实施例中,构建了载体pfcpA(包含编码用作选择标记的ble基因

产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在三角褐指藻中表达进行了密码子优化以反映根据表69A-D的三角褐指藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的三角褐指藻fcpA基因启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的三角褐指藻fcpA基因3'UTR/终止子。该转化构建体可以另外地包含与三角褐指藻基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。本领域的技术人员可以鉴定这样的在三角褐指藻基因组的序列内的同源区域(引用于Bowler等人的出版物,《自然》(Nature),第456卷(2008),第239-244页中)。通过熟知的包括微粒轰击的转化技术或其他已知方法实现用该转化载体稳定转化三角褐指藻。ble基因产物的活性可以用作用于在(但不限于)包含巴龙霉素的LDM培养基中选择用该转化载体转化的三角褐指藻的标记。适合于三角褐指藻脂质产生的生长培养基包括但不限于如由Radokovits等人报道的f/2培养基。可以通过在此描述的标准脂质提取和分析方法来评定三角褐指藻脂质的脂肪酸谱的评估。

[0945] 实例33:将角毛藻种类工程化

[0946] 可以通过修改如在此讨论的由Yamaguchi等人传授的方法和载体来完成根据本发明的重组基因在角毛藻种类中的表达。简言之,Yamaguchi等人,《藻类学研究》(Phycological Research),第59:2卷(2011),第113-119页报道了用质粒DNA对角毛藻种类的稳定核转化。使用微粒轰击转化方法,Yamaguchi将质粒pTpfcp/nat引入到角毛藻种类中。pTpfcp/nat包含诺尔丝菌素抗性盒,该抗性盒包含编码诺尔丝菌素乙酰转移酶(nat)基因产物(GenBank登录号AAC60439)的序列,该序列可操作地连接至该nat蛋白编码区的上游的假微型海链藻藻褐素叶绿素a/c结合蛋白基因(fcp)启动子并且可操作地连接至在3'区域(该nat蛋白编码区的下游)的假微型海链藻fcp基因3'UTR/终止子。该nat基因产物赋予对抗生素诺尔丝菌素的抗性。在用pTpfcp/nat转化之前,角毛藻种类不能在包含500μg/ml诺尔丝菌素的培养基上增殖。用pTpfcp/nat转化后,获得角毛藻种类的转化子,这些转化子能在包含500μg/ml诺尔丝菌素的选择性培养基中增殖。nat基因产物在角毛藻种类中的表达使得能够在500μg/ml诺尔丝菌素存在的情况下增殖,由此建立了诺尔丝菌素抗生素抗性盒作为用于在角毛藻种类中使用的选择标记的效用。通过Southern分析来进行稳定转化子的基因组DNA的评估。Yamaguchi报道了在包含具有500μg/ml诺尔丝菌素的f/2培养基(如由Guilard,R.R.,Culture of Phytoplankton for feeding marine invertebrates,In Culture of Marine Invertebrate Animals(在培养海洋无脊椎动物中,培养浮游植物用于饲喂海洋无脊椎动物),Smith和Chanley(编辑)1975,Plenum Press(普莱南出版社),纽约,第26-60页报道)的琼脂平板上进行的转化的角毛藻种类的选择和维持。如Yamaguchi所进行的,在含有500mg/L诺尔丝菌素的f/2培养基上进行了角毛藻种类转化子的液体增殖。已经报道了角毛藻种类在另外的培养基中的增殖(例如在Napolitano等人,Journal of the World Aquaculture Society(《全球水产业学会杂志》),第21:2卷(1990),第122-130页,以及Volkman等人,Journal of Experimental Marine Biology and Ecology(《实验海洋学与生态学杂志》),第128:3卷(1989),第219-240页中)。已经由Yamaguchi等人在相同的

报道中报道了适合于使异源基因能够在角毛藻种类中表达的另外的质粒、启动子、3'UTR/终止子、以及选择标记。Yamaguchi报道了质粒pTpfcp/nat、假微型海链藻fcp启动子和3'UTR/终止子适合于使外源基因能够在角毛藻种类中表达。另外,Yamaguchi报道了在pTpfcp/nat上编码的诺尔丝菌素抗性盒适合用作角毛藻种类中的选择标记。

[0947] 在本发明的一个实施例中,构建了载体pTpfcp/nat(包含编码用作选择标记的nat基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在密切相关的扁形角毛藻中表达进行了密码子优化以反映根据表69A-D的扁形角毛藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的假微型海链藻fcp基因启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的假微型海链藻fcp基因3'UTR/终止子。该转化构建体可以另外地包含与角毛藻种类基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过包括微粒轰击的熟知的转化或其他已知方法实现用该转化载体稳定转化角毛藻种类。nat基因产物的活性可以用作用于在(但不限于)包含诺尔丝菌素的f/2琼脂培养基中选择用该转化载体转化的角毛藻种类的选择标记。适合于角毛藻种类脂质产生的生长培养基包括但不限于f/2培养基、以及由Napolitano等人和Volkman等人讨论的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定角毛藻种类脂质的脂肪酸谱的评估。

[0948] 实例34:将梭形筒柱藻工程化

[0949] 可以通过修改如在此讨论的由Poulsen和Kroger等人传授的方法和载体来完成根据本发明的重组基因在梭形筒柱藻(Cylindrotheca fusiformis)中的表达。简言之,Poulsen和Kroger等人,《FEBS杂志》(FEBS Journal),第272卷(2005),第3413-3423页报道了用质粒DNA对梭形筒柱藻的转化。使用微粒轰击转化方法,Poulsen和Kroger将pCF-ble质粒引入到梭形筒柱藻中。质粒pCF-ble包含博来霉素抗性盒,该抗性盒包含编码针对抗生素吉欧霉素和腐草霉素的抗性的印度斯坦链异壁菌博来霉素结合蛋白(ble)的序列,该序列可操作地连接至该ble蛋白编码区的上游的梭形筒柱藻藻褐素叶绿素a/c结合蛋白基因(fcpA,GenBank登录号AY125580)的启动子并且可操作地连接至在3'区域(该ble蛋白编码区的下游)的梭形筒柱藻fcpA基因3'UTR/终止子。在用pCF-ble转化之前,梭形筒柱藻不能在包含1mg/ml吉欧霉素的培养基上增殖。用pCF-ble质粒转化后,获得梭形筒柱藻的转化子,这些转化子能在包含1mg/ml吉欧霉素的选择性培养基中增殖。ble基因产物在梭形筒柱藻中的表达使得能够在1mg/ml吉欧霉素存在的情况下增殖,由此建立了博来霉素抗生素抗性盒作为用于在梭形筒柱藻中使用的选择标记的效用。Poulsen和Kroger报道了在包含具有1mg/ml吉欧霉素的人工海水培养基的琼脂平板上进行的转化的梭形筒柱藻的选择和维持。Poulsen和Kroger报道了在含有1mg/ml吉欧霉素的人工海水培养基中的梭形筒柱藻转化子的液体增殖。已经讨论了梭形筒柱藻在另外的培养基中的增殖(例如在Liang等人,《应用藻类学杂志》(Journal of Applied Phycology),第17:1卷(2005),第61-65页,以及Orcutt和Patterson,《脂质》(Lipids),第9:12卷(1974),第1000-1003页中)。已经由Poulsen和Kroger在相同的报道中报道了用于使异源基因能够在梭形筒柱藻中表达的另外

的质粒、启动子、3'UTR/终止子、以及选择标记。Poulsen和Kroger报道了质粒pCF-ble以及梭形筒柱藻fcp启动子和3'UTR/终止子适合于使外源基因能够在梭形筒柱藻中表达。另外，Poulsen和Kroger报道了在pCF-ble上编码的博来霉素抗性盒适合用作梭形筒柱藻中的选择标记。

[0950] 在本发明的一个实施例中，构建了载体pCF-ble(包含编码用作选择标记的ble基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列，由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白，对每种蛋白编码序列针对在梭形筒柱藻中表达进行了密码子优化以反映根据表69A-D的梭形筒柱藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白，密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的梭形筒柱藻fcp基因启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的梭形筒柱藻fcp基因3'UTR/终止子。该转化构建体可以另外地包含与梭形筒柱藻基因组同源的区域，用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括微粒轰击的转化技术或其他已知方法实现用该转化载体稳定转化梭形筒柱藻。ble基因产物的活性可以用作用于在(但不限于)包含吉欧霉素的人工海水琼脂培养基中选择用该转化载体转化的梭形筒柱藻的选择标记。适合于梭形筒柱藻脂质产生的生长培养基包括但不限于人工海水以及由Liang等人和Orcutt与Patterson报道的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定梭形筒柱藻脂质的脂肪酸谱的评估。

[0951] 实例35:将前沟藻种类(Amphidinium sp.)工程化

[0952] 可以通过修改如在此讨论的由ten Lohuis和Miller等人传授的方法和载体来完成根据本发明的重组基因在前沟藻种类中的表达。简言之，ten Lohuis和Miller等人，《植物杂志》(The Plant Journal)，第13:3卷(1998)，第427-435页报道了用质粒DNA对前沟藻种类的稳定转化。使用在碳化硅晶须存在的情况下搅动的转化技术，ten Lohuis将质粒pMT NPT/GUS引入到前沟藻种类中。pMT NPT/GUS包含新霉素抗性盒，该抗性盒包含编码新霉素磷酸转移酶II(nptII)基因产物(GenBank登录号AAL92039)的序列，该序列可操作地连接至该nptII蛋白编码区的上游或5'的根癌土壤杆菌胭脂碱合成酶(nos)基因启动子并且可操作地连接至在3'区域(该nptII蛋白编码区的下游)的nos基因3'UTR/终止子。该nptII基因产物赋予对抗生素G418的抗性。pMT NPT/GUS质粒进一步包含编码β-葡糖醛酸酶(GUS)报告基因产物的序列，该序列可操作地连接至CaMV 35S启动子并且进一步地可操作地连接至CaMV35S 3'UTR/终止子。在用pMT NPT/GUS转化之前，前沟藻种类不能在包含3mg/ml G418的培养基上增殖。用pMT NPT/GUS质粒转化后，获得前沟藻种类的转化子，这些转化子能在包含3mg/ml G418的选择性培养基中增殖。nptII基因产物在前沟藻种类中的表达使得能够在3mg/ml G418存在的情况下增殖，由此建立了新霉素抗生素抗性盒作为用于在前沟藻种类中使用的选择标记的效用。GUS报告基因的可检测活性表明CaMV 35S启动子和3'UTR适合于使基因能够在前沟藻种类中表达。通过Southern分析来进行稳定转化子的基因组DNA的评估。ten Lohuis和Miller报道了在包含补充有F/2富集溶液(由供应商西格玛公司提供)和3mg/ml G418的海水的培养基中的前沟藻种类的液体增殖，以及在包含补充有F/2富集溶液和3mg/ml G418的海水的琼脂培养基上的前沟藻种类的选择和维持。已经报道了前沟藻

种类在另外的培养基中的增殖(例如在Mansour等人,《应用藻类学杂志》(Journal of Applied Phycology),第17:4卷(2005)第287-v300页中)。已经由ten Lohuis和Miller在相同的报道中报道了用于使异源基因能够在前沟藻种类中表达的另外的质粒(包含另外的启动子、3'UTR/终止子、以及选择标记)。ten Lohuis和Miller报道了质粒pMT NPT/GUS、nos的的启动子和3'UTR/终止子以及CaMV 35S基因适合于使外源基因能够在前沟藻种类中表达。另外,ten Lohuis和Miller报道了在pMT NPT/GUS上编码的新霉素抗性盒适合用作前沟藻种类中的选择标记。

[0953] 在本发明的一个实施例中,构建了载体pMT NPT/GUS(包含编码用作选择标记的nptII基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在前沟藻种类中表达进行了密码子优化以反映根据表69A-D的近缘种(closely-related species)强壮前沟藻(Amphidinium carterae)的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的根癌土壤杆菌胭脂碱合成酶(nos)基因启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的nos 3'UTR/终止子。该转化构建体可以另外地包含与前沟藻种类基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过包括熟知的硅纤维介导的显微注射的转化技术或其他已知方法实现用该转化载体稳定转化前沟藻种类。nptII基因产物的活性可以用作用于在(但不限于)包含G418的海水琼脂培养基上选择用该转化载体转化的前沟藻种类的选择标记。适合于前沟藻种类脂质产生的生长培养基包括但不限于人工海水以及由Mansour等人以及ten Lohuis和Miller报道的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定前沟藻种类脂质的脂肪酸谱的评估。

[0954] 实例36:将Symbiodinium microadriacticum工程化

[0955] 可以通过修改如在此讨论的由ten Lohuis和Miller等人传授的方法和载体来完成根据本发明的重组基因在Symbiodinium microadriacticum中的表达。简言之,ten Lohuis和Miller等人,《植物杂志》(The Plant Journal),第13:3卷(1998),第427-435页报道了用质粒DNA对Symbiodinium microadriacticum的稳定转化。使用硅纤维介导的显微注射的转化技术,ten Lohuis将质粒pMT NPT/GUS引入到Symbiodinium microadriacticum中。pMT NPT/GUS包含新霉素抗性盒,该抗性盒包含编码新霉素磷酸转移酶II(nptII)基因产物(GenBank登录号AAL92039)的序列,该序列可操作地连接至该nptII蛋白编码区的上游或5'的根癌土壤杆菌胭脂碱合成酶(nos)基因启动子并且可操作地连接至在3'区域(该nptII蛋白编码区的下游)的nos基因3'UTR/终止子。该nptII基因产物赋予对抗生素G418的抗性。pMT NPT/GUS质粒进一步包含编码β-葡糖醛酸酶(GUS)报告基因产物的序列,该序列可操作地连接至CaMV 35S启动子并且进一步地可操作地连接至CaMV 35S 3'UTR/终止子。在用pMT NPT/GUS转化之前,Symbiodinium microadriacticum不能在包含3mg/ml G418的培养基上增殖。用pMT NPT/GUS质粒转化后,获得Symbiodinium microadriacticum的转化子,这些转化子能在包含3mg/ml G418的选择性培养基中增殖。nptII基因产物在Symbiodinium microadriacticum中的表达使得能够在3mg/ml G418存在的情况下增殖,由

此建立了新霉素抗生素抗性盒作为用于在Symbiodinium microadriacticum中使用的选择标记的效用。GUS报告基因的可检测活性表明CaMV35S启动子和3'UTR适合于使基因能够在Symbiodinium microadriacticum中表达。通过Southern分析来进行稳定转化子的基因组DNA的评估。ten Lohuis和Miller报道了在包含补充有F/2富集溶液(由供应商西格玛公司提供)与和3mg/ml G418的海水的培养基中的Symbiodinium microadriacticum的液体增殖,以及在包含补充有F/2富集溶液和3mg/ml G418的海水的琼脂培养基中的Symbiodinium microadriacticum的选择和维持。已经讨论了Symbiodinium microadriacticum在另外的培养基中的增殖(例如在Iglesias-Prieto等人,《美国科学院院报》(Proceedings of the National Academy of Sciences),第89:21卷(1992)第10302-10305页中)。已经由ten Lohuis和Miller在相同报道中讨论了用于使异源基因能够在Symbiodinium microadriacticum中表达的另外的质粒(包含另外的启动子、3'UTR/终止子、以及选择标记)。ten Lohuis和Miller报道了质粒pMT NPT/GUS、nos的启动子和3'UTR/终止子以及CaMV 35S基因的适合于使外源基因能够在Symbiodinium microadriacticum中表达。另外,ten Lohuis和Miller报道了在pMT NPT/GUS上编码的新霉素抗性盒适合用作Symbiodinium microadriacticum中的选择标记。

[0956] 在本发明的一个实施例中,构建了载体pMT NPT/GUS(包含编码用作选择标记的nptII基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在Symbiodinium microadriacticum中表达进行了密码子优化以反映根据表69A-D的Symbiodinium microadriacticum的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的根癌土壤杆菌胭脂碱合成酶(nos)基因启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的nos 3'UTR/终止子。该转化构建体可以另外地包含与Symbiodinium microadriacticum基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过包括熟知的硅纤维介导的显微注射的转化技术或其他已知方法实现用该转化载体稳定转化Symbiodinium microadriacticum。nptII基因产物的活性可以用作用于在(但不限于)包含G418的海水琼脂培养基中选择用该转化载体转化的Symbiodinium microadriacticum的选择标记。适合于Symbiodinium microadriacticum脂质产生的生长培养基包括但不限于人工海水以及由Iglesias-Prieto等人以及ten Lohuis和Miller报道的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定Symbiodinium microadriacticum脂质的脂肪酸谱的评估。

[0957] 实例37:将微拟球藻种类工程化

[0958] 可以通过修改如在此讨论的由Kilian等人传授的方法和载体来完成根据本发明的重组基因在微拟球藻种类W2J3B中的表达。简言之,Kilian等人,《美国科学院院报》(Proceedings of the National Academy of Sciences),第108:52卷(2011)第21265-21269页报道了用一种转化构建体对微拟球藻属的稳定核转化。使用电穿孔转化方法,Kilian将转化构建体C2引入到微拟球藻种类W2J3B中。C2转化构建体包含博来霉素抗性盒,该抗性盒包含编码针对抗生素腐草霉素和吉欧霉素的抗性的印度斯坦链异壁菌博来霉素

结合蛋白(ble)的序列,该序列可操作地连接至该ble蛋白编码区的上游的微拟球藻种类W2J3B紫黄质/叶绿素a-结合蛋白基因VCP2的启动子并且可操作地连接至该ble蛋白编码区的下游的微拟球藻种类W2J3B紫黄质/叶绿素a-结合蛋白基因VCP1的3'UTR/终止子。在用C2转化之前,微拟球藻种类W2J3B不能在包含2μg/ml吉欧霉素的培养基上增殖。用C2转化后,获得微拟球藻种类W2J3B的转化子,这些转化子能在包含2μg/ml吉欧霉素的选择性培养基中增殖。ble基因产物在微拟球藻种类W2J3B中的表达使得能够在2μg/ml吉欧霉素存在的情况下增殖,由此建立了博来霉素抗生素抗性盒作为用于在微绿球藻属中使用的选择标记的效用。通过PCR来进行稳定转化子的基因组DNA的评估。Kilian报道了在包含五倍水平的痕量金属、维生素、以及磷酸盐溶液并且进一步包含2μg/ml吉欧霉素的F/2培养基(由Guilard和Ryther,《加拿大微生物学杂志》(Canadian Journal of Microbiology),第8卷(1962),第229-239页报道)中的微拟球藻种类W2J3B转化子的液体增殖。Kilian还报道了在包含人工海水2mg/ml吉欧霉素的琼脂F/2培养基上的微拟球藻种类W2J3B转化子的选择和维持。已经讨论了微拟球藻属在另外的培养基中的增殖(例如在Chiu等人,《生物资源技术》(Bioresour Technol.),第100:2卷(2009),第833-838页以及Pal等人,《应用微生物学与生物技术》(Applied Microbiology and Biotechnology),第90:4卷(2011),第1429-1441页中)。已经由Kilian在相同的报道中描述了另外的转化载体,这些转化载体包含用于使异源基因能够在微拟球藻种类W2J3B中表达的另外的启动子和3'UTR/终止子以及用于转化子选择的选择标记。Kilian报道了转化载体C2以及微拟球藻种类W2J3B紫黄质/叶绿素a-结合蛋白基因VCP2的启动子以及微拟球藻种类W2J3B紫黄质/叶绿素a-结合蛋白基因VCP1的3'UTR/终止子适合于使外源基因能够在微拟球藻种类W2J3B中表达。另外,Kilian报道了在C2上编码的博来霉素抗性盒适合用作微拟球藻种类W2J3B中的选择标记。

[0959] 在本发明的一个实施例中,构建了转化构建体C2(包含编码用作选择标记的ble基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在微拟球藻种类W2J3B中表达进行了密码子优化以反映根据表69A-D的微拟球藻种类的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的微拟球藻种类W2J3B VCP2基因启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的微拟球藻种类W2J3B VCP1基因3'UTR/终止子。该转化构建体可以另外地包含与微拟球藻种类W2J3B基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括电穿孔的转化技术或其他已知方法实现用该转化载体稳定转化微拟球藻种类W2J3B。ble基因产物的活性可以用作用于在(但不限于)包含吉欧霉素的F/2培养基中选择用该转化载体转化的微拟球藻种类W2J3B的选择标记。适合于微拟球藻种类W2J3B脂质产生的生长培养基包括但不限于F/2培养基以及由Chiu等人以及Pal等人报道的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定微拟球藻种类W2J3B脂质的脂肪酸谱的评估。

[0960] 实例38:将隐秘小环藻(Cyclotella cryptica)工程化

[0961] 可以通过修改如在此讨论的由Dunahay等人传授的方法和载体来完成根据本发明的重组基因在隐秘小环藻中的表达。简言之,Dunahay等人,《藻类学杂志》(Journal of

Phycology)，第31卷(1995)，第1004-1012页报道了用质粒DNA对隐秘小环藻的稳定转化。使用微粒轰击转化方法，Dunahay将质粒pACCNPT5.1引入到隐秘小环藻中。质粒pACCNPT5.1包含新霉素抗性盒，该抗性盒包含新霉素磷酸转移酶II(nptII)基因产物的编码序列，该序列可操作地连接至该nptII蛋白编码区的上游的隐秘小环藻乙酰辅酶A羧化酶(ACCase)基因(GenBank登录号L20784)的启动子并且可操作地连接至在3'区域(该nptII蛋白编码区的下游)的隐秘小环藻ACCase基因的3'UTR/终止子。该nptII基因产物赋予对抗生素G418的抗性。在用pACCNPT5.1转化之前，隐秘小环藻不能在包含100μg/ml G418的50％人工海水培养基上增殖。用pACCNPT5.1转化后，获得隐秘小环藻的转化子，这些转化子能在包含100μg/ml G418的选择性50％人工海水培养基中增殖。nptII基因产物在隐秘小环藻中的表达使得能够在100μg/ml G418存在的情况下增殖，由此建立了新霉素抗生素抗性盒作为用于在隐秘小环藻中使用的选择标记的效用。通过Southern分析来进行稳定转化子的基因组DNA的评估。Dunahay报道了在补充有1.07mM硅酸钠以及100μg/ml G418的人工海水培养基(ASW，如由Brown，L.，《藻类学》(Phycologia)，第21卷(1982)，第408-410页讨论)中的隐秘小环藻的液体增殖。Dunahay还报道了在包含具有100μg/ml G418的ASW培养基的琼脂平板上的隐秘小环藻转化子的选择和维持。已经讨论了隐秘小环藻在另外的培养基中的增殖(例如在Sriharan等人，Applied Biochemistry and Biotechnology(《应用生物化学与生物技术》)，第28-29:1卷(1991)，第317-326页以及Pahl等人，《生物科学与生物工程杂志》(Journal of Bioscience and Bioengineering)，第109:3卷(2010)，第235-239页中)。Dunahay报道了质粒pACCNPT5.1以及隐秘小环藻乙酰辅酶A羧化酶(ACCase)基因的启动子适合于使外源基因能够在隐秘小环藻中表达。另外，Dunahay报道了在pACCNPT5.1上编码的新霉素抗性盒适合用作隐秘小环藻中的选择标记。

[0962] 在本发明的一个实施例中，构建了载体pACCNPT5.1(包含编码用作选择标记的nptII基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列，由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白，对每种蛋白编码序列针对在隐秘小环藻中表达进行了密码子优化以反映根据表69A-D的隐秘小环藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白，密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的隐秘小环藻ACCase启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的隐秘小环藻ACCase 3'UTR/终止子。该转化构建体可以另外地包含与隐秘小环藻基因组同源的区域，用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括微粒轰击的转化技术或其他已知方法实现用该转化载体稳定转化隐秘小环藻。nptII基因产物的活性可以用作用于在(但不限于)包含G418的琼脂ASW培养基中选择用该转化载体转化的隐秘小环藻的标记。适合于隐秘小环藻脂质产生的生长培养基包括但不限于ASW培养基以及由Sriharan等人(1991)以及Pahl等人报道的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定隐秘小环藻脂质的脂肪酸谱的评估。

[0963] 实例39:将腐生舟形藻(Navicula saprophila)工程化

[0964] 可以通过修改如在此讨论的由Dunahay等人传授的方法和载体来完成根据本发明的重组基因在腐生舟形藻中的表达。简言之，Dunahay等人，《藻类学杂志》(Journal of

Phycology)，第31卷(1995)，第1004-1012页报道了用质粒DNA对腐生舟形藻的稳定转化。使用微粒轰击转化方法，Dunahay将质粒pACCNPT5.1引入到腐生舟形藻中。质粒pACCNPT5.1包含新霉素抗性盒，该抗性盒包含新霉素磷酸转移酶II(nptII)基因产物的编码序列，该序列可操作地连接至该nptII蛋白编码区的上游的隐秘小环藻乙酰辅酶A羧化酶(ACCase)基因(GenBank登录号L20784)的启动子并且可操作地连接至在3'区域(该nptII蛋白编码区的下游)的隐秘小环藻ACCase基因的3'UTR/终止子。该nptII基因产物赋予对抗生素G418的抗性。在用pACCNPT5.1转化之前，腐生舟形藻不能在包含100μg/ml G418的人工海水培养基上增殖。用pACCNPT5.1转化后，获得腐生舟形藻的转化子，这些转化子能在包含100μg/ml G418的选择性人工海水培养基中增殖。nptII基因产物在腐生舟形藻中的表达使得能够在G418存在的情况下增殖，由此建立了新霉素抗生素抗性盒作为用于在腐生舟形藻中使用的选择标记的效用。通过Southern分析来进行稳定转化子的基因组DNA的评估。Dunahay报道了在补充有1.07mM硅酸钠和100μg/ml G418的人工海水培养基(ASW，如由Brown,L.,《藻类学》(Phycologia)，第21卷(1982)，第408-410页所讨论)中的腐生舟形藻的液体增殖。Dunahay还报道了在包含具有100μg/ml G418的ASW培养基的琼脂平板上的腐生舟形藻转化子的选择和维持。已经讨论了腐生舟形藻在另外的培养基中的增殖(例如在Tadros与Johansen,《藻类学杂志》(Journal of Phycology)，第24:4卷(1988)，第445-452页以及Sriharan等人，《应用生物化学与生物技术》(Applied Biochemistry and Biotechnology)，第20-21:1卷(1989)，第281-291页中)。Dunahay报道了质粒pACCNPT5.1和隐秘小环藻乙酰辅酶A羧化酶(ACCase)基因的启动子适合于使外源基因能够在腐生舟形藻中表达。另外，Dunahay报道在pACCNPT5.1上编码的新霉素抗性盒适合用作腐生舟形藻中的选择标记。

[0965] 在本发明的一个实施例中，构建了载体pACCNPT5.1(包含编码用作选择标记的nptII基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列，由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白，对每种蛋白编码序列针对在腐生舟形藻中表达进行了密码子优化以反映根据表69A-D的近缘的羽纹硅藻(Navicula pelliculosa)的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白，密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的隐秘小环藻ACCase基因启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的隐秘小环藻ACCase基因3'UTR/终止子。该转化构建体可以另外地包含与腐生舟形藻基因组同源的区域，用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。通过熟知的包括微粒轰击的转化技术或其他已知方法实现用该转化载体稳定转化腐生舟形藻。nptII基因产物的活性可以用作用于在(但不限于)包含G418的琼脂ASW培养基中选择用该转化载体转化的腐生舟形藻的选择标记。适合于腐生舟形藻脂质产生的生长培养基包括但不限于ASW培养基以及由Sriharan等人(1989)以及Tadros和Johansen报道的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定腐生舟形藻脂质的脂肪酸谱的评估。

[0966] 实例40:将假微型海链藻工程化

[0967] 可以通过修改如在此讨论的由Poulsen等人传授的方法和载体来完成根据本发明的重组基因在假微型海链藻中的表达。简言之，Poulsen等人，《藻类学杂志》(Journal of

Phycology)，第42卷(2006)，第1059-1065页报道了用质粒DNA对假微型海链藻的稳定转化。使用微粒轰击转化方法，Poulsen将质粒pTpfcp/nat引入到假微型海链藻中。pTpfcp/nat包含诺尔丝菌素抗性盒，该抗性盒包含编码诺尔丝菌素乙酰转移酶(nat)基因产物(GenBank登录号AAC60439)的序列，该序列可操作地连接至该nat蛋白编码区的上游的假微型海链藻藻褐素叶绿素a/c结合蛋白基因(fcp)启动子并且可操作地连接至在3'区域(该nat蛋白编码区的下游)的假微型海链藻fcp基因3'UTR/终止子。该nat基因产物赋予对抗生素诺尔丝菌素的抗性。在用pTpfcp/nat转化之前，假微型海链藻不能在包含10μg/ml诺尔丝菌素的培养基上增殖。用pTpfcp/nat转化后，获得假微型海链藻的转化子，这些转化子能在包含100μg/ml诺尔丝菌素的选择性培养基中增殖。nat基因产物在假微型海链藻中的表达使得能够在100μg/ml诺尔丝菌素存在的情况下增殖，由此建立了诺尔丝菌素抗生素抗性盒作为用于在假微型海链藻中使用的选择标记的效用。通过Southern分析来进行稳定转化子的基因组DNA的评估。Poulsen报道了在包含具有100μg/ml诺尔丝菌素的改良ESAW培养基(如由Harrison等人，《藻类学杂志》(Journal of Phycology)，第16卷(1980)，第28-35页所讨论)的液体培养中进行的转化的假微型海链藻的选择和维持。已经讨论了假微型海链藻在另外的培养基中的增殖(例如在Volkman等人，《实验海洋生物学与生态学杂志》(Journal of Experimental Marine Biology and Ecology)，第128:3卷(1989)，第219-240页中)。已经由Poulsen在相同的报道中讨论了另外的质粒(包含适合于在假微型海链藻中使用的另外的选择标记)。Poulsen报道了质粒pTpfcp/nat、假微型海链藻fcp启动子以及3'UTR/终止子适合于使外源基因能够在假微型海链藻中表达。另外，Poulsen报道了在pTpfcp/nat上编码的诺尔丝菌素抗性盒适合用作假微型海链藻中的选择标记。

[0968] 在本发明的一个实施例中，构建了载体pTpfcp/nat(包含编码用作选择标记的nat基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列，由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白，对每种蛋白编码序列针对在假微型海链藻中表达进行了密码子优化以反映根据表69A-D的假微型海链藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白，密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的假微型海链藻fcp基因启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的假微型海链藻fcp基因3'UTR/终止子。该转化构建体可以另外地包含与假微型海链藻基因组同源的区域，用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。本领域的技术人员可以鉴定这样的在假微型海链藻基因组的序列内的同源区域(引用于Armbrust等人的出版物，《科学》(Science)，第306:5693卷(2004):第79-86页中)。通过熟知的包括微粒轰击的转化技术或其他已知方法实现用该转化载体稳定转化假微型海链藻。nat基因产物的活性可以用作用于在(但不限于)包含诺尔丝菌素的ESAW琼脂培养基中选择用该转化载体转化的假微型海链藻的标记。适合于假微型海链藻脂质产生的生长培养基包括但不限于ESAW培养基以及由Volkman等人和Harrison等人讨论的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定假微型海链藻脂质的脂肪酸谱的评估。

[0969] 实例41:将莱茵衣藻工程化

[0970] 可以通过修改如在此讨论的由Cerutti等人传授的方法和载体来完成根据本发明

的重组基因在莱茵衣藻中的表达。简言之,Cerutti等人,《遗传学》(Genetics),第145:1卷(1997),第97-110页报道了用转化载体对莱茵衣藻的稳定核转化。使用微粒轰击转化方法,Cerutti将转化构建体P引入到莱茵衣藻中。构建体P包含壮观霉素抗性盒,该抗性盒包含编码氨基糖甙3”-腺苷转移酶(aadA)基因产物的序列,该序列可操作地连接至该aadA蛋白编码区的上游的莱茵衣藻核酮糖-1,5-二磷酸羧化酶/加氧酶小亚基基因(RbcS2,GenBank登录号X04472)启动子并且可操作地连接至在3’区域(该aadA蛋白编码区的下游)的莱茵衣藻RbcS2基因3’UTR/终止子。该aadA基因产物赋予对抗生素壮观霉素的抗性。在用P转化之前,莱茵衣藻不能在包含90μg/ml壮观霉素的培养基上增殖。用P转化后,得到莱茵衣藻的转化子,这些转化子能在包含90μg/ml壮观霉素的选择性培养基中增殖,由此建立了壮观霉素抗生素抗性盒作为用于在莱茵衣藻中使用的选择标记的效用。通过Southern分析来进行稳定转化子的基因组DNA的评估。Cerutti报道了在包含具有90μg/ml壮观霉素的Tris-乙酸盐-磷酸盐培养基(TAP,如由Harris,《衣藻属资料手册》(The Chlamydomonas Sourcebook),学术出版社,圣地亚哥,1989所描述)的琼脂平板上进行的转化的莱茵衣藻的选择和维持。Cerutti另外报道了莱茵衣藻在具有90μg/ml壮观霉素的TAP液体培养中的增殖。已经讨论了莱茵衣藻在可替代的培养基中的增殖(例如在Dent等人,《非洲微生物学研究杂志》(African Journal of Microbiology Research),第5:3卷(2011),第260-270页以及Yantao等人,《生物技术与生物工程》(Biotechnology and Bioengineering),第107:2卷(2010),第258-268页中)。适合用于促进异源基因在莱茵衣藻中的表达的另外的构建体(包含适合于在莱茵衣藻中使用的另外的选择标记以及很多调节序列(包括启动子与3’UTR))在本领域是已知的并且已经讨论(为了回顾,参见Radakovits等人,《真核细胞》(Eurkaryotic Cell),第9:4卷(2010),第486-501页)。Poulsen报道了转化载体P以及莱茵衣藻启动子和3’UTR/终止子适合于使外源基因能够在莱茵衣藻中表达。另外,Cerutti报道了在P上编码的壮观霉素抗性盒适合用作莱茵衣藻中的选择标记。

[0971] 在本发明的一个实施例中,构建了载体P(包含编码用作选择标记的aadA基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在莱茵衣藻中表达进行了密码子优化以反映根据表69A-D的莱茵衣藻的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的莱茵衣藻RbcS2启动子并且可操作地连接至该蛋白编码序列的在3’区域或下游的莱茵衣藻RbcS2 3’UTR/终止子。该转化构建体可以另外地包含与莱茵衣藻基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。本领域的技术人员可以鉴定这样的在莱茵衣藻基因组的序列内的同源区域(引用于Merchant等人的出版物,《科学》(Science)第318:5848卷(2007):第245-250页中)。通过熟知的包括微粒轰击的转化技术或其他已知方法实现用该转化载体稳定转化莱茵衣藻。aadA基因产物的活性可以用作用于在(但不限于)包含壮观霉素的TAP琼脂培养基上选择用该转化载体转化的莱茵衣藻的标记。适合于莱茵衣藻脂质产生的生长培养基包括但不限于ESAW培养基以及由Yantao等人和Dent等人讨论的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定莱茵衣藻脂质的脂肪酸谱的评估。

[0972] 实例42:将解脂耶罗威亚酵母工程化

[0973] 可以通过修改如在此讨论的由Fickers等人传授的方法和载体来完成根据本发明的重组基因在解脂耶罗威亚酵母中的表达。简言之,Fickers等人,《微生物学方法杂志》(Journal of Microbiological Methods),第55卷(2003),第727-737页报道了用质粒DNA对解脂耶罗威亚酵母的稳定核转化。使用乙酸锂转化方法,Fickers将质粒JMP123引入到解脂耶罗威亚酵母中。质粒JMP123包含潮霉素B抗性盒,该抗性盒包含编码潮霉素B磷酸转移酶基因产物(hph)的序列,该序列可操作地连接至该hph蛋白编码区的上游的解脂耶罗威亚酵母LIP2基因启动子(GenBank登录号AJ012632)并且可操作地连接至该hph蛋白编码区的下游的解脂耶罗威亚酵母LIP2基因3'UTR/终止子。在用JMP123转化前,解脂耶罗威亚酵母不能在包含100μg/ml潮霉素的培养基上增殖。用JMP123转化后,得到了转化的解脂耶罗威亚酵母,它们能在包含100μg/ml潮霉素的培养基上增殖,由此建立了潮霉素B抗生素抗性盒作为用于在解脂耶罗威亚酵母中使用的选择标记的效用。该核苷酸序列在JMP123上提供了用作用于将hph编码序列同源重组到LIP2基因座的供体序列的解脂耶罗威亚酵母LIP2基因的启动子和3'UTR/终止子。通过Southern来进行稳定转化子的基因组DNA的评估。Fickers报道了在包含具有100μg/ml潮霉素的标准YPD培养基(酵母浸出粉胨葡萄糖琼脂)的琼脂平板上进行的转化的解脂耶罗威亚酵母的选择和维持。在具有潮霉素的YPD培养基中进行转化的解脂耶罗威亚酵母的液体培养。已经报道了用于培养解脂耶罗威亚酵母的其他培养基和技术并且已经报道了用于在解脂耶罗威亚酵母中使用的许多其他质粒、启动子、3'UTR、以及选择标记(例如参见Pignede等人,《应用与环境生物学》(Applied and Environmental Biology),第66:8卷(2000),第3283-3289页,Chuang等人,《新生物技术》(New Biotechnology),第27:4卷(2010),第277-282页以及Barth和Gaillardin,(1996),In:K,W.(编辑),《生物技术中的非常规酵母》(Nonconventional Yeasts in Biotecnology),Sprinter-Verlag,柏林-海德尔堡,第313-388页)。Fickers报道了转化载体JMP123以及解脂耶罗威亚酵母LIP2基因启动子和3'UTR/终止子适合于使异源基因能够在解脂耶罗威亚酵母中表达。另外,Fickers报道了在JMP123上编码的潮霉素抗性盒适合用作解脂耶罗威亚酵母中的选择标记。

[0974] 在本发明的一个实施例中,构建了载体JMP123(包含编码用作选择标记的hph基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在解脂耶罗威亚酵母中表达进行了密码子优化以反映根据表69A-D的解脂耶罗威亚酵母的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的解脂耶罗威亚酵母LIP2基因启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的解脂耶罗威亚酵母LIP2基因3'UTR/终止子。该转化构建体可以另外地包含与解脂耶罗威亚酵母基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。本领域的技术人员可以鉴定这样的在解脂耶罗威亚酵母基因组的序列内的同源区域(引用于Dujun等人的出版物,《自然》(Nature),第430卷(2004),第35-44页中)。通过熟知的包括乙酸锂转化的转化技术或其他已知方法实现用该转化载体稳定转化解脂耶罗威亚酵母。hph基因产物的活性可以用作

用于在(但不限于)包含潮霉素的YPD培养基上选择用该转化载体转化的解脂耶罗威亚酵母的标记。适合于解脂耶罗威亚酵母脂质产生的生长培养基包括但不限于YPD培养基以及由Chuang等人描述的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定解脂耶罗威亚酵母脂质的脂肪酸谱的评估。

[0975] 实例43:将高山被孢霉工程化

[0976] 可以通过修改如在此讨论的由Mackenzie等人传授的方法和载体来完成根据本发明的重组基因在高山被孢霉中的表达。简言之,Mackenzie等人,《应用与环境微生物学》(Applied and Environmental Microbiology),第66卷(2000),第4655-4661页报道了用质粒DNA对高山被孢霉的稳定核转化。使用原生质体转化方法,MacKenzie将质粒pD4引入到高山被孢霉中。质粒pD4包含潮霉素B抗性盒,该抗性盒包含编码潮霉素B磷酸转移酶基因产物(hpt)的序列,该序列可操作地连接至该hpt蛋白编码区的上游的高山被孢霉组蛋白H4.1基因启动子(GenBank登录号AJ249812)并且可操作地连接至该hpt蛋白编码区的下游的构巢曲霉(Aspergillus nidulans)N-(5'-磷酸核糖)邻氨基苯甲酸异构酶(phophoribosyl) anthranilate isomerase)(trpC)基因3'UTR/终止子。在用pD4转化前,高山被孢霉不能在包含300μg/ml潮霉素的培养基上增殖。用pD4转化后,得到了转化的高山被孢霉,它们能在包含300μg/ml潮霉素的培养基上增殖,由此建立了潮霉素B抗生素抗性盒作为用于在高山被孢霉中使用的选择标记的效用。通过Southern来进行稳定转化子的基因组DNA的评估。Mackenzie报道了在包含潮霉素的PDA(马铃薯葡萄糖琼脂)培养基上进行的转化的高山被孢霉的选择和维持。由Mackenzie在具有潮霉素的PDA培养基中或S2GYE培养基中(包含5%葡萄糖、0.5%酵母提取物、0.18%NH_4SO_4、0.02%$MgSO_4$-$7H_2O$、0.0001%$FeCl_3$-$6H_2O$、0.1%微量元素、10mM K_2HPO_4-NaH_2PO_4)上进行了转化的高山被孢霉的液体培养。已经报道了用于培养高山被孢霉的其他培养基和技术并且已经报道了用于在高山被孢霉中使用的其他质粒、启动子、3'UTR、以及选择标记(例如参见Ando等人,《应用与环境生物学》(Applied and Environmental Biology),第75:17卷(2009)第5529-35页以及Lu等人,《应用生物化学与生物技术》(Applied Biochemistry and Biotechnology),第164:7卷(2001),第979-90页)。Mackenzie报道了转化载体pD4以及高山被孢霉组蛋白H4.1启动子和构巢曲霉trpC基因3'UTR/终止子适合于使异源基因能够在高山被孢霉中表达。另外,Mackenzie报道了在pD4上编码的潮霉素抗性盒适合用作高山被孢霉中的选择标记。

[0977] 在本发明的一个实施例中,构建了载体pD4(包含编码用作选择标记的hpt基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在高山被孢霉中表达进行了密码子优化以反映根据表69A-D的高山被孢霉的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的高山被孢霉组蛋白H4.1基因启动子并且可操作地连接至该蛋白编码序列的在3'区域或下游的构巢曲霉trpC 3'UTR/终止子。该转化构建体可以另外地包含与高山被孢霉基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。本领域的技术人员可以鉴定这样的在高山被孢霉基因组的序列内的同源区域(引用于Wang等人的出版物,PLOS One,第6:12卷(2011)中)。通过熟知的包括原

生质体转化的转化技术或其他已知方法实现用该转化载体稳定转化高山被孢霉。hpt基因产物的活性可以用作用于在(但不限于)包含潮霉素的PDA培养基上选择用该转化载体转化的高山被孢霉的标记。适合于高山被孢霉脂质产生的生长培养基包括但不限于S2GYE培养基以及由Lu等人和Ando等人描述的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定高山被孢霉脂质的脂肪酸谱的评估。

[0978] 实例44:将混浊红球菌PD630工程化

[0979] 可以通过修改如在此讨论的由Kalscheuer等人传授的方法和载体来完成根据本发明的重组基因在混浊红球菌PD630中的表达。简言之,Kalscheuer等人,《应用与环境微生物学》(Applied and Environmental Microbiology),第52卷(1999),第508-515页报道了用质粒DNA对混浊红球菌的稳定转化。使用电穿孔转化方法,Kalscheuer将质粒pNC9501引入到混浊红球菌PD63中。质粒pNC9501包含硫链丝菌素抗性($thio^r$)盒,该抗性盒包含远青链霉菌(Streptomyces azureus)23S rRNA A1067甲基转移酶基因的完全核苷酸序列,包括该基因的启动子和3'终止子序列。用pNC9501转化之前,混浊红球菌不能在包含1mg/ml硫链丝菌素的培养基上增殖。用pNC9501对混浊红球菌PD630转化后,得到了转化子,这些转化子能在包含1mg/ml硫链丝菌素的培养基上增殖,由此建立了硫链丝菌素抗性盒作为混浊红球菌PD630中的选择标记的用途。由Kalscheuer描述的一种第二质粒pAK68包含抗性$thio^r$盒以及真氧产碱杆菌(Ralstonia eutropha)β-酮硫解酶(phaB)、乙酰乙酸-CoA还原酶(phaA)、和针对聚羟基脂肪酸酯生物合成的聚3-羟基脂肪酸合成酶(phaC)基因的基因序列,由lacZ启动子驱动。pAK68对聚羟基脂肪酸酯生物合成缺陷型混浊红球菌PD630株转化后,得到转化的混浊红球菌PD630,其比未转化株生产更高量的聚羟基脂肪酸酯。被引入的真氧产碱杆菌phaB、phaA、以及phaC酶的可检测活性表明在pAK68质粒上编码的调节元件适合于在混浊红球菌PD630中的异源基因表达。Kalscheuer报道了在包含硫链丝菌素的标准Luria Broth(LB)培养基、营养肉汤(NB)、或矿物盐培养基(MSM)上进行的转化的混浊红球菌PD630的选择和维持。已经描述了用于培养混浊红球菌PD630的其他培养基和技术(例如参见Kurosawa等人,《生物技术杂志》(Journal of Biotechnology),第147:3-4卷(2010),第212-218页以及Alverez等人,《应用微生物与生物技术》(Applied Microbial and Biotechnology),第54:2卷(2000),第218-223页)。Kalscheuer报道了转化载体pNC9501和pAK68、远青链霉菌23S rRNA A1067甲基转移酶基因和lacZ基因的启动子适合于使异源基因能够在混浊红球菌PD630中表达。另外,Kalscheuer报道了在pNC9501和pAK68上编码的$thio^r$盒适合用作混浊红球菌PD630中的选择标记。

[0980] 在本发明的一个实施例中,构建了载体pNC9501(包含编码用作选择标记的$thio^r$基因产物的核苷酸序列)并且将其进行修饰以进一步包含脂质生物合成途径表达盒序列,由此产生了一种转化载体。该脂质生物合成途径表达盒编码一种或多种选自表70的脂质生物合成途径蛋白,对每种蛋白编码序列针对在混浊红球菌PD630中表达进行了密码子优化以反映根据表69A-D的混浊红球菌的核基因中的固有密码子偏倚。对于表70的每种脂质生物合成途径蛋白,密码子优化的基因序列可以单独地可操作地连接至该蛋白编码序列上游的lacZ基因启动子。该转化构建体可以另外地包含与混浊红球菌PD630基因组同源的区域,用于该转化载体的靶向基因组整合。可以选择同源区域以破坏内源脂质生物合成途径基因的一个或多个基因组位点。本领域的技术人员可以鉴定这样的在混浊红球菌PD630基因组

的序列内的同源区域(引用于Holder等人的出版物,《PLOS遗传学》(PLOS Genetics),第7:9卷(2011)中)。通过熟知的包括电穿孔的转化技术或其他已知方法实现用该转化载体稳定转化混浊红球菌PD630。远青链霉菌23S rRNA A1067甲基转移酶基因产物的活性可以用作用于在(但不限于)包含硫链丝菌素的LB培养基上选择用该转化载体转化的混浊红球菌PD630的标记。适合于混浊红球菌PD630脂质产生的生长培养基包括但不限于由Kurosawa等人和Alvarez等人讨论的那些培养基。可以通过在此描述的标准脂质提取和分析方法来评定混浊红球菌PD630脂质的脂肪酸谱的评估。

[0981] 在此引证的所有参考文献,包括专利、专利申请、以及出版物,包括Genbank登录号,特此通过引用以其整体而结合,无论先前是否明确地结合。出于描述和披露可以与本发明一起使用的试剂、方法学和构思的目的而引证在此所提到的出版物。在此任何内容均不得解释为承认这些参考文献相对于在此所述的发明是现有技术。具体而言,出于所有目的将以下专利申请特此通过引用以其整体而结合:PCT申请号PCT/US2008/065563,2008年6月2日提交,名称为“微生物中油的生产(Production of Oil in Microorganisms)”,PCT申请号PCT/US2010/31108,2010年4月14日提交,名称为“微生物油提取和分离方法(Methods of Microbial Oil Extraction and Separation)”、以及PCT申请号PCT/US2009/066142,2009年11月30日提交,名称为“异养微生物中定制油的生产(Production of Tailored Oils in Heterotrophic Microorganisms)”。

序 列 表

<110> SOLAZYME, INC.
<120> 产自重组产油微生物的定制油
<130> 057518-415312 (068PCT)
<140> PCT/US12/023696
<141> 2012-02-02
<150> US 61/438,969
<151> 2011-02-02
<150> US 61/476,691
<151> 2011-04-18
<150> US 61/484,458
<151> 2011-05-10
<150> US 61/548,616
<151> 2011-10-18
<160> 256
<170> PatentIn 版本3.5
来自小球藻属的HUP启动子(GeneBank登录号X55349的子序列)
<210> 1
<211> 1187
<212> DNA
<213> 未知
<220>
<223> 小球藻属
<400> 1

gatcagacgg gcctgacctg cgagataatc aagtgctcgt aggcaaccaa ctcagcagct 60
gcttggtgtt gggtctgcag gatagtgttg cagggcccca aggacagcag gggaacttac 120
accttgtccc cgacccagtt ttatggagtg cattgcctca agagcctagc cggagcgcta 180
ggctacatac ttgccgcacc ggtatgaggg gatatagtac tcgcactgcg ctgtctagtg 240
agatgggcag tgctgcccat aaacaactgg ctgctcagcc atttgttggc ggaccattct 300
gggggggcca gcaatgcctg actttcgggt agggtgaaaa ctgaacaaag actaccaaaa 360
cagaatttct tcctccttgg aggtaagcgc aggccggccc gcctgcgccc acatggcgct 420
ccgaacacct ccatagctgt aagggcgcaa acatggccgg actgttgtca gcactctttc 480
atggccatac aaggtcatgt cgagattagt gctgagtaag acactatcac cccatgttcg 540
attgaagccg tgacttcatg ccaacctgcc cctgggcgta gcagacgtat gccatcatga 600
ccactagccg acatgcgctg tcttttgcca ccaaaacaac tggtacaccg ctcgaagtcg 660
tgccgcacac ctccgggagt gagtccggcg actcctcccc ggcgggccgc ggccctacct 720
gggtagggtc gccatacgcc cacgaccaaa cgacgcagga ggggattggg gtagggaatc 780
ccaaccagcc taaccaagac ggcacctata ataataggtg gggggactaa cagccctata 840

tcgcaagctt tgggtgccta tcttgagaag cacgagttgg agtggctgtg tacggtcgac 900
cctaaggtgg gtgtgccgca gcctgaaaca aagcgtctag cagctgcttc tataatgtgt 960
cagccgttgt gtttcagtta tattgtatgc tattgtttgt tcgtgctagg gtggcgcagg 1020
cccacctact gtggcgggcc attggttggt gcttgaattg cctcaccatc taaggtctga 1080
acgctcactc aaacgccttt gtacaactgc agaactttcc ttggcgctgc aactacagtg 1140
tgcaaaccag cacatagcac tcccttacat cacccagcag tacaaca 1187

来自AY307383的椭圆小球藻硝酸盐还原酶启动子

<210> 2
<211> 1414
<212> DNA
<213> 椭圆小球藻
<400> 2

cgctgcgcac cagggccgcc agctcgctga tgtcgctcca aatgcggtcc cccgattttt 60
tgttcttcat cttctccacc ttggtggcct tcttggccag ggccttcagc tgcatgcgca 120
cagaccgttg agctcctgat cagcatcctc aggaggccct ttgacaagca agcccctgtg 180
caagcccatt cacggggtac cagtggtgct gaggtagatg ggtttgaaaa ggattgctcg 240
gtcgattgct gctcatggaa ttggcatgtg catgcatgtt cacaatatgc caccaggctt 300
tggagcaaga gagcatgaat gccttcaggc aggttgaaag ttcctggggg tgaagaggca 360
gggccgagga ttggaggagg aaagcatcaa gtcgtcgctc atgctcatgt tttcagtcag 420
agtttgccaa gctcacagga gcagagacaa gactggctgc tcaggtgttg catcgtgtgt 480
gtggtggggg gggggggttt aatacggtac gaaatgcact tggaattccc acctcatgcc 540
agcggaccca catgcttgaa ttcgaggcct gtggggtgag aaatgctcac tctgccctcg 600
ttgctgaggt acttcaggcc gctgagctca aagtcgatgc cctgctcgtc tatcagggcc 660
tgcacctctg ggctgaccgg ctcagcctcc ttcgcgggca tggagtaggc gccggcagcg 720
ttcatgtccg ggcccagggc agcggtggtg ccataaatgt cggtgatggt ggggaggggg 780
gccgtcgcca caccattgcc gttgctggct gacgcatgca catgtggcct ggctggcacc 840
ggcagcactg gtctccagcc agccagcaag tggctgttca ggaaagcggc catgttgttg 900
gtccctgcgc atgtaattcc ccagatcaaa ggagggaaca gcttggattt gatgtagtgc 960
ccaaccggac tgaatgtgcg atggcaggtc cctttgagtc tcccgaatta ctagcagggc 1020
actgtgacct aacgcagcat gccaaccgca aaaaaatgat tgacagaaaa tgaagcggtg 1080
tgtcaatatt tgctgtattt attcgtttta atcagcaacc aagttcgaaa cgcaactatc 1140
gtggtgatca agtgaacctc atcagactta cctcgttcgg caaggaaacg gaggcaccaa 1200
attccaattt gatattatcg cttgccaagc tagagctgat ctttgggaaa ccaactgcca 1260
gacagtggac tgtgatggag tgccccgagt ggtggagcct cttcgattcg gttagtcatt 1320
actaacgtga accctcagtg aagggaccat cagaccagaa agaccagatc tcctcctcga 1380
caccgagaga gtgttgcggc agtaggacga caag 1414

酵母蔗糖转化酶

<210> 3
<211> 512

<212> PRT
<213> 未知
<220>
<223> 酵母序列
<400> 3

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

Phe Asn Thr Asp Pro Thr Tyr Gly Ser Ala Leu Gly Ile Ala Trp Ala
275 280 285
Ser Asn Trp Glu Tyr Ser Ala Phe Val Pro Thr Asn Pro Trp Arg Ser
290 295 300
Ser Met Ser Leu Val Arg Lys Phe Ser Leu Asn Thr Glu Tyr Gln Ala
305 310 315 320
Asn Pro Glu Thr Glu Leu Ile Asn Leu Lys Ala Glu Pro Ile Leu Asn
325 330 335
Ile Ser Asn Ala Gly Pro Trp Ser Arg Phe Ala Thr Asn Thr Thr Leu
340 345 350
Thr Lys Ala Asn Ser Tyr Asn Val Asp Leu Ser Asn Ser Thr Gly Thr
355 360 365
Leu Glu Phe Glu Leu Val Tyr Ala Val Asn Thr Thr Gln Thr Ile Ser
370 375 380
Lys Ser Val Phe Ala Asp Leu Ser Leu Trp Phe Lys Gly Leu Glu Asp
385 390 395 400
Pro Glu Glu Tyr Leu Arg Met Gly Phe Glu Val Ser Ala Ser Ser Phe
405 410 415
Phe Leu Asp Arg Gly Asn Ser Lys Val Lys Phe Val Lys Glu Asn Pro
420 425 430
Tyr Phe Thr Asn Arg Met Ser Val Asn Asn Gln Pro Phe Lys Ser Glu
435 440 445
Asn Asp Leu Ser Tyr Tyr Lys Val Tyr Gly Leu Leu Asp Gln Asn Ile
450 455 460
Leu Glu Leu Tyr Phe Asn Asp Gly Asp Val Val Ser Thr Asn Thr Tyr
465 470 475 480
Phe Met Thr Thr Gly Asn Ala Leu Gly Ser Val Asn Met Thr Thr Gly
485 490 495
Val Asp Asn Leu Phe Tyr Ile Asp Lys Phe Gln Val Arg Glu Val Lys
500 505 510

酵母分泌信号

<210> 4

<211> 20

<212> PRT

<213> 未知

<220>

<223> 酵母序列

<400> 4

Met Leu Leu Gln Ala Phe Leu Phe Leu Leu Ala Gly Phe Ala Ala Lys

1 5 10 15
Ile Ser Ala Ser
20
高等植物分泌信号
<210> 5
<211> 18
<212> PRT
<213> 未知
<220>
<223> 高等植物分泌信号
<400> 5
Met Ala Asn Lys Ser Leu Leu Leu Leu Leu Leu Leu Gly Ser Leu Ala
1 5 10 15
Ser Gly
共有真核分泌信号
<210> 6
<211> 10
<212> PRT
<213> 人工序列
<220>
<223> 合成的共有序列
<400> 6
Met Ala Arg Leu Pro Leu Ala Ala Leu Gly
1 5 10
高等植物/真核分泌信号组合
<210> 7
<211> 20
<212> PRT
<213> 未知
<220>
<223> 合成肽
<400> 7
Met Ala Asn Lys Leu Leu Leu Leu Leu Leu Leu Leu Leu Leu Pro Leu
1 5 10 15
Ala Ala Ser Gly
20
酿酒酵母蔗糖转化酶NP_012104
<210> 8
<211> 2615

<212> DNA

<213> 酿酒酵母

<400> 8

gaattcccca acatggtgga gcacgacact ctcgtctact ccaagaatat caaagataca 60
gtctcagaag accaaagggc tattgagact tttcaacaaa gggtaatatc gggaaacctc 120
ctcggattcc attgcccagc tatctgtcac ttcatcaaaa ggacagtaga aaaggaaggt 180
ggcacctaca aatgccatca ttgcgataaa ggaaaggcta tcgttcaaga tgcctctgcc 240
gacagtggtc ccaaagatgg acccccaccc acgaggagca tcgtggaaaa agaagacgtt 300
ccaaccacgt cttcaaagca agtggattga tgtgaacatg gtggagcacg acactctcgt 360
ctactccaag aatatcaaag atacagtctc agaagaccaa agggctattg agacttttca 420
acaaagggta atatcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat 480
caaaaggaca gtagaaaagg aaggtggcac ctacaaatgc catcattgcg ataaaggaaa 540
ggctatcgtt caagatgcct ctgccgacag tggtcccaaa gatggacccc cacccacgag 600
gagcatcgtg gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga 660
tatctccact gacgtaaggg atgacgcaca atcccactat ccttcgcaag acccttcctc 720
tatataagga agttcatttc atttggagag gacacgctga aatcaccagt ctctctctac 780
aaatctatct ctggcgcgcc atatcaatgc ttcttcaggc ctttcttttt cttcttgctg 840
gttttgctgc caagatcagc gcctctatga cgaacgaaac ctcggataga ccacttgtgc 900
actttacacc aaacaagggc tggatgaatg accccaatgg actgtggtac gacgaaaaag 960
atgccaagtg gcatctgtac tttcaataca acccgaacga tactgtctgg gggacgccat 1020
tgttttgggg ccacgccacg tccgacgacc tgaccaattg ggaggaccaa ccaatagcta 1080
tcgctccgaa gaggaacgac tccggagcat tctcgggttc catggtggtt gactacaaca 1140
atacttccgg ctttttcaac gataccattg acccgagaca acgctgcgtg gccatatgga 1200
cttacaacac accggagtcc gaggagcagt acatctcgta tagcctggac ggtggataca 1260
cttttacaga gtatcagaag aaccctgtgc ttgctgcaaa ttcgactcag ttccgagatc 1320
cgaaggtctt ttggtacgag ccctcgcaga agtggatcat gacagcggca aagtcacagg 1380
actacaagat cgaaatttac tcgtctgacg accttaaatc ctggaagctc gaatccgcgt 1440
tcgcaaacga gggctttctc ggctaccaat acgaatgccc aggcctgata gaggtcccaa 1500
cagagcaaga tcccagcaag tcctactggg tgatgtttat ttccattaat ccaggagcac 1560
cggcaggagg ttcttttaat cagtacttcg tcggaagctt taacggaact catttcgagg 1620
catttgataa ccaatcaaga gtagttgatt ttggaaagga ctactatgcc ctgcagactt 1680
tcttcaatac tgacccgacc tatgggagcg ctcttggcat tgcgtgggct tctaactggg 1740
agtattccgc attcgttcct acaaaccctt ggaggtcctc catgtcgctc gtgaggaaat 1800
tctctctcaa cactgagtac caggccaacc cggaaaccga actcataaac ctgaaagccg 1860
aaccgatcct gaacattagc aacgctggcc cctggagccg gtttgcaacc aacaccacgt 1920
tgacgaaagc caacagctac aacgtcgatc tttcgaatag caccggtaca cttgaatttg 1980
aactggtgta tgccgtcaat accacccaaa cgatctcgaa gtcggtgttc gcggacctct 2040
ccctctggtt taaaggcctg gaagaccccg aggagtacct cagaatgggt ttcgaggttt 2100
ctgcgtcctc cttcttcctt gatcgcggga acagcaaagt aaaatttgtt aaggagaacc 2160

catattttac caacaggatg agcgttaaca accaaccatt caagagcgaa aacgacctgt 2220
cgtactacaa agtgtatggt ttgcttgatc aaaatatcct ggaactctac ttcaacgatg 2280
gtgatgtcgt gtccaccaac acatacttca tgacaaccgg gaacgcactg ggctccgtga 2340
acatgacgac gggtgtggat aacctgttct acatcgacaa attccaggtg agggaagtca 2400
agtgagatct gtcgatcgac aagctcgagt ttctccataa taatgtgtga gtagttccca 2460
gataagggaa ttagggttcc tatagggttt cgctcatgtg ttgagcatat aagaaaccct 2520
tagtatgtat ttgtatttgt aaaatacttc tatcaataaa atttctaatt cctaaaacca 2580
aaatccagta ctaaaatcca gatcccccga attaa 2615

<210> 9
<211> 22
<212> DNA
<213> Prototheca moriformis
<400> 9
tgttgaagaa tgagccggcg ac 22

<210> 10
<211> 20
<212> DNA
<213> Prototheca moriformis
<400> 10
cagtgagcta ttacgcactc 20

UTEX 329 Prototheca kruegani

<210> 11
<211> 541
<212> DNA
<213> Prototheca kruegani
<400> 11
tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagaaaa tactctggag 60
ccatagcgaa agcaagttta gtaagcttag gtcattcttt ttagacccga aaccgagtga 120
tctacccatg atcagggtga agtgttagta aaataacatg gaggcccgaa ccgactaatg 180
ttgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcag tacaaataga ggggtaaagc 300
actgtttctt ttgtgggctt cgaaagttgt acctcaaagt ggcaaactct gaatactcta 360
tttagatatc tactagtgag accttggggg ataagctcct tggtcaaaag ggaaacagcc 420
cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gggtatgtca 480
aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
g 541

UTEX 1440魏氏无绿藻

<210> 12
<211> 573

<212> DNA
<213> 魏氏无绿藻
<400> 12
tgttgaagaa tgagccggcg acttaaaata aatggcaggc taagagattt aataactcga 60
aacctaagcg aaagcaagtc ttaatagggc gtcaatttaa caaaacttta aataaattat 120
aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg 180
gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt 240
tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc 300
gcagcaatat atctcgtcta tctaggggta aagcactgtt tcggtgcggg ctatgaaaat 360
ggtaccaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg 420
gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaaatg 480
ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggaggttgg cttagaagca 540
gccatccttt aaagagtgcg taatagctca ctg 573

UTEX 1442雍滞无绿藻
<210> 13
<211> 541
<212> DNA
<213> 雍滞无绿藻
<400> 13
tgttgaagaa tgagccggcg agttaaaaaa aatggcatgg ttaaagatat ttctctgaag 60
ccatagcgaa agcaagtttt acaagctata gtcatttttt ttagacccga aaccgagtga 120
tctacccatg atcagggtga agtgttggtc aaataacatg gaggcccgaa ccgactaatg 180
gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300
actgtttctt ttgtgggctt cgaaagttgt acctcaaagt ggcaaactct gaatactcta 360
tttagatatc tactagtgag accttggggg ataagctcct tggtcaaaag ggaaacagcc 420
cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggacgt gagtatgtca 480
aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
g 541

UTEX 288 Prototheca moriformis
<210> 14
<211> 541
<212> DNA
<213> Prototheca moriformis
<400> 14
tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagataa ttctctggag 60
ccatagcgaa agcaagttta acaagctaaa gtcacccttt ttagacccga aaccgagtga 120
tctacccatg atcagggtga agtgttggta aaataacatg gaggcccgaa ccgactaatg 180
gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240

ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300
actgtttctt ttgtgggctt cgaaagttgt acctcaaagt ggcaaactct gaatactcta 360
tttagatatc tactagtgag accttggggg ataagctcct tggtcaaaag ggaaacagcc 420
cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gggtatgtta 480
aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
g 541

UTEX 1439、UTEX 1441、UTEX 1435、UTEX 1437 Prototheca moriformis

<210> 15
<211> 573
<212> DNA
<213> Prototheca moriformis
<400> 15

tgttgaagaa tgagccggcg acttaaaata aatggcaggc taagagaatt aataactcga 60
aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaaacatta aataaaatct 120
aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg 180
gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt 240
tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc 300
gcagcaatat atctcgtcta tctaggggta aagcactgtt tcggtgcggg ctatgaaaat 360
ggtaccaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg 420
gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaaatg 480
ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggaggttgg cttagaagca 540
gccatccttt aaagagtgcg taatagctca ctg 573

UTEX 1533魏氏无绿藻

<210> 16
<211> 573
<212> DNA
<213> 魏氏无绿藻
<400> 16

tgttgaagaa tgagccgtcg acttaaaata aatggcaggc taagagaatt aataactcga 60
aacctaagcg aaagcaagtc ttaatagggc gctaatttaa caaaacatta aataaaatct 120
aaagtcattt attttagacc cgaacctgag tgatctaacc atggtcagga tgaaacttgg 180
gtgacaccaa gtggaagtcc gaaccgaccg atgttgaaaa atcggcggat gaactgtggt 240
tagtggtgaa ataccagtcg aactcagagc tagctggttc tccccgaaat gcgttgaggc 300
gcagcaatat atctcgtcta tctaggggta aagcactgtt tcggtgcggg ctatgaaaat 360
ggtaccaaat cgtggcaaac tctgaatact agaaatgacg atatattagt gagactatgg 420
gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaaatg 480
ataatgaagt ggtaaaggag gtgaaaatgc aaatacaacc aggaggttgg cttagaagca 540
gccatccttt aaagagtgcg taatagctca ctg 573

UTEX 1434 Prototheca moriformis

<210> 17
<211> 541
<212> DNA
<213> Prototheca moriformis
<400> 17
tgttgaagaa tgagccggcg agttaaaaag agtggcgtgg ttaaagaaaa ttctctggaa 60
ccatagcgaa agcaagttta acaagcttaa gtcacttttt ttagacccga aaccgagtga 120
tctacccatg atcagggtga agtgttggta aaataacatg gaggcccgaa ccgactaatg 180
gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300
actgtttctt ttgtgggctc cgaaagttgt acctcaaagt ggcaaactct gaatactcta 360
tttagatatc tactagtgag accttggggg ataagctcct tggtcgaaag ggaaacagcc 420
cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gagtatgtca 480
aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
g 541
UTEX 1438祖菲无绿藻
<210> 18
<211> 541
<212> DNA
<213> 祖菲无绿藻
<400> 18
tgttgaagaa tgagccggcg agttaaaaag agtggcatgg ttaaagaaaa ttctctggag 60
ccatagcgaa agcaagttta acaagcttaa gtcacttttt ttagacccga aaccgagtga 120
tctacccatg atcagggtga agtgttggta aaataacatg gaggcccgaa ccgactaatg 180
gtgaaaaatt agcggatgaa ttgtgggtag gggcgaaaaa ccaatcgaac tcggagttag 240
ctggttctcc ccgaaatgcg tttaggcgca gcagtagcaa cacaaataga ggggtaaagc 300
actgtttctt tcgtgggctt cgaaagttgt acctcaaagt ggcaaactct gaatactcta 360
tttagatatc tactagtgag accttggggg ataagctcct tggtcaaaag ggaaacagcc 420
cagatcacca gttaaggccc caaaatgaaa atgatagtga ctaaggatgt gagtatgtca 480
aaacctccag caggttagct tagaagcagc aatcctttca agagtgcgta atagctcact 540
g 541
UTEX 1436 Prototheca moriformis
<210> 19
<211> 565
<212> DNA
<213> Prototheca moriformis
<400> 19
tgttgaagaa tgagccggcg acttagaaaa ggtggcatgg ttaaggaaat attccgaagc 60
cgtagcaaaa gcgagtctga atagggcgat aaaatatatt aatatttaga atctagtcat 120

ttttttctaga cccgaacccg ggtgatctaa ccatgaccag gatgaagctt gggtgatacc 180
aagtgaaggt ccgaaccgac cgatgttgaa aaatcggcgg atgagttgtg gttagcggtg 240
aaataccagt cgaacccgga gctagctggt tctccccgaa atgcgttgag gcgcagcagt 300
acatctagtc tatctagggg taaagcactg tttcggtgcg ggctgtgaga acggtaccaa 360
atcgtggcaa actctgaata ctagaaatga cgatgtagta gtgagactgt gggggataag 420
ctccattgtc aagagggaaa cagcccagac caccagctaa ggccccaaaa tggtaatgta 480
gtgacaaagg aggtgaaaat gcaaatacaa ccaggaggtt ggcttagaag cagccatcct 540
ttaaagagtg cgtaatagct cactg 565

菊苣转化酶:Genbank登录号Y11124

<210> 20
<211> 550
<212> PRT
<213> 菊苣
<400> 20

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

195 200 205

Gly Thr Gly Met Trp Glu Cys Val Asp Phe Phe Pro Val Trp Val Asp

210 215 220

Ser Thr Asn Gly Val Asp Thr Ser Ile Ile Asn Pro Ser Asn Arg Val

225 230 235 240

Lys His Val Leu Lys Leu Gly Ile Gln Asp His Gly Lys Asp Cys Tyr

245 250 255

Leu Ile Gly Lys Tyr Ser Ala Asp Lys Glu Asn Tyr Val Pro Glu Asp

260 265 270

Glu Leu Thr Leu Ser Thr Leu Arg Leu Asp Tyr Gly Met Tyr Tyr Ala

275 280 285

Ser Lys Ser Phe Phe Asp Pro Val Lys Asn Arg Arg Ile Met Thr Ala

290 295 300

Trp Val Asn Glu Ser Asp Ser Glu Ala Asp Val Ile Ala Arg Gly Trp

305 310 315 320

Ser Gly Val Gln Ser Phe Pro Arg Ser Leu Trp Leu Asp Lys Asn Gln

325 330 335

Lys Gln Leu Leu Gln Trp Pro Ile Glu Glu Ile Glu Met Leu His Gln

340 345 350

Asn Glu Val Ser Phe His Asn Lys Lys Leu Asp Gly Gly Ser Ser Leu

355 360 365

Glu Val Leu Gly Ile Thr Ala Ser Gln Ala Asp Val Lys Ile Ser Phe

370 375 380

Lys Leu Ala Asn Leu Glu Glu Ala Glu Glu Leu Asp Pro Ser Trp Val

385 390 395 400

Asp Pro Gln Leu Ile Cys Ser Glu Asn Asp Ala Ser Lys Lys Gly Lys

405 410 415

Phe Gly Pro Phe Gly Leu Leu Ala Leu Ala Ser Ser Asp Leu Arg Glu

420 425 430

Gln Thr Ala Ile Phe Phe Arg Val Phe Arg Lys Asn Gly Arg Tyr Val

435 440 445

Val Leu Met Cys Ser Asp Gln Ser Arg Ser Ser Met Lys Asn Gly Ile

450 455 460

Glu Lys Arg Thr Tyr Gly Ala Phe Val Asp Ile Asp Pro Gln Gln Asp

465 470 475 480

Glu Ile Ser Leu Arg Thr Leu Ile Asp His Ser Ile Val Glu Ser Phe

485 490 495

Gly Gly Arg Gly Lys Thr Cys Ile Thr Thr Arg Val Tyr Pro Thr Leu

500 505 510

Ala Ile Gly Glu Gln Ala Arg Leu Phe Ala Phe Asn His Gly Thr Glu
515 520 525
Ser Val Glu Ile Ser Glu Leu Ser Ala Trp Ser Met Lys Lys Ala Gln
530 535 540
Met Lys Val Glu Glu Pro
545 550

粟酒裂殖酵母转化酶:Genbank登录号AB011433

<210> 21
<211> 581
<212> PRT
<213> 粟酒裂殖酵母
<400> 21

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

210 215 220

Gln Phe Arg Asp Pro Lys Val Ile Trp Asp Phe Asp Ala Asn Arg Trp
225 230 235 240

Val Met Ile Val Ala Met Ser Gln Asn Tyr Gly Ile Ala Phe Tyr Ser
245 250 255

Ser Tyr Asp Leu Ile His Trp Thr Glu Leu Ser Val Phe Ser Thr Ser
260 265 270

Gly Tyr Leu Gly Leu Gln Tyr Glu Cys Pro Gly Met Ala Arg Val Pro
275 280 285

Val Glu Gly Thr Asp Glu Tyr Lys Trp Val Leu Phe Ile Ser Ile Asn
290 295 300

Pro Gly Ala Pro Leu Gly Gly Ser Val Val Gln Tyr Phe Val Gly Asp
305 310 315 320

Trp Asn Gly Thr Asn Phe Val Pro Asp Asp Gly Gln Thr Arg Phe Val
325 330 335

Asp Leu Gly Lys Asp Phe Tyr Ala Ser Ala Leu Tyr His Ser Ser Ser
340 345 350

Ala Asn Ala Asp Val Ile Gly Val Gly Trp Ala Ser Asn Trp Gln Tyr
355 360 365

Thr Asn Gln Ala Pro Thr Gln Val Phe Arg Ser Ala Met Thr Val Ala
370 375 380

Arg Lys Phe Thr Leu Arg Asp Val Pro Gln Asn Pro Met Thr Asn Leu
385 390 395 400

Thr Ser Leu Ile Gln Thr Pro Leu Asn Val Ser Leu Leu Arg Asp Glu
405 410 415

Thr Leu Phe Thr Ala Pro Val Ile Asn Ser Ser Ser Ser Leu Ser Gly
420 425 430

Ser Pro Ile Thr Leu Pro Ser Asn Thr Ala Phe Glu Phe Asn Val Thr
435 440 445

Leu Ser Ile Asn Tyr Thr Glu Gly Cys Thr Thr Gly Tyr Cys Leu Gly
450 455 460

Arg Ile Ile Ile Asp Ser Asp Asp Pro Tyr Arg Leu Gln Ser Ile Ser
465 470 475 480

Val Asp Val Asp Phe Ala Ala Ser Thr Leu Val Ile Asn Arg Ala Lys
485 490 495

Ala Gln Met Gly Trp Phe Asn Ser Leu Phe Thr Pro Ser Phe Ala Asn
500 505 510

Asp Ile Tyr Ile Tyr Gly Asn Val Thr Leu Tyr Gly Ile Val Asp Asn
515 520 525

Gly Leu Leu Glu Leu Tyr Val Asn Asn Gly Glu Lys Thr Tyr Thr Asn
530 535 540
Asp Phe Phe Phe Leu Gln Gly Ala Thr Pro Gly Gln Ile Ser Phe Ala
545 550 555 560
Ala Phe Gln Gly Val Ser Phe Asn Asn Val Thr Val Thr Pro Leu Lys
565 570 575
Thr Ile Trp Asn Cys
580

异常毕赤酵母β-呋喃果糖苷酶(转化酶):Genbank登录号X80640

<210> 22
<211> 550
<212> PRT
<213> 异常毕赤酵母
<400> 22

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

195 200 205

Lys Ser Asn Phe Thr Ser Asn Gly Tyr Leu Gly Phe Gln Tyr Glu Cys

210 215 220

Pro Gly Leu Phe Lys Leu Pro Ile Glu Asn Pro Leu Asn Asp Thr Val

225 230 235 240

Thr Ser Lys Trp Val Leu Leu Leu Ala Ile Asn Pro Gly Ser Pro Leu

245 250 255

Gly Gly Ser Ile Asn Glu Tyr Phe Ile Gly Asp Phe Asp Gly Thr Thr

260 265 270

Phe His Pro Asp Asp Gly Ala Thr Arg Phe Met Asp Ile Gly Lys Asp

275 280 285

Phe Tyr Ala Phe Gln Ser Phe Asp Asn Thr Glu Pro Glu Asp Gly Ala

290 295 300

Leu Gly Leu Ala Trp Ala Ser Asn Trp Gln Tyr Ala Asn Thr Val Pro

305 310 315 320

Thr Glu Asn Trp Arg Ser Ser Met Ser Leu Val Arg Asn Tyr Thr Leu

325 330 335

Lys Tyr Val Asp Val Asn Pro Glu Asn Tyr Gly Leu Thr Leu Ile Gln

340 345 350

Lys Pro Val Tyr Asp Thr Lys Glu Thr Arg Leu Asn Glu Thr Leu Lys

355 360 365

Thr Leu Glu Thr Ile Asn Glu Tyr Glu Val Asn Asp Leu Lys Leu Asp

370 375 380

Lys Ser Ser Phe Val Ala Thr Asp Phe Asn Thr Glu Arg Asn Ala Thr

385 390 395 400

Gly Val Phe Glu Phe Asp Leu Lys Phe Thr Gln Thr Asp Leu Lys Met

405 410 415

Gly Tyr Ser Asn Met Thr Thr Gln Phe Gly Leu Tyr Ile His Ser Gln

420 425 430

Thr Val Lys Gly Ser Gln Glu Thr Leu Gln Leu Val Phe Asp Thr Leu

435 440 445

Ser Thr Thr Trp Tyr Ile Asp Arg Thr Thr Gln His Ser Phe Gln Arg

450 455 460

Asn Ser Pro Val Phe Thr Glu Arg Ile Ser Thr Tyr Val Glu Lys Ile

465 470 475 480

Asp Thr Thr Asp Gln Gly Asn Val Tyr Thr Leu Tyr Gly Val Val Asp

485 490 495

Arg Asn Ile Leu Glu Leu Tyr Phe Asn Asp Gly Ser Ile Ala Met Thr

500 505 510

Asn Thr Phe Phe Phe Arg Glu Gly Lys Ile Pro Thr Ser Phe Glu Val
515 520 525
Val Cys Asp Ser Glu Lys Ser Phe Ile Thr Ile Asp Glu Leu Ser Val
530 535 540
Arg Glu Leu Ala Arg Lys
545 550

西洋德巴利酵母转化酶:Genbank登录号X17604

<210> 23
<211> 533
<212> PRT
<213> 西洋德巴利酵母
<400> 23

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

210 215 220

Tyr Gly Asn Gln Tyr Gly Met Ser Arg Leu Ile Glu Val Pro Ile Glu
225 230 235 240

Asn Ser Asp Lys Ser Lys Trp Val Met Phe Leu Ala Ile Asn Pro Gly
245 250 255

Ser Pro Leu Gly Gly Ser Ile Asn Gln Tyr Phe Val Gly Asp Phe Asp
260 265 270

Gly Phe Gln Phe Val Pro Asp Asp Ser Gln Thr Arg Phe Val Asp Ile
275 280 285

Gly Lys Asp Phe Tyr Ala Phe Gln Thr Phe Ser Glu Val Glu His Gly
290 295 300

Val Leu Gly Leu Ala Trp Ala Ser Asn Trp Gln Tyr Ala Asp Gln Val
305 310 315 320

Pro Thr Asn Pro Trp Arg Ser Ser Thr Ser Leu Ala Arg Asn Tyr Thr
325 330 335

Leu Arg Tyr Val Ile Gln Met Leu Lys Leu Thr Ala Asn Ile Asp Lys
340 345 350

Ser Val Leu Pro Asp Ser Ile Asn Val Val Asp Lys Leu Lys Lys Lys
355 360 365

Asn Val Lys Leu Thr Asn Lys Lys Pro Ile Lys Thr Asn Phe Lys Gly
370 375 380

Ser Thr Gly Leu Phe Asp Phe Asn Ile Thr Phe Lys Val Leu Asn Leu
385 390 395 400

Asn Val Ser Pro Gly Lys Thr His Phe Asp Ile Leu Ile Asn Ser Gln
405 410 415

Glu Leu Asn Ser Ser Val Asp Ser Ile Lys Ile Gly Phe Asp Ser Ser
420 425 430

Gln Ser Leu Phe Tyr Ile Asp Arg His Ile Pro Asn Val Glu Phe Pro
435 440 445

Arg Lys Gln Phe Phe Thr Asp Lys Leu Ala Ala Tyr Leu Glu Pro Leu
450 455 460

Asp Tyr Asp Gln Asp Leu Arg Val Phe Ser Leu Tyr Gly Ile Val Asp
465 470 475 480

Lys Asn Ile Ile Glu Leu Tyr Phe Asn Asp Gly Thr Val Ala Met Thr
485 490 495

Asn Thr Phe Phe Met Gly Glu Gly Lys Tyr Pro His Asp Ile Gln Ile
500 505 510

Val Thr Asp Thr Glu Glu Pro Leu Phe Glu Leu Glu Ser Val Ile Ile
515 520 525

Arg Glu Leu Asn Lys
530

水稻转化酶:Genbank登录号AF019113

<210> 24

<211> 654

<212> PRT

<213> 水稻

<400> 24

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

245 250 255
Asp Pro Thr Thr Ala Gly Thr Ala Gly Met Gln Asn Gly Gln Arg Leu
260 265 270
Val Thr Ile Gly Ser Lys Val Gly Lys Thr Gly Ile Ser Leu Val Tyr
275 280 285
Glu Thr Thr Asn Phe Thr Thr Phe Lys Leu Leu Tyr Gly Val Leu His
290 295 300
Ala Val Pro Gly Thr Gly Met Trp Glu Cys Val Asp Leu Tyr Pro Val
305 310 315 320
Ser Thr Thr Gly Glu Asn Gly Leu Asp Thr Ser Val Asn Gly Leu Gly
325 330 335
Val Lys His Val Leu Lys Thr Ser Leu Asp Asp Asp Lys His Asp Tyr
340 345 350
Tyr Ala Leu Gly Thr Tyr Asp Pro Val Lys Asn Lys Trp Thr Pro Asp
355 360 365
Asn Pro Asp Leu Asp Val Gly Ile Gly Leu Arg Leu Asp Tyr Gly Lys
370 375 380
Tyr Tyr Ala Ala Arg Thr Phe Tyr Asp Gln Asn Lys Gln Arg Arg Ile
385 390 395 400
Leu Trp Gly Trp Ile Gly Glu Thr Asp Leu Glu Ala Val Asp Leu Met
405 410 415
Lys Gly Trp Ala Ser Leu Gln Ala Ile Pro Arg Thr Ile Val Phe Asp
420 425 430
Lys Lys Thr Gly Thr Asn Val Leu Gln Arg Pro Glu Glu Glu Val Glu
435 440 445
Ser Trp Ser Ser Gly Asp Pro Ile Thr Gln Arg Arg Ile Phe Glu Pro
450 455 460
Gly Ser Val Val Pro Ile His Val Ser Gly Ala Thr Gln Leu Asp Ile
465 470 475 480
Thr Ala Ser Phe Glu Val Asp Glu Thr Leu Leu Glu Thr Thr Ser Glu
485 490 495
Ser His Asp Ala Gly Tyr Asp Cys Ser Asn Ser Gly Gly Ala Gly Thr
500 505 510
Arg Gly Ser Leu Gly Pro Phe Gly Leu Leu Val Val Ala Asp Glu Lys
515 520 525
Leu Ser Glu Leu Thr Pro Val Tyr Leu Tyr Val Ala Lys Gly Gly Asp
530 535 540
Gly Lys Ala Lys Ala His Leu Cys Ala Tyr Gln Thr Arg Ser Ser Met
545 550 555 560

Ala Ser Gly Val Glu Lys Glu Val Tyr Gly Ser Ala Val Pro Val Leu
565 570 575
Asp Gly Glu Asn Tyr Ser Ala Arg Ile Leu Ile Asp His Ser Ile Val
580 585 590
Glu Ser Phe Ala Gln Ala Gly Arg Thr Cys Val Arg Ser Arg Asp Tyr
595 600 605
Pro Thr Lys Asp Ile Tyr Gly Ala Ala Arg Cys Phe Phe Phe Asn Asn
610 615 620
Ala Thr Glu Ala Ser Val Arg Ala Ser Leu Lys Ala Trp Gln Met Lys
625 630 635 640
Ser Phe Ile Arg Pro Tyr Pro Phe Ile Pro Asp Gln Lys Ser
645 650

洋葱转化酶:Genbank登录号AJ006067

<210> 25
<211> 690
<212> PRT
<213> 洋葱
<400> 25

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

165 170 175

Phe Phe Tyr Gln Tyr Asn Pro Glu Gly Ala Val Trp Gly Asn Ile Ala
180 185 190

Trp Gly His Ala Val Ser Arg Asp Leu Val His Trp Thr His Leu Pro
195 200 205

Leu Ala Met Val Pro Asp Gln Trp Tyr Asp Ile Asn Gly Val Trp Thr
210 215 220

Gly Ser Ala Thr Ile Leu Pro Asp Gly Gln Ile Val Met Leu Tyr Thr
225 230 235 240

Gly Ala Thr Asn Glu Ser Val Gln Val Gln Asn Leu Ala Val Pro Ala
245 250 255

Asp Gln Ser Asp Thr Leu Leu Leu Arg Trp Lys Lys Ser Glu Ala Asn
260 265 270

Pro Ile Leu Val Pro Pro Pro Gly Ile Gly Asp Lys Asp Phe Arg Asp
275 280 285

Pro Thr Thr Ala Trp Tyr Glu Pro Ser Asp Asp Thr Trp Arg Ile Val
290 295 300

Ile Gly Ser Lys Asp Ser Ser His Ser Gly Ile Ala Ile Val Tyr Ser
305 310 315 320

Thr Lys Asp Phe Ile Asn Tyr Lys Leu Ile Pro Gly Ile Leu His Ala
325 330 335

Val Glu Arg Val Gly Met Trp Glu Cys Val Asp Phe Tyr Pro Val Ala
340 345 350

Thr Ala Asp Ser Ser His Ala Asn His Gly Leu Asp Pro Ser Ala Arg
355 360 365

Pro Ser Pro Ala Val Lys His Val Leu Lys Ala Ser Met Asp Asp Asp
370 375 380

Arg His Asp Tyr Tyr Ala Ile Gly Thr Tyr Asp Pro Ala Gln Asn Thr
385 390 395 400

Trp Val Pro Asp Asp Ala Ser Val Asp Val Gly Ile Gly Leu Arg Tyr
405 410 415

Asp Trp Gly Lys Phe Tyr Ala Ser Lys Thr Phe Tyr Asp His Ala Lys
420 425 430

Lys Arg Arg Ile Leu Trp Ser Trp Ile Gly Glu Thr Asp Ser Glu Thr
435 440 445

Ala Asp Ile Ala Lys Gly Trp Ala Ser Leu Gln Gly Val Pro Arg Thr
450 455 460

Val Leu Leu Asp Val Lys Thr Gly Ser Asn Leu Ile Thr Trp Pro Val
465 470 475 480

Val Glu Ile Glu Ser Leu Arg Thr Arg Pro Arg Asp Phe Ser Gly Ile
485 490 495
Thr Val Asp Ala Gly Ser Thr Phe Lys Leu Asp Val Gly Gly Ala Ala
500 505 510
Gln Leu Asp Ile Glu Ala Glu Phe Lys Ile Ser Ser Glu Glu Leu Glu
515 520 525
Ala Val Lys Glu Ala Asp Val Ser Tyr Asn Cys Ser Ser Ser Gly Gly
530 535 540
Ala Ala Glu Arg Gly Val Leu Gly Pro Phe Gly Leu Leu Val Leu Ala
545 550 555 560
Asn Gln Asp Leu Thr Glu Gln Thr Ala Thr Tyr Phe Tyr Val Ser Arg
565 570 575
Gly Met Asp Gly Gly Leu Asn Thr His Phe Cys Gln Asp Glu Lys Arg
580 585 590
Ser Ser Lys Ala Ser Asp Ile Val Lys Arg Ile Val Gly His Ser Val
595 600 605
Pro Val Leu Asp Gly Glu Ser Phe Ala Leu Arg Ile Leu Val Asp His
610 615 620
Ser Ile Val Glu Ser Phe Ala Gln Gly Gly Arg Ala Ser Ala Thr Ser
625 630 635 640
Arg Val Tyr Pro Thr Glu Ala Ile Tyr Asn Asn Ala Arg Val Phe Val
645 650 655
Phe Asn Asn Ala Thr Gly Ala Lys Val Thr Ala Gln Ser Leu Lys Val
660 665 670
Trp His Met Ser Thr Ala Ile Asn Glu Ile Tyr Asp Pro Ala Thr Ser
675 680 685
Val Met
690

栽培甜菜转化酶:Genbank登录号AJ278531

<210> 26
<211> 501
<212> PRT
<213> 甜菜
<220>
<221> misc_特征
<222> (382)..(382)
<223> Xaa可以是任何自然发生的氨基酸
<400> 26
Leu Phe Tyr Gln Tyr Asn Pro Asn Gly Val Ile Trp Gly Pro Pro Val

1 5 10 15
Trp Gly His Ser Thr Ser Lys Asp Leu Val Asn Trp Val Pro Gln Pro
20 25 30
Leu Thr Met Glu Pro Glu Met Ala Ala Asn Ile Asn Gly Ser Trp Ser
35 40 45
Gly Ser Ala Thr Ile Leu Pro Gly Asn Lys Pro Ala Ile Leu Phe Thr
50 55 60
Gly Leu Asp Pro Lys Tyr Glu Gln Val Gln Val Leu Ala Tyr Pro Lys
65 70 75 80
Asp Thr Ser Asp Pro Asn Leu Lys Glu Trp Phe Leu Ala Pro Gln Asn
85 90 95
Pro Val Met Phe Pro Thr Pro Gln Asn Gln Ile Asn Ala Thr Ser Phe
100 105 110
Arg Asp Pro Thr Thr Ala Trp Arg Leu Pro Asp Gly Val Trp Arg Leu
115 120 125
Leu Ile Gly Ser Lys Arg Gly Gln Arg Gly Leu Ser Leu Leu Phe Arg
130 135 140
Ser Arg Asp Phe Val His Trp Val Gln Ala Lys His Pro Leu Tyr Ser
145 150 155 160
Asp Lys Leu Ser Gly Met Trp Glu Cys Pro Asp Phe Phe Pro Val Tyr
165 170 175
Ala Asn Gly Asp Gln Met Gly Val Asp Thr Ser Ile Ile Gly Ser His
180 185 190
Val Lys His Val Leu Lys Asn Ser Leu Asp Ile Thr Lys His Asp Ile
195 200 205
Tyr Thr Ile Gly Asp Tyr Asn Ile Lys Lys Asp Ala Tyr Thr Pro Asp
210 215 220
Ile Gly Tyr Met Asn Asp Ser Ser Leu Arg Tyr Asp Tyr Gly Lys Tyr
225 230 235 240
Tyr Ala Ser Lys Thr Phe Phe Asp Asp Ala Lys Lys Glu Arg Ile Leu
245 250 255
Leu Gly Trp Ala Asn Glu Ser Ser Ser Val Glu Asp Asp Ile Lys Lys
260 265 270
Gly Trp Ser Gly Ile His Thr Ile Pro Arg Lys Ile Trp Leu Asp Lys
275 280 285
Leu Gly Lys Gln Leu Ile Gln Trp Pro Ile Ala Asn Ile Glu Lys Leu
290 295 300
Arg Gln Lys Pro Val Asn Ile Tyr Arg Lys Val Leu Lys Gly Gly Ser
305 310 315 320

Gln Ile Glu Val Ser Gly Ile Thr Ala Ala Gln Ala Asp Val Glu Ile
325 330 335
Ser Phe Lys Ile Lys Asp Leu Lys Asn Val Glu Lys Phe Asp Ala Ser
340 345 350
Trp Thr Ser Pro Gln Leu Leu Cys Ser Lys Lys Gly Ala Ser Val Lys
355 360 365
Gly Gly Leu Gly Pro Phe Gly Leu Leu Thr Leu Ala Ser Xaa Gly Leu
370 375 380
Glu Glu Tyr Thr Ala Val Phe Phe Arg Ile Phe Lys Ala Tyr Asp Asn
385 390 395 400
Lys Phe Val Val Leu Met Cys Ser Asp Gln Ser Arg Ser Ser Leu Asn
405 410 415
Pro Thr Asn Asp Lys Thr Thr Tyr Gly Thr Phe Val Asp Val Asn Pro
420 425 430
Ile Arg Glu Gly Leu Ser Leu Arg Val Leu Ile Asp His Ser Val Val
435 440 445
Glu Ser Phe Gly Ala Lys Gly Lys Asn Val Ile Thr Ala Arg Val Tyr
450 455 460
Pro Thr Leu Ala Ile Asn Glu Lys Ala His Leu Tyr Val Phe Asn Arg
465 470 475 480
Gly Thr Ser Asn Val Glu Ile Thr Gly Leu Thr Ala Trp Ser Met Lys
485 490 495
Lys Ala Asn Ile Ala
500

短双歧杆菌UCC2003β-呋喃果糖苷酶(转化酶):Genbank登录号AAT28190

<210> 27
<211> 518
<212> PRT
<213> 短双歧杆菌
<400> 27

Met Thr Asp Phe Thr Pro Glu Thr Pro Val Leu Thr Pro Ile Arg Asp
1 5 10 15
His Ala Ala Glu Leu Ala Lys Ala Glu Ala Gly Val Ala Glu Met Ala
20 25 30
Ala Lys Arg Asn Asn Arg Trp Tyr Pro Lys Tyr His Ile Ala Ser Asn
35 40 45
Gly Gly Trp Ile Asn Asp Pro Asn Gly Leu Cys Phe Tyr Lys Gly Arg
50 55 60
Trp His Val Phe Tyr Gln Leu His Pro Tyr Gly Thr Gln Trp Gly Pro

65 70 75 80
Met His Trp Gly His Val Ser Ser Thr Asp Met Leu Asn Trp Lys Arg
85 90 95
Glu Pro Ile Met Phe Ala Pro Ser Leu Glu Gln Glu Lys Asp Gly Val
100 105 110
Phe Ser Gly Ser Ala Val Ile Asp Asp Asn Gly Asp Leu Arg Phe Tyr
115 120 125
Tyr Thr Gly His Arg Trp Ala Asn Gly His Asp Asn Thr Gly Gly Asp
130 135 140
Trp Gln Val Gln Met Thr Ala Leu Pro Asp Asn Asp Glu Leu Thr Ser
145 150 155 160
Ala Thr Lys Gln Gly Met Ile Ile Asp Cys Pro Thr Asp Lys Val Asp
165 170 175
His His Tyr Arg Asp Pro Lys Val Trp Lys Thr Gly Asp Thr Trp Tyr
180 185 190
Met Thr Phe Gly Val Ser Ser Glu Asp Lys Arg Gly Gln Met Trp Leu
195 200 205
Phe Ser Ser Lys Asp Met Val Arg Trp Glu Tyr Glu Arg Val Leu Phe
210 215 220
Gln His Pro Asp Pro Asp Val Phe Met Leu Glu Cys Pro Asp Phe Phe
225 230 235 240
Pro Ile Lys Asp Lys Asp Gly Asn Glu Lys Trp Val Ile Gly Phe Ser
245 250 255
Ala Met Gly Ser Lys Pro Ser Gly Phe Met Asn Arg Asn Val Asn Asn
260 265 270
Ala Gly Tyr Met Ile Gly Thr Trp Glu Pro Gly Gly Glu Phe Lys Pro
275 280 285
Glu Thr Glu Phe Arg Leu Trp Asp Cys Gly His Asn Tyr Tyr Ala Pro
290 295 300
Gln Ser Phe Asn Val Asp Gly Arg Gln Ile Val Tyr Gly Trp Met Ser
305 310 315 320
Pro Phe Val Gln Pro Ile Pro Met Glu Asp Asp Gly Trp Cys Gly Gln
325 330 335
Leu Thr Leu Pro Arg Glu Ile Thr Leu Asp Asp Asp Gly Asp Val Val
340 345 350
Thr Ala Pro Val Ala Glu Met Glu Gly Leu Arg Glu Asp Thr Leu Asp
355 360 365
His Gly Ser Ile Thr Leu Asp Met Asp Gly Glu Gln Val Ile Ala Asp
370 375 380

Asp Ala Glu Ala Val Glu Ile Glu Met Thr Ile Asp Leu Ala Ala Ser
385 390 395 400
Thr Ala Asp Arg Ala Gly Leu Lys Ile His Ala Thr Glu Asp Gly Ala
405 410 415
Tyr Thr Tyr Val Ala Tyr Asp Asp Gln Ile Gly Arg Val Val Val Asp
420 425 430
Arg Gln Ala Met Ala Asn Gly Asp His Gly Tyr Arg Ala Ala Pro Leu
435 440 445
Thr Asp Ala Glu Leu Ala Ser Gly Lys Leu Asp Leu Arg Val Phe Val
450 455 460
Asp Arg Gly Ser Val Glu Val Tyr Val Asn Gly Gly His Gln Val Leu
465 470 475 480
Ser Ser Tyr Ser Tyr Ala Ser Glu Gly Pro Arg Ala Ile Lys Leu Val
485 490 495
Ala Glu Phe Gly Asn Leu Lys Val Glu Ser Leu Lys Leu His His Met
500 505 510
Lys Ser Ile Gly Leu Glu
515

酿酒酵母转化酶:Genbank登录号NP_012104

<210> 28
<211> 532
<212> PRT
<213> 酿酒酵母
<400> 28

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

115 120 125

Ala Ile Trp Thr Tyr Asn Thr Pro Glu Ser Glu Glu Gln Tyr Ile Ser
130 135 140

Tyr Ser Leu Asp Gly Gly Tyr Thr Phe Thr Glu Tyr Gln Lys Asn Pro
145 150 155 160

Val Leu Ala Ala Asn Ser Thr Gln Phe Arg Asp Pro Lys Val Phe Trp
165 170 175

Tyr Glu Pro Ser Gln Lys Trp Ile Met Thr Ala Ala Lys Ser Gln Asp
180 185 190

Tyr Lys Ile Glu Ile Tyr Ser Ser Asp Asp Leu Lys Ser Trp Lys Leu
195 200 205

Glu Ser Ala Phe Ala Asn Glu Gly Phe Leu Gly Tyr Gln Tyr Glu Cys
210 215 220

Pro Gly Leu Ile Glu Val Pro Thr Glu Gln Asp Pro Ser Lys Ser Tyr
225 230 235 240

Trp Val Met Phe Ile Ser Ile Asn Pro Gly Ala Pro Ala Gly Gly Ser
245 250 255

Phe Asn Gln Tyr Phe Val Gly Ser Phe Asn Gly Thr His Phe Glu Ala
260 265 270

Phe Asp Asn Gln Ser Arg Val Val Asp Phe Gly Lys Asp Tyr Tyr Ala
275 280 285

Leu Gln Thr Phe Phe Asn Thr Asp Pro Thr Tyr Gly Ser Ala Leu Gly
290 295 300

Ile Ala Trp Ala Ser Asn Trp Glu Tyr Ser Ala Phe Val Pro Thr Asn
305 310 315 320

Pro Trp Arg Ser Ser Met Ser Leu Val Arg Lys Phe Ser Leu Asn Thr
325 330 335

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

Pro Ile Leu Asn Ile Ser Asn Ala Gly Pro Trp Ser Arg Phe Ala Thr
355 360 365

Asn Thr Thr Leu Thr Lys Ala Asn Ser Tyr Asn Val Asp Leu Ser Asn
370 375 380

Ser Thr Gly Thr Leu Glu Phe Glu Leu Val Tyr Ala Val Asn Thr Thr
385 390 395 400

Gln Thr Ile Ser Lys Ser Val Phe Ala Asp Leu Ser Leu Trp Phe Lys
405 410 415

Gly Leu Glu Asp Pro Glu Glu Tyr Leu Arg Met Gly Phe Glu Val Ser
420 425 430

Ala Ser Ser Phe Phe Leu Asp Arg Gly Asn Ser Lys Val Lys Phe Val
435 440 445
Lys Glu Asn Pro Tyr Phe Thr Asn Arg Met Ser Val Asn Asn Gln Pro
450 455 460
Phe Lys Ser Glu Asn Asp Leu Ser Tyr Tyr Lys Val Tyr Gly Leu Leu
465 470 475 480
Asp Gln Asn Ile Leu Glu Leu Tyr Phe Asn Asp Gly Asp Val Val Ser
485 490 495
Thr Asn Thr Tyr Phe Met Thr Thr Gly Asn Ala Leu Gly Ser Val Asn
500 505 510
Met Thr Thr Gly Val Asp Asn Leu Phe Tyr Ile Asp Lys Phe Gln Val
515 520 525
Arg Glu Val Lys
530

运动发酵单胞菌转化酶A:Genbank登录号AY171597

<210> 29
<211> 512
<212> PRT
<213> 运动发酵单胞菌
<400> 29

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

145 150 155 160
Val Ala His Phe Arg Asp Pro Arg Val Trp Lys Glu Asn Asp His Trp
165 170 175
Phe Met Val Val Gly Tyr Arg Thr Asp Asp Glu Lys His Gln Gly Ile
180 185 190
Gly His Val Ala Leu Tyr Arg Ser Glu Asn Leu Lys Asp Trp Ile Phe
195 200 205
Val Lys Thr Leu Leu Gly Asp Asn Ser Gln Leu Pro Leu Gly Lys Arg
210 215 220
Ala Phe Met Trp Glu Cys Pro Asp Phe Phe Ser Leu Gly Asn Arg Ser
225 230 235 240
Val Leu Met Phe Ser Pro Gln Gly Leu Lys Ala Ser Gly Tyr Lys Asn
245 250 255
Arg Asn Leu Phe Gln Asn Gly Tyr Ile Leu Gly Lys Trp Gln Ala Pro
260 265 270
Gln Phe Thr Pro Glu Thr Ser Phe Gln Glu Leu Asp Tyr Gly His Asp
275 280 285
Phe Tyr Ala Ala Gln Arg Phe Glu Ala Lys Asp Gly Arg Gln Ile Leu
290 295 300
Ile Ala Trp Phe Asp Met Trp Glu Asn Gln Lys Pro Ser Gln Arg Asp
305 310 315 320
Gly Trp Ala Gly Cys Met Thr Leu Pro Arg Lys Leu Asp Leu Ile Asp
325 330 335
Asn Lys Ile Val Met Thr Pro Val Arg Glu Met Glu Ile Leu Arg Gln
340 345 350
Ser Glu Lys Ile Glu Ser Val Val Thr Leu Ser Asp Ala Glu His Pro
355 360 365
Phe Thr Met Asp Ser Pro Leu Gln Glu Ile Glu Leu Ile Phe Asp Leu
370 375 380
Glu Lys Ser Ser Ala Tyr Gln Ala Gly Leu Ala Leu Arg Cys Asn Gly
385 390 395 400
Lys Gly Gln Glu Thr Leu Leu Tyr Ile Asp Arg Ser Gln Asn Arg Ile
405 410 415
Ile Leu Asp Arg Asn Arg Ser Gly Gln Asn Val Lys Gly Ile Arg Ser
420 425 430
Cys Pro Leu Pro Asn Thr Ser Lys Val Arg Leu His Ile Phe Leu Asp
435 440 445
Arg Ser Ser Ile Glu Ile Phe Val Gly Asp Asp Gln Thr Gln Gly Leu
450 455 460

Tyr Ser Ile Ser Ser Arg Ile Phe Pro Asp Lys Asp Ser Leu Lys Gly
465 470 475 480
Arg Leu Phe Ala Ile Glu Gly Tyr Ala Val Phe Asp Ser Phe Lys Arg
485 490 495
Trp Thr Leu Gln Asp Ala Asn Leu Ala Ala Phe Ser Ser Asp Ala Cys
500 505 510

具有Prototheca moriformis Δ12脂肪酸去饱和酶转运肽的樟树FATB1(Genbank Q39473)氨基酸序列

〈210〉 30
〈211〉 360
〈212〉 PRT
〈213〉 人工序列
〈220〉
〈223〉 合成的构建体
〈400〉 30

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

Cys Thr Ser Leu Ser Val Met Met Asn Thr Arg Thr Arg Arg Leu Ser
195 200 205
Lys Ile Pro Glu Glu Val Arg Gly Glu Ile Gly Pro Ala Phe Ile Asp
210 215 220
Asn Val Ala Val Lys Asp Glu Glu Ile Lys Lys Pro Gln Lys Leu Asn
225 230 235 240
Asp Ser Thr Ala Asp Tyr Ile Gln Gly Gly Leu Thr Pro Arg Trp Asn
245 250 255
Asp Leu Asp Ile Asn Gln His Val Asn Asn Ile Lys Tyr Val Asp Trp
260 265 270
Ile Leu Glu Thr Val Pro Asp Ser Ile Phe Glu Ser His His Ile Ser
275 280 285
Ser Phe Thr Ile Glu Tyr Arg Arg Glu Cys Thr Met Asp Ser Val Leu
290 295 300
Gln Ser Leu Thr Thr Val Ser Gly Gly Ser Ser Glu Ala Gly Leu Val
305 310 315 320
Cys Glu His Leu Leu Gln Leu Glu Gly Gly Ser Glu Val Leu Arg Ala
325 330 335
Lys Thr Glu Trp Arg Pro Lys Leu Thr Asp Ser Phe Arg Gly Ile Ser
340 345 350
Val Ile Pro Ala Glu Ser Ser Val
355 360

具有Prototheca moriformis Δ12脂肪酸去饱和酶转运肽的樟树FATB1 cDNA的相关密码子优化的表达构建体

<210> 31
<211> 4653
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 31

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480

gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600
cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020
cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080
atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260
gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320
tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500
cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860
gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980
ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040
cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100
cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280
cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340
ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400
ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460
tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580
gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060
cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag 3120
cctccgttca cgatcgggac gctgcgcaag gccatccccg cgcactgttt cgagcgctcg 3180
gcgcttcgtg ggcgcgcccc cgactggtcc atgctgttcg ccgtgatcac caccatcttc 3240
tccgccgccg agaagcagtg gaccaacctg gagtggaagc ccaagcccaa ccccccccag 3300
ctgctggacg accacttcgg cccccacggc ctggtgttcc gccgcacctt cgccatccgc 3360
agctacgagg tgggccccga ccgctccacc agcatcgtgg ccgtgatgaa ccacctgcag 3420
gaggccgccc tgaaccacgc caagtccgtg ggcatcctgg gcgacggctt cggcaccacc 3480
ctggagatgt ccaagcgcga cctgatctgg gtggtgaagc gcacccacgt ggccgtggag 3540
cgctaccccg cctggggcga caccgtggag gtggagtgct gggtgggcgc ctccggcaac 3600
aacggccgcc gccacgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc 3660
tgcacctccc tgagcgtgat gatgaacacc cgcacccgcc gcctgagcaa gatccccgag 3720
gaggtgcgcg gcgagatcgg ccccgccttc atcgacaacg tggccgtgaa ggacgaggag 3780
atcaagaagc cccagaagct gaacgactcc accgccgact acatccaggg cggcctgacc 3840
ccccgctgga acgacctgga catcaaccag cacgtgaaca acatcaagta cgtggactgg 3900
atcctggaga ccgtgcccga cagcatcttc gagagccacc acatctcctc cttcaccatc 3960
gagtaccgcc gcgagtgcac catggacagc gtgctgcagt ccctgaccac cgtgagcggc 4020
ggctcctccg aggccggcct ggtgtgcgag cacctgctgc agctggaggg cggcagcgag 4080
gtgctgcgcg ccaagaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc 4140
gtgatccccg ccgagtccag cgtgatggac tacaaggacc acgacggcga ctacaaggac 4200
cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg 4260
atagtatcga cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg 4320
ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt 4380
gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcgaatacc acccccagca 4440
tccccttccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc 4500
tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct 4560
ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 4620
aagtagtggg atgggaacac aaatggaaag ctt 4653

具有Prototheca moriformis Δ12脂肪酸去饱和酶转运肽的萼距花FATB2(Genbank AAC49269)氨基酸序列

<210> 32
<211> 366
<212> PRT
<213> 人工序列
<220>
<223> 合成的构建体

<400> 32

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

Ser Val Leu Glu Ser Val Thr Ala Met Asp Pro Ser Lys Val Gly Val
305 310 315 320
Arg Ser Gln Tyr Gln His Leu Leu Arg Leu Glu Asp Gly Thr Ala Ile
325 330 335
Val Asn Gly Ala Thr Glu Trp Arg Pro Lys Asn Ala Gly Ala Asn Gly
340 345 350
Ala Ile Ser Thr Gly Lys Thr Ser Asn Gly Asn Ser Val Ser
355 360 365

具有Prototheca moriformis Δ12脂肪酸去饱和酶转运肽的蓖距花FATB2 cDNA的相关密码子优化的表达构建体

〈210〉 33
〈211〉 4668
〈212〉 DNA
〈213〉 人工序列
〈220〉
〈223〉 合成的构建体
〈400〉 33

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600
cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020
cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080
atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260
gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320

tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500
cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860
gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980
ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040
cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100
cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280
cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340
ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400
ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460
tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580
gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060
cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag 3120
cctccgttca cgatcgggac gctgcgcaag gccatccccg cgcactgttt cgagcgctcg 3180
gcgcttcgtg ggcgcgccca gctgcccgac tggagccgcc tgctgaccgc catcaccacc 3240
gtgttcgtga agtccaagcg ccccgacatg cacgaccgca agtccaagcg ccccgacatg 3300
ctggtggaca gcttcggcct ggagtccacc gtgcaggacg gcctggtgtt ccgccagtcc 3360
ttctccatcc gctcctacga gatcggcacc gaccgcaccg ccagcatcga gaccctgatg 3420
aaccacctgc aggagacctc cctgaaccac tgcaagagca ccggcatcct gctggacggc 3480
ttcggccgca ccctggagat gtgcaagcgc gacctgatct gggtggtgat caagatgcag 3540
atcaaggtga accgctaccc cgcctggggc gacaccgtgg agatcaacac ccgcttcagc 3600
cgcctgggca agatcggcat gggccgcgac tggctgatct ccgactgcaa caccggcgag 3660

atcctggtgc gcgccaccag cgcctacgcc atgatgaacc agaagacccg ccgcctgtcc 3720
aagctgccct acgaggtgca ccaggagatc gtgcccctgt tcgtggacag ccccgtgatc 3780
gaggactccg acctgaaggt gcacaagttc aaggtgaaga ccggcgacag catccagaag 3840
ggcctgaccc ccggctggaa cgacctggac gtgaaccagc acgtgtccaa cgtgaagtac 3900
atcggctgga tcctggagag catgcccacc gaggtgctgg agacccagga gctgtgctcc 3960
ctggccctgg agtaccgccg cgagtgcggc cgcgactccg tgctggagag cgtgaccgcc 4020
atggacccca gcaaggtggg cgtgcgctcc cagtaccagc acctgctgcg cctggaggac 4080
ggcaccgcca tcgtgaacgg cgccaccgag tggcgcccca agaacgccgg cgccaacggc 4140
gccatctcca ccggcaagac cagcaacggc aactccgtgt ccatggacta caaggaccac 4200
gacggcgact acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgactcgag 4260
gcagcagcag ctcggatagt atcgacacac tctggacgct ggtcgtgtga tggactgttg 4320
ccgccacact tgctgccttg acctgtgaat atccctgccg cttttatcaa acagcctcag 4380
tgtgtttgat cttgtgtgta cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga 4440
ataccacccc cagcatcccc ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta 4500
tctacgctgt cctgctatcc ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca 4560
gccttggttt gggctccgcc tgtattctcc tggtactgca acctgtaaac cagcactgca 4620
atgctgatgc acgggaagta gtgggatggg aacacaaatg gaaagctt 4668

具有Prototheca moriformis Δ12脂肪酸去饱和酶转运肽的加州月桂FATB1(Genbank Q41635)氨基酸序列

<210> 34
<211> 360
<212> PRT
<213> 人工序列
<220>
<223> 合成的构建体
<400> 34

Met Ala Ile Lys Thr Asn Arg Gln Pro Val Glu Lys Pro Pro Phe Thr
1 5 10 15
Ile Gly Thr Leu Arg Lys Ala Ile Pro Ala His Cys Phe Glu Arg Ser
20 25 30
Ala Leu Arg Gly Arg Ala Pro Asp Trp Ser Met Leu Phe Ala Val Ile
35 40 45
Thr Thr Ile Phe Ser Ala Ala Glu Lys Gln Trp Thr Asn Leu Glu Trp
50 55 60
Lys Pro Lys Pro Lys Leu Pro Gln Leu Leu Asp Asp His Phe Gly Leu
65 70 75 80
His Gly Leu Val Phe Arg Arg Thr Phe Ala Ile Arg Ser Tyr Glu Val
85 90 95
Gly Pro Asp Arg Ser Thr Ser Ile Leu Ala Val Met Asn His Met Gln

100 105 110

Glu Ala Thr Leu Asn His Ala Lys Ser Val Gly Ile Leu Gly Asp Gly
115 120 125

Phe Gly Thr Thr Leu Glu Met Ser Lys Arg Asp Leu Met Trp Val Val
130 135 140

Arg Arg Thr His Val Ala Val Glu Arg Tyr Pro Thr Trp Gly Asp Thr
145 150 155 160

Val Glu Val Glu Cys Trp Ile Gly Ala Ser Gly Asn Asn Gly Met Arg
165 170 175

Arg Asp Phe Leu Val Arg Asp Cys Lys Thr Gly Glu Ile Leu Thr Arg
180 185 190

Cys Thr Ser Leu Ser Val Leu Met Asn Thr Arg Thr Arg Arg Leu Ser
195 200 205

Thr Ile Pro Asp Glu Val Arg Gly Glu Ile Gly Pro Ala Phe Ile Asp
210 215 220

Asn Val Ala Val Lys Asp Asp Glu Ile Lys Lys Leu Gln Lys Leu Asn
225 230 235 240

Asp Ser Thr Ala Asp Tyr Ile Gln Gly Gly Leu Thr Pro Arg Trp Asn
245 250 255

Asp Leu Asp Val Asn Gln His Val Asn Asn Leu Lys Tyr Val Ala Trp
260 265 270

Val Phe Glu Thr Val Pro Asp Ser Ile Phe Glu Ser His His Ile Ser
275 280 285

Ser Phe Thr Leu Glu Tyr Arg Arg Glu Cys Thr Arg Asp Ser Val Leu
290 295 300

Arg Ser Leu Thr Thr Val Ser Gly Gly Ser Ser Glu Ala Gly Leu Val
305 310 315 320

Cys Asp His Leu Leu Gln Leu Glu Gly Gly Ser Glu Val Leu Arg Ala
325 330 335

Arg Thr Glu Trp Arg Pro Lys Leu Thr Asp Ser Phe Arg Gly Ile Ser
340 345 350

Val Ile Pro Ala Glu Pro Arg Val
355 360

具有Prototheca moriformis Δ12脂肪酸去饱和酶转运肽的加州月桂FATB1 cDNA的相关密码子优化的表达构建体

〈210〉 35

〈211〉 4653

〈212〉 DNA

〈213〉 人工序列

<220>

<223> 合成的构建体

<400> 35

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600
cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020
cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080
atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260
gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320
tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500
cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860
gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980
ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040
cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100
cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160

atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280
cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340
ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400
ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460
tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580
gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060
cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag 3120
cctccgttca cgatcgggac gctgcgcaag gccatccccg cgcactgttt cgagcgctcg 3180
gcgcttcgtg ggcgcgcccc cgactggtcc atgctgttcg ccgtgatcac caccatcttc 3240
agcgccgccg agaagcagtg gaccaacctg gagtggaagc ccaagcccaa gctgccccag 3300
ctgctggacg accacttcgg cctgcacggc ctggtgttcc gccgcacctt cgccatccgc 3360
tcctacgagg tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag 3420
gaggccaccc tgaaccacgc caagagcgtg ggcatcctgg gcgacggctt cggcaccacc 3480
ctggagatgt ccaagcgcga cctgatgtgg gtggtgcgcc gcacccacgt ggccgtggag 3540
cgctacccca cctggggcga caccgtggag gtggagtgct ggatcggcgc cagcggcaac 3600
aacggcatgc gccgcgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc 3660
tgcacctccc tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac 3720
gaggtgcgcg gcgagatcgg ccccgccttc atcgacaacg tggccgtgaa ggacgacgag 3780
atcaagaagc tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc 3840
ccccgctgga acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg 3900
gtgttcgaga ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcaccctg 3960
gagtaccgcc gcgagtgcac ccgcgactcc gtgctgcgca gcctgaccac cgtgagcggc 4020
ggcagctccg aggccggcct ggtgtgcgac cacctgctgc agctggaggg cggcagcgag 4080
gtgctgcgcg cccgcaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc 4140
gtgatccccg ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac 4200
cacgacatcg actacaagga cgacgacgac aagtgatgac tcgaggcagc agcagctcgg 4260
atagtatcga cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg 4320
ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt 4380
gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcgaatacc acccccagca 4440
tcccccttccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc 4500

tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct 4560
ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 4620
aagtagtggg atgggaacac aaatggaaag ctt 4653

具有Prototheca moriformis Δ12脂肪酸去饱和酶转运肽的湿地萼距花C8偏好型硫酯酶(Genbank AAC49179)氨基酸序列

<210> 36
<211> 393
<212> PRT
<213> 人工序列
<220>
<223> 合成的构建体
<400> 36

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

210 215 220

Glu Ile Leu Val Arg Ala Thr Ser Val Tyr Ala Met Met Asn Gln Lys
225 230 235 240
Thr Arg Arg Phe Ser Lys Leu Pro His Glu Val Arg Gln Glu Phe Ala
245 250 255
Pro His Phe Leu Asp Ser Pro Pro Ala Ile Glu Asp Asn Asp Gly Lys
260 265 270
Leu Gln Lys Phe Asp Val Lys Thr Gly Asp Ser Ile Arg Lys Gly Leu
275 280 285
Thr Pro Gly Trp Tyr Asp Leu Asp Val Asn Gln His Val Ser Asn Val
290 295 300
Lys Tyr Ile Gly Trp Ile Leu Glu Ser Met Pro Thr Glu Val Leu Glu
305 310 315 320
Thr Gln Glu Leu Cys Ser Leu Thr Leu Glu Tyr Arg Arg Glu Cys Gly
325 330 335
Arg Asp Ser Val Leu Glu Ser Val Thr Ser Met Asp Pro Ser Lys Val
340 345 350
Gly Asp Arg Phe Gln Tyr Arg His Leu Leu Arg Leu Glu Asp Gly Ala
355 360 365
Asp Ile Met Lys Gly Arg Thr Glu Trp Arg Pro Lys Asn Ala Gly Thr
370 375 380
Asn Gly Ala Ile Ser Thr Gly Lys Thr
385 390

具有Prototheca moriformis Δ12脂肪酸去饱和酶转运肽的湿地萼距花C8偏好型硫酯酶cDNA的相关密码子优化的表达构建体

<210> 37
<211> 4686
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 37

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420

gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600
cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020
cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080
atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260
gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320
tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500
cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gcccccttcga 1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860
gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980
ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040
cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100
cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280
cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340
ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400
ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460
tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580
gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760

gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060
cagtcacaac ccgcaaacac tagtatggct atcaagacga acaggcagcc tgtggagaag 3120
cctccgttca cgatcgggac gctgcgcaag gccatccccg cgcactgttt cgagcgctcg 3180
gcgcttcgtg ggcgcgcccc cgcgaacggc agcgcggtga ccctgaagtc gggctccctg 3240
aacacccagg aggacacgct gagctcgtcc ccccccccccc gcgcgttctt caaccagctg 3300
cccgactgga gcatgctgct gaccgcgatc accacggtct tcgtggcgcc cgagaagcgc 3360
tggaccatgt tcgaccgcaa gtcgaagcgc cccaacatgc tgatggactc cttcggcctg 3420
gagcgcgtgg tccaggacgg cctggtgttc cgccagagct tctcgatccg ctcctacgag 3480
atctgcgcgg accgcaccgc gagcatcgag acggtgatga accacgtcca ggagacctcg 3540
ctgaaccagt gcaagtccat cggcctgctg gacgacggct tcggccgcag ccccgagatg 3600
tgcaagcgcg acctgatctg ggtggtcacc cgcatgaaga tcatggtgaa ccgctacccc 3660
acgtggggcg acaccatcga ggtctcgacg tggctgtccc agagcggcaa gatcggcatg 3720
ggccgcgact ggctgatctc ggactgcaac accggcgaga tcctggtgcg cgcgacgtcc 3780
gtctacgcga tgatgaacca gaagacccgc cgcttcagca agctgcccca cgaggtgcgc 3840
caggagttcg cgccccactt cctggactcg ccccccgcga tcgaggacaa cgacggcaag 3900
ctgcagaagt tcgacgtcaa gacgggcgac tccatccgca agggcctgac ccccggctgg 3960
tacgacctgg acgtgaacca gcacgtgagc aacgtcaagt acatcggctg gatcctggag 4020
tcgatgccca ccgaggtcct ggagacgcag gagctgtgct ccctgaccct ggagtaccgc 4080
cgcgagtgcg gccgcgactc ggtgctggag agcgtcacca gcatggaccc ctcgaaggtg 4140
ggcgaccgct tccagtaccg ccacctgctg cgcctggagg acggcgcgga catcatgaag 4200
ggccgcaccg agtggcgccc caagaacgcg ggcacgaacg gcgcgatctc caccggcaag 4260
acgtgactcg aggcagcagc agctcggata gtatcgacac actctggacg ctggtcgtgt 4320
gatggactgt tgccgccaca cttgctgcct tgacctgtga atatccctgc cgcttttatc 4380
aaacagcctc agtgtgtttg atcttgtgtg tacgcgcttt tgcgagttgc tagctgcttg 4440
tgctatttgc gaataccacc cccagcatcc ccttccctcg tttcatatcg cttgcatccc 4500
aaccgcaact tatctacgct gtcctgctat ccctcagcgc tgctcctgct cctgctcact 4560
gcccctcgca cagccttggt ttgggctccg cctgtattct cctggtactg caacctgtaa 4620
accagcactg caatgctgat gcacgggaag tagtgggatg ggaacacaaa tggaaagctt 4680
gagctc 4686

具有Prototheca moriformis Δ12脂肪酸去饱和酶转运肽的披针叶萼距花C10偏好型硫酯酶(Genbank CAB60830)氨基酸序列

<210> 38
<211> 408
<212> PRT
<213> 人工序列

<220>

<223> 合成的构建体

<400> 38

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

Pro Arg Trp Asn Asp Leu Asp Val Asn Gln His Val Ser Asn Val Lys
290 295 300
Tyr Ile Gly Trp Ile Leu Glu Ser Met Pro Ile Glu Val Leu Glu Thr
305 310 315 320
Gln Glu Leu Cys Ser Leu Thr Val Glu Tyr Arg Arg Glu Cys Gly Met
325 330 335
Asp Ser Val Leu Glu Ser Val Thr Ala Val Asp Pro Ser Glu Asn Gly
340 345 350
Gly Arg Ser Gln Tyr Lys His Leu Leu Arg Leu Glu Asp Gly Thr Asp
355 360 365
Ile Val Lys Ser Arg Thr Glu Trp Arg Pro Lys Asn Ala Gly Thr Asn
370 375 380
Gly Ala Ile Ser Thr Ser Thr Ala Lys Thr Ser Asn Gly Asn Ser Ala
385 390 395 400
Ser Asp Asp Asp Asp Lys Leu Gly
405

具有Prototheca moriformis Δ12脂肪酸去饱和酶转运肽的披针叶蓼距花C10偏好型硫酯酶的相关密码子优化的编码区

<210> 39
<211> 1239
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 39

actagtatgg ctatcaagac gaacaggcag cctgtggaga agcctccgtt cacgatcggg 60
acgctgcgca aggccatccc cgcgcactgt ttcgagcgct cggcgcttcg tgggcgcgcc 120
cccgcgaacg gcagcgcggt gaacctgaag tcgggctccc tgaacaccca ggaggacacg 180
agctcgtccc ccccccccg cgcgttcctg aaccagctgc ccgactggag catgctgctg 240
accgcgatca ccaccgtctt cgtggcggcg gagaagcagt ggacgatgct ggaccgcaag 300
tcgaagcgcc ccgacatgct ggtggactcc gtcggcctga agagcatcgt gcgcgacggc 360
ctggtctcgc gccagtcctt cctgatccgc agctacgaga tcggcgcgga ccgcaccgcg 420
tcgatcgaga ccctgatgaa ccacctgcag gagacgtcca tcaaccactg caagagcctg 480
ggcctgctga acgacggctt cggccgcacc cccggcatgt gcaagaacga cctgatctgg 540
gtgctgacca agatgcagat catggtcaac cgctacccca cgtggggcga caccgtcgag 600
atcaacacgt ggttctcgca gtccggcaag atcggcatgg cgagcgactg gctgatctcg 660
gactgcaaca ccggcgagat cctgatccgc gcgacctccg tgtgggcgat gatgaaccag 720
aagacgcgcc gcttcagccg cctgccctac gaggtccgcc aggagctgac cccccacttc 780
gtggactcgc cccacgtcat cgaggacaac gaccagaagc tgcacaagtt cgacgtgaag 840

accggcgact ccatccgcaa gggcctgacg ccccgctgga acgacctgga cgtcaaccag 900
cacgtgtcga acgtgaagta catcggctgg atcctggagt ccatgcccat cgaggtcctg 960
gagacccagg agctgtgctc gctgaccgtg gagtaccgcc gcgagtgcgg catggactcc 1020
gtgctggagt cggtcacggc ggtggacccc agcgagaacg gcggccgcag ccagtacaag 1080
cacctgctgc gcctggagga cggcaccgac atcgtcaagt cgcgcaccga gtggcgcccc 1140
aagaacgcgg gcacgaacgg cgcgatctcc accagcaccg cgaagacgtc gaacggcaac 1200
tccgcgagcg atgacgatga caagctggga tgactcgag 1239

具有原始小球藻硬脂酰ACP去饱和酶叶绿体转运肽的德国鸢尾C14偏好型硫酯酶(Genbank AAG43858.1)氨基酸序列

<210> 40
<211> 399
<212> PRT
<213> 人工序列
<220>
<223> 合成的构建体
<400> 40

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

180 185 190

Val Glu His Tyr Pro Ser Trp Gly Asp Val Ile Glu Val Asp Thr Trp

195 200 205

Ala Ala Ala Ser Gly Lys Asn Gly Met Arg Arg Asp Trp His Val Arg

210 215 220

Asp Trp Gln Thr Gly Gln Thr Ile Met Arg Ala Ser Ser Asn Trp Val

225 230 235 240

Met Met Asn Gln Asn Thr Arg Arg Leu Ser Lys Phe Pro Glu Glu Val

245 250 255

Arg Ala Glu Ile Glu Pro Tyr Phe Met Glu Arg Ala Pro Val Ile Asp

260 265 270

Asp Asp Asn Arg Lys Leu Pro Lys Leu Asp Asp Asp Thr Ala Asp His

275 280 285

Val Arg Asn Gly Leu Thr Pro Arg Trp Ser Asp Leu Asp Val Asn Gln

290 295 300

His Val Lys Asn Val Lys Tyr Ile Gly Trp Ile Leu Glu Ser Ala Pro

305 310 315 320

Ile Ser Ile Leu Glu Ser His Glu Leu Ala Ser Met Thr Leu Glu Tyr

325 330 335

Arg Arg Glu Cys Gly Arg Asp Ser Val Leu Gln Ser Leu Thr Ser Val

340 345 350

Ser Asn Asn Cys Thr Asp Gly Ser Glu Glu Leu Pro Ile Glu Cys Gln

355 360 365

His Leu Leu Arg Asn Glu Gly Gly Ser Glu Ile Val Lys Gly Arg Thr

370 375 380

Glu Trp Arg Pro Lys Lys Cys Gly Pro Phe Gly Ala Gly Arg Pro

385 390 395

具有原始小球藻硬脂酰ACP去饱和酶转运肽的德国鸢尾C14偏好型硫酯酶的相关密码子优化的编码区

〈210〉 41

〈211〉 1212

〈212〉 DNA

〈213〉 人工序列

〈220〉

〈223〉 合成的构建体

〈400〉 41

actagtatgg ccaccgcatc cactttctcg gcgttcaatg cccgctgcgg cgacctgcgt 60

cgctcggcgg gctccgggcc ccggcgccca gcgaggcccc tccccgtgcg cgggcgcgcc 120

gcccaggcgg ccacccgcgt gaacggcagc aaggtgggcc tgaagaccga caccaacaag 180

ctggaggacg cgcccttcat cccctcgtcc gccccccgca ccttctacaa ccagctgccc 240
gactggagcg tcctgctggc ggccatcacc accatcttcc tggcggccga gaagcagtgg 300
accctgatcg actggaagcg cggcggcccc gacatgctgt cggacgcgtt cggcctgccc 360
aagatcatcg agaacggcct gctgtaccgc cagaagttct ccatccgcag ctacgagatc 420
ggcgccgacc agaccgcctc gatcgagacc ctgatgaacc acctgcagga gaccgcgctg 480
aaccacgtca agtgcgccgg cctgctgggc aacggcttcg gctccacccc cgagatgagc 540
aagatgaacc tgatctgggt ggtcaccaag atgcaggtgc tggtcgagca ctacccctcg 600
tggggcgacg tgatcgaggt ggacacctgg gcggccgcgt ccggcaagaa cggcatgcgc 660
cgcgactggc acgtccgcga ctggcagacc ggccagacca tcatgcgcgc cagctcgaac 720
tgggtgatga tgaaccagaa cacccgccgc ctgtccaagt tccccgagga ggtccgcgcc 780
gagatcgagc cctacttcat ggagcgcgcc cccgtgatcg acgacgacaa ccgcaagctg 840
cccaagctgg acgacgacac cgcggaccac gtgcgcaacg gcctgacccc ccgctggagc 900
gacctggacg tgaaccagca cgtcaagaac gtgaagtaca tcggctggat cctggagtcg 960
gcccccatct ccatcctgga gagccacgag ctggcctcga tgaccctgga gtaccgccgc 1020
gagtgcggcc gcgactccgt cctgcagagc ctgacctcgg tgtccaacaa ctgcaccgac 1080
ggcagcgagg agctgcccat cgagtgccag cacctgctgc gcaacgaggg cggctcggag 1140
atcgtcaagg gccgcaccga gtggcgcccc aagaagtgcg gcccccttcgg cgccggccgc 1200
ccctgactcg ag 1212

具有Prototheca moriformis Δ12脂肪酸去饱和酶叶绿体转运肽的肉豆蔻脂肪酰基硫酯酶(Genbank AAB717291.1)氨基酸序列

<210> 42
<211> 415
<212> PRT
<213> 人工序列
<220>
<223> 合成的构建体
<400> 42

Met Ala Ile Lys Thr Asn Arg Gln Pro Val Glu Lys Pro Pro Phe Thr
1 5 10 15
Ile Gly Thr Leu Arg Lys Ala Ile Pro Ala His Cys Phe Glu Arg Ser
20 25 30
Ala Leu Arg Gly Arg Ala Ala Asn Ala His Thr Val Pro Lys Ile Asn
35 40 45
Gly Asn Lys Ala Gly Leu Leu Thr Pro Met Glu Ser Thr Lys Asp Glu
50 55 60
Asp Ile Val Ala Ala Pro Thr Val Ala Pro Lys Arg Thr Phe Ile Asn
65 70 75 80
Gln Leu Pro Asp Trp Ser Met Leu Leu Ala Ala Ile Thr Thr Ile Phe
85 90 95

Leu Ala Ala Glu Lys Gln Trp Thr Asn Leu Asp Trp Lys Pro Arg Arg
100 105 110
Pro Asp Met Leu Val Asp Phe Asp Pro Phe Ser Leu Gly Arg Phe Val
115 120 125
Gln Asp Gly Leu Ile Phe Arg Gln Asn Phe Ser Ile Arg Ser Tyr Glu
130 135 140
Ile Gly Ala Asp Arg Thr Ala Ser Ile Glu Thr Leu Met Asn His Leu
145 150 155 160
Gln Glu Thr Ala Leu Asn His Val Arg Cys Ile Gly Leu Leu Asp Asp
165 170 175
Gly Phe Gly Ser Thr Pro Glu Met Thr Arg Arg Asp Leu Ile Trp Val
180 185 190
Val Thr Arg Met Gln Val Leu Val Asp Arg Tyr Pro Ser Trp Gly Asp
195 200 205
Val Ile Glu Val Asp Ser Trp Val Thr Pro Ser Gly Lys Asn Gly Met
210 215 220
Lys Arg Glu Trp Phe Leu Arg Asp Cys Lys Thr Gly Glu Ile Leu Thr
225 230 235 240
Arg Ala Thr Ser Val Trp Val Met Met Asn Lys Arg Thr Arg Arg Leu
245 250 255
Ser Lys Ile Pro Glu Glu Val Arg Val Glu Ile Glu Pro Tyr Phe Val
260 265 270
Glu His Gly Val Leu Asp Glu Asp Ser Arg Lys Leu Pro Lys Leu Asn
275 280 285
Asp Asn Thr Ala Asn Tyr Ile Arg Arg Gly Leu Ala Pro Arg Trp Ser
290 295 300
Asp Leu Asp Val Asn Gln His Val Asn Asn Val Lys Tyr Ile Gly Trp
305 310 315 320
Ile Leu Glu Ser Val Pro Ser Ser Leu Leu Glu Ser His Glu Leu Tyr
325 330 335
Gly Met Thr Leu Glu Tyr Arg Lys Glu Cys Gly Lys Asp Gly Leu Leu
340 345 350
Gln Ser Leu Thr Ala Val Ala Ser Asp Tyr Gly Gly Gly Ser Leu Glu
355 360 365
Ala Gly Val Glu Cys Asp His Leu Leu Arg Leu Glu Asp Gly Ser Glu
370 375 380
Ile Met Arg Gly Lys Thr Glu Trp Arg Pro Lys Arg Ala Ala Asn Thr
385 390 395 400
Thr Tyr Phe Gly Ser Val Asp Asp Ile Pro Pro Ala Asn Asn Ala

405 410 415

具有Prototheca moriformis Δ12脂肪酸去饱和酶叶绿体转运肽的肉豆蔻脂肪酰基硫酯酶的相关密码子优化的编码区

<210> 43

<211> 1260

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 43

actagtatgg ctatcaagac gaacaggcag cctgtggaga agcctccgtt cacgatcggg 60
acgctgcgca aggccatccc cgcgcactgt ttcgagcgct cggcgcttcg tgggcgcgcc 120
gccaacgccc acaccgtgcc caagatcaac ggcaacaagg ccggcctgct gacccccatg 180
gagagcacca aggacgagga catcgtcgcg gcccccaccg tggcgcccaa gcgcaccttc 240
atcaaccagc tgcccgactg gtcgatgctg ctggccgcga tcaccaccat cttcctggcg 300
gccgagaagc agtggaccaa cctggactgg aagccccgcc gccccgacat gctggtcgac 360
ttcgacccct tctccctggg ccgcttcgtg caggacggcc tgatcttccg ccagaacttc 420
agcatccgct cgtacgagat cggcgcggac cgcaccgcct ccatcgagac cctgatgaac 480
cacctgcagg agaccgcgct gaaccacgtc cgctgcatcg gcctgctgga cgacggcttc 540
ggcagcaccc ccgagatgac ccgccgcgac ctgatctggg tggtcacccg catgcaggtc 600
ctggtggacc gctacccctc gtggggcgac gtgatcgagg tcgactcctg ggtgaccccc 660
agcggcaaga acggcatgaa gcgcgagtgg ttcctgcgcg actgcaagac cggcgagatc 720
ctgacccgcg ccacctcggt ctgggtgatg atgaacaagc gcacccgccg cctgtccaag 780
atccccgagg aggtccgcgt ggagatcgag ccctacttcg tcgagcacgg cgtgctggac 840
gaggactcgc gcaagctgcc caagctgaac gacaacaccg ccaactacat ccgccgcggc 900
ctggcgcccc gctggtccga cctggacgtc aaccagcacg tgaacaacgt caagtacatc 960
ggctggatcc tggagagcgt gcccagcagc ctgctggagt cgcacgagct gtacggcatg 1020
accctggagt accgcaagga gtgcggcaag gacggcctgc tgcagtccct gaccgccgtc 1080
gccagcgact acggcggcgg ctcgctggag gccggcgtgg agtgcgacca cctgctgcgc 1140
ctggaggacg gctccgagat catgcgcggc aagaccgagt ggcgccccaa gcgcgccgcg 1200
aacaccacct acttcggcag cgtcgacgac atccccccg ccaacaacgc gtgactcgag 1260

具有原始小球藻硬脂酰ACP去饱和酶转运肽的湿地萼距花C14偏好型硫酯酶(Genbank AAC49180)氨基酸序列

<210> 44

<211> 362

<212> PRT

<213> 人工序列

<220>

<223> 合成的构建体

<400> 44

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

Ser Val Thr Ala Met Asp Pro Ser Lys Glu Gly Asp Arg Ser Leu Tyr
305 310 315 320
Gln His Leu Leu Arg Leu Glu Asp Gly Ala Asp Ile Val Lys Gly Arg
325 330 335
Thr Glu Trp Arg Pro Lys Asn Ala Gly Ala Lys Gly Ala Ile Leu Thr
340 345 350
Gly Lys Thr Ser Asn Gly Asn Ser Ile Ser
355 360

具有原始小球藻硬脂酰ACP去饱和酶转运肽的湿地萼距花C14偏好型硫酯酶的相关密码子优化的编码区

〈210〉 45
〈211〉 1101
〈212〉 DNA
〈213〉 人工序列
〈220〉
〈223〉 合成的构建体
〈400〉 45

actagtatgg ccaccgcatc cactttctcg gcgttcaatg cccgctgcgg cgacctgcgt 60
cgctcggcgg gctccgggcc ccggcgccca gcgaggcccc tccccgtgcg cgggcgcgcc 120
agcatgctgc tgtcggcggt gaccacggtc ttcggcgtgg ccgagaagca gtggcccatg 180
ctggaccgca agtccaagcg ccccgacatg ctggtcgagc ccctgggcgt ggaccgcatc 240
gtctacgacg gcgtgagctt ccgccagtcg ttctccatcc gcagctacga gatcggcgcc 300
gaccgcaccg cctcgatcga gacgctgatg aacatgttcc aggagacctc cctgaaccac 360
tgcaagatca tcggcctgct gaacgacggc ttcggccgca cgcccgagat gtgcaagcgc 420
gacctgatct gggtcgtgac caagatgcag atcgaggtga accgctaccc cacgtggggc 480
gacaccatcg aggtcaacac gtgggtgagc gcctcgggca agcacggcat gggccgcgac 540
tggctgatct ccgactgcca caccggcgag atcctgatcc gcgcgacgag cgtctgggcg 600
atgatgaacc agaagacccg ccgcctgtcg aagatcccct acgaggtgcg ccaggagatc 660
gagccccagt tcgtcgactc cgcccccgtg atcgtggacg accgcaagtt ccacaagctg 720
gacctgaaga cgggcgacag catctgcaac ggcctgaccc cccgctggac ggacctggac 780
gtgaaccagc acgtcaacaa cgtgaagtac atcggctgga tcctgcagtc ggtccccacc 840
gaggtgttcg agacgcagga gctgtgcggc ctgaccctgg agtaccgccg cgagtgcggc 900
cgcgactccg tgctggagag cgtcacggcc atggacccct cgaaggaggg cgaccgctcc 960
ctgtaccagc acctgctgcg cctggaggac ggcgcggaca tcgtgaaggg ccgcaccgag 1020
tggcgcccca agaacgccgg cgccaagggc gccatcctga cgggcaagac cagcaacggc 1080
aactcgatct cctgactcga g 1101

具有原始小球藻硬脂酰ACP去饱和酶转运肽的美洲榆广谱特异性硫酯酶(Genbank AAB71731)氨基酸序列

〈210〉 46

<211> 369
<212> PRT
<213> 人工序列
<220>
<223> 合成的构建体
<400> 46

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

260 265 270

Asn Val Lys Tyr Ile Gly Trp Leu Leu Glu Ser Ala Pro Pro Glu Ile
275 280 285

His Glu Ser His Glu Ile Ala Ser Leu Thr Leu Glu Tyr Arg Arg Glu
290 295 300

Cys Gly Arg Asp Ser Val Leu Asn Ser Ala Thr Lys Val Ser Asp Ser
305 310 315 320

Ser Gln Leu Gly Lys Ser Ala Val Glu Cys Asn His Leu Val Arg Leu
325 330 335

Gln Asn Gly Gly Glu Ile Val Lys Gly Arg Thr Val Trp Arg Pro Lys
340 345 350

Arg Pro Leu Tyr Asn Asp Gly Ala Val Val Asp Val Pro Ala Lys Thr
355 360 365

Ser

具有原始小球藻硬脂酰ACP去饱和酶转运肽的美洲榆广谱特异性硫酯酶的相关密码子优化的编码区

〈210〉 47

〈211〉 1122

〈212〉 DNA

〈213〉 人工序列

〈220〉

〈223〉 合成的构建体

〈400〉 47

actagtatgg ccaccgcatc cactttctcg gcgttcaatg cccgctgcgg cgacctgcgt 60

cgctcggcgg gctccgggcc ccggcgccca gcgaggcccc tccccgtgcg cgggcgcgcc 120

cagctgcccg actggagcat gctgctggcc gcgatcacca ccctgttcct ggcggccgag 180

aagcagtgga tgatgctgga ctggaagccc aagcgccccg acatgctggt ggacccttc 240

ggcctgggcc gcttcgtgca ggacggcctg gtgttccgca acaacttcag catccgcagc 300

tacgagatcg gcgcggaccg caccgccagc atcgagaccc tgatgaacca cctgcaggag 360

accgccctga accacgtgaa gagcgtgggc ctgctggagg acggcctggg cagcacccgc 420

gagatgagcc tgcgcaacct gatctgggtg gtgaccaaga tgcaggtggc ggtggaccgc 480

taccccacct ggggcgacga ggtgcaggtg agcagctggg cgaccgccat cggcaagaac 540

ggcatgcgcc gcgagtggat cgtgaccgac ttccgcaccg gcgagaccct gctgcgcgcc 600

accagcgtgt gggtgatgat gaacaagctg acccgccgca tcagcaagat ccccgaggag 660

gtgtggcacg agatcggccc cagcttcatc gacgcgcccc ccctgcccac cgtggaggac 720

gacggccgca agctgacccg cttcgacgag agcagcgccg acttcatccg caagggcctg 780

acccccgct ggagcgacct ggacatcaac cagcacgtga acaacgtgaa gtacatcggc 840

tggctgctgg agagcgcgcc ccccgagatc cacgagagcc acgagatcgc cagcctgacc 900

ctggagtacc gccgcgagtg cggccgcgac agcgtgctga acagcgccac caaggtgagc 960

gacagcagcc agctgggcaa gagcgccgtg gagtgcaacc acctggtgcg cctgcagaac 1020
ggcggcgaga tcgtgaaggg ccgcaccgtg tggcgcccca agcgcccccc gtacaacgac 1080
ggcgccgtgg tggacgtgcc cgccaagacc agctgactcg ag 1122

密码子优化的Prototheca moriformis(UTEX 1435)Δ12脂肪酸去饱和酶转运肽cDNA序列

<210> 48
<211> 120
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 48

actagtatgg ctatcaagac gaacaggcag cctgtggaga agcctccgtt cacgatcggg 60
acgctgcgca aggccatccc cgcgcactgt ttcgagcgct cggcgcttcg tgggcgcgcc 120

密码子优化的原始小球藻(UTEX 250)硬脂酰ACP去饱和酶转运肽cDNA序列

<210> 49
<211> 120
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 49

actagtatgg ccaccgcatc cactttctcg gcgttcaatg cccgctgcgg cgacctgcgt 60
cgctcggcgg gctccgggcc ccggcgccca gcgaggcccc tccccgtgcg cgggcgcgcc 120

美洲榆广谱特异性硫酯酶的密码子优化的编码区的相关同源重组表达构建体

<210> 50
<211> 6207
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 50

gctcttcggc cgccgccact cctgctcgag cgcgcccgac tcgcgctccg cctgcgcccg 60
cgcgtgcgcc gccagcgcct tggccttttc gccgcgctcg tgcgcgtcgc tgatgtccat 120
caccaggtcc atgaggtctg ccttgcgccg gctgagccac tgcttcgtcc gggcggccaa 180
gaggagcatg agggaggact cctggtccag ggtcctgacg tggtcgcggc tctgggagcg 240
ggccagcatc atctggctct gccgcaccga ggccgcctcc aactggtcct ccagcagccg 300
cagtcgccgc cgaccctggc agaggaagac aggtgagggg ggtatgaatt gtacagaaca 360
accacgagcc ttgtctaggc agaatcccta ccagtcatgg ctttacctgg atgacggcct 420

gcgaacagct gtccagcgac cctcgctgcc gccgcttctc ccgcacgctt ctttccagca 480
ccgtgatggc gcgagccagc gccgcacgct ggcgctgcgc ttcgccgatc tgaggacagt 540
cggggaactc tgatcagtct aaaccccctt gcgcgttagt gttgccatcc tttgcagacc 600
ggtgagagcc gacttgttgt gcgccacccc ccacaccacc tcctcccaga ccaattctgt 660
caccttttg gcgaaggcat cggcctcggc ctgcagagag gacagcagtg cccagccgct 720
gggggttggc ggatgcacgc tcaggtaccc tttcttgcgc tatgacactt ccagcaaaag 780
gtagggcggg ctgcgagacg gcttcccggc gctgcatgca acaccgatga tgcttcgacc 840
ccccgaagct ccttcggggc tgcatgggcg ctccgatgcc gctccagggc gagcgctgtt 900
taaatagcca ggcccccgat tgcaaagaca ttatagcgag ctaccaaagc catattcaaa 960
cacctagatc actaccactt ctacacaggc cactcgagct tgtgatcgca ctccgctaag 1020
ggggcgcctc ttcctcttcg tttcagtcac aacccgcaaa cggcgcgcca tgctgctgca 1080
ggccttcctg ttcctgctgg ccggcttcgc cgccaagatc agcgcctcca tgacgaacga 1140
gacgtccgac cgccccctgg tgcacttcac ccccaacaag ggctggatga acgaccccaa 1200
cggcctgtgg tacgacgaga aggacgccaa gtggcacctg tacttccagt acaacccgaa 1260
cgacaccgtc tgggggacgc ccttgttctg gggccacgcc acgtccgacg acctgaccaa 1320
ctgggaggac cagcccatcg ccatcgcccc gaagcgcaac gactccggcg ccttctccgg 1380
ctccatggtg gtggactaca acaacacctc cggcttcttc aacgacacca tcgacccgcg 1440
ccagcgctgc gtggccatct ggacctacaa caccccggag tccgaggagc agtacatctc 1500
ctacagcctg gacggcggct acaccttcac cgagtaccag aagaaccccg tgctggccgc 1560
caactccacc cagttccgcg acccgaaggt cttctggtac gagccctccc agaagtggat 1620
catgaccgcg gccaagtccc aggactacaa gatcgagatc tactcctccg acgacctgaa 1680
gtcctggaag ctggagtccg cgttcgccaa cgagggcttc ctcggctacc agtacgagtg 1740
ccccggcctg atcgaggtcc ccaccgagca ggaccccagc aagtcctact gggtgatgtt 1800
catctccatc aaccccggcg ccccggccgg cggctccttc aaccagtact tcgtcggcag 1860
cttcaacggc acccacttcg aggccttcga caaccagtcc cgcgtggtgg acttcggcaa 1920
ggactactac gccctgcaga ccttcttcaa caccgacccg acctacggga gcgccctggg 1980
catcgcgtgg gcctccaact gggagtactc cgccttcgtg cccaccaacc cctggcgctc 2040
ctccatgtcc ctcgtgcgca agttctccct caacaccgag taccaggcca acccggagac 2100
ggagctgatc aacctgaagg ccgagccgat cctgaacatc agcaacgccg gcccctggag 2160
ccggttcgcc accaacacca cgttgacgaa ggccaacagc tacaacgtcg acctgtccaa 2220
cagcaccggc accctggagt tcgagctggt gtacgccgtc aacaccaccc agacgatctc 2280
caagtccgtg ttcgcggacc tctccctctg gttcaagggc ctggaggacc ccgaggagta 2340
cctccgcatg ggcttcgagg tgtccgcgtc ctccttcttc ctggaccgcg ggaacagcaa 2400
ggtgaagttc gtgaaggaga acccctactt caccaaccgc atgagcgtga acaaccagcc 2460
cttcaagagc gagaacgacc tgtcctacta caaggtgtac ggcttgctgg accagaacat 2520
cctggagctg tacttcaacg acggcgacgt cgtgtccacc aacacctact tcatgaccac 2580
cgggaacgcc ctgggctccg tgaacatgac gacgggggtg gacaacctgt tctacatcga 2640
caagttccag gtgcgcgagg tcaagtgaca attggcagca gcagctcgga tagtatcgac 2700
acactctgga cgctggtcgt gtgatggact gttgccgcca cacttgctgc cttgacctgt 2760

gaatatccct gccgctttta tcaaacagcc tcagtgtgtt tgatcttgtg tgtacgcgct 2820
tttgcgagtt gctagctgct tgtgctattt gcgaatacca cccccagcat ccccttccct 2880
cgtttcatat cgcttgcatc ccaaccgcaa cttatctacg ctgtcctgct atccctcagc 2940
gctgctcctg ctcctgctca ctgcccctcg cacagccttg gtttgggctc cgcctgtatt 3000
ctcctggtac tgcaacctgt aaaccagcac tgcaatgctg atgcacggga agtagtggga 3060
tgggaacaca aatggaggat cccgcgtctc gaacagagcg cgcagaggaa cgctgaaggt 3120
ctcgcctctg tcgcacctca gcgcggcata caccacaata accacctgac gaatgcgctt 3180
ggttcttcgt ccattagcga agcgtccggt tcacacacgt gccacgttgg cgaggtggca 3240
ggtgacaatg atcggtggag ctgatggtcg aaacgttcac agcctaggga tatcgaattc 3300
ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 3360
cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc 3420
gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc aggcccccga ttgcaaagac 3480
attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 3540
ccactcgagc ttgtgatcgc actccgctaa gggggcgcct cttcctcttc gtttcagtca 3600
caacccgcaa acactagtat ggccaccgca tccactttct cggcgttcaa tgcccgctgc 3660
ggcgacctgc gtcgctcggc gggctccggg ccccggcgcc cagcgaggcc cctccccgtg 3720
cgcgggcgcg cccagctgcc cgactggagc atgctgctgg ccgcgatcac caccctgttc 3780
ctggcggccg agaagcagtg gatgatgctg gactggaagc ccaagcgccc cgacatgctg 3840
gtggaccccct tcggcctggg ccgcttcgtg caggacggcc tggtgttccg caacaacttc 3900
agcatccgca gctacgagat cggcgcggac cgcaccgcca gcatcgagac cctgatgaac 3960
cacctgcagg agaccgccct gaaccacgtg aagagcgtgg gcctgctgga ggacggcctg 4020
ggcagcaccc gcgagatgag cctgcgcaac ctgatctggg tggtgaccaa gatgcaggtg 4080
gcggtggacc gctaccccac ctggggcgac gaggtgcagg tgagcagctg ggcgaccgcc 4140
atcggcaaga acggcatgcg ccgcgagtgg atcgtgaccg acttccgcac cggcgagacc 4200
ctgctgcgcg ccaccagcgt gtgggtgatg atgaacaagc tgacccgccg catcagcaag 4260
atccccgagg aggtgtggca cgagatcggc cccagcttca tcgacgcgcc ccccctgccc 4320
accgtggagg acgacggccg caagctgacc cgcttcgacg agagcagcgc cgacttcatc 4380
cgcaagggcc tgacccccccg ctggagcgac ctggacatca accagcacgt gaacaacgtg 4440
aagtacatcg gctggctgct ggagagcgcg ccccccgaga tccacgagag ccacgagatc 4500
gccagcctga ccctggagta ccgccgcgag tgcggccgcg acagcgtgct gaacagcgcc 4560
accaaggtga gcgacagcag ccagctgggc aagagcgccg tggagtgcaa ccacctggtg 4620
cgcctgcaga acggcggcga gatcgtgaag ggccgcaccg tgtggcgccc caagcgcccc 4680
ctgtacaacg acggcgccgt ggtggacgtg cccgccaaga ccagcgatga cgatgacaag 4740
ctgggatgac tcgagttaat taactcgagg cagcagcagc tcggatagta tcgacacact 4800
ctggacgctg gtcgtgtgat ggactgttgc cgccacactt gctgccttga cctgtgaata 4860
tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac gcgcttttgc 4920
gagttgctag ctgcttgtgc tatttgcgaa taccaccccc agcatcccct tccctcgttt 4980
catatcgctt gcatcccaac cgcaacttat ctacgctgtc ctgctatccc tcagcgctgc 5040
tcctgctcct gctcactgcc cctcgcacag ccttggtttg ggctccgcct gtattctcct 5100

ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca cgggaagtag tgggatggga 5160
acacaaatgg aaagctgtag agctccttgt tttccagaag gagttgctcc ttgagccttt 5220
cattctcagc ctcgataacc tccaaagccg ctctaattgt ggaggggttt cgaatttaaa 5280
agcttggaat gttggttcgt gcgtctggaa caagcccaga cttgttgctc actgggaaaa 5340
ggaccatcag ctccaaaaaa cttgccgctc aaaccgcgta cctctgcttt cgcgcaatct 5400
gccctgttga aatcgccacc acattcatat tgtgacgctt gagcagtctg taattgcctc 5460
agaatgtgga atcatctgcc ccctgtgcga gcccatgcca ggcatgtcgc gggcgaggac 5520
acccgccact cgtacagcag accattatgc tacctcacaa tagttcataa cagtgaccat 5580
atttctcgaa gctccccaac gagcacctcc atgctctgag tggccacccc ccggccctgg 5640
tgcttgcgga gggcaggtca accggcatgg ggctaccgaa atccccgacc ggatcccacc 5700
acccccgcga tgggaagaat ctctccccgg gatgtgggcc caccaccagc acaacctgct 5760
ggcccaggcg agcgtcaaac cataccacac aaatatcctt ggcatcggcc ctgaattcct 5820
tctgccgctc tgctacccgg tgcttctgtc cgaagcaggg gttgctaggg atcgctccga 5880
gtccgcaaac ccttgtcgcg tggcggggct tgttcgagct tgttcgagct tgaagagcct 5940
ctagagtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct gtttcctgtg 6000
tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat aaagtgtaaa 6060
gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgctc actgcccgct 6120
ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgcggggaga 6180
ggcggtttgc gtattgggcg ctcttcc 6207

樟树C14偏好型硫酯酶的密码子优化的编码区的相关同源重组表达构建体

<210> 51
<211> 5576
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 51

gaattcgccc tcccgtgatc acacaggtgc cttgcgagcg tgatcacact attttggggg 60
tcctacagta ctgaaatggt gagaagtcgt actgaaatca aggatgaaca atgaaaatgg 120
tgctgtggtg gcttctcaaa ggtcaagaat cagtcgctcg cgtcaggaaa tcgcggcgtc 180
aaccagcgtg ggcgcggtca gtggccccgc actggtcacc atagcctctc ctgccacagt 240
agcgatcccc tgggcgttca ctctcagcag cggctgtact gcctcccaga ttttcttctt 300
ctggacctgc gggcgtgaga ggatgagcag ggtgggccaa gggctcaatc ctgaacggcc 360
ctcattcggt ttccaatccc acaacacata cccacagcag gtcagaccac gcattccacc 420
atgcgcacca ataacgtgtc cttacctgat tgggtgtggc aggctccgtg gacaggagtg 480
cctcgtcccc cgcccagacc cgctcccccg tcacggcggc gtccgggacc cgcagcggct 540
ccaccgcggt gtgatccgcg ttggcggcgc agagcagcat cccagccgat ttgaccccgc 600
gcatgctccg aggcttgagg ttggccagca ccaccacccg ccggccgaca aggtcctcca 660
gggtcacgtg ccggaccagg ccactcacga tggtgcgagg gccccccctcc tcgccgaggt 720

cgatctgctc gacgtacaga ctgcgacatg cgtggcgagt ggtcatcaga aggaagcagg 780
tgtgcagaag gggcacgtgg ttggtattga gagtagccaa agctttgtgc caatcagaaa 840
gtcaacgcag ctgcctgcct ggctcgcgta caattccttt cttgcgctat gacacttcca 900
gcaaaaggta gggcgggctg cgagacggct tcccggcgct gcatgcaaca ccgatgatgc 960
ttcgaccccc cgaagctcct tcggggctgc atgggcgctc cgatgccgct ccagggcgag 1020
cgctgtttaa atagccaggc ccccgattgc aaagacatta tagcgagcta ccaaagcata 1080
ttcaaacacc tagatcacta ccacttctac acaggccact cgagcttgtg atcgcactcc 1140
gctaaggggg cgcctcttcc cttcgtttca gtcacaaccc gcaaacggcg cgccatgctg 1200
ctgcaggcct tcctgttcct gctggccggc ttcgccgcca agatcagcgc ctccatgacg 1260
aacgagacgt ccgaccgccc cctggtgcac ttcaccccca acaagggctg gatgaacgac 1320
cccaacggcc tgtggtacga cgagaaggac gccaagtggc acctgtactt ccagtacaac 1380
ccgaacgaca ccgtctgggg gacgcccttg ttctggggcc acgccacgtc cgacgacctg 1440
accaactggg aggaccagcc catcgccatc gccccgaagc gcaacgactc cggcgccttc 1500
tccggctcca tggtggtgga ctacaacaac acctccggct tcttcaacga caccatcgac 1560
ccgcgccagc gctgcgtggc catctggacc tacaacaccc cggagtccga ggagcagtac 1620
atctcctaca gcctggacgg cggctacacc ttcaccgagt accagaagaa ccccgtgctg 1680
gccgccaact ccacccagtt ccgcgacccg aaggtcttct ggtacgagcc ctcccagaag 1740
tggatcatga ccgcggccaa gtcccaggac tacaagatcg agatctactc ctccgacgac 1800
ctgaagtcct ggaagctgga gtccgcgttc gccaacgagg gcttcctcgg ctaccagtac 1860
gagtgccccg gcctgatcga ggtccccacc gagcaggacc ccagcaagtc ctactgggtg 1920
atgttcatct ccatcaaccc cggcgccccg gccggcggct ccttcaacca gtacttcgtc 1980
ggcagcttca acggcaccca cttcgaggcc ttcgacaacc agtcccgcgt ggtggacttc 2040
ggcaaggact actacgccct gcagaccttc ttcaacaccg acccgaccta cgggagcgcc 2100
ctgggcatcg cgtgggcctc caactgggag tactccgcct tcgtgcccac caacccctgg 2160
cgctcctcca tgtccctcgt gcgcaagttc tccctcaaca ccgagtacca ggccaacccg 2220
gagacggagc tgatcaacct gaaggccgag ccgatcctga acatcagcaa cgccggcccc 2280
tggagccggt tcgccaccaa caccacgttg acgaaggcca acagctacaa cgtcgacctg 2340
tccaacagca ccggcaccct ggagttcgag ctggtgtacg ccgtcaacac cacccagacg 2400
atctccaagt ccgtgttcgc ggacctctcc ctctggttca agggcctgga ggaccccgag 2460
gagtacctcc gcatgggctt cgaggtgtcc gcgtcctcct tcttcctgga ccgcgggaac 2520
agcaaggtga agttcgtgaa ggagaacccc tacttcacca accgcatgag cgtgaacaac 2580
cagcccttca agagcgagaa cgacctgtcc tactacaagg tgtacggctt gctggaccag 2640
aacatcctgg agctgtactt caacgacggc gacgtcgtgt ccaccaacac ctacttcatg 2700
accaccggga acgccctggg ctccgtgaac atgacgacgg gggtggacaa cctgttctac 2760
atcgacaagt tccaggtgcg cgaggtcaag tgattaatta actcgaggca gcagcagctc 2820
ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg ccacacttgc 2880
tgccttgacc tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt gtttgatctt 2940
gtgtgtacgc gcttttgcga gttgctagct gcttgtgcta tttgcgaata ccacccccag 3000
catcccctcc ctcgtttcat atcgcttgca tcccaaccgc aacttatcta cgctgtcctg 3060

ctatccctca gcgctgctcc tgctcctgct cactgcccct cgcacagcct tggtttgggc 3120
tccgcctgta ttctcctggt actgcaacct gtaaaccagc actgcaatgc tgatgcacgg 3180
gaagtagtgg gatgggaaca caaatggaaa gcttgagctc ctttcttgcg ctatgacact 3240
tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg cgctgcatgc aacaccgatg 3300
atgcttcgac cccccgaagc tccttcgggg ctgcatgggc gctccgatgc cgctccaggg 3360
cgagcgctgt ttaaatagcc aggcccccga ttgcaaagac attatagcga gctaccaaag 3420
ccatattcaa acacctagat cactaccact tctacacagg ccactcgagc ttgtgatcgc 3480
actccgctaa gggggcgcct cttcctcttc gtttcagtca caacccgcaa acactagtat 3540
gacgttcggg gtcgccctcc cggccatggg ccgcggtgtc tcccttcccc ggcccagggt 3600
cgcggtgcgc gcccagtcgg cgagtcaggt tttggagagc gggcgcgccc ccgactggtc 3660
catgctgttc gccgtgatca ccaccatctt ctccgccgcc gagaagcagt ggaccaacct 3720
ggagtggaag cccaagccca acccccccca gctgctggac gaccacttcg gcccccacgg 3780
cctggtgttc cgccgcacct tcgccatccg cagctacgag gtgggccccg accgctccac 3840
cagcatcgtg gccgtgatga accacctgca ggaggccgcc ctgaaccacg ccaagtccgt 3900
gggcatcctg ggcgacggct tcggcaccac cctggagatg tccaagcgcg acctgatctg 3960
ggtggtgaag cgcacccacg tggccgtgga gcgctacccc gcctggggcg acaccgtgga 4020
ggtggagtgc tgggtgggcg cctccggcaa caacggccgc cgccacgact tcctggtgcg 4080
cgactgcaag accggcgaga tcctgacccg ctgcacctcc ctgagcgtga tgatgaacac 4140
ccgcacccgc cgcctgagca agatccccga ggaggtgcgc ggcgagatcg gccccgcctt 4200
catcgacaac gtggccgtga aggacgagga gatcaagaag ccccagaagc tgaacgactc 4260
caccgccgac tacatccagg gcggcctgac cccccgctgg aacgacctgg acatcaacca 4320
gcacgtgaac aacatcaagt acgtggactg gatcctggag accgtgcccg acagcatctt 4380
cgagagccac cacatctcct ccttcaccat cgagtaccgc cgcgagtgca ccatggacag 4440
cgtgctgcag tccctgacca ccgtgagcgg cggctcctcc gaggccggcc tggtgtgcga 4500
gcacctgctg cagctggagg gcggcagcga ggtgctgcgc gccaagaccg agtggcgccc 4560
caagctgacc gactccttcc gcggcatcag cgtgatcccc gccgagtcca gcgtgatgga 4620
ctacaaggac cacgacggcg actacaagga ccacgacatc gactacaagg acgacgacga 4680
caagtgactc gaggcagcag cagctcggat agtatcgaca cactctggac gctggtcgtg 4740
tgatggactg ttgccgccac acttgctgcc ttgacctgtg aatatccctg ccgcttttat 4800
caaacagcct cagtgtgttt gatcttgtgt gtacgcgctt ttgcgagttg ctagctcttg 4860
tgctatttgc gaataccacc cccagcatcc ccttccctcg tttcatatcg cttgcatccc 4920
aaccgcaact tatctacgct gtcctgctat ccctcagcgc tgctcctgct cctgctcact 4980
gcccctcgca cagccttggt ttgggctccg cctgtattct cctggtactg caacctgtaa 5040
accagcactg caatgctgat gcacgggaag tagtgggatg ggaacacaaa tggaaagctg 5100
gtacccgtac ccatcagcat ccgggtgaat cttggcctcc aagatatggc caatcctcac 5160
atccagcttg gcaaaatcga ctagactgtc tgcaagtggg aatgtggagc acaaggttgc 5220
ttgtagcgat cgacagactg gtggggtaca ttgacaggtg ggcagcgccg catccatcgt 5280
gcctgacgcg agcgccgccg gttgctcgcc cgtgcctgcc gtcaaagagc ggcagagaaa 5340
tcgggaaccg aaaacgtcac attgcctgat gttgttacat gctggactag actttcttgg 5400

cgtgggtctg ctcctcgcca ggtgcgcgac gcctcggggc tgggtgcgag ggagccgtgc 5460
ggccacgcat ttgacaagac ccaaagctcg catctcagac ggtcaaccgt tcgtattata 5520
cattcaacat atggtacata cgcaaaaagc atgccaacga tgacataggc gaattc 5576

具有原始小球藻硬脂酰ACP去饱和酶转运肽的蓖距花C10偏好型硫酯酶的密码子优化的编码区的相关表达构建体

〈210〉 52
〈211〉 4674
〈212〉 DNA
〈213〉 人工序列
〈220〉
〈223〉 合成的构建体
〈400〉 52

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600
cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020
cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080
atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260
gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320
tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500
cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620

gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860
gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980
ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040
cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100
cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280
cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340
ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400
ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460
tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580
gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060
cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc 3120
cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccagc gaggcccctc 3180
cccgtgcgcg ggcgcgccca gctgcccgac tggagccgcc tgctgaccgc catcaccacc 3240
gtgttcgtga agtccaagcg ccccgacatg cacgaccgca agtccaagcg ccccgacatg 3300
ctggtggaca gcttcggcct ggagtccacc gtgcaggacg gcctggtgtt ccgccagtcc 3360
ttctccatcc gctcctacga gatcggcacc gaccgcaccg ccagcatcga gaccctgatg 3420
aaccacctgc aggagacctc cctgaaccac tgcaagagca ccggcatcct gctggacggc 3480
ttcggccgca ccctggagat gtgcaagcgc gacctgatct gggtggtgat caagatgcag 3540
atcaaggtga accgctaccc cgcctggggc gacaccgtgg agatcaacac ccgcttcagc 3600
cgcctgggca agatcggcat gggccgcgac tggctgatct ccgactgcaa caccggcgag 3660
atcctggtgc gcgccaccag cgcctacgcc atgatgaacc agaagacccg ccgcctgtcc 3720
aagctgccct acgaggtgca ccaggagatc gtgcccctgt tcgtggacag ccccgtgatc 3780
gaggactccg acctgaaggt gcacaagttc aaggtgaaga ccggcgacag catccagaag 3840
ggcctgaccc ccggctggaa cgacctggac gtgaaccagc acgtgtccaa cgtgaagtac 3900
atcggctgga tcctggagag catgcccacc gaggtgctgg agacccagga gctgtgctcc 3960

ctggccctgg agtaccgccg cgagtgcggc cgcgactccg tgctggagag cgtgaccgcc 4020
atggacccca gcaaggtggg cgtgcgctcc cagtaccagc acctgctgcg cctggaggac 4080
ggcaccgcca tcgtgaacgg cgccaccgag tggcgcccca agaacgccgg cgccaacggc 4140
gccatctcca ccggcaagac cagcaacggc aactccgtgt ccatggacta caaggaccac 4200
gacggcgact acaaggacca cgacatcgac tacaaggacg acgacgacaa gtgactcgag 4260
gcagcagcag ctcggatagt atcgacacac tctggacgct ggtcgtgtga tggactgttg 4320
ccgccacact tgctgccttg acctgtgaat atccctgccg cttttatcaa acagcctcag 4380
tgtgtttgat cttgtgtgta cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga 4440
ataccacccc cagcatcccc ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta 4500
tctacgctgt cctgctatcc ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca 4560
gccttggttt gggctccgcc tgtattctcc tggtactgca acctgtaaac cagcactgca 4620
atgctgatgc acgggaagta gtgggatggg aacacaaatg gaaagcttga gctc 4674

具有原始小球藻硬脂酰ACP去饱和酶转运肽的加州月桂C12偏好型硫酯酶的密码子优化的编码区的相关表达构建体

<210> 53
<211> 4656
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 53

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540
tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600
cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020
cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080

atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260
gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320
tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500
cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860
gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980
ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040
cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100
cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280
cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340
ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400
ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460
tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580
gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060
cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc 3120
cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccagc gaggcccctc 3180
cccgtgcgcg ggcgcgcccc cgactggtcc atgctgttcg ccgtgatcac caccatcttc 3240
agcgccgccg agaagcagtg gaccaacctg gagtggaagc ccaagcccaa gctgccccag 3300
ctgctggacg accacttcgg cctgcacggc ctggtgttcc gccgcacctt cgccatccgc 3360
tcctacgagg tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag 3420

gaggccaccc tgaaccacgc caagagcgtg ggcatcctgg gcgacggctt cggcaccacc 3480
ctggagatgt ccaagcgcga cctgatgtgg gtggtgcgcc gcacccacgt ggccgtggag 3540
cgctacccca cctggggcga caccgtggag gtggagtgct ggatcggcgc cagcggcaac 3600
aacggcatgc gccgcgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc 3660
tgcacctccc tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac 3720
gaggtgcgcg gcgagatcgg ccccgccttc atcgacaacg tggccgtgaa ggacgacgag 3780
atcaagaagc tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc 3840
ccccgctgga acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg 3900
gtgttcgaga ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcaccctg 3960
gagtaccgcc gcgagtgcac ccgcgactcc gtgctgcgca gcctgaccac cgtgagcggc 4020
ggcagctccg aggccggcct ggtgtgcgac cacctgctgc agctggaggg cggcagcgag 4080
gtgctgcgcg cccgcaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc 4140
gtgatccccg ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac 4200
cacgacatcg actacaagga cgacgacgac aagtgactcg aggcagcagc agctcggata 4260
gtatcgacac actctggacg ctggtcgtgt gatggactgt tgccgccaca cttgctgcct 4320
tgacctgtga atatccctgc cgcttttatc aaacagcctc agtgtgtttg atcttgtgtg 4380
tacgcgcttt tgcgagttgc tagctgcttg tgctatttgc gaataccacc cccagcatcc 4440
ccttccctcg tttcatatcg cttgcatccc aaccgcaact tatctacgct gtcctgctat 4500
ccctcagcgc tgctcctgct cctgctcact gcccctcgca cagccttggt ttgggctccg 4560
cctgtattct cctggtactg caacctgtaa accagcactg caatgctgat gcacgggaag 4620
tagtgggatg ggaacacaaa tggaaagctt gagctc 4656

具有原始小球藻硬脂酰ACP去饱和酶转运肽的美洲榆广谱特异性硫酯酶的密码子优化的编码区的相关表达构建体

<210> 54
<211> 4635
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 54

ggtacccgcc tgcaacgcaa gggcagccac agccgctccc acccgccgct gaaccgacac 60
gtgcttgggc gcctgccgcc tgcctgccgc atgcttgtgc tggtgaggct gggcagtgct 120
gccatgctga ttgaggcttg gttcatcggg tggaagctta tgtgtgtgct gggcttgcat 180
gccgggcaat gcgcatggtg gcaagagggc ggcagcactt gctggagctg ccgcggtgcc 240
tccaggtggt tcaatcgcgg cagccagagg gatttcagat gatcgcgcgt acaggttgag 300
cagcagtgtc agcaaaggta gcagtttgcc agaatgatcg gttcagctgt taatcaatgc 360
cagcaagaga aggggtcaag tgcaaacacg ggcatgccac agcacgggca ccggggagtg 420
gaatggcacc accaagtgtg tgcgagccag catcgccgcc tggctgtttc agctacaacg 480
gcaggagtca tccaacgtaa ccatgagctg atcaacactg caatcatcgg gcgggcgtga 540

tgcaagcatg cctggcgaag acacatggtg tgcggatgct gccggctgct gcctgctgcg 600
cacgccgttg agttggcagc aggctcagcc atgcactgga tggcagctgg gctgccactg 660
caatgtggtg gataggatgc aagtggagcg aataccaaac cctctggctg cttgctgggt 720
tgcatggcat cgcaccatca gcaggagcgc atgcgaaggg actggcccca tgcacgccat 780
gccaaaccgg agcgcaccga gtgtccacac tgtcaccagg cccgcaagct ttgcagaacc 840
atgctcatgg acgcatgtag cgctgacgtc ccttgacggc gctcctctcg ggtgtgggaa 900
acgcaatgca gcacaggcag cagaggcggc ggcagcagag cggcggcagc agcggcgggg 960
gccacccttc ttgcggggtc gcgccccagc cagcggtgat gcgctgatcc caaacgagtt 1020
cacattcatt tgcatgcctg gagaagcgag gctggggcct ttgggctggt gcagcccgca 1080
atggaatgcg ggaccgccag gctagcagca aaggcgcctc ccctactccg catcgatgtt 1140
ccatagtgca ttggactgca tttgggtggg gcggccggct gtttctttcg tgttgcaaaa 1200
cgcgccagct cagcaacctg tcccgtgggt cccccgtgcc gatgaaatcg tgtgcacgcc 1260
gatcagctga ttgcccggct cgcgaagtag gcgccctcct ttctgctcgc cctctctccg 1320
tcccgcctct agaatatcaa tgatcgagca ggacggcctc cacgccggct cccccgccgc 1380
ctgggtggag cgcctgttcg gctacgactg ggcccagcag accatcggct gctccgacgc 1440
cgccgtgttc cgcctgtccg cccagggccg ccccgtgctg ttcgtgaaga ccgacctgtc 1500
cggcgccctg aacgagctgc aggacgaggc cgcccgcctg tcctggctgg ccaccaccgg 1560
cgtgccctgc gccgccgtgc tggacgtggt gaccgaggcc ggccgcgact ggctgctgct 1620
gggcgaggtg cccggccagg acctgctgtc ctcccacctg gcccccgccg agaaggtgtc 1680
catcatggcc gacgccatgc gccgcctgca caccctggac cccgccacct gccccttcga 1740
ccaccaggcc aagcaccgca tcgagcgcgc ccgcacccgc atggaggccg gcctggtgga 1800
ccaggacgac ctggacgagg agcaccaggg cctggccccc gccgagctgt tcgcccgcct 1860
gaaggcccgc atgcccgacg gcgaggacct ggtggtgacc cacggcgacg cctgcctgcc 1920
caacatcatg gtggagaacg gccgcttctc cggcttcatc gactgcggcc gcctgggcgt 1980
ggccgaccgc taccaggaca tcgccctggc cacccgcgac atcgccgagg agctgggcgg 2040
cgagtgggcc gaccgcttcc tggtgctgta cggcatcgcc gcccccgact cccagcgcat 2100
cgccttctac cgcctgctgg acgagttctt ctgacaattg gcagcagcag ctcggatagt 2160
atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact tgctgccttg 2220
acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat cttgtgtgta 2280
cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc cagcatcccc 2340
ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt cctgctatcc 2400
ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt gggctccgcc 2460
tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc acgggaagta 2520
gtgggatggg aacacaaatg gaggatcccg cgtctcgaac agagcgcgca gaggaacgct 2580
gaaggtctcg cctctgtcgc acctcagcgc ggcatacacc acaataacca cctgacgaat 2640
gcgcttggtt cttcgtccat tagcgaagcg tccggttcac acacgtgcca cgttggcgag 2700
gtggcaggtg acaatgatcg gtggagctga tggtcgaaac gttcacagcc tagggatatc 2760
gaattccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 2820
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 2880

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 2940
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 3000
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 3060
cagtcacaac ccgcaaacac tagtatggcc accgcatcca ctttctcggc gttcaatgcc 3120
cgctgcggcg acctgcgtcg ctcggcgggc tccgggcccc ggcgcccagc gaggcccctc 3180
cccgtgcgcg ggcgcgccca gctgcccgac tggagcatgc tgctggccgc gatcaccacc 3240
ctgttcctgg cggccgagaa gcagtggatg atgctggact ggaagcccaa gcgccccgac 3300
atgctggtgg acccccttcgg cctgggccgc ttcgtgcagg acggcctggt gttccgcaac 3360
aacttcagca tccgcagcta cgagatcggc gcggaccgca ccgccagcat cgagaccctg 3420
atgaaccacc tgcaggagac cgccctgaac cacgtgaaga gcgtgggcct gctggaggac 3480
ggcctgggca gcacccgcga gatgagcctg cgcaacctga tctgggtggt gaccaagatg 3540
caggtggcgg tggaccgcta ccccacctgg ggcgacgagg tgcaggtgag cagctgggcg 3600
accgccatcg gcaagaacgg catgcgccgc gagtggatcg tgaccgactt ccgcaccggc 3660
gagaccctgc tgcgcgccac cagcgtgtgg gtgatgatga acaagctgac ccgccgcatc 3720
agcaagatcc ccgaggaggt gtggcacgag atcggcccca gcttcatcga cgcgcccccc 3780
ctgcccaccg tggaggacga cggccgcaag ctgacccgct tcgacgagag cagcgccgac 3840
ttcatccgca agggcctgac cccccgctgg agcgacctgg acatcaacca gcacgtgaac 3900
aacgtgaagt acatcggctg gctgctggag agcgcgcccc ccgagatcca cgagagccac 3960
gagatcgcca gcctgaccct ggagtaccgc cgcgagtgcg gccgcgacag cgtgctgaac 4020
agcgccacca aggtgagcga cagcagccag ctgggcaaga gcgccgtgga gtgcaaccac 4080
ctggtgcgcc tgcagaacgg cggcgagatc gtgaagggcc gcaccgtgtg gcgccccaag 4140
cgccccctgt acaacgacgg cgccgtggtg gacgtgcccg ccaagaccag cgatgacgat 4200
gacaagctgg gatgactcga ggcagcagca gctcggatag tatcgacaca ctctggacgc 4260
tggtcgtgtg atggactgtt gccgccacac ttgctgcctt gacctgtgaa tatccctgcc 4320
gcttttatca aacagcctca gtgtgtttga tcttgtgtgt acgcgctttt gcgagttgct 4380
agctgcttgt gctatttgcg aataccaccc ccagcatccc cttccctcgt ttcatatcgc 4440
ttgcatccca accgcaactt atctacgctg tcctgctatc cctcagcgct gctcctgctc 4500
ctgctcactg cccctcgcac agccttggtt tgggctccgc ctgtattctc ctggtactgc 4560
aacctgtaaa ccagcactgc aatgctgatg cacgggaagt agtgggatgg gaacacaaat 4620
ggaaagcttg agctc 4635

Prototheca moriformis FatB/A启动子/5'UTR

<210> 55
<211> 997
<212> DNA
<213> Prototheca moriformis
<400> 55

cctgtcgatc gaagagaagg agacatgtgt acattattgg tgtgagggcg ctgaatcggc 60
catttttaa aatgatcacg ctcatgccaa tagacgcggc acataacgac gttcaaaccc 120
ccgccaaagc cgcggacaac cccatccctc cacacccccc acacaaagaa cccgccaccg 180

cttaccttgc ccacgaggta ggcctttcgt tgcgcaaaac cggcctcggt gatgaatgca 240
tgcccgttcc tgacgagcgc tgcccgggcc aacacgctct tttgctgcgt ctcctcaggc 300
ttgggggcct ccttgggctt gggtgccgcc atgatctgcg cgcatcagag aaacgttgct 360
ggtaaaaagg agcgcccggc tgcgcaatat atatataggc atgccaacac agcccaacct 420
cactcgggag cccgtcccac cacccccaag tcgcgtgcct tgacggcata ctgctgcaga 480
agcttcatga gaatgatgcc gaacaagagg ggcacgagga cccaatcccg gacatccttg 540
tcgataatga tctcgtgagt ccccatcgtc cgcccgacgc tccggggagc ccgccgatgc 600
tcaagacgag agggccctcg accaggaggg gctggcccgg gcgggcactg gcgtcgaagg 660
tgcgcccgtc gttcgcctgc agtcctatgc cacaaaacaa gtcttctgac ggggtgcgtt 720
tgctcccgtg cgggcaggca acagaggtat tcaccctggt catggggaga tcggcgatcg 780
agctgggata agagatactt ctggcaagca atgacaactt gtcaggaccg gaccgtgcca 840
tatatttctc acctagcgcc gcaaaaccta acaatttggg agtcactgtg ccactgagtt 900
cgactggtag ctgaatggag tcgctgctcc actaaacgaa ttgtcagcac cgccagccgg 960
ccgaggaccc gagtcatagc gagggtagta gcgcgcc 997

Prototheca moriformis NRAMP金属转运蛋白启动子/5'UTR

<210> 56
<211> 753
<212> DNA
<213> Prototheca moriformis
<400> 56

actaattgca atcgtgcagt aatcatcgat atggtcacaa gtagatcccc tactgacacc 60
ctctcgtaca tgtaggcaat gtcatcggcg ccgtcctgct gaccgatgcc gacgtagcag 120
agcagacccg ggccgatctg ggatacgagc cggccctcca cctgcgctcg aggtggaatc 180
aagtaaataa ccaatacact tttcgacacc acacagagtt gcacggacgg tggcgtacct 240
ctacgctcgc gctcttcacg cgctggacga ccgcacgcat gagcccgggt ggcttggtct 300
gggctgcaaa aatgcacaac aaacaagtat cagacgctca tggatgcaca cgcgctccca 360
agcacgctca gactaaatat tacagtagct cgtatctgat aagatatcga gacataccgc 420
tcaactcacc cgcaaactgc gccccgccag gtgatgcgca cagggcccca ccatgcgatc 480
catcgcatcg ctcctcgagg gcgctatcac gtggccggag agcgttcaca gcgtacgcca 540
ctgtatctgg gcggtatgcg gtccgtcaac atggagacag atacccgcac caccaccttg 600
caagctcttc catattggaa gtagaaaatt gtaattgtat catcgcacga ggggccaact 660
tgccgtcggc gagctgggcg acgaacacca cctggacgtt gtcgagactc gctcgtgccg 720
tgcgccgggc cgctgggtat ccagaccgtc gcc 753

Prototheca moriformis FLAP鞭毛相关蛋白启动子/5'UTR

<210> 57
<211> 1122
<212> DNA
<213> Prototheca moriformis
<400> 57

caacgacaac cagcaggcaa ctcggtcagc gacccaacac gcgagtcaaa ttgttgcgtg 60
ttcttgcctt gtctatttac tgtgatagca agactgtcgg tcagtcaata ccgcggtgcg 120
cacgtcgggg tgccaagcct agcagagcac gggacggctg gtgctgtgcg ccagctcagc 180
tcgcttcgcg accaattgta ggaccggcaa agtcaccaaa acatgccagc ggtgcgattc 240
aattggtcat gagctctaca aaattgtttt gtgcgtcgcg caggtatcca acggcgcggc 300
agagaaagtt tgacagctct cgatttcatc tcggaaaaat ggggagaatt tatgacacac 360
aagtgcgcag gcggcccagg cggccagcat attctggcgt gacctgggcc gcccacaaaa 420
tgcttggatg cactctaaaa taattatatt tgccatgaac aagggaagag ttaccgcacc 480
cagccctaga cttgggcgcc cgagcaaggt tacgtcaagc caccttcgcc catcgcccaa 540
ctccgtattc cccgacagcc gcacgtggcc ctcgccggaa tgaaccctga atcggcatca 600
cgccacgcgt tcgccaatcg ttccgctctc tggcttcatc ggcctgcgcc ttcacgtcgt 660
ggtcacgaca gtgcattcat acttccattt gcacctcggc acacactttt acgcatcgcc 720
taccttgct gcggcagtct agggtcactt tgcagccatg ggacagtgct acaccaccgt 780
cggtgcgcaa agctatttca agtgaaccgt gggcggaaaa aaggaatgta cactgtctca 840
accgactcct acaattgttt accatgcaga tcagagctcg acggccatca tcgagcaggt 900
gtggggcctt ggtggcgcgg cgcggggccc cagggcgtcg caggcattga tggcactctg 960
agactttcgc acgcgcatga gggaccccat caagagaaga gtgtgtcttt atgtccccat 1020
tcatgatgat gtatcttgtg attgtcgcag tttggcaagt ttaaccggat cgccgctcca 1080
ggtgtggcgt ggcggatttt tctaggggtg cttgagcagt cg 1122

Prototheca moriformis SulfRed亚硫酸盐还原酶启动子/5'UTR

<210> 58
<211> 574
<212> DNA
<213> Prototheca moriformis
<400> 58

ggcccagggc cctgcggatg gcccacacca gatctagcct ctcttatgcc atgcccgcct 60
cgctgcccgt cgtatccccc cgccgatccg cgcgtagggg accgcggcct gacccacgcc 120
acgaaagagc tttgctcctc aatttctcgc caacagaacc gtatcaaacg ctcaacgcct 180
atcccgaaca atccgtattc acaccaaatc gagtataccg gactggtttg cctagtcttg 240
aaggaaatga tcccgtccat gctcggaagg gggagcgggc ggaggatcct actcatctct 300
gaaatgggat tggtccgaag atgggttggg caagcacgtg ccaaacccca gcgagttgct 360
gacgagcagg ctcatccaat cccccggcga atcctccctc acgccccgca tgcatacaag 420
tccctcccac acgccccctc ccatccattt tcgcctggtc cgaacgcgag cggcgtcgag 480
gcggaccact tgctccgcag cgccgtctgg gtctccaccc cacagcggct ttgctgccag 540
aggcaccccc cttgccccac ctcctcttgc agcc 574

Prototheca moriformis SugT糖转运蛋白启动子/5'UTR

<210> 59
<211> 1096
<212> DNA

<213> Prototheca moriformis

<400> 59

ccaggcaggc ggtagggttg ccgattgctt gagcgaattg gaagatataa ttttttgtgg 60
tgtccctgga cgctgtttgt ggcgctcctt tttggagaag attgcgtggg ggagctttcc 120
atgtaccacg cttccttctg aaaggattct ggccgagtcc tgatgagccc aaagaaaaca 180
cctgcctttc agtgctggca ctctgaaaac gtcaacagat gattatacat gtcacaaaag 240
gcagccgatt aggaacggga gctctggccg ttcgtttggc tgcctgggct gattgaagtg 300
atccaccctg ttcgaatgaa ggcggtcgag tcgaattatc gaccggagct gtcgggaagg 360
cgtccggggc agagtgaggt gctgcggcct ggttgtcgtt caaaaagacc ccggtagccc 420
aacaatcacg aacgaaagga atataattgc ttgcatacta tacattcagt ttctatgtgg 480
cgggtagaca agtctcatgg gcttctaaag gctgtccctt gaaggctact tataaaaact 540
tgctgcgcca tggcacggat cgcgcttgcg caggctgcaa ccctgcgcgc aaggtcaaat 600
acacagcaaa agatactaac agaatttcta aaaacattta aatatttgtt tcgaccagcc 660
aattgtggtc gtaggcacgc aaaagacttt gttttgcgcc caccgagcat ccacgctggc 720
agtcaagcca gtccgatgtg cattgcgtgg cagcatcgag gagcatcaaa aacctcgtgc 780
acgcttttct gtcaatcatc atcaaccact ccaccatgta tacccgatgc atcgcggtgc 840
gcagcgcgcc acgcgtccca gacccgccca aaaacccagc agcggcgaaa gcaaatcttc 900
acttgcccga aaccccgagc agcggcattc acacgtgggc gaaaacccca cttgccctaa 960
caggcgtatg tctgctgtca cgatgcctga caacggtatt atagatatac actgattaat 1020
gtttgagtgt gtgcgagtcg cgaatcagga atgaattgct agtaggcact ccgaccgggc 1080
gggggccgag ggacca 1096

Prototheca moriformis Amt03-铵转运蛋白启动子/5'UTR

<210> 60

<211> 1075

<212> DNA

<213> Prototheca moriformis

<400> 60

ggccgacagg acgcgcgtca aaggtgctgg gcgtgtatgc cctggtcggc aggtcgttgc 60
tgttgctgcg ctcgtggttc cgcaaccctg attttggcgt cttattctgg cgtggcaagc 120
gctgacgccc gcgagccggg ccggcggcga tgcggtgtct cacggctgcc gagctccaag 180
ggaggcaaga gcgcccggat cagctgaagg gctttacacg caaggtacag ccgctcctgc 240
aaggctgcgt ggtggacttg aacctgtagg tcctctgctg aagttcctcc actacctcac 300
caggcccagc agaccaaagc acaggctttt caggtccgtg tcatccactc taaaacactc 360
gactacgacc tactgatggc cctagattct tcatcaacaa tgcctgagac acttgctcag 420
aattgaaact ccctgaaggg accaccagag gccctgagtt gttccttccc cccgtggcga 480
gctgccagcc aggctgtacc tgtgatcgag gctggcggga aaataggctt cgtgtgctca 540
ggtcatggga ggtgcaggac agctcatgaa acgccaacaa tcgcacaatt catgtcaagc 600
taatcagcta tttcctcttc acgagctgta attgtcccaa aattctggtc taccgggggt 660
gatccttcgt gtacgggccc ttccctcaac cctaggtatg cgcgcatgcg gtcgccgcgc 720

aactcgcgcg agggccgagg gtttgggacg ggccgtcccg aaatgcagtt gcacccggat 780
gcgcggcgcc tttcttgcga taatttatgc aatggactgc tctgcaaatt tctgggtctg 840
tcgccaaccc taggatcagc ggcgtaggat ttcgtaatca ttcgtcctga tggggagcta 900
ccgactaccc taatatcagc ccggctgcct gacgccagcg tccacttttg cgtacacatt 960
ccattcgtgc ccaagacatt tcattgtggt gcgaagcgtc cccagttacg ctcacctgtt 1020
tcccgacctc cttactgttc tgtcgacaga gcgggcccac aggccggtcg cagcc 1075

Prototheca moriformis Amt02-铵转运蛋白启动子/5'UTR

<210> 61
<211> 772
<212> DNA
<213> Prototheca moriformis
<400> 61

tcaccagcgg acaaagcacc ggtgtatcag gtccgtgtca tccactctaa agagctcgac 60
tacgacctac tgatggccct agattcttca tcaaaaacgc ctgagacact tgcccaggat 120
tgaaactccc tgaagggacc accaggggcc ctgagttgtt ccttcccccc gtggcgagct 180
gccagccagg ctgtacctgt gatcggggct ggcgggaaaa caggcttcgt gtgctcaggt 240
tatgggaggt gcaggacagc tcattaaacg ccaacaatcg cacaattcat ggcaagctaa 300
tcagttattt cccattaacg agctataatt gtcccaaaat tctggtctac cgggggtgat 360
ccttcgtgta cgggcccttc cctcaaccct aggtatgcgc acatgcggtc gccgcgcaac 420
gcgcgcgagg gccgagggtt tgggacgggc cgtcccgaaa tgcagttgca cccggatgcg 480
tggcaccttt tttgcgataa tttatgcaat ggactgctct gcaaaattct ggctctgtcg 540
ccaaccctag gatcagcggt gtaggatttc gtaatcattc gtcctgatgg ggagctaccg 600
actgccctag tatcagcccg actgcctgac gccagcgtcc acttttgtgc acacattcca 660
ttcgtgccca agacatttca ttgtggtgcg aagcgtcccc agttacgctc acctgatccc 720
caacctcctt attgttctgt cgacagagtg ggcccagagg ccggtcgcag cc 772

Prototheca moriformis Aat-01氨基酸转运蛋白启动子/5'UTR

<210> 62
<211> 991
<212> DNA
<213> Prototheca moriformis
<400> 62

cgaaggggtc tgcatcgatt cgcgcggtct ggaggccagc gtgactgctc gcgaaaatgc 60
tctgccgtgt cgggctctgg ctggggcggc cagagatctc accgtgccac acgcaactgc 120
cgcactctgt gcccgccacc tggcgcgcac atgcgacctc ttccccgtca taccctctcc 180
tcatgtgatc tttccacacg agtgacgcag gtgcgcggag tggagggaat caggacgttt 240
tcaaggtacc tgctcgagcc gtaccaacag ctgccgcccg gcaaggaaga gatcgaggca 300
gagattgccc ggctggaggc ccggataacg gagctcaaga gcaagctgtc cgagtgagac 360
cgcccaggtg cacgtgtcga ctcgctatga catgtactcg acacaacatg aggaattcat 420
cgaatttgta ggaagcgggc attggtacgg gagtgggaaa gcgaaaaaac ctccctccgg 480

cagtgccatc tgccggagtc gaacgttgat agggttctcg tgacagggtg tgacctctca 540
gccttgcatc aattaaacgc tatagacatt atcagtaacc gtgaatcccg cattggatgc 600
cacccgcgcg accattgggg acctgcatta cagatctagg tgagatgaca gcgaggcaac 660
ttcggcccgc ggcccagctt gcggcgcacc aatattggtc acgggaagcc acacaccgac 720
cataaatgaa tacttgtaag ctatgtcaac cgatcaatgg cgtcgaaagt gtgccacgag 780
gatccatctg gcggggcggc gtggcgcaca agcgcagtcg caatttctcg gacccatctg 840
acctaggccc agcgccgcgg gagaaatccc cggcgggtcc tccacgcagt aaccctaatg 900
agtatcgagc gccgaccatt tacaccatcg cccccgaaat ccttccgaca ttattattat 960
cttttagatc ttggaacaga ctctgccaac c 991

Prototheca moriformis Aat02-氨基酸转运蛋白启动子/5'UTR

<210> 63
<211> 1347
<212> DNA
<213> Prototheca moriformis
<400> 63

agagagcgga ggtggggttg tgaggtgggg ttgctgacca ggagctcgcg tcgccgagcg 60
cgactcgcac acggtccagt taccccccc tccgcccaaa cgcaagcctc ccatcttgat 120
gcctttccgg ccacctatac tatttcttag ttcgctgtaa catccagacc gtcctgaata 180
ataacaatgc cctgtgtcaa gtgcattcct aaaaaaattc tgtcccaacc aacaatccca 240
cctgaaatac caccagccct gcccagtaca ctcttccaat accatctccc tacctccacg 300
cgcaagcgac ccccatgcgc gaccaggctc gaaagtgatt tatgacttga gacgagcgag 360
tggcggcgcg gtcgactgcc ttttcatcac gtgccgtacg tcggcgaccg ctagggcttt 420
gcacggcaac gcacggcttc gccaacccga ccagccagga cctcgactac tctaccgcga 480
attcgcctca agaagtcgcc aaatgtgcca tacaccattc cttacagcac tgttcaaact 540
tgatgccaat tttgacattc gggttgctcg ttggctgcgc ccacatcggc cgtgagtgca 600
gcaggcggga tcggacacgg aggacgcggc gtcacgcccc gaacgcagcc cgtaactcta 660
catcaacacg acgtgttgcg taatcccgcc cggctgcgca tcgtgccaac ccattcgcga 720
tggatggtcg gaaaatggtg tgccaactgc cctgagggag gctctcgcga aacgggcacg 780
tccctgaaac cgaaactgtg gccttgtcgt cggccacgca agcacgtgga ccctaaacac 840
caagaaaatc agtaaacaag gttgacatcc tctacgggcg aattgtttgc ccaacccttc 900
atcgcacact gccattataa tgcatctagc tcggcgacaa gtttagaaaa ggcaggctgc 960
attgttccat ttcgccgtgg cggcgtgggt gcccatttta cgaggtttgg gctcccgggc 1020
agcgaccgag ccaggtcgag tccctctcgc ccgtcgacaa tgttgcgaac cccacaagcg 1080
gctaacaaca acttgatggt acctgtacac tgccaattcc ttcttccccg gccgaggttt 1140
acacgtgatg gccatggctt cgcattcagg ccgacttccc attccgactt tccagagggt 1200
ccgcggacgc tgggggttgg ctgcctgagg cccacccttt gttccccgcg tcccgacaaa 1260
cacaattgcg ttacataagg gggagccgcc cccgttcaga gtgcagaaat ctttcactat 1320
attttccagt cgtcagcgaa atcaagt 1347

Prototheca moriformis Aat03-氨基酸转运蛋白启动子/5'UTR

<210> 64
<211> 1180
<212> DNA
<213> Prototheca moriformis
<400> 64
gatggtgggg tgtctgcctt gggctgggtg atggaggctg gtggtgcgcg ggtttcctga 60
tgcattctat ctacgcagtg tcatggtgtc cattccacac accagtacac ccttacacta 120
aggatccatc cctccttccc tcttcaggac tacatggacc ccacgagcta ccgaccgggc 180
tttctcaaaa acgtcaaggt catgtttgac atgcgggacg tggtggacga cgtgcaaggt 240
gcgtccggag tgcgcgcaaa tgagcaagtc gggcaatgtg tcggggtggg caccggggct 300
ggagatccgc gatccccgag aaaacgccgt accacccccc gcgctattcc ctcgattgcg 360
cgcagatgtg gtgaccgaca cgggggacaa cctggcggac atggggcgcc ggacctggaa 420
gcacgccaag tcgcacacgg ggaggctcgt gcagtccccc ccatcgtacc tcaagggtct 480
ctttggtcgc gatccaaagt acgctggtgg catggcatgc ccgaaatgaa catcatgtgt 540
gatctccgat tgccaatggc cacctccacg gaccaccttg caggcggaag cgcaatccag 600
ggcccgagcc tgacgaggac ggagactcct cgtccagcgc ggggtccccg acccgacgca 660
gcagccgacc cctgctaacc cggcaacgat cggaccagca accttgctgt agttccgatc 720
cgtgatgacg ggcattgccg ccgctcgatc cgctttgatg actgtctatt atttgcgcgg 780
agccccctcg gaaccctacc ccgctcttgc aagccccttg catcggagat cctcgtgcgc 840
ccgccatgac cccactggat tgcccaacat ccttctttat cgtgtaaaat gtgattcctc 900
ggctgcaatc gactggcctt cgcttctggc cccaagaggg ctcgaacgtg cggcagcgag 960
ggcgctgaca cacccaagcc ctagggcttt caacgtcggc tgccaggccg gataggggga 1020
tcgcctcctt tccaccaccc acctacgagg gattcgagtc ggcttccagc tcagctattc 1080
ggccgcgccc ccggccctgc agacgtcctc cagtttccga acaggtcgct ctcagaacac 1140
ctgccgcggc tgcgatacgg caggctctca aagcgtcgac 1180
Prototheca moriformis Aat04-氨基酸转运蛋白启动子/5'UTR
<210> 65
<211> 1263
<212> DNA
<213> Prototheca moriformis
<400> 65
cgcgtggagc ggtgcgtgcg gatgccgcgc gcctgccaag gccttttgta tgcctggcct 60
gggaagtttc ctgactgaag catcttcaag atgctctctc acgaccagcg acaccaacac 120
cgtcactttt tgcccctcct gccgcaggtg ccactttcta ctttgacgtc ttctccaggc 180
ggtacattgc gggactgagc gccaattcgg ccaagaacag cgctgtcgac ttgaggaggc 240
aggggtccgt cgactctgcc gagtgacacg ccttcgaccc gactgtacta cggcctgctg 300
aagagtgggt ctcgccggcc ggcgtgaccg gccctgtgcc cacaatcgac catctattcg 360
ctccttgtca tctggcgccg tcaattgccc gcgacttgac ggcaactggc tcgatcgagt 420
cgtattgaaa aagcacgttt tgtcctacag ggccgcggtc cgttaccaac gtggttctcg 480

ttaggttttc gtcgggcggt ggtgcgcgaa ctgtccgatg ccatcccggc aaaccccagc 540
aaggtcgcca gtctggttct gacgcaatag agtgcgtttt gggccagtct aaaaattcgt 600
ctggcatgac gtggctccac atcgtacccg gagcctgcct tggtaatgtg aggcaccggt 660
gccaactcca ttatggcagg catcgagcgc gcaggtgagt acatgacctt ccgtgaattg 720
ggaaggcgag cttgtgtaac gcctgcgatc gtgccagtga ggcatcgtaa actcaaaata 780
ttttgtagaa agtgtctgat gcctggtgag gctgcgtagg gcaagggcaa gcccttggca 840
gatgggtaat gggtccggac ctcacaacag caaccccgcg tccccttag ggcccctgag 900
gctcgatggc agggccagcg agcccgcggc caaagggcgc catcccacgg tcgcccaacg 960
actccacggg tcctatacct catcttgaat ggcactaaaa actatagaat atcgggcact 1020
ggtgggcgtc tggggtacag ctggccgagc gcagtggcaa accctaggtc ccgcctcaag 1080
ggcgattccc gggtcaatga cacgcaagca agatcacatg gcgcggtccg cctcgcggct 1140
ccacacccag gccctagttt cgcaacccat aaatatcgcc ccgataccat cataagccag 1200
caaataattt tttatcagag ttccaaacct cctcagctgt gggaaaccag cccactctga 1260
acg 1263

Prototheca moriformis Aat05-氨基酸转运蛋白启动子/5'UTR

<210> 66
<211> 1400
<212> DNA
<213> Prototheca moriformis
<400> 66

ccgagcagtt catggccaag tacaaggact agagaccgga ggtcggtagg ctgaatggag 60
ctggcgtcgt cgtgcgcgac gtgcacgcga tgcgatacta cgaccccaca aacgcatgcc 120
tcccatcttg atgcctttcc ggccatttat actatttctc atttcgctgt aacatcttga 180
ataatagaat tgccctgtgt caagtggatt ccaagaaata ttctgtccca acaaaacaac 240
ccaacctgaa aacaacctca aataccacca gccctgccca cctgcccagt acacttttcc 300
aataccatct ccctaccttc acgcgcaagc ggcacccatg cgcgaccagg ctcgaaagga 360
tttcacgact caggacgagc gagtggcggc gcgaccgcct gcctgttcgt cacgtgccgt 420
acgtcggcga ccgctagagc tttgcctggc aacccccggc ttcgtcaacc cggccagcca 480
ggatctcgac cactctaccg cgaaatcgcc tcaagaagtc gccaaaagtg ccgtacacca 540
tgcttcgcag cgctgttcaa acttgatgcc aatcttgaca atcaggttgc tcgttggctg 600
cgtccacatc ggccgtgatt gcagcaggcg gggatcggac acggaggacg cggcgtcacg 660
ccgcgaacgc agcccgtaac tctacatcaa cgcgatatgt tgcgtaatcc cgcccggctg 720
cgcattgtga caacccattc gcgatggatg gtcggaaaat ggtgtgccaa ctgccctgag 780
ggactctctc gcgaaacggg cacgtccctg tatccgaaac tgtggcatgg ccttgtcgac 840
cacgcaagca cgtggaccct aacaccacga aaataagtaa aaaaggttga catcctctac 900
gagcgaattg tttgctcgac ccttcatcgc acactgtcat tataatgcat ctagctcggc 960
gacaagttta aaaaaggcag gctgcattat tccattttgc cgtggcggca tgggtgccca 1020
ttttatgagg tttgggctct tgggcagcga ccgagccagg ttgagtccct ctcgcccgtc 1080
gacaacgttc caaagcccat aagtggctaa taaacaactt gatggtacct gtacactgcc 1140

agttccttct tccccggccg aggtttacac gtgatggcca tggcttcgcg tttcaggctg 1200
acttcccatt ccgactttcc agagggtccg cggacgccgg gggttggctg cgtgaggccc 1260
acccccttgtt ccccgcgtcc cgacaaacac aattgcgtta cataaggggg aagccgcccc 1320
ccgttcagag tgcaaacatc tttcattata tttttcagtc gtcagcgaaa tcaagtatgt 1380
cgctgacagg catgaaggcc 1400

用于测试假定的Prototheca moriformis Aat01启动子/5'UTR的报告基因构建体的相关部分

<210> 67
<211> 6490
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 67

gctcttcggc cgccgccact cctgctcgag cgcgcccgac tcgcgctccg cctgcgcccg 60
cgcgtgcgcc gccagcgcct tggccttttc gccgcgctcg tgcgcgtcgc tgatgtccat 120
caccaggtcc atgaggtctg ccttgcgccg gctgagccac tgcttcgtcc gggcggccaa 180
gaggagcatg agggaggact cctggtccag ggtcctgacg tggtcgcggc tctgggagcg 240
ggccagcatc atctggctct gccgcaccga ggccgcctcc aactggtcct ccagcagccg 300
cagtcgccgc cgaccctggc agaggaagac aggtgagggg ggtatgaatt gtacagaaca 360
accacgagcc ttgtctaggc agaatcccta ccagtcatgg ctttacctgg atgacggcct 420
gcgaacagct gtccagcgac cctcgctgcc gccgcttctc ccgcacgctt ctttccagca 480
ccgtgatggc gcgagccagc gccgcacgct ggcgctgcgc ttcgccgatc tgaggacagt 540
cggggaactc tgatcagtct aaaccccctt gcgcgttagt gttgccatcc tttgcagacc 600
ggtgagagcc gacttgttgt gcgccacccc ccacaccacc tcctcccaga ccaattctgt 660
cacctttttg gcgaaggcat cggcctcggc ctgcagagag gacagcagtg cccagccgct 720
gggggttggc ggatgcacgc tcaggtaccc tttcttgcgc tatgacactt ccagcaaaag 780
gtagggcggg ctgcgagacg gcttcccggc gctgcatgca acaccgatga tgcttcgacc 840
ccccgaagct ccttcggggc tgcatgggcg ctccgatgcc gctccagggc gagcgctgtt 900
taaatagcca ggccccccgat tgcaaagaca ttatagcgag ctaccaaagc catattcaaa 960
cacctagatc actaccactt ctacacaggc cactcgagct tgtgatcgca ctccgctaag 1020
ggggcgcctc ttcctcttcg tttcagtcac aacccgcaaa cggcgcgcca tgctgctgca 1080
ggccttcctg ttcctgctgg ccggcttcgc cgccaagatc agcgcctcca tgacgaacga 1140
gacgtccgac cgccccctgg tgcacttcac ccccaacaag ggctggatga acgaccccaa 1200
cggcctgtgg tacgacgaga aggacgccaa gtggcacctg tacttccagt acaacccgaa 1260
cgacaccgtc tgggggacgc ccttgttctg gggccacgcc acgtccgacg acctgaccaa 1320
ctgggaggac cagcccatcg ccatcgcccc gaagcgcaac gactccggcg ccttctccgg 1380
ctccatggtg gtggactaca acaacacctc cggcttcttc aacgacacca tcgacccgcg 1440
ccagcgctgc gtggccatct ggacctacaa caccccggag tccgaggagc agtacatctc 1500

ctacagcctg gacggcggct acaccttcac cgagtaccag aagaaccccg tgctggccgc 1560
caactccacc cagttccgcg acccgaaggt cttctggtac gagccctccc agaagtggat 1620
catgaccgcg gccaagtccc aggactacaa gatcgagatc tactcctccg acgacctgaa 1680
gtcctggaag ctggagtccg cgttcgccaa cgagggcttc ctcggctacc agtacgagtg 1740
ccccggcctg atcgaggtcc ccaccgagca ggaccccagc aagtcctact gggtgatgtt 1800
catctccatc aaccccggcg ccccggccgg cggctccttc aaccagtact tcgtcggcag 1860
cttcaacggc acccacttcg aggccttcga caaccagtcc cgcgtggtgg acttcggcaa 1920
ggactactac gccctgcaga ccttcttcaa caccgacccg acctacggga gcgccctggg 1980
catcgcgtgg gcctccaact gggagtactc cgccttcgtg cccaccaacc cctggcgctc 2040
ctccatgtcc ctcgtgcgca agttctccct caacaccgag taccaggcca acccggagac 2100
ggagctgatc aacctgaagg ccgagccgat cctgaacatc agcaacgccg gcccctggag 2160
ccggttcgcc accaacacca cgttgacgaa ggccaacagc tacaacgtcg acctgtccaa 2220
cagcaccggc accctggagt tcgagctggt gtacgccgtc aacaccaccc agacgatctc 2280
caagtccgtg ttcgcggacc tctccctctg gttcaagggc ctggaggacc ccgaggagta 2340
cctccgcatg ggcttcgagg tgtccgcgtc ctccttcttc ctggaccgcg ggaacagcaa 2400
ggtgaagttc gtgaaggaga acccctactt caccaaccgc atgagcgtga acaaccagcc 2460
cttcaagagc gagaacgacc tgtcctacta caaggtgtac ggcttgctgg accagaacat 2520
cctggagctg tacttcaacg acggcgacgt cgtgtccacc aacacctact tcatgaccac 2580
cgggaacgcc ctgggctccg tgaacatgac gacgggggtg gacaacctgt tctacatcga 2640
caagttccag gtgcgcgagg tcaagtgaca attggcagca gcagctcgga tagtatcgac 2700
acactctgga cgctggtcgt gtgatggact gttgccgcca cacttgctgc cttgacctgt 2760
gaatatccct gccgctttta tcaaacagcc tcagtgtgtt tgatcttgtg tgtacgcgct 2820
tttgcgagtt gctagctgct tgtgctattt gcgaatacca cccccagcat ccccttccct 2880
cgtttcatat cgcttgcatc ccaaccgcaa cttatctacg ctgtcctgct atccctcagc 2940
gctgctcctg ctcctgctca ctgcccctcg cacagccttg gtttgggctc cgcctgtatt 3000
ctcctggtac tgcaacctgt aaaccagcac tgcaatgctg atgcacggga agtagtggga 3060
tgggaacaca aatggaggat cccgcgtctc gaacagagcg cgcagaggaa cgctgaaggt 3120
ctcgcctctg tcgcacctca gcgcggcata caccacaata accacctgac gaatgcgctt 3180
ggttcttcgt ccattagcga agcgtccggt tcacacacgt gccacgttgg cgaggtggca 3240
ggtgacaatg atcggtggag ctgatggtcg aaacgttcac agcctaggga tatcgaattc 3300
atcgaatttg taggaagcgg gcattggtac gggagtggga aagcgaaaaa acctccctcc 3360
ggcagtgcca tctgccggag tcgaacgttg atagggttct cgtgacaggg tgtgacctct 3420
cagccttgca tcaattaaac gctatagaca ttatcagtaa ccgtgaatcc cgcattggat 3480
gccacccgcg cgaccattgg ggacctgcat tacagatcta ggtgagatga cagcgaggca 3540
acttcggccc gcggcccagc ttgcggcgca ccaatattgg tcacgggaag ccacacaccg 3600
accataaatg aatacttgta agctatgtca accgatcaat ggcgtcgaaa gtgtgccacg 3660
aggatccatc tggcggggcg gcgtggcgca caagcgcagt cgcaatttct cggacccatc 3720
tgacctaggc ccagcgccgc gggagaaatc cccggcgggt cctccacgca gtaaccctaa 3780
tgagtatcga gcgccgacca tttacaccat cgccccccga aatccttccg acattattat 3840

tatcttttag atcttggaac agactctgcc aaccactagt atggccaccg catccacttt 3900
ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg gcgggctccg ggccccggcg 3960
cccagcgagg cccctccccg tgcgcgggcg cgcccccgac tggtccatgc tgttcgccgt 4020
gatcaccacc atcttctccg ccgccgagaa gcagtggacc aacctggagt ggaagcccaa 4080
gcccaacccc ccccagctgc tggacgacca cttcggcccc cacggcctgg tgttccgccg 4140
caccttcgcc atccgcagct acgaggtggg ccccgaccgc tccaccagca tcgtggccgt 4200
gatgaaccac ctgcaggagg ccgccctgaa ccacgccaag tccgtgggca tcctgggcga 4260
cggcttcggc accaccctgg agatgtccaa gcgcgacctg atctgggtgg tgaagcgcac 4320
ccacgtggcc gtggagcgct accccgcctg gggcgacacc gtggaggtgg agtgctgggt 4380
gggcgcctcc ggcaacaacg gccgccgcca cgacttcctg gtgcgcgact gcaagaccgg 4440
cgagatcctg acccgctgca cctccctgag cgtgatgatg aacacccgca cccgccgcct 4500
gagcaagatc cccgaggagg tgcgcggcga gatcggcccc gccttcatcg acaacgtggc 4560
cgtgaaggac gaggagatca agaagcccca gaagctgaac gactccaccg ccgactacat 4620
ccagggcggc ctgacccccc gctggaacga cctggacatc aaccagcacg tgaacaacat 4680
caagtacgtg gactggatcc tggagaccgt gcccgacagc atcttcgaga gccaccacat 4740
ctcctccttc accatcgagt accgccgcga gtgcaccatg gacagcgtgc tgcagtccct 4800
gaccaccgtg agcggcggct cctccgaggc cggcctggtg tgcgagcacc tgctgcagct 4860
ggagggcggc agcgaggtgc tgcgcgccaa gaccgagtgg cgccccaagc tgaccgactc 4920
cttccgcggc atcagcgtga tccccgccga gtccagcgtg atggactaca aggaccacga 4980
cggcgactac aaggaccacg acatcgacta caaggacgac gacgacaagt gactcgagtt 5040
aattaactcg aggcagcagc agctcggata gtatcgacac actctggacg ctggtcgtgt 5100
gatggactgt tgccgccaca cttgctgcct tgacctgtga atatccctgc cgcttttatc 5160
aaacagcctc agtgtgtttg atcttgtgtg tacgcgcttt tgcgagttgc tagctgcttg 5220
tgctatttgc gaataccacc cccagcatcc ccttccctcg tttcatatcg cttgcatccc 5280
aaccgcaact tatctacgct gtcctgctat ccctcagcgc tgctcctgct cctgctcact 5340
gcccctcgca cagccttggt ttgggctccg cctgtattct cctggtactg caacctgtaa 5400
accagcactg caatgctgat gcacgggaag tagtgggatg ggaacacaaa tggaaagctg 5460
tagagctcct tgttttccag aaggagttgc tccttgagcc tttcattctc agcctcgata 5520
acctccaaag ccgctctaat tgtggagggg gttcgaattt aaaagcttgg aatgttggtt 5580
cgtgcgtctg gaacaagccc agacttgttg ctcactggga aaaggaccat cagctccaaa 5640
aaacttgccg ctcaaaccgc gtacctctgc tttcgcgcaa tctgccctgt tgaaatcgcc 5700
accacattca tattgtgacg cttgagcagt ctgtaattgc ctcagaatgt ggaatcatct 5760
gccccctgtg cgagcccatg ccaggcatgt cgcgggcgag gacacccgcc actcgtacag 5820
cagaccatta tgctacctca caatagttca taacagtgac catatttctc gaagctcccc 5880
aacgagcacc tccatgctct gagtggccac cccccggccc tggtgcttgc ggagggcagg 5940
tcaaccggca tggggctacc gaaatccccg accggatccc accacccccg cgatgggaag 6000
aatctctccc cgggatgtgg gcccaccacc agcacaacct gctggcccag gcgagcgtca 6060
aaccatacca cacaaatatc cttggcatcg gccctgaatt ccttctgccg ctctgctacc 6120
cggtgcttct gtccgaagca ggggttgcta gggatcgctc cgagtccgca aacccttgtc 6180

gcgtggcggg gcttgttcga gcttgttcga gcttgaagag cctctagagt cgacctgcag 6240
gcatgcaagc ttggcgtaat catggtcata gctgtttcct gtgtgaaatt gttatccgct 6300
cacaattcca cacaacatac gagccggaag cataaagtgt aaagcctggg gtgcctaatg 6360
agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct 6420
gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg 6480
gcgctcttcc 6490

樟树硫酯酶正向引物

<210> 68
<211> 20
<212> DNA
<213> 樟树
<400> 68
taccccgcct ggggcgacac 20

樟树硫酯酶反向引物

<210> 69
<211> 20
<212> DNA
<213> 樟树
<400> 69
cttgctcagg cggcgggtgc 20

cd189正向引物

<210> 70
<211> 20
<212> DNA
<213> Prototheca moriformis
<400> 70
ccggatctcg gccagggcta 20

cd189反向引物

<210> 71
<211> 20
<212> DNA
<213> Prototheca moriformis
<400> 71
tcgatgtcgt gcaccgtcgc 20

Prototheca moriformis Δ12FAD敲除同源重组靶向构建体的5'供体DNA序列

<210> 72
<211> 1092
<212> DNA
<213> Prototheca moriformis

<400> 72

gctcttcggg tttgctcacc cgcgaggtcg acgcccagca tggctatcaa gacgaacagg 60
cagcctgtgg agaagcctcc gttcacgatc gggacgctgc gcaaggccat ccccgcgcac 120
tgtttcgagc gctcggcgct tcgtagcagc atgtacctgg cctttgacat cgcggtcatg 180
tccctgctct acgtcgcgtc gacgtacatc gaccctgcgc cggtgcctac gtgggtcaag 240
tatggcgtca tgtggccgct ctactggttc ttccaggtgt gtgtgagggt tgtggttgcc 300
cgtatcgagg tcctggtggc gcgcatgggg gagaaggcgc ctgtcccgct gaccccccg 360
gctaccctcc cggcaccttc cagggcgcct tcggcacggg tgtctgggtg tgcgcgcacg 420
agtgcggcca ccaggccttt tcctccagcc aggccatcaa cgacggcgtg ggcctggtgt 480
tccacagcct gctgctggtg ccctactact cctggaagca ctcgcaccgc cgccaccact 540
ccaacacggg gtgcctggac aaggacgagg tgtttgtgcc gccgcaccgc gcagtggcgc 600
acgagggcct ggagtgggag gagtggctgc ccatccgcat gggcaaggtg ctggtcaccc 660
tgaccctggg ctggccgctg tacctcatgt tcaacgtcgc ctcgcggccg tacccgcgct 720
tcgccaacca ctttgacccg tggtcgccca tcttcagcaa gcgcgaggta ccctttcttg 780
cgctatgaca cttccagcaa aaggtagggc gggctgcgag acggcttccc ggcgctgcat 840
gcaacaccga tgatgcttcg acccccccgaa gctccttcgg ggctgcatgg gcgctccgat 900
gccgctccag ggcgagcgct gtttaaatag ccaggccccc gattgcaaag acattatagc 960
gagctaccaa agccatattc aaacacctag atcactacca cttctacaca ggccactcga 1020
gcttgtgatc gcactccgct aaggggggcgc ctcttcctct tcgtttcagt cacaacccgc 1080
aaacggcgcg cc 1092

Prototheca moriformis Δ12FAD敲除同源重组靶向构建体的3'供体DNA序列

<210> 73

<211> 1400

<212> DNA

<213> Prototheca moriformis

<400> 73

caattggcag cagcagctcg gatagtatcg acacactctg gacgctggtc gtgtgatgga 60
ctgttgccgc cacacttgct gccttgacct gtgaatatcc ctgccgcttt tatcaaacag 120
cctcagtgtg tttgatcttg tgtgtacgcg cttttgcgag ttgctagctg cttgtgctat 180
ttgcgaatac cacccccagc atccccttcc ctcgtttcat atcgcttgca tcccaaccgc 240
aacttatcta cgctgtcctg ctatccctca gcgctgctcc tgctcctgct cactgccccct 300
cgcacagcct tggtttgggc tccgcctgta ttctcctggt actgcaacct gtaaaccagc 360
actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca caaatggagc atcgaggtgg 420
tcatctccga cctcgcgttg gtggcggtgc tcagcgggct cagcgtgctg ggccgcacca 480
tgggctgggc ctggctggtc aagacctacg tggtgcccta catgatcgtg aacatgtggc 540
tggtgctcat cacgctgctc cagcacacgc acccggccct gccgcactac ttcgagaagg 600
actgggactg gctacgcggc gccatggcca ccgtcgaccg ctccatgggc ccgcccttca 660
tggacagcat cctgcaccac atctccgaca cccacgtgct gcaccacctc ttcagcacca 720
tcccgcacta ccacgccgag gaggcctccg ccgccatccg gcccatcctg ggcaagtact 780

accaatccga cagccgctgg gtcggccgcg ccctgtggga ggactggcgc gactgccgct 840
acgtcgtccc cgacgcgccc gaggacgact ccgcgctctg gttccacaag tgagcgcgcc 900
tgcgcgagga cgcagaacaa cgctgccgcc gtgtcttttg cacgcgcgac tccggcgctt 960
cgctggtggc acccccataa agaaaccctc aattctgttt gtggaagaca cggtgtaccc 1020
ccacccaccc acctgcacct ctattattgg tattattgac gcgggagtgg gcgttgtacc 1080
ctacaacgta gcttctctag ttttcagctg gctcccacca ttgtaaagag cctctagagt 1140
cgacctgcag gcatgcaagc ttggcgtaat catggtcata gctgtttcct gtgtgaaatt 1200
gttatccgct cacaattcca cacaacatac gagccggaag cataaagtgt aaagcctggg 1260
gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc gctttccagt 1320
cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg agaggcggtt 1380
tgcgtattgg gcgctcttcc 1400

Prototheca moriformis Δ12FAD敲除同源重组靶向构建体

<210> 74

<211> 4091

<212> DNA

<213> Prototheca moriformis

<400> 74

gctcttcggg tttgctcacc cgcgaggtcg acgcccagca tggctatcaa gacgaacagg 60
cagcctgtgg agaagcctcc gttcacgatc gggacgctgc gcaaggccat ccccgcgcac 120
tgtttcgagc gctcggcgct tcgtagcagc atgtacctgg cctttgacat cgcggtcatg 180
tccctgctct acgtcgcgtc gacgtacatc gaccctgcgc cggtgcctac gtgggtcaag 240
tatggcgtca tgtggccgct ctactggttc ttccaggtgt gtgtgagggt tgtggttgcc 300
cgtatcgagg tcctggtggc gcgcatgggg gagaaggcgc ctgtcccgct gaccccccg 360
gctaccctcc cggcaccttc cagggcgcct tcggcacggg tgtctgggtg tgcgcgcacg 420
agtgcggcca ccaggccttt tcctccagcc aggccatcaa cgacggcgtg ggcctggtgt 480
tccacagcct gctgctggtg ccctactact cctggaagca ctcgcaccgc cgccaccact 540
ccaacacggg gtgcctggac aaggacgagg tgtttgtgcc gccgcaccgc gcagtggcgc 600
acgagggcct ggagtgggag gagtggctgc ccatccgcat gggcaaggtg ctggtcaccc 660
tgaccctggg ctggccgctg tacctcatgt tcaacgtcgc ctcgcggccg tacccgcgct 720
tcgccaacca ctttgacccg tggtcgccca tcttcagcaa gcgcgaggta ccctttcttg 780
cgctatgaca cttccagcaa aaggtagggc gggctgcgag acggcttccc ggcgctgcat 840
gcaacaccga tgatgcttcg acccccccgaa gctccttcgg ggctgcatgg gcgctccgat 900
gccgctccag ggcgagcgct gtttaaatag ccaggccccc gattgcaaag acattatagc 960
gagctaccaa agccatattc aaacacctag atcactacca cttctacaca ggccactcga 1020
gcttgtgatc gcactccgct aaggggcgc ctcttcctct tcgtttcagt cacaacccgc 1080
aaacggcgcg ccatgctgct gcaggccttc ctgttcctgc tggccggctt cgccgccaag 1140
atcagcgcct ccatgacgaa cgagacgtcc gaccgccccc tggtgcactt cacccccaac 1200
aagggctgga tgaacgaccc caacggcctg tggtacgacg agaaggacgc caagtggcac 1260
ctgtacttcc agtacaaccc gaacgacacc gtctgggggа cgcccttgtt ctggggccac 1320

gccacgtccg acgacctgac caactgggag gaccagccca tcgccatcgc cccgaagcgc 1380
aacgactccg gcgccttctc cggctccatg gtggtggact acaacaacac ctccggcttc 1440
ttcaacgaca ccatcgaccc gcgccagcgc tgcgtggcca tctggaccta caacaccccg 1500
gagtccgagg agcagtacat ctcctacagc ctggacggcg gctacacctt caccgagtac 1560
cagaagaacc ccgtgctggc cgccaactcc acccagttcc gcgacccgaa ggtcttctgg 1620
tacgagccct cccagaagtg gatcatgacc gcggccaagt cccaggacta caagatcgag 1680
atctactcct ccgacgacct gaagtcctgg aagctggagt ccgcgttcgc caacgagggc 1740
ttcctcggct accagtacga gtgccccggc ctgatcgagg tccccaccga gcaggacccc 1800
agcaagtcct actgggtgat gttcatctcc atcaaccccg gcgccccggc cggcggctcc 1860
ttcaaccagt acttcgtcgg cagcttcaac ggcacccact tcgaggcctt cgacaaccag 1920
tcccgcgtgg tggacttcgg caaggactac tacgccctgc agaccttctt caacaccgac 1980
ccgacctacg ggagcgccct gggcatcgcg tgggcctcca actgggagta ctccgccttc 2040
gtgcccacca acccctggcg ctcctccatg tccctcgtgc gcaagttctc cctcaacacc 2100
gagtaccagg ccaacccgga gacggagctg atcaacctga aggccgagcc gatcctgaac 2160
atcagcaacg ccggcccctg gagccggttc gccaccaaca ccacgttgac gaaggccaac 2220
agctacaacg tcgacctgtc caacagcacc ggcaccctgg agttcgagct ggtgtacgcc 2280
gtcaacacca cccagacgat ctccaagtcc gtgttcgcgg acctctccct ctggttcaag 2340
ggcctggagg accccgagga gtacctccgc atgggcttcg aggtgtccgc gtcctccttc 2400
ttcctggacc gcgggaacag caaggtgaag ttcgtgaagg agaaccccta cttcaccaac 2460
cgcatgagcg tgaacaacca gcccttcaag agcgagaacg acctgtccta ctacaaggtg 2520
tacggcttgc tggaccagaa catcctggag ctgtacttca acgacggcga cgtcgtgtcc 2580
accaacacct acttcatgac caccgggaac gccctgggct ccgtgaacat gacgacgggg 2640
gtggacaacc tgttctacat cgacaagttc caggtgcgcg aggtcaagtg acaattggca 2700
gcagcagctc ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg 2760
ccacacttgc tgccttgacc tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt 2820
gtttgatctt gtgtgtacgc gcttttgcga gttgctagct gcttgtgcta tttgcgaata 2880
ccacccccag catccccttc cctcgtttca tatcgcttgc atcccaaccg caacttatct 2940
acgctgtcct gctatccctc agcgctgctc ctgctcctgc tcactgcccc tcgcacagcc 3000
ttggtttggg ctccgcctgt attctcctgg tactgcaacc tgtaaaccag cactgcaatg 3060
ctgatgcacg ggaagtagtg ggatgggaac acaaatggag catcgaggtg gtcatctccg 3120
acctcgcgtt ggtggcggtg ctcagcgggc tcagcgtgct gggccgcacc atgggctggg 3180
cctggctggt caagacctac gtggtgccct acatgatcgt gaacatgtgg ctggtgctca 3240
tcacgctgct ccagcacacg cacccggccc tgccgcacta cttcgagaag gactgggact 3300
ggctacgcgg cgccatggcc accgtcgacc gctccatggg cccgcccttc atggacagca 3360
tcctgcacca catctccgac acccacgtgc tgcaccacct cttcagcacc atcccgcact 3420
accacgccga ggaggcctcc gccgccatcc ggcccatcct gggcaagtac taccaatccg 3480
acagccgctg ggtcggccgc gccctgtggg aggactggcg cgactgccgc tacgtcgtcc 3540
ccgacgcgcc cgaggacgac tccgcgctct ggttccacaa gtgagcgcgc ctgcgcgagg 3600
acgcagaaca acgctgccgc cgtgtctttt gcacgcgcga ctccggcgct tcgctggtgg 3660

caccccccata aagaaaccct caattctgtt tgtggaagac acggtgtacc cccacccacc 3720
cacctgcacc tctattattg gtattattga cgcgggagtg ggcgttgtac cctacaacgt 3780
agcttctcta gttttcagct ggctcccacc attgtaaaga gcctctagag tcgacctgca 3840
ggcatgcaag cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc 3900
tcacaattcc acacaacata cgagccggaa gcataaagtg taaagcctgg ggtgcctaat 3960
gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 4020
tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg 4080
ggcgctcttc c 4091

Prototheca moriformis SAD2A敲除同源重组靶向构建体的5'供体DNA序列

<210> 75
<211> 510
<212> DNA
<213> Prototheca moriformis
<400> 75

gctcttccgc ctggagctgg tgcagagcat gggtcagttt gcggaggaga gggtgctccc 60
cgtgctgcac cccgtggaca agctgtggca gccgcaggac ttcctgcccg accccgagtc 120
gcccgacttc gaggaccagg tggcggagct gcgcgcgcgc gccaaggacc tgcccgacga 180
gtactttgtg gtgctggtgg gcgacatgat cacggaggag gcgctgccga cctacatggc 240
catgctcaac accttggacg gtgtgcgcga cgacacgggc gcggctgacc acccgtgggc 300
gcgctggacg cggcagtggg tggccgagga gaaccggcac ggcgacctgc tgaacaagta 360
ctgttggctg acggggcgcg tcaacatgcg ggccgtggag gtgaccatca acaacctgat 420
caagagcggc atgaacccgc agacggacaa caacccttac ttgggcttcg tctacacctc 480
cttccaggag cgcgccacca agtaggtacc 510

Prototheca moriformis SAD2A敲除同源重组靶向构建体的3'供体DNA序列

<210> 76
<211> 1186
<212> DNA
<213> Prototheca moriformis
<400> 76

caattggcag cagcagctcg gatagtatcg acacactctg gacgctggtc gtgtgatgga 60
ctgttgccgc cacacttgct gccttgacct gtgaatatcc ctgccgcttt tatcaaacag 120
cctcagtgtg tttgatcttg tgtgtacgcg cttttgcgag ttgctagctg cttgtgctat 180
ttgcgaatac cacccccagc atcccccttcc ctcgtttcat atcgcttgca tcccaaccgc 240
aacttatcta cgctgtcctg ctatccctca gcgctgctcc tgctcctgct cactgccccct 300
cgcacagcct tggtttgggc tccgcctgta ttctcctggt actgcaacct gtaaaccagc 360
actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca caaatggaag gatcgtagag 420
ctccagccac ggcaacaccg cgcgcctggc ggccgagcac ggcgacaagg gcctgagcaa 480
gatctgcggg ctgatcgcca gcgacgaggg ccggcacgag atcgcctaca cgcgcatcgt 540
ggacgagttc ttccgcctcg accccgaggg cgccgtcgcc gcctacgcca acatgatgcg 600

caagcagatc accatgcccg cgcacctcat ggacgacatg ggccacggcg aggccaaccc 660
gggccgcaac ctcttcgccg acttctccgc cgtcgccgag aagatcgacg tctacgacgc 720
cgaggactac tgccgcatcc tggagcacct caacgcgcgc tggaaggtgg acgagcgcca 780
ggtcagcggc caggccgccg cggaccagga gtacgttctg ggcctgcccc agcgcttccg 840
gaaactcgcc gagaagaccg ccgccaagcg caagcgcgtc gcgcgcaggc ccgtcgcctt 900
ctcctggaga gaagagcctc tagagtcgac ctgcaggcat gcaagcttgg cgtaatcatg 960
gtcatagctg tttcctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc 1020
cggaagcata aagtgtaaag cctggggtgc ctaatgagtg agctaactca cattaattgc 1080
gttgcgctca ctgcccgctt tccagtcggg aaacctgtcg tgccagctgc attaatgaat 1140
cggccaacgc gcggggagag gcggtttgcg tattgggcgc tcttcc 1186

Prototheca moriformis SAD2A敲除同源重组靶向构建体

<210> 77
<211> 3615
<212> DNA
<213> Prototheca moriformis
<400> 77

gctcttccgc ctggagctgg tgcagagcat gggtcagttt gcggaggaga gggtgctccc 60
cgtgctgcac cccgtggaca agctgtggca gccgcaggac ttcctgcccg accccgagtc 120
gcccgacttc gaggaccagg tggcggagct gcgcgcgcgc gccaaggacc tgcccgacga 180
gtactttgtg gtgctggtgg gcgacatgat cacggaggag gcgctgccga cctacatggc 240
catgctcaac accttggacg gtgtgcgcga cgacacgggc gcggctgacc acccgtgggc 300
gcgctggacg cggcagtggg tggccgagga gaaccggcac ggcgacctgc tgaacaagta 360
ctgttggctg acggggcgcg tcaacatgcg ggccgtggag gtgaccatca acaacctgat 420
caagagcggc atgaacccgc agacggacaa caacccttac ttgggcttcg tctacacctc 480
cttccaggag cgcgccacca agtaggtacc ctttcttgcg ctatgacact tccagcaaaa 540
ggtagggcgg gctgcgagac ggcttcccgg cgctgcatgc aacaccgatg atgcttcgac 600
cccccgaagc tccttcgggg ctgcatgggc gctccgatgc cgctccaggg cgagcgctgt 660
ttaaatagcc aggccccccga ttgcaaagac attatagcga gctaccaaag ccatattcaa 720
acacctagat cactaccact tctacacagg ccactcgagc ttgtgatcgc actccgctaa 780
gggggcgcct cttcctcttc gtttcagtca caacccgcaa acggcgcgcc atgctgctgc 840
aggccttcct gttcctgctg gccggcttcg ccgccaagat cagcgcctcc atgacgaacg 900
agacgtccga ccgccccctg gtgcacttca cccccaacaa gggctggatg aacgacccca 960
acggcctgtg gtacgacgag aaggacgcca agtggcacct gtacttccag tacaacccga 1020
acgacaccgt ctgggggacg cccttgttct ggggccacgc cacgtccgac gacctgacca 1080
actgggagga ccagcccatc gccatcgccc cgaagcgcaa cgactccggc gccttctccg 1140
gctccatggt ggtggactac aacaacacct ccggcttctt caacgacacc atcgacccgc 1200
gccagcgctg cgtggccatc tggacctaca acaccccgga gtccgaggag cagtacatct 1260
cctacagcct ggacggcggc tacaccttca ccgagtacca gaagaacccc gtgctggccg 1320
ccaactccac ccagttccgc gacccgaagg tcttctggta cgagccctcc cagaagtgga 1380

tcatgaccgc ggccaagtcc caggactaca agatcgagat ctactcctcc gacgacctga 1440
agtcctggaa gctggagtcc gcgttcgcca acgagggctt cctcggctac cagtacgagt 1500
gccccggcct gatcgaggtc cccaccgagc aggaccccag caagtcctac tgggtgatgt 1560
tcatctccat caaccccggc gccccggccg gcggctcctt caaccagtac ttcgtcggca 1620
gcttcaacgg cacccacttc gaggccttcg acaaccagtc ccgcgtggtg gacttcggca 1680
aggactacta cgccctgcag accttcttca acaccgaccc gacctacggg agcgccctgg 1740
gcatcgcgtg ggcctccaac tgggagtact ccgccttcgt gcccaccaac ccctggcgct 1800
cctccatgtc cctcgtgcgc aagttctccc tcaacaccga gtaccaggcc aacccggaga 1860
cggagctgat caacctgaag gccgagccga tcctgaacat cagcaacgcc ggcccctgga 1920
gccggttcgc caccaacacc acgttgacga aggccaacag ctacaacgtc gacctgtcca 1980
acagcaccgg caccctggag ttcgagctgg tgtacgccgt caacaccacc cagacgatct 2040
ccaagtccgt gttcgcggac ctctccctct ggttcaaggg cctggaggac cccgaggagt 2100
acctccgcat gggcttcgag gtgtccgcgt cctccttctt cctggaccgc gggaacagca 2160
aggtgaagtt cgtgaaggag aacccctact tcaccaaccg catgagcgtg aacaaccagc 2220
ccttcaagag cgagaacgac ctgtcctact acaaggtgta cggcttgctg gaccagaaca 2280
tcctggagct gtacttcaac gacggcgacg tcgtgtccac caacacctac ttcatgacca 2340
ccgggaacgc cctgggctcc gtgaacatga cgacgggggt ggacaacctg ttctacatcg 2400
acaagttcca ggtgcgcgag gtcaagtgac aattggcagc agcagctcgg atagtatcga 2460
cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg ccttgacctg 2520
tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc 2580
ttttgcgagt tgctagctgc ttgtgctatt tgcgaatacc acccccagca tccccttccc 2640
tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc tatccctcag 2700
cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct ccgcctgtat 2760
tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg aagtagtggg 2820
atgggaacac aaatggaagg atcgtagagc tccagccacg gcaacaccgc gcgcctggcg 2880
gccgagcacg gcgacaaggg cctgagcaag atctgcgggc tgatcgccag cgacgagggc 2940
cggcacgaga tcgcctacac gcgcatcgtg gacgagttct tccgcctcga ccccgagggc 3000
gccgtcgccg cctacgccaa catgatgcgc aagcagatca ccatgcccgc gcacctcatg 3060
gacgacatgg gccacggcga ggccaacccg ggccgcaacc tcttcgccga cttctccgcc 3120
gtcgccgaga agatcgacgt ctacgacgcc gaggactact gccgcatcct ggagcacctc 3180
aacgcgcgct ggaaggtgga cgagcgccag gtcagcggcc aggccgccgc ggaccaggag 3240
tacgttctgg gcctgcccca gcgcttccgg aaactcgccg agaagaccgc cgccaagcgc 3300
aagcgcgtcg cgcgcaggcc cgtcgccttc tcctggagag aagagcctct agagtcgacc 3360
tgcaggcatg caagcttggc gtaatcatgg tcatagctgt ttcctgtgtg aaattgttat 3420
ccgctcacaa ttccacacaa catacgagcc ggaagcataa agtgtaaagc ctggggtgcc 3480
taatgagtga gctaactcac attaattgcg ttgcgctcac tgcccgcttt ccagtcggga 3540
aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt 3600
attgggcgct cttcc 3615

Prototheca moriformis SAD2B敲除同源重组靶向构建体的5’供体DNA序列

<210> 78
<211> 511
<212> DNA
<213> Prototheca moriformis
<400> 78
gctcttcccg cctggagctg gtgcagagca tggggcagtt tgcggaggag agggtgctcc 60
ccgtgctgca ccccgtggac aagctgtggc agccgcagga cttcctgccc gaccccgagt 120
cgcccgactt cgaggaccag gtggcggagc tgcgcgcgcg cgccaaggac ctgcccgacg 180
agtactttgt ggtgctggtg ggcgacatga tcacggagga ggcgctgccg acctacatgg 240
ccatgctcaa caccttggac ggtgtgcgcg acgacacggg cgcggctgac cacccgtggg 300
cgcgctggac gcggcagtgg gtggccgagg agaaccggca cggcgacctg ctgaacaagt 360
actgttggct gacggggcgc gtcaacatgc gggccgtgga ggtgaccatc aacaacctga 420
tcaagagcgg catgaacccg cagacggaca acaaccctta cttgggcttc gtctacacct 480
ccttccagga gcgcgccacc aagtaggtac c 511
Prototheca moriformis SAD2B敲除同源重组靶向构建体的3'供体DNA序列
<210> 79
<211> 761
<212> DNA
<213> Prototheca moriformis
<400> 79
cagccacggc aacaccgcgc gccttgcggc cgagcacggc gacaagaacc tgagcaagat 60
ctgcgggctg atcgccagcg acgagggccg gcacgagatc gcctacacgc gcatcgtgga 120
cgagttcttc cgcctcgacc ccgagggcgc cgtcgccgcc tacgccaaca tgatgcgcaa 180
gcagatcacc atgcccgcgc acctcatgga cgacatgggc cacggcgagg ccaacccggg 240
ccgcaacctc ttcgccgact tctccgcggt cgccgagaag atcgacgtct acgacgccga 300
ggactactgc cgcatcctgg agcacctcaa cgcgcgctgg aaggtggacg agcgccaggt 360
cagcggccag gccgccgcgg accaggagta cgtcctgggc ctgccccagc gcttccggaa 420
actcgccgag aagaccgccg ccaagcgcaa gcgcgtcgcg cgcaggcccg tcgccttctc 480
ctggagaaga gcctctagag tcgacctgca ggcatgcaag cttggcgtaa tcatggtcat 540
agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata cgagccggaa 600
gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta attgcgttgc 660
gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc 720
aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc c 761
Prototheca moriformis SAD2B敲除同源重组靶向构建体
<210> 80
<211> 3599
<212> DNA
<213> Prototheca moriformis
<400> 80

gctcttcccg cctggagctg gtgcagagca tggggcagtt tgcggaggag agggtgctcc 60
ccgtgctgca ccccgtggac aagctgtggc agccgcagga cttcctgccc gaccccgagt 120
cgcccgactt cgaggaccag gtggcggagc tgcgcgcgcg cgccaaggac ctgcccgacg 180
agtactttgt ggtgctggtg ggcgacatga tcacggagga ggcgctgccg acctacatgg 240
ccatgctcaa caccttggac ggtgtgcgcg acgacacggg cgcggctgac cacccgtggg 300
cgcgctggac gcggcagtgg gtggccgagg agaaccggca cggcgacctg ctgaacaagt 360
actgttggct gacggggcgc gtcaacatgc gggccgtgga ggtgaccatc aacaacctga 420
tcaagagcgg catgaacccg cagacggaca acaaccctta cttgggcttc gtctacacct 480
ccttccagga gcgcgccacc aagtaggtac cctttcttgc gctatgacac ttccagcaaa 540
aggtagggcg ggctgcgaga cggcttcccg gcgctgcatg caacaccgat gatgcttcga 600
ccccccgaag ctccttcggg gctgcatggg cgctccgatg ccgctccagg gcgagcgctg 660
tttaaatagc caggcccccg attgcaaaga cattatagcg agctaccaaa gccatattca 720
aacacctaga tcactaccac ttctacacag gccactcgag cttgtgatcg cactccgcta 780
agggggcgcc tcttcctctt cgtttcagtc acaacccgca aacggcgcgc catgctgctg 840
caggccttcc tgttcctgct ggccggcttc gccgccaaga tcagcgcctc catgacgaac 900
gagacgtccg accgcccccc ggtgcacttc acccccaaca agggctggat gaacgacccc 960
aacggcctgt ggtacgacga gaaggacgcc aagtggcacc tgtacttcca gtacaacccg 1020
aacgacaccg tctgggggac gcccttgttc tggggccacg ccacgtccga cgacctgacc 1080
aactgggagg accagcccat cgccatcgcc ccgaagcgca acgactccgg cgccttctcc 1140
ggctccatgg tggtggacta caacaacacc tccggcttct tcaacgacac catcgacccg 1200
cgccagcgct gcgtggccat ctggacctac aacaccccgg agtccgagga gcagtacatc 1260
tcctacagcc tggacggcgg ctacaccttc accgagtacc agaagaaccc cgtgctggcc 1320
gccaactcca cccagttccg cgacccgaag gtcttctggt acgagccctc ccagaagtgg 1380
atcatgaccg cggccaagtc ccaggactac aagatcgaga tctactcctc cgacgacctg 1440
aagtcctgga agctggagtc cgcgttcgcc aacgagggct tcctcggcta ccagtacgag 1500
tgccccggcc tgatcgaggt ccccaccgag caggacccca gcaagtccta ctgggtgatg 1560
ttcatctcca tcaaccccgg cgccccggcc ggcggctcct tcaaccagta cttcgtcggc 1620
agcttcaacg gcacccactt cgaggccttc gacaaccagt cccgcgtggt ggacttcggc 1680
aaggactact acgccctgca gaccttcttc aacaccgacc cgacctacgg gagcgccctg 1740
ggcatcgcgt gggcctccaa ctgggagtac tccgccttcg tgcccaccaa cccctggcgc 1800
tcctccatgt ccctcgtgcg caagttctcc ctcaacaccg agtaccaggc caacccggag 1860
acggagctga tcaacctgaa ggccgagccg atcctgaaca tcagcaacgc cggcccctgg 1920
agccggttcg ccaccaacac cacgttgacg aaggccaaca gctacaacgt cgacctgtcc 1980
aacagcaccg gcaccctgga gttcgagctg gtgtacgccg tcaacaccac ccagacgatc 2040
tccaagtccg tgttcgcgga cctctccctc tggttcaagg gcctggagga ccccgaggag 2100
tacctccgca tgggcttcga ggtgtccgcg tcctccttct tcctggaccg cgggaacagc 2160
aaggtgaagt tcgtgaagga gaacccctac ttcaccaacc gcatgagcgt gaacaaccag 2220
cccttcaaga gcgagaacga cctgtcctac tacaaggtgt acggcttgct ggaccagaac 2280
atcctggagc tgtacttcaa cgacggcgac gtcgtgtcca ccaacaccta cttcatgacc 2340

accgggaacg ccctgggctc cgtgaacatg acgacggggg tggacaacct gttctacatc 2400
gacaagttcc aggtgcgcga ggtcaagtga caattggcag cagcagctcg gatagtatcg 2460
acacactctg gacgctggtc gtgtgatgga ctgttgccgc cacacttgct gccttgacct 2520
gtgaatatcc ctgccgcttt tatcaaacag cctcagtgtg tttgatcttg tgtgtacgcg 2580
cttttgcgag ttgctagctg cttgtgctat ttgcgaatac cacccccagc atccccttcc 2640
ctcgtttcat atcgcttgca tcccaaccgc aacttatcta cgctgtcctg ctatccctca 2700
gcgctgctcc tgctcctgct cactgcccct cgcacagcct tggtttgggc tccgcctgta 2760
ttctcctggt actgcaacct gtaaaccagc actgcaatgc tgatgcacgg gaagtagtgg 2820
gatgggaaca caaatggaca gccacggcaa caccgcgcgc cttgcggccg agcacggcga 2880
caagaacctg agcaagatct gcgggctgat cgccagcgac gagggccggc acgagatcgc 2940
ctacacgcgc atcgtggacg agttcttccg cctcgacccc gagggcgccg tcgccgccta 3000
cgccaacatg atgcgcaagc agatcaccat gcccgcgcac ctcatggacg acatgggcca 3060
cggcgaggcc aacccgggcc gcaacctctt cgccgacttc tccgcggtcg ccgagaagat 3120
cgacgtctac gacgccgagg actactgccg catcctggag cacctcaacg cgcgctggaa 3180
ggtggacgag cgccaggtca gcggccaggc cgccgcggac caggagtacg tcctgggcct 3240
gccccagcgc ttccggaaac tcgccgagaa gaccgccgcc aagcgcaagc gcgtcgcgcg 3300
caggcccgtc gccttctcct ggagaagagc ctctagagtc gacctgcagg catgcaagct 3360
tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac 3420
acaacatacg agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac 3480
tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc 3540
tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttcc 3599

P.moriformis Aat02启动子/5'UTR: :原始小球藻硬脂酰ACP去饱和酶 PTS: : 樟树C14硫酯酶: : 普通小球藻硝酸盐还原酶3″UTR

<210> 81
<211> 2933
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 81

agagagcgga ggtggggttg tgaggtgggg ttgctgacca ggagctcgcg tcgccgagcg 60
cgactcgcac acggtccagt tacccccccc tccgcccaaa cgcaagcctc ccatcttgat 120
gcctttccgg ccacctatac tatttcttag ttcgctgtaa catccagacc gtcctgaata 180
ataacaatgc cctgtgtcaa gtgcattcct aaaaaaattc tgtcccaacc aacaatccca 240
cctgaaatac caccagccct gcccagtaca ctcttccaat accatctccc tacctccacg 300
cgcaagcgac ccccatgcgc gaccaggctc gaaagtgatt tatgacttga gacgagcgag 360
tggcggcgcg gtcgactgcc ttttcatcac gtgccgtacg tcggcgaccg ctagggcttt 420
gcacggcaac gcacggcttc gccaacccga ccagccagga cctcgactac tctaccgcga 480
attcgcctca agaagtcgcc aaatgtgcca tacaccattc cttacagcac tgttcaaact 540

tgatgccaat tttgacattc gggttgctcg ttggctgcgc ccacatcggc cgtgagtgca 600
gcaggcggga tcggacacgg aggacgcggc gtcacgcccc gaacgcagcc cgtaactcta 660
catcaacacg acgtgttgcg taatcccgcc cggctgcgca tcgtgccaac ccattcgcga 720
tggatggtcg gaaaatggtg tgccaactgc cctgagggag gctctcgcga aacgggcacg 780
tccctgaaac cgaaactgtg gccttgtcgt cggccacgca agcacgtgga ccctaaacac 840
caagaaaatc agtaaacaag gttgacatcc tctacgggcg aattgtttgc ccaacccttc 900
atcgcacact gccattataa tgcatctagc tcggcgacaa gtttagaaaa ggcaggctgc 960
attgttccat ttcgccgtgg cggcgtgggt gcccatttta cgaggtttgg gctcccgggc 1020
agcgaccgag ccaggtcgag tccctctcgc ccgtcgacaa tgttgcgaac cccacaagcg 1080
gctaacaaca acttgatggt acctgtacac tgccaattcc ttcttccccg gccgaggttt 1140
acacgtgatg gccatggctt cgcattcagg ccgacttccc attccgactt tccagagggt 1200
ccgcggacgc tgggggttgg ctgcctgagg cccacccttt gttccccgcg tcccgacaaa 1260
cacaattgcg ttacataagg gggagccgcc cccgttcaga gtgcagaaat ctttcactat 1320
attttccagt cgtcagcgaa atcaagtact agtatggcca ccgcatccac tttctggcg 1380
ttcaatgccc gctgcggcga cctgcgtcgc tcggcgggct ccgggccccg gcgcccagcg 1440
aggcccctcc ccgtgcgcgg gcgcgccccc gactggtcca tgctgttcgc cgtgatcacc 1500
accatcttct ccgccgccga gaagcagtgg accaacctgg agtggaagcc caagcccaac 1560
ccccccccagc tgctggacga ccacttcggc ccccacggcc tggtgttccg ccgcaccttc 1620
gccatccgca gctacgaggt gggccccgac cgctccacca gcatcgtggc cgtgatgaac 1680
cacctgcagg aggccgccct gaaccacgcc aagtccgtgg gcatcctggg cgacggcttc 1740
ggcaccaccc tggagatgtc caagcgcgac ctgatctggg tggtgaagcg cacccacgtg 1800
gccgtggagc gctaccccgc ctggggcgac accgtggagg tggagtgctg ggtgggcgcc 1860
tccggcaaca acggccgccg ccacgacttc ctggtgcgcg actgcaagac cggcgagatc 1920
ctgacccgct gcacctccct gagcgtgatg atgaacaccc gcacccgccg cctgagcaag 1980
atccccgagg aggtgcgcgg cgagatcggc cccgccttca tcgacaacgt ggccgtgaag 2040
gacgaggaga tcaagaagcc ccagaagctg aacgactcca ccgccgacta catccagggc 2100
ggcctgaccc cccgctggaa cgacctggac atcaaccagc acgtgaacaa catcaagtac 2160
gtggactgga tcctggagac cgtgcccgac agcatcttcg agagccacca catctcctcc 2220
ttcaccatcg agtaccgccg cgagtgcacc atggacagcg tgctgcagtc cctgaccacc 2280
gtgagcggcg gctcctccga ggccggcctg gtgtgcgagc acctgctgca gctggagggc 2340
ggcagcgagg tgctgcgcgc caagaccgag tggcgcccca agctgaccga ctccttccgc 2400
ggcatcagcg tgatccccgc cgagtccagc gtgatggact acaaggacca cgacggcgac 2460
tacaaggacc acgacatcga ctacaaggac gacgacgaca agtgactcga gttaattaac 2520
tcgaggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg tgtgatggac 2580
tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt atcaaacagc 2640
ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt 2700
tgcgaatacc acccccagca tccccttccc tcgtttcata tcgcttgcat cccaaccgca 2760
acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc actgcccctc 2820
gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg taaaccagca 2880

ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggaaag ctt 2933

5’6S区域基因组供体序列

<210> 82

<211> 726

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 82

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60

ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120

tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180

ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240

gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300

ggaagacagg tgagggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360

atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420

cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480

gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540

ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600

ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660

cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720

ggtacc 726

樟树硫酯酶的相关表达构建体(βtub::neo::nitred::βtub::C.camphora TE::nitred)

<210> 83

<211> 3635

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 83

ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60

cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc 120

gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc aggcccccga ttgcaaagac 180

attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240

ccactcgagc ttgtgatcgc actccgctaa gggggcgcct cttcctcttc gtttcagtca 300

caacccgcaa actctagaat atcaatgatc gagcaggacg gcctccacgc cggctccccc 360

gccgcctggg tggagcgcct gttcggctac gactgggccc agcagaccat cggctgctcc 420

gacgccgccg tgttccgcct gtccgcccag ggccgccccg tgctgttcgt gaagaccgac 480

ctgtccggcg ccctgaacga gctgcaggac gaggccgccc gcctgtcctg gctggccacc 540
accggcgtgc cctgcgccgc cgtgctggac gtggtgaccg aggccggccg cgactggctg 600
ctgctgggcg aggtgcccgg ccaggacctg ctgtcctccc acctggcccc cgccgagaag 660
gtgtccatca tggccgacgc catgcgccgc ctgcacaccc tggaccccgc cacctgcccc 720
ttcgaccacc aggccaagca ccgcatcgag cgcgcccgca cccgcatgga ggccggcctg 780
gtggaccagg acgacctgga cgaggagcac cagggcctgg cccccgccga gctgttcgcc 840
cgcctgaagg cccgcatgcc cgacggcgag gacctggtgg tgacccacgg cgacgcctgc 900
ctgcccaaca tcatggtgga gaacggccgc ttctccggct tcatcgactg cggccgcctg 960
ggcgtggccg accgctacca ggacatcgcc ctggccaccc gcgacatcgc cgaggagctg 1020
ggcggcgagt gggccgaccg cttcctggtg ctgtacggca tcgccgcccc cgactcccag 1080
cgcatcgcct tctaccgcct gctggacgag ttcttctgac aattggcagc agcagctcgg 1140
atagtatcga cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg 1200
ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt 1260
gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcgaatacc acccccagca 1320
tccccttccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc 1380
tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct 1440
ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 1500
aagtagtggg atgggaacac aaatggagga tcccgcgtct cgaacagagc gcgcagagga 1560
acgctgaagg tctcgcctct gtcgcacctc agcgcggcat acaccacaat aaccacctga 1620
cgaatgcgct tggttcttcg tccattagcg aagcgtccgg ttcacacacg tgccacgttg 1680
gcgaggtggc aggtgacaat gatcggtgga gctgatggtc gaaacgttca cagcctaggg 1740
atatcgaatt cctttcttgc gctatgacac ttccagcaaa aggtagggcg ggctgcgaga 1800
cggcttcccg gcgctgcatg caacaccgat gatgcttcga ccccccgaag ctccttcggg 1860
gctgcatggg cgctccgatg ccgctccagg gcgagcgctg tttaaatagc caggcccccg 1920
attgcaaaga cattatagcg agctaccaaa gccatattca aacacctaga tcactaccac 1980
ttctacacag gccactcgag cttgtgatcg cactccgcta agggggcgcc tcttcctctt 2040
cgtttcagtc acaacccgca aacactagta tggccaccgc atccactttc tcggcgttca 2100
atgcccgctg cggcgacctg cgtcgctcgg cgggctccgg gccccggcgc ccagcgaggc 2160
ccctccccgt gcgcgggcgc gcccccgact ggtccatgct gttcgccgtg atcaccacca 2220
tcttctccgc cgccgagaag cagtggacca acctggagtg gaagcccaag cccaacccc 2280
cccagctgct ggacgaccac ttcggccccc acggcctggt gttccgccgc accttcgcca 2340
tccgcagcta cgaggtgggc cccgaccgct ccaccagcat cgtggccgtg atgaaccacc 2400
tgcaggaggc cgccctgaac cacgccaagt ccgtgggcat cctgggcgac ggcttcggca 2460
ccaccctgga gatgtccaag cgcgacctga tctgggtggt gaagcgcacc cacgtggccg 2520
tggagcgcta ccccgcctgg ggcgacaccg tggaggtgga gtgctgggtg ggcgcctccg 2580
gcaacaacgg ccgccgccac gacttcctgg tgcgcgactg caagaccggc gagatcctga 2640
cccgctgcac ctccctgagc gtgatgatga acacccgcac ccgccgcctg agcaagatcc 2700
ccgaggaggt gcgcggcgag atcggccccg ccttcatcga caacgtggcc gtgaaggacg 2760
aggagatcaa gaagccccag aagctgaacg actccaccgc cgactacatc cagggcggcc 2820

tgacccccg ctggaacgac ctggacatca accagcacgt gaacaacatc aagtacgtgg 2880
actggatcct ggagaccgtg cccgacagca tcttcgagag ccaccacatc tcctccttca 2940
ccatcgagta ccgccgcgag tgcaccatgg acagcgtgct gcagtccctg accaccgtga 3000
gcggcggctc ctccgaggcc ggcctggtgt gcgagcacct gctgcagctg gagggcggca 3060
gcgaggtgct gcgcgccaag accgagtggc gccccaagct gaccgactcc ttccgcggca 3120
tcagcgtgat ccccgccgag tccagcgtga tggactacaa ggaccacgac ggcgactaca 3180
aggaccacga catcgactac aaggacgacg acgacaagtg actcgaggca gcagcagctc 3240
ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg ccacacttgc 3300
tgccttgacc tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt gtttgatctt 3360
gtgtgtacgc gcttttgcga gttgctagct gcttgtgcta tttgcgaata ccacccccag 3420
catccccttc cctcgtttca tatcgcttgc atcccaaccg caacttatct acgctgtcct 3480
gctatccctc agcgctgctc ctgctcctgc tcactgcccc tcgcacagcc ttggtttggg 3540
ctccgcctgt attctcctgg tactgcaacc tgtaaaccag cactgcaatg ctgatgcacg 3600
ggaagtagtg ggatgggaac acaaatggaa agctt 3635

3’6S区域基因组供体序列

<210> 84
<211> 749
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 84

gagctccttg ttttccagaa ggagttgctc cttgagcctt tcattctcag cctcgataac 60
ctccaaagcc gctctaattg tggaggggtt tcgaatttaa aagcttggaa tgttggttcg 120
tgcgtctgga acaagcccag acttgttgct cactgggaaa aggaccatca gctccaaaaa 180
acttgccgct caaaccgcgt acctctgctt tcgcgcaatc tgccctgttg aaatcgccac 240
cacattcata ttgtgacgct tgagcagtct gtaattgcct cagaatgtgg aatcatctgc 300
cccctgtgcg agcccatgcc aggcatgtcg cgggcgagga cacccgccac tcgtacagca 360
gaccattatg ctacctcaca atagttcata acagtgacca tatttctcga agctccccaa 420
cgagcacctc catgctctga gtggccaccc cccggccctg gtgcttgcgg agggcaggtc 480
aaccggcatg gggctaccga aatccccgac cggatcccac cacccccgcg atgggaagaa 540
tctctccccg ggatgtgggc ccaccaccag cacaacctgc tggcccaggc gagcgtcaaa 600
ccataccaca caaatatcct tggcatcggc cctgaattcc ttctgccgct ctgctacccg 660
gtgcttctgt ccgaagcagg ggttgctagg gatcgctccg agtccgcaaa cccttgtcgc 720
gtggcggggc ttgttcgagc ttgaagagc 749

Amt03正向引物

<210> 85
<211> 30
<212> DNA

<213> 人工序列
<220>
<223> 合成的引物
<400> 85
ggaggaattc ggccgacagg acgcgcgtca 30
Amt03反向引物
<210> 86
<211> 28
<212> DNA
<213> 人工序列
<220>
<223> 合成的引物
<400> 86
ggagactagt ggctgcgacc ggcctgtg 28
Amt02正向引物
<210> 87
<211> 31
<212> DNA
<213> 人工序列
<220>
<223> 合成的引物
<400> 87
ggaggaattc tcaccagcgg acaaagcacc g 31
Amt02反向引物
<210> 88
<211> 29
<212> DNA
<213> 人工序列
<220>
<223> 合成的引物
<400> 88
ggagactagt ggctgcgacc ggcctctgg 29
amt03启动子/UTR序列
<210> 89
<211> 1065
<212> DNA
<213> 人工序列
<220>
<223> 合成的启动子

<400> 89

ggccgacagg acgcgcgtca aaggtgctgg tcgtgtatgc cctggccggc aggtcgttgc 60
tgctgctggt tagtgattcc gcaaccctga ttttggcgtc ttattttggc gtggcaaacg 120
ctggcgcccg cgagccgggc cggcggcgat gcggtgcccc acggctgccg gaatccaagg 180
gaggcaagag cgcccgggtc agttgaaggg ctttacgcgc aaggtacagc cgctcctgca 240
aggctgcgtg gtggaattgg acgtgcaggt cctgctgaag ttcctccacc gcctcaccag 300
cggacaaagc accggtgtat caggtccgtg tcatccactc taaagagctc gactacgacc 360
tactgatggc cctagattct tcatcaaaaa cgcctgagac acttgcccag gattgaaact 420
ccctgaaggg accaccaggg gccctgagtt gttccttccc cccgtggcga gctgccagcc 480
aggctgtacc tgtgatcgag gctggcggga aaataggctt cgtgtgctca ggtcatggga 540
ggtgcaggac agctcatgaa acgccaacaa tcgcacaatt catgtcaagc taatcagcta 600
tttcctcttc acgagctgta attgtcccaa aattctggtc taccgggggt gatccttcgt 660
gtacgggccc ttccctcaac cctaggtatg cgcgcatgcg gtcgccgcgc aactcgcgcg 720
agggccgagg gtttgggacg ggccgtcccg aaatgcagtt gcacccggat gcgtggcacc 780
ttttttgcga taatttatgc aatggactgc tctgcaaaat tctggctctg tcgccaaccc 840
taggatcagc ggcgtaggat ttcgtaatca ttcgtcctga tggggagcta ccgactaccc 900
taatatcagc ccgactgcct gacgccagcg tccacttttg tgcacacatt ccattcgtgc 960
ccaagacatt tcattgtggt gcgaagcgtc cccagttacg ctcacctgtt tcccgacctc 1020
cttactgttc tgtcgacaga gcgggcccac aggccggtcg cagcc 1065

amt启动子/UTR::樟树硫酯酶表达构建体的相关序列

<210> 90
<211> 6981
<212> DNA
<213> 樟树
<400> 90

gctcttcggc cgccgccact cctgctcgag cgcgcccgac tcgcgctccg cctgcgcccg 60
cgcgtgcgcc gccagcgcct tggccttttc gccgcgctcg tgcgcgtcgc tgatgtccat 120
caccaggtcc atgaggtctg ccttgcgccg gctgagccac tgcttcgtcc gggcggccaa 180
gaggagcatg agggaggact cctggtccag ggtcctgacg tggtcgcggc tctgggagcg 240
ggccagcatc atctggctct gccgcaccga ggccgcctcc aactggtcct ccagcagccg 300
cagtcgccgc cgaccctggc agaggaagac aggtgagggg ggtatgaatt gtacagaaca 360
accacgagcc ttgtctaggc agaatcccta ccagtcatgg ctttacctgg atgacggcct 420
gcgaacagct gtccagcgac cctcgctgcc gccgcttctc ccgcacgctt ctttccagca 480
ccgtgatggc gcgagccagc gccgcacgct ggcgctgcgc ttcgccgatc tgaggacagt 540
cggggaactc tgatcagtct aaaccccctt gcgcgttagt gttgccatcc tttgcagacc 600
ggtgagagcc gacttgttgt gcgccacccc ccacaccacc tcctcccaga ccaattctgt 660
caccttttg gcgaaggcat cggcctcggc ctgcagagag gacagcagtg cccagccgct 720
gggggttggc ggatgcacgc tcaggtaccc tttcttgcgc tatgacactt ccagcaaaag 780
gtagggcggg ctgcgagacg gcttcccggc gctgcatgca acaccgatga tgcttcgacc 840

ccccgaagct ccttcggggc tgcatgggcg ctccgatgcc gctccagggc gagcgctgtt 900
taaatagcca ggccccccgat tgcaaagaca ttatagcgag ctaccaaagc catattcaaa 960
cacctagatc actaccactt ctacacaggc cactcgagct tgtgatcgca ctccgctaag 1020
ggggcgcctc ttcctcttcg tttcagtcac aacccgcaaa cggcgcgcca tgctgctgca 1080
ggccttcctg ttcctgctgg ccggcttcgc cgccaagatc agcgcctcca tgacgaacga 1140
gacgtccgac cgccccctgg tgcacttcac ccccaacaag ggctggatga acgaccccaa 1200
cggcctgtgg tacgacgaga aggacgccaa gtggcacctg tacttccagt acaacccgaa 1260
cgacaccgtc tgggggacgc ccttgttctg gggccacgcc acgtccgacg acctgaccaa 1320
ctgggaggac cagcccatcg ccatcgcccc gaagcgcaac gactccggcg ccttctccgg 1380
ctccatggtg gtggactaca acaacacctc cggcttcttc aacgacacca tcgacccgcg 1440
ccagcgctgc gtggccatct ggacctacaa caccccggag tccgaggagc agtacatctc 1500
ctacagcctg gacggcggct acaccttcac cgagtaccag aagaaccccg tgctggccgc 1560
caactccacc cagttccgcg acccgaaggt cttctggtac gagccctccc agaagtggat 1620
catgaccgcg gccaagtccc aggactacaa gatcgagatc tactcctccg acgacctgaa 1680
gtcctggaag ctggagtccg cgttcgccaa cgagggcttc ctcggctacc agtacgagtg 1740
ccccggcctg atcgaggtcc ccaccgagca ggaccccagc aagtcctact gggtgatgtt 1800
catctccatc aaccccggcg ccccggccgg cggctccttc aaccagtact tcgtcggcag 1860
cttcaacggc acccacttcg aggccttcga caaccagtcc cgcgtggtgg acttcggcaa 1920
ggactactac gccctgcaga ccttcttcaa caccgacccg acctacggga gcgccctggg 1980
catcgcgtgg gcctccaact gggagtactc cgccttcgtg cccaccaacc cctggcgctc 2040
ctccatgtcc ctcgtgcgca agttctccct caacaccgag taccaggcca acccggagac 2100
ggagctgatc aacctgaagg ccgagccgat cctgaacatc agcaacgccg gcccctggag 2160
ccggttcgcc accaacacca cgttgacgaa ggccaacagc tacaacgtcg acctgtccaa 2220
cagcaccggc accctggagt tcgagctggt gtacgccgtc aacaccaccc agacgatctc 2280
caagtccgtg ttcgcggacc tctccctctg gttcaagggc ctggaggacc ccgaggagta 2340
cctccgcatg ggcttcgagg tgtccgcgtc ctccttcttc ctggaccgcg ggaacagcaa 2400
ggtgaagttc gtgaaggaga acccctactt caccaaccgc atgagcgtga acaaccagcc 2460
cttcaagagc gagaacgacc tgtcctacta caaggtgtac ggcttgctgg accagaacat 2520
cctggagctg tacttcaacg acggcgacgt cgtgtccacc aacacctact tcatgaccac 2580
cgggaacgcc ctgggctccg tgaacatgac gacgggggtg gacaacctgt tctacatcga 2640
caagttccag gtgcgcgagg tcaagtgaca attggcagca gcagctcgga tagtatcgac 2700
acactctgga cgctggtcgt gtgatggact gttgccgcca cacttgctgc cttgacctgt 2760
gaatatccct gccgctttta tcaaacagcc tcagtgtgtt tgatcttgtg tgtacgcgct 2820
tttgcgagtt gctagctgct tgtgctattt gcgaatacca cccccagcat ccccttccct 2880
cgtttcatat cgcttgcatc ccaaccgcaa cttatctacg ctgtcctgct atccctcagc 2940
gctgctcctg ctcctgctca ctgcccctcg cacagccttg gtttgggctc cgcctgtatt 3000
ctcctggtac tgcaacctgt aaaccagcac tgcaatgctg atgcacggga agtagtggga 3060
tgggaacaca aatggaggat cccgcgtctc gaacagagcg cgcagaggaa cgctgaaggt 3120
ctcgcctctg tcgcacctca gcgcggcata caccacaata accacctgac gaatgcgctt 3180

ggttcttcgt ccattagcga agcgtccggt tcacacacgt gccacgttgg cgaggtggca 3240
ggtgacaatg atcggtggag ctgatggtcg aaacgttcac agcctaggga tatcgaattc 3300
ggccgacagg acgcgcgtca aaggtgctgg tcgtgtatgc cctggccggc aggtcgttgc 3360
tgctgctggt tagtgattcc gcaaccctga ttttggcgtc ttattttggc gtggcaaacg 3420
ctggcgcccg cgagccgggc cggcggcgat gcggtgcccc acggctgccg gaatccaagg 3480
gaggcaagag cgcccgggtc agttgaaggg ctttacgcgc aaggtacagc cgctcctgca 3540
aggctgcgtg gtggaattgg acgtgcaggt cctgctgaag ttcctccacc gcctcaccag 3600
cggacaaagc accggtgtat caggtccgtg tcatccactc taaagagctc gactacgacc 3660
tactgatggc cctagattct tcatcaaaaa cgcctgagac acttgcccag gattgaaact 3720
ccctgaaggg accaccaggg gccctgagtt gttccttccc cccgtggcga gctgccagcc 3780
aggctgtacc tgtgatcgag gctggcggga aaataggctt cgtgtgctca ggtcatggga 3840
ggtgcaggac agctcatgaa acgccaacaa tcgcacaatt catgtcaagc taatcagcta 3900
tttcctcttc acgagctgta attgtcccaa aattctggtc taccgggggt gatccttcgt 3960
gtacgggccc ttccctcaac cctaggtatg cgcgcatgcg gtcgccgcgc aactcgcgcg 4020
agggccgagg gtttgggacg ggccgtcccg aaatgcagtt gcacccggat gcgtggcacc 4080
ttttttgcga taatttatgc aatggactgc tctgcaaaat tctggctctg tcgccaaccc 4140
taggatcagc ggcgtaggat ttcgtaatca ttcgtcctga tggggagcta ccgactaccc 4200
taatatcagc ccgactgcct gacgccagcg tccacttttg tgcacacatt ccattcgtgc 4260
ccaagacatt tcattgtggt gcgaagcgtc cccagttacg ctcacctgtt tcccgacctc 4320
cttactgttc tgtcgacaga gcgggcccac aggccggtcg cagccactag tatggccacc 4380
gcatccactt tctcggcgtt caatgcccgc tgcggcgacc tgcgtcgctc ggcgggctcc 4440
gggcccccggc gcccagcgag gcccctcccc gtgcgcgggc gcgcccccga ctggtccatg 4500
ctgttcgccg tgatcaccac catcttctcc gccgccgaga agcagtggac caacctggag 4560
tggaagccca agcccaaccc cccccagctg ctggacgacc acttcggccc ccacggcctg 4620
gtgttccgcc gcaccttcgc catccgcagc tacgaggtgg gccccgaccg ctccaccagc 4680
atcgtggccg tgatgaacca cctgcaggag gccgccctga accacgccaa gtccgtgggc 4740
atcctgggcg acggcttcgg caccaccctg gagatgtcca agcgcgacct gatctgggtg 4800
gtgaagcgca cccacgtggc cgtggagcgc taccccgcct ggggcgacac cgtggaggtg 4860
gagtgctggg tgggcgcctc cggcaacaac ggccgccgcc acgacttcct ggtgcgcgac 4920
tgcaagaccg gcgagatcct gacccgctgc acctccctga gcgtgatgat gaacacccgc 4980
acccgccgcc tgagcaagat ccccgaggag gtgcgcggcg agatcggccc cgccttcatc 5040
gacaacgtgg ccgtgaagga cgaggagatc aagaagcccc agaagctgaa cgactccacc 5100
gccgactaca tccagggcgg cctgaccccc cgctggaacg acctggacat caaccagcac 5160
gtgaacaaca tcaagtacgt ggactggatc ctggagaccg tgcccgacag catcttcgag 5220
agccaccaca tctcctcctt caccatcgag taccgccgcg agtgcaccat ggacagcgtg 5280
ctgcagtccc tgaccaccgt gagcggcggc tcctccgagg ccggcctggt gtgcgagcac 5340
ctgctgcagc tggagggcgg cagcgaggtg ctgcgcgcca agaccgagtg gcgccccaag 5400
ctgaccgact ccttccgcgg catcagcgtg atccccgccg agtccagcgt gatggactac 5460
aaggaccacg acggcgacta caaggaccac gacatcgact acaaggacga cgacgacaag 5520

tgactcgagt taattaactc gaggcagcag cagctcggat agtatcgaca cactctggac 5580
gctggtcgtg tgatggactg ttgccgccac acttgctgcc ttgacctgtg aatatccctg 5640
ccgcttttat caaacagcct cagtgtgttt gatcttgtgt gtacgcgctt ttgcgagttg 5700
ctagctgctt gtgctatttg cgaataccac ccccagcatc cccttccctc gtttcatatc 5760
gcttgcatcc caaccgcaac ttatctacgc tgtcctgcta tccctcagcg ctgctcctgc 5820
tcctgctcac tgcccctcgc acagccttgg tttgggctcc gcctgtattc tcctggtact 5880
gcaacctgta aaccagcact gcaatgctga tgcacgggaa gtagtgggat gggaacacaa 5940
atggaaagct gtagagctcc ttgttttcca gaaggagttg ctccttgagc ctttcattct 6000
cagcctcgat aacctccaaa gccgctctaa ttgtggaggg ggttcgaatt taaaagcttg 6060
gaatgttggt tcgtgcgtct ggaacaagcc cagacttgtt gctcactggg aaaaggacca 6120
tcagctccaa aaaacttgcc gctcaaaccg cgtacctctg ctttcgcgca atctgccctg 6180
ttgaaatcgc caccacattc atattgtgac gcttgagcag tctgtaattg cctcagaatg 6240
tggaatcatc tgccccctgt gcgagcccat gccaggcatg tcgcgggcga ggacacccgc 6300
cactcgtaca gcagaccatt atgctacctc acaatagttc ataacagtga ccatatttct 6360
cgaagctccc caacgagcac ctccatgctc tgagtggcca ccccccggcc ctggtgcttg 6420
cggagggcag gtcaaccggc atggggctac cgaaatcccc gaccggatcc caccaccccc 6480
gcgatgggaa gaatctctcc ccgggatgtg ggcccaccac cagcacaacc tgctggccca 6540
ggcgagcgtc aaaccatacc acacaaatat ccttggcatc ggccctgaat tccttctgcc 6600
gctctgctac ccggtgcttc tgtccgaagc aggggttgct agggatcgct ccgagtccgc 6660
aaacccttgt cgcgtggcgg ggcttgttcg agcttgttcg agcttgaaga gcctctagag 6720
tcgacctgca ggcatgcaag cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat 6780
tgttatccgc tcacaattcc acacaacata cgagccggaa gcataaagtg taaagcctgg 6840
ggtgcctaat gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag 6900
tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt 6960
ttgcgtattg ggcgctcttc c 6981

密码子优化的樟树硫酯酶序列

<210> 91
<211> 1152
<212> DNA
<213> 樟树
<400> 91

atggccaccg catccacttt ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg 60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgcccccgac 120
tggtccatgc tgttcgccgt gatcaccacc atcttctccg ccgccgagaa gcagtggacc 180
aacctggagt ggaagcccaa gcccaacccc ccccagctgc tggacgacca cttcggcccc 240
cacggcctgg tgttccgccg caccttcgcc atccgcagct acgaggtggg ccccgaccgc 300
tccaccagca tcgtggccgt gatgaaccac ctgcaggagg ccgccctgaa ccacgccaag 360
tccgtgggca tcctgggcga cggcttcggc accaccctgg agatgtccaa gcgcgacctg 420
atctgggtgg tgaagcgcac ccacgtggcc gtggagcgct accccgcctg gggcgacacc 480

gtggaggtgg agtgctgggt gggcgcctcc ggcaacaacg gccgccgcca cgacttcctg 540
gtgcgcgact gcaagaccgg cgagatcctg acccgctgca cctccctgag cgtgatgatg 600
aacacccgca cccgccgcct gagcaagatc cccgaggagg tgcgcggcga gatcggcccc 660
gccttcatcg acaacgtggc cgtgaaggac gaggagatca agaagcccca gaagctgaac 720
gactccaccg ccgactacat ccagggcggc ctgacccccc gctggaacga cctggacatc 780
aaccagcacg tgaacaacat caagtacgtg gactggatcc tggagaccgt gcccgacagc 840
atcttcgaga gccaccacat ctcctccttc accatcgagt accgccgcga gtgcaccatg 900
gacagcgtgc tgcagtccct gaccaccgtg agcggcggct cctccgaggc cggcctggtg 960
tgcgagcacc tgctgcagct ggagggcggc agcgaggtgc tgcgcgccaa gaccgagtgg 1020
cgccccaagc tgaccgactc cttccgcggc atcagcgtga tccccgccga gtccagcgtg 1080
atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 1140
gacgacaagt ga 1152

密码子优化的加州月桂硫酯酶序列

<210> 92
<211> 1052
<212> DNA
<213> 加州月桂
<400> 92

ggcgcgcccc cgactggtcc atgctgttcg ccgtgatcac caccatcttc agcgccgccg 60
agaagcagtg gaccaacctg gagtggaagc ccaagcccaa gctgccccag ctgctggacg 120
accacttcgg cctgcacggc ctggtgttcc gccgcacctt cgccatccgc tcctacgagg 180
tgggccccga ccgcagcacc tccatcctgg ccgtgatgaa ccacatgcag gaggccaccc 240
tgaaccacgc caagagcgtg ggcatcctgg gcgacggctt cggcaccacc ctggagatgt 300
ccaagcgcga cctgatgtgg gtggtgcgcc gcacccacgt ggccgtggag cgctacccca 360
cctggggcga caccgtggag gtggagtgct ggatcggcgc cagcggcaac aacggcatgc 420
gccgcgactt cctggtgcgc gactgcaaga ccggcgagat cctgacccgc tgcacctccc 480
tgagcgtgct gatgaacacc cgcacccgcc gcctgagcac catccccgac gaggtgcgcg 540
gcgagatcgg ccccgccttc atcgacaacg tggccgtgaa ggacgacgag atcaagaagc 600
tgcagaagct gaacgactcc accgccgact acatccaggg cggcctgacc ccccgctgga 660
acgacctgga cgtgaaccag cacgtgaaca acctgaagta cgtggcctgg gtgttcgaga 720
ccgtgcccga cagcatcttc gagtcccacc acatcagctc cttcaccctg gagtaccgcc 780
gcgagtgcac ccgcgactcc gtgctgcgca gcctgaccac cgtgagcggc ggcagctccg 840
aggccggcct ggtgtgcgac cacctgctgc agctggaggg cggcagcgag gtgctgcgcg 900
cccgcaccga gtggcgcccc aagctgaccg actccttccg cggcatcagc gtgatccccg 960
ccgagccccg cgtgatggac tacaaggacc acgacggcga ctacaaggac cacgacatcg 1020
actacaagga cgacgacgac aagtgactcg ag 1052

密码子优化的美洲榆硫酯酶序列

<210> 93
<211> 1031

<212> DNA
<213> 美洲榆
<400> 93

ggcgcgccca gctgcccgac tggagcatgc tgctggccgc gatcaccacc ctgttcctgg 60
cggccgagaa gcagtggatg atgctggact ggaagcccaa gcgccccgac atgctggtgg 120
acccccttcgg cctgggccgc ttcgtgcagg acggcctggt gttccgcaac aacttcagca 180
tccgcagcta cgagatcggc gcggaccgca ccgccagcat cgagaccctg atgaaccacc 240
tgcaggagac cgccctgaac cacgtgaaga gcgtgggcct gctggaggac ggcctgggca 300
gcacccgcga gatgagcctg cgcaacctga tctgggtggt gaccaagatg caggtggcgg 360
tggaccgcta ccccacctgg ggcgacgagg tgcaggtgag cagctgggcg accgccatcg 420
gcaagaacgg catgcgccgc gagtggatcg tgaccgactt ccgcaccggc gagaccctgc 480
tgcgcgccac cagcgtgtgg gtgatgatga acaagctgac ccgccgcatc agcaagatcc 540
ccgaggaggt gtggcacgag atcggcccca gcttcatcga cgcgcccccc ctgcccaccg 600
tggaggacga cggccgcaag ctgacccgct tcgacgagag cagcgccgac ttcatccgca 660
agggcctgac cccccgctgg agcgacctgg acatcaacca gcacgtgaac aacgtgaagt 720
acatcggctg gctgctggag agcgcgcccc ccgagatcca cgagagccac gagatcgcca 780
gcctgaccct ggagtaccgc cgcgagtgcg gccgcgacag cgtgctgaac agcgccacca 840
aggtgagcga cagcagccag ctgggcaaga gcgccgtgga gtgcaaccac ctggtgcgcc 900
tgcagaacgg cggcgagatc gtgaagggcc gcaccgtgtg gcgccccaag cgccccctgt 960
acaacgacgg cgccgtggtg gacgtgcccg ccaagaccag cgatgacgat gacaagctgg 1020
gatgactcga g 1031

密码子优化的萼距花硫酯酶序列

<210> 94
<211> 1182
<212> DNA
<213> 萼距花
<400> 94

actagtatgg ctatcaagac gaacaggcag cctgtggaga agcctccgtt cacgatcggg 60
acgctgcgca aggccatccc cgcgcactgt ttcgagcgct cggcgcttcg tgggcgcgcc 120
cagctgcccg actggagccg cctgctgacc gccatcacca ccgtgttcgt gaagtccaag 180
cgccccgaca tgcacgaccg caagtccaag cgccccgaca tgctggtgga cagcttcggc 240
ctggagtcca ccgtgcagga cggcctggtg ttccgccagt ccttctccat ccgctcctac 300
gagatcggca ccgaccgcac cgccagcatc gagaccctga tgaaccacct gcaggagacc 360
tccctgaacc actgcaagag caccggcatc ctgctggacg gcttcggccg caccctggag 420
atgtgcaagc gcgacctgat ctgggtggtg attaagatgc agatcaaggt gaaccgctac 480
cccgcctggg gcgacaccgt ggagatcaac acccgcttca gccgcctggg caagatcggc 540
atgggccgcg actggctgat ctccgactgc aacaccggcg agatcctggt gcgcgccacc 600
agcgcctacg ccatgatgaa ccagaagacc cgccgcctgt ccaagctgcc ctacgaggtg 660
caccaggaga tcgtgccccc gttcgtggac agccccgtga tcgaggactc cgacctgaag 720

gtgcacaagt tcaaggtgaa gaccggcgac agcatccaga agggcctgac ccccggctgg 780
aacgacctgg acgtgaacca gcacgtgtcc aacgtgaagt acatcggctg gatcctggag 840
agcatgccca ccgaggtgct ggagacccag gagctgtgct ccctggccct ggagtaccgc 900
cgcgagtgcg gccgcgactc cgtgctggag agcgtgaccg ccatggaccc cagcaaggtg 960
ggcgtgcgct cccagtacca gcacctgctg cgcctggagg acggcaccgc catcgtgaac 1020
ggcgccaccg agtggcgccc caagaacgcc ggcgccaacg gcgccatctc caccggcaag 1080
accagcaacg gcaactccgt gtccatggac tacaaggacc acgacggcga ctacaaggac 1140
cacgacatcg actacaagga cgacgacgac aagtgactcg ag 1182

加州月桂硫酯酶正向引物

<210> 95
<211> 20
<212> DNA
<213> 加州月桂
<400> 95

ctgggcgacg gcttcggcac 20

加州月桂硫酯酶反向引物

<210> 96
<211> 20
<212> DNA
<213> 加州月桂
<400> 96

aagtcgcggc gcatgccgtt 20

美洲榆硫酯酶正向引物

<210> 97
<211> 20
<212> DNA
<213> 美洲榆
<400> 97

cccagctgct cacctgcacc 20

美洲榆硫酯酶反向引物

<210> 98
<211> 26
<212> DNA
<213> 美洲榆
<400> 98

cacccaaggc caacggcagc gccgtg 26

萼距花硫酯酶正向引物

<210> 99
<211> 20

<212> DNA
<213> 萼距花
<400> 99
taccccgcct ggggcgacac 20
萼距花硫酯酶反向引物
<210> 100
<211> 20
<212> DNA
<213> 萼距花
<400> 100
agcttggaca ggcggcgggt 20
FAD引物1
<210> 101
<211> 20
<212> DNA
<213> Prototheca moriformis
<400> 101
tcacttcatg ccggcggtcc 20
FAD引物2
<210> 102
<211> 20
<212> DNA
<213> Prototheca moriformis
<400> 102
gcgctcctgc ttggctcgaa 20
pSZ1124(FAD2B)5'基因组靶向序列
<210> 103
<211> 733
<212> DNA
<213> 人工序列
<220>
<223> 合成的多核苷酸
<400> 103
gctcttcgag acgtggtctg aatcctccag gcgggtttcc ccgagaaaga aagggtgccg 60
atttcaaagc agagccatgt gccgggccct gtggcctgtg ttggcgccta tgtagtcacc 120
ccccctcacc caattgtcgc cagtttgcgc aatccataaa ctcaaaactg cagcttctga 180
gctgcgctgt tcaagaacac ctctggggtt tgctcacccg cgaggtcgac gcccagcatg 240
gctatcaaga cgaacaggca gcctgtggag aagcctccgt tcacgatcgg gacgctgcgc 300
aaggccatcc ccgcgcactg tttcgagcgc tcggcgcttc gtagcagcat gtacctggcc 360

tttgacatcg cggtcatgtc cctgctctac gtcgcgtcga cgtacatcga ccctgcgccg 420
gtgcctacgt gggtcaagta tggcgtcatg tggccgctct actggttctt ccaggtgtgt 480
gtgagggttg tggttgcccg tatcgaggtc ctggtggcgc gcatggggga gaaggcgcct 540
gtcccgctga ccccccggc taccctcccg gcaccttcca gggcgccttc ggcacgggtg 600
tctgggtgtg cgcgcacgag tgcggccacc aggccttttc ctccagccag gccatcaacg 660
acggcgtggg cctggtgttc cacagcctgc tgctggtgcc ctactactcc tggaagcact 720
cgcaccgggt acc 733

pSZ1124(FAD2B)3'基因组靶向序列

<210> 104
<211> 739
<212> DNA
<213> 人工序列
<220>
<223> 合成的多核苷酸
<400> 104

ccgccaccac tccaacacgg ggtgcctgga caaggacgag gtgtttgtgc cgccgcaccg 60
cgcagtggcg cacgagggcc tggagtggga ggagtggctg cccatccgca tgggcaaggt 120
gctggtcacc ctgaccctgg gctggccgct gtacctcatg ttcaacgtcg cctcgcggcc 180
gtacccgcgc ttcgccaacc actttgaccc gtggtcgccc atcttcagca agcgcgagcg 240
catcgaggtg gtcatctccg acctggcgct ggtggcggtg ctcagcgggc tcagcgtgct 300
gggccgcacc atgggctggg cctggctggt caagacctac gtggtgccct acctgatcgt 360
gaacatgtgg ctcgtgctca tcacgctgct ccagcacacg cacccggcgc tgccgcacta 420
cttcgagaag gactgggact ggctgcgcgg cgccatggcc accgtggacc gctccatggg 480
cccgcccttc atggacaaca tcctgcacca catctccgac acccacgtgc tgcaccacct 540
cttcagcacc atcccgcact accacgccga ggaggcctcc gccgccatca ggcccatcct 600
gggcaagtac taccagtccg acagccgctg ggtcggccgc gccctgtggg aggactggcg 660
cgactgccgc tacgtcgtcc cggacgcgcc cgaggacgac tccgcgctct ggttccacaa 720
gtgagtgagt gagaagagc 739

酿酒酵母suc2盒

<210> 105
<211> 2327
<212> DNA
<213> 酿酒酵母
<400> 105

ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60
cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc 120
gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc aggcccccga ttgcaaagac 180
attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240
ccactcgagc ttgtgatcgc actccgctaa gggggcgcct cttcctcttc gtttcagtca 300

caacccgcaa acggcgcgcc atgctgctgc aggccttcct gttcctgctg gccggcttcg 360
ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg gtgcacttca 420
cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag aaggacgcca 480
agtggcacct gtacttccag tacaacccga acgacaccgt ctgggggacg cccttgttct 540
ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc gccatcgccc 600
cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac aacaacacct 660
ccggcttctt caacgacacc atcgacccgc gccagcgctg cgtggccatc tggacctaca 720
acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc tacaccttca 780
ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc gacccgaagg 840
tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc caggactaca 900
agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc gcgttcgcca 960
acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc cccaccgagc 1020
aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc gccccggccg 1080
gcggctcctt caaccagtac ttcgtcggca gcttcaacgg cacccacttc gaggccttcg 1140
acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag accttcttca 1200
acaccgaccc gacctacggg agcgccctgg gcatcgcgtg ggcctccaac tgggagtact 1260
ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc aagttctccc 1320
tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag gccgagccga 1380
tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc acgttgacga 1440
aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag ttcgagctgg 1500
tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac ctctccctct 1560
ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag gtgtccgcgt 1620
cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag aacccctact 1680
tcaccaaccg catgagcgtg aacaaccagc ccttcaagag cgagaacgac ctgtcctact 1740
acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac gacggcgacg 1800
tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc gtgaacatga 1860
cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag gtcaagtgac 1920
aattggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg tgtgatggac 1980
tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt atcaaacagc 2040
ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt 2100
tgcgaatacc acccccagca tccccttccc tcgtttcata tcgcttgcat cccaaccgca 2160
acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc actgcccctc 2220
gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg taaaccagca 2280
ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatgga 2327

pSZ1125(FAD2C)5'基因组靶向序列

<210> 106

<211> 735

<212> DNA

<213> 人工序列

<220>
<223> 合成的多核苷酸
<400> 106
gctcttcgag gggctggtct gaatccttca ggcgggtgtt acccgagaaa gaaagggtgc 60
cgatttcaaa gcagacccat gtgccgggcc ctgtggcctg tgttggcgcc tatgtagtca 120
ccccccctca cccaattgtc gccagtttgc gcactccata aactcaaaac agcagcttct 180
gagctgcgct gttcaagaac acctctgggg tttgctcacc cgcgaggtcg acgcccagca 240
tggctatcaa gacgaacagg cagcctgtgg agaagcctcc gttcacgatc gggacgctgc 300
gcaaggccat ccccgcgcac tgtttcgagc gctcggcgct tcgtagcagc atgtacctgg 360
cctttgacat cgcggtcatg tccctgctct acgtcgcgtc gacgtacatc gaccctgcac 420
cggtgcctac gtgggtcaag tacggcatca tgtggccgct ctactggttc ttccaggtgt 480
gtttgagggt tttggttgcc cgtattgagg tcctggtggc gcgcatggag gagaaggcgc 540
ctgtcccgct gaccccccg gctaccctcc cggcaccttc cagggcgcct tcggcacggg 600
tgtctgggtg tgcgcgcacg agtgcggcca ccaggccttt tcctccagcc aggccatcaa 660
cgacggcgtg ggcctggtgt tccacagcct gctgctggtg ccctactact cctggaagca 720
ctcgcaccgg gtacc 735
pSZ1125(FAD2C)3'基因组靶向序列
<210> 107
<211> 739
<212> DNA
<213> 人工序列
<220>
<223> 合成的多核苷酸
<400> 107
ccgccaccac tccaacacgg ggtgcctgga caaggacgag gtgtttgtgc cgccgcaccg 60
cgcagtggcg cacgagggcc tggagtggga ggagtggctg cccatccgca tgggcaaggt 120
gctggtcacc ctgaccctgg gctggccgct gtacctcatg ttcaacgtcg cctcgcggcc 180
gtacccgcgc ttcgccaacc actttgaccc gtggtcgccc atcttcagca agcgcgagcg 240
catcgaggtg gtcatctccg acctggcgct ggtggcggtg ctcagcgggc tcagcgtgct 300
gggccgcacc atgggctggg cctggctggt caagacctac gtggtgccct acctgatcgt 360
gaacatgtgg ctcgtgctca tcacgctgct ccagcacacg cacccggcgc tgccgcacta 420
cttcgagaag gactgggact ggctgcgcgg cgccatggcc accgtggacc gctccatggg 480
cccgcccttc atggacaaca tcctgcacca catctccgac acccacgtgc tgcaccacct 540
cttcagcacc atcccgcact accacgccga ggaggcctcc gccgccatca ggcccatcct 600
gggcaagtac taccagtccg acagccgctg ggtcggccgc gccctgtggg aggactggcg 660
cgactgccgc tacgtcgtcc cggacgcgcc cgaggacgac tccgcgctct ggttccacaa 720
gtgagtgagt gagaagagc 739
卡尔酵母MEL1氨基酸序列
<210> 108

<211> 471
<212> PRT
<213> 卡尔酵母
<400> 108

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

275 280 285
Pro Leu Ile Ile Gly Ala Asn Val Asn Asn Leu Lys Ala Ser Ser Tyr
290 295 300
Ser Ile Tyr Ser Gln Ala Ser Val Ile Ala Ile Asn Gln Asp Ser Asn
305 310 315 320
Gly Ile Pro Ala Thr Arg Val Trp Arg Tyr Tyr Val Ser Asp Thr Asp
325 330 335
Glu Tyr Gly Gln Gly Glu Ile Gln Met Trp Ser Gly Pro Leu Asp Asn
340 345 350
Gly Asp Gln Val Val Ala Leu Leu Asn Gly Gly Ser Val Ser Arg Pro
355 360 365
Met Asn Thr Thr Leu Glu Glu Ile Phe Phe Asp Ser Asn Leu Gly Ser
370 375 380
Lys Lys Leu Thr Ser Thr Trp Asp Ile Tyr Asp Leu Trp Ala Asn Arg
385 390 395 400
Val Asp Asn Ser Thr Ala Ser Ala Ile Leu Gly Arg Asn Lys Thr Ala
405 410 415
Thr Gly Ile Leu Tyr Asn Ala Thr Glu Gln Ser Tyr Lys Asp Gly Leu
420 425 430
Ser Lys Asn Asp Thr Arg Leu Phe Gly Gln Lys Ile Gly Ser Leu Ser
435 440 445
Pro Asn Ala Ile Leu Asn Thr Thr Val Pro Ala His Gly Ile Ala Phe
450 455 460
Tyr Arg Leu Arg Pro Ser Ser
465 470

卡尔酵母MEL1天然信号肽氨基酸序列

<210> 109
<211> 18
<212> PRT
<213> 卡尔酵母
<400> 109
Met Phe Ala Phe Tyr Phe Leu Thr Ala Cys Ile Ser Leu Lys Gly Val
1 5 10 15
Phe Gly

卡尔酵母MEL1表达盒的相关序列

<210> 110
<211> 3183
<212> DNA
<213> 卡尔酵母

〈400〉 110

gcggccgcgt ggacgagggc tacaaccccg cctacggcgc gcgcccgctg cgccgcgcca 60
tcatgcgcct gctggaggac gcgctggccg agcgcatgct cgccggcgac gtcaaggagg 120
gcgactcggt catcatggac gtcgatggcg atggcgccat cagcgtcctc aacggagacc 180
gcacccacac caccaccatc gactcctccc cggcgggcat ctcgtagacg cgagagggag 240
ggagggagga ttttcagagg gagatgagac gaggactggg ccgggggcct tggggcgctg 300
ctggagcgtg gtgagagcgc ggcggacgtg ccttttcttc ttccgtgcgc gcgctcttgg 360
ccattgatcc ccgattcgcg cccgcatccc cccactgccc ccatcatctt gcctgttgtc 420
gtggcactga cataaacccc ctgcgctgcg ctgctccgct actattgata taggtctcac 480
gcgccaatct tttttgctcc gggtaaccgt ctggacgcca gaattccttt cttgcgctat 540
gacacttcca gcaaaaggta gggcgggctg cgagacggct tcccggcgct gcatgcaaca 600
ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc atgggcgctc cgatgccgct 660
ccagggcgag cgctgtttaa atagccaggc ccccgattgc aaagacatta tagcgagcta 720
ccaaagccat attcaaacac ctagatcact accacttcta cacaggccac tcgagcttgt 780
gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt cagtcacaac ccgcaaacac 840
tagtatgttc gcgttctact tcctgacggc ctgcatctcc ctgaagggcg tgttcggcgt 900
ctccccctcc tacaacggcc tgggcctgac gccccagatg ggctgggaca actggaacac 960
gttcgcctgc gacgtctccg agcagctgct gctggacacg gccgaccgca tctccgacct 1020
gggcctgaag gacatgggct acaagtacat catcctggac gactgctggt cctccggccg 1080
cgactccgac ggcttcctgg tcgccgacga gcagaagttc cccaacggca tgggccacgt 1140
cgccgaccac ctgcacaaca actccttcct gttcggcatg tactcctccg cgggcgagta 1200
cacgtgcgcc ggctaccccg gctccctggg ccgcgaggag gaggacgccc agttcttcgc 1260
gaacaaccgc gtggactacc tgaagtacga caactgctac aacaagggcc agttcggcac 1320
gcccgagatc tcctaccacc gctacaaggc catgtccgac gccctgaaca agacgggccg 1380
ccccatcttc tactccctgt gcaactgggg ccaggacctg accttctact ggggctccgg 1440
catcgcgaac tcctggcgca tgtccggcga cgtcacggcg gagttcacgc gccccgactc 1500
ccgctgcccc tgcgacggcg acgagtacga ctgcaagtac gccggcttcc actgctccat 1560
catgaacatc ctgaacaagg ccgcccccat gggccagaac gcgggcgtcg gcggctggaa 1620
cgacctggac aacctggagg tcggcgtcgg caacctgacg gacgacgagg agaaggcgca 1680
cttctccatg tgggccatgg tgaagtcccc cctgatcatc ggcgcgaacg tgaacaacct 1740
gaaggcctcc tcctactcca tctactccca ggcgtccgtc atcgccatca accaggactc 1800
caacggcatc cccgccacgc gcgtctggcg ctactacgtg tccgacacgg acgagtacgg 1860
ccagggcgag atccagatgt ggtccggccc cctggacaac ggcgaccagg tcgtggcgct 1920
gctgaacggc ggctccgtgt cccgccccat gaacacgacc ctggaggaga tcttcttcga 1980
ctccaacctg ggctccaaga agctgacctc cacctgggac atctacgacc tgtgggcgaa 2040
ccgcgtcgac aactccacgg cgtccgccat cctgggccgc aacaagaccg ccaccggcat 2100
cctgtacaac gccaccgagc agtcctacaa ggacggcctg tccaagaacg acacccgcct 2160
gttcggccag aagatcggct ccctgtcccc caacgcgatc ctgaacacga ccgtccccgc 2220
ccacggcatc gcgttctacc gcctgcgccc ctcctcctga caattgaagc agcagcagct 2280

cggatagtat cgacacactc tggacgctgg tcgtgtgatg gactgttgcc gccacacttg 2340
ctgccttgac ctgtgaatat ccctgccgct tttatcaaac agcctcagtg tgtttgatct 2400
tgtgtgtacg cgcttttgcg agttgctagc tgcttgtgct atttgcgaat accaccccca 2460
gcatcccctt ccctcgtttc atatcgcttg catcccaacc gcaacttatc tacgctgtcc 2520
tgctatccct cagcgctgct cctgctcctg ctcactgccc ctcgcacagc cttggtttgg 2580
gctccgcctg tattctcctg gtactgcaac ctgtaaacca gcactgcaat gctgatgcac 2640
gggaagtagt gggatgggaa cacaaatgga aagcttgagc tcagaatagt atcgggtgat 2700
gcgaagtcag aaccaggcag ggcctgtcgc ctgaggtggc aacgatggga agcaatcaat 2760
ctgggtacag tcgtccgcac gatcccgtga tctcccccac cgacacctat ccccgcccat 2820
cccggcccac cctttcagtc ccctcagcat gcattgtgca ccgcgacaaa gcatgtctgc 2880
tcgtgcactg gttcaggcca cggcgcaccg agtcctcgcc cttcgcagag tgatcaccct 2940
ccccggaacc agccacgctc gctgctgcgg gccgatcagc cgcgcgcact ccctgcaact 3000
agggacaact caggcaacca cgcgcctcac aagcatggcc gccgtggcat ccaacccgct 3060
cgtgacggtg ggtgcgcaag tgccaggggc ctcgtcgtca cggcgtgcat cctcgaggga 3120
tgcgatccgg caactatatg tcgtttatct ccccaccaat cacaggatga gcccctgtct 3180
aga 3183

密码子优化的卡尔酵母MEL1序列

<210> 111

<211> 1416

<212> DNA

<213> 卡尔酵母

<400> 111

atgttcgcgt tctacttcct gacggcctgc atctccctga agggcgtgtt cggcgtctcc 60
ccctcctaca acggcctggg cctgacgccc cagatgggct gggacaactg gaacacgttc 120
gcctgcgacg tctccgagca gctgctgctg gacacggccg accgcatctc cgacctgggc 180
ctgaaggaca tgggctacaa gtacatcatc ctggacgact gctggtcctc cggccgcgac 240
tccgacggct tcctggtcgc cgacgagcag aagttcccca acggcatggg ccacgtcgcc 300
gaccacctgc acaacaactc cttcctgttc ggcatgtact cctccgcggg cgagtacacg 360
tgcgccggct accccggctc cctgggccgc gaggaggagg acgcccagtt cttcgcgaac 420
aaccgcgtgg actacctgaa gtacgacaac tgctacaaca agggccagtt cggcacgccc 480
gagatctcct accaccgcta caaggccatg tccgacgccc tgaacaagac gggccgcccc 540
atcttctact ccctgtgcaa ctggggccag gacctgacct tctactgggg ctccggcatc 600
gcgaactcct ggcgcatgtc cggcgacgtc acggcggagt tcacgcgccc cgactcccgc 660
tgcccctgcg acggcgacga gtacgactgc aagtacgccg gcttccactg ctccatcatg 720
aacatcctga acaaggccgc ccccatgggc cagaacgcgg gcgtcggcgg ctggaacgac 780
ctggacaacc tggaggtcgg cgtcggcaac ctgacggacg acgaggagaa ggcgcacttc 840
tccatgtggg ccatggtgaa gtccccccctg atcatcggcg cgaacgtgaa caacctgaag 900
gcctcctcct actccatcta ctcccaggcg tccgtcatcg ccatcaacca ggactccaac 960
ggcatccccg ccacgcgcgt ctggcgctac tacgtgtccg acacggacga gtacggccag 1020

ggcgagatcc agatgtggtc cggccccctg gacaacggcg accaggtcgt ggcgctgctg 1080
aacggcggct ccgtgtcccg ccccatgaac acgaccctgg aggagatctt cttcgactcc 1140
aacctgggct ccaagaagct gacctccacc tgggacatct acgacctgtg ggcgaaccgc 1200
gtcgacaact ccacggcgtc cgccatcctg ggccgcaaca agaccgccac cggcatcctg 1260
tacaacgcca ccgagcagtc ctacaaggac ggcctgtcca agaacgacac ccgcctgttc 1320
ggccagaaga tcggctccct gtcccccaac gcgatcctga acacgaccgt ccccgcccac 1380
ggcatcgcgt tctaccgcct gcgcccctcc tcctga 1416

5’Clp同源重组靶向序列

<210> 112
<211> 526
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 112

gcggccgcgt ggacgagggc tacaaccccg cctacggcgc gcgcccgctg cgccgcgcca 60
tcatgcgcct gctggaggac gcgctggccg agcgcatgct cgccggcgac gtcaaggagg 120
gcgactcggt catcatggac gtcgatggcg atggcgccat cagcgtcctc aacggagacc 180
gcacccacac caccaccatc gactcctccc cggcgggcat ctcgtagacg cgagagggag 240
ggagggagga ttttcagagg gagatgagac gaggactggg ccgggggcct tggggcgctg 300
ctggagcgtg gtgagagcgc ggcggacgtg ccttttcttc ttccgtgcgc gcgctcttgg 360
ccattgatcc ccgattcgcg cccgcatccc cccactgccc ccatcatctt gcctgttgtc 420
gtggcactga cataaacccc ctgcgctgcg ctgctccgct actattgata taggtctcac 480
gcgccaatct tttttgctcc gggtaaccgt ctggacgcca gaattc 526

3’Clp同源重组靶向序列

<210> 113
<211> 507
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 113

gagctcagaa tagtatcggg tgatgcgaag tcagaaccag gcagggcctg tcgcctgagg 60
tggcaacgat gggaagcaat caatctgggt acagtcgtcc gcacgatccc gtgatctccc 120
ccaccgacac ctatccccgc ccatcccggc ccacccttc agtcccctca gcatgcattg 180
tgcaccgcga caaagcatgt ctgctcgtgc actggttcag gccacggcgc accgagtcct 240
cgcccttcgc agagtgatca ccctcccccgg aaccagccac gctcgctgct gcgggccgat 300
cagccgcgcg cactccctgc aactagggac aactcaggca accacgcgcc tcacaagcat 360
ggccgccgtg gcatccaacc cgctcgtgac ggtgggtgcg caagtgccag gggcctcgtc 420

gtcacggcgt gcatcctcga gggatgcgat ccggcaacta tatgtcgttt atctccccac 480
caatcacagg atgagcccct gtctaga 507

莱茵衣藻TUB2启动子/5’UTR

<210> 114
<211> 312
<212> DNA
<213> 莱茵衣藻
<400> 114

ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60
cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc 120
gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc aggccccccga ttgcaaagac 180
attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240
ccactcgagc ttgtgatcgc actccgctaa gggggcgcct cttcctcttc gtttcagtca 300
caacccgcaa ac 312

普通小球藻硝酸盐还原酶3’UTR

<210> 115
<211> 408
<212> DNA
<213> 普通小球藻
<400> 115

gcagcagcag ctcggatagt atcgacacac tctggacgct ggtcgtgtga tggactgttg 60
ccgccacact tgctgccttg acctgtgaat atccctgccg cttttatcaa acagcctcag 120
tgtgtttgat cttgtgtgta cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga 180
ataccacccc cagcatcccc ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta 240
tctacgctgt cctgctatcc ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca 300
gccttggttt gggctccgcc tgtattctcc tggtactgca acctgtaaac cagcactgca 360
atgctgatgc acgggaagta gtgggatggg aacacaaatg gaaagctt 408

黑曲霉AlgC氨基酸序列

<210> 116
<211> 747
<212> PRT
<213> 黑曲霉
<400> 116

Met Ile Gly Ser Ser His Ala Val Val Ala Leu Gly Leu Phe Thr Leu
1 5 10 15
Tyr Gly His Ser Ala Ala Ala Pro Ala Ile Gly Ala Ser Asn Ser Gln
20 25 30
Thr Ile Val Thr Asn Gly Thr Ser Phe Ala Leu Asn Gly Asp Asn Val
35 40 45

Ser Tyr Arg Phe His Val Asn Ser Ser Thr Gly Asp Leu Ile Ser Asp
50 55 60
His Phe Gly Gly Val Val Ser Gly Thr Ile Pro Ser Pro Val Glu Pro
65 70 75 80
Ala Val Asn Gly Trp Val Gly Met Pro Gly Arg Ile Arg Arg Glu Phe
85 90 95
Pro Asp Gln Gly Arg Gly Asp Phe Arg Ile Pro Ala Val Arg Ile Arg
100 105 110
Glu Ser Ala Gly Tyr Thr Val Ser Asp Leu Gln Tyr Val Ser His Glu
115 120 125
Val Ile Glu Gly Lys Tyr Ala Leu Pro Gly Leu Pro Ala Thr Phe Gly
130 135 140
Asp Ala Gln Asp Ala Thr Thr Leu Val Val His Leu Tyr Asp Asn Tyr
145 150 155 160
Ser Ser Val Ala Ala Asp Leu Ser Tyr Ser Ile Phe Pro Lys Tyr Asp
165 170 175
Ala Ile Val Arg Ser Val Asn Val Thr Asn Gln Gly Pro Gly Asn Ile
180 185 190
Thr Ile Glu Ala Leu Ala Ser Ile Ser Ile Asp Phe Pro Tyr Glu Asp
195 200 205
Leu Asp Met Val Ser Leu Arg Gly Asp Trp Ala Arg Glu Ala Asn Val
210 215 220
Gln Arg Ser Lys Val Gln Tyr Gly Val Gln Gly Phe Gly Ser Ser Thr
225 230 235 240
Gly Tyr Ser Ser His Leu His Asn Pro Phe Leu Ala Ile Val Asp Pro
245 250 255
Ala Thr Thr Glu Ser Gln Gly Glu Ala Trp Gly Phe Asn Leu Val Tyr
260 265 270
Thr Gly Ser Phe Ser Ala Gln Val Glu Lys Gly Ser Gln Gly Phe Thr
275 280 285
Arg Ala Leu Leu Gly Phe Asn Pro Asp Gln Leu Ser Trp Asn Leu Gly
290 295 300
Pro Gly Glu Thr Leu Thr Ser Pro Glu Cys Val Ala Val Tyr Ser Asp
305 310 315 320
Lys Gly Leu Gly Ser Val Ser Arg Lys Phe His Arg Leu Tyr Arg Asn
325 330 335
His Leu Met Lys Ser Lys Phe Ala Thr Ser Asp Arg Pro Val Leu Leu
340 345 350
Asn Ser Trp Glu Gly Val Tyr Phe Asp Tyr Asn Gln Ser Ser Ile Glu

355 360 365

Thr Leu Ala Glu Glu Ser Ala Ala Leu Gly Val His Leu Phe Val Met

370 375 380

Asp Asp Gly Trp Phe Gly Asp Lys Tyr Pro Arg Val Ser Asp Asn Ala

385 390 395 400

Gly Leu Gly Asp Trp Met Pro Asn Pro Ala Arg Phe Pro Asp Gly Leu

405 410 415

Thr Pro Val Val Gln Asp Ile Thr Asn Leu Thr Val Asn Gly Thr Glu

420 425 430

Ser Thr Lys Leu Arg Phe Gly Ile Trp Val Glu Pro Glu Met Val Asn

435 440 445

Pro Asn Ser Thr Leu Tyr His Glu His Pro Glu Trp Ala Leu His Ala

450 455 460

Gly Pro Tyr Pro Arg Thr Glu Arg Arg Asn Gln Leu Val Leu Asn Leu

465 470 475 480

Ala Leu Pro Ala Val Gln Asp Phe Ile Ile Asp Phe Met Thr Asn Leu

485 490 495

Leu Gln Asp Thr Gly Ile Ser Tyr Val Lys Trp Asp Asn Asn Arg Gly

500 505 510

Ile His Glu Thr Pro Ser Pro Ser Thr Asp His Gln Tyr Met Leu Gly

515 520 525

Leu Tyr Arg Val Phe Asp Thr Leu Thr Thr Arg Phe Pro Asp Val Leu

530 535 540

Trp Glu Gly Cys Ala Ser Gly Gly Gly Arg Phe Asp Ala Gly Met Leu

545 550 555 560

Gln Tyr Val Pro Gln Ile Trp Thr Ser Asp Asn Thr Asp Ala Ile Asp

565 570 575

Arg Ile Thr Ile Gln Phe Gly Thr Ser Leu Ala Tyr Pro Pro Ser Ala

580 585 590

Met Gly Ala His Leu Ser Ala Val Pro Asn Ala Gln Thr Gly Arg Thr

595 600 605

Val Pro Phe Thr Phe Arg Ala His Val Ala Met Met Gly Gly Ser Phe

610 615 620

Gly Leu Glu Leu Asp Pro Ala Thr Val Glu Gly Asp Glu Ile Val Pro

625 630 635 640

Glu Leu Leu Ala Leu Ala Glu Lys Val Asn Pro Ile Ile Leu Asn Gly

645 650 655

Asp Leu Tyr Arg Leu Arg Leu Pro Gln Asp Ser Gln Trp Pro Ala Ala

660 665 670

Leu Phe Val Ser Gln Asp Gly Ala Gln Ala Val Leu Phe Tyr Phe Gln
675 680 685
Val Gln Pro Asn Val Asn His Ala Val Pro Trp Val Arg Leu Gln Gly
690 695 700
Leu Asp Pro Lys Ala Asp Tyr Thr Val Asp Gly Asp Gln Thr Tyr Ser
705 710 715 720
Gly Ala Thr Leu Met Asn Leu Gly Leu Gln Tyr Ser Phe Asp Thr Glu
725 730 735
Tyr Gly Ser Lys Val Val Phe Leu Glu Arg Gln
740 745

黑曲霉AlgC天然信号肽氨基酸序列

<210> 117
<211> 28
<212> PRT
<213> 黑曲霉
<400> 117

Met Ile Gly Ser Ser His Ala Val Val Ala Leu Gly Leu Phe Thr Leu
1 5 10 15
Tyr Gly His Ser Ala Ala Ala Pro Ala Ile Gly Ala
20 25

密码子优化的黑曲霉AlgC编码序列

<210> 118
<211> 2244
<212> DNA
<213> 黑曲霉
<400> 118

atgatcggct cctcccacgc ggtcgtcgcc ctgggcctgt tcaccctgta cggccactcc 60
gccgccgcgc ccgccatcgg cgcctccaac tcccagacca tcgtcacgaa cggcacctcc 120
ttcgccctga acggcgacaa cgtgtcctac cgcttccacg tgaactcctc cacgggcgac 180
ctgatctccg accacttcgg cggcgtggtg tccggcacca tcccctcccc cgtggagccc 240
gcggtcaacg gctgggtggg catgcccggc cgcatccgcc gcgagttccc cgaccagggc 300
cgcggcgact tccgcatccc cgcggtgcgc atccgcgagt ccgccggcta caccgtctcc 360
gacctgcagt acgtgtccca cgaggtgatc gagggcaagt acgcgctgcc cggcctgccc 420
gccacgttcg gcgacgccca ggacgccacc accctggtgg tgcacctgta cgacaactac 480
tcctccgtcg cggccgacct gtcctactcc atcttcccca agtacgacgc gatcgtccgc 540
tccgtgaacg tgaccaacca gggccccggc aacatcacca tcgaggcgct ggcctccatc 600
tccatcgact tcccctacga ggacctggac atggtgtccc tgcgcggcga ctgggcccgc 660
gaggcgaacg tgcagcgctc caaggtccag tacggcgtgc agggcttcgg ctcctccacc 720
ggctactcct cccacctgca caacccccttc ctggcgatcg tcgaccccgc gaccaccgag 780

tcccagggcg aggcctgggg cttcaacctg gtctacaccg gctccttctc cgcccaggtc 840
gagaagggct cccagggctt cacgcgcgcc ctgctgggct tcaaccccga ccagctgtcc 900
tggaacctgg gccccggcga gacgctgacg tcccccgagt gcgtcgccgt ctactccgac 960
aagggcctgg gctccgtctc ccgcaagttc caccgcctgt accgcaacca cctgatgaag 1020
tccaagttcg ccacgtccga ccgccccgtg ctgctgaact cctgggaggg cgtctacttc 1080
gactacaacc agtcctccat cgagacgctg gcggaggagt ccgccgccct gggcgtgcac 1140
ctgttcgtca tggacgacgg ctggttcggc gacaagtacc cccgcgtgtc cgacaacgcc 1200
ggcctgggcg actggatgcc caaccccgcc cgcttccccg acggcctgac gcccgtcgtg 1260
caggacatca ccaacctgac cgtcaacggc accgagtcca ccaagctgcg cttcggcatc 1320
tgggtggagc ccgagatggt gaaccccaac tccaccctgt accacgagca ccccgagtgg 1380
gcgctgcacg cgggccccta cccccgcacc gagcgccgca accagctggt cctgaacctg 1440
gccctgcccg cggtccagga cttcatcatc gacttcatga ccaacctgct gcaggacacc 1500
ggcatctcct acgtcaagtg ggacaacaac cgcggcatcc acgagacgcc ctccccctcc 1560
acggaccacc agtacatgct gggcctgtac cgcgtgttcg acacgctgac cacgcgcttc 1620
cccgacgtcc tgtgggaggg ctgcgcgtcc ggcggcggcc ggttcgacgc cggcatgctg 1680
cagtacgtcc cccagatctg gacgtccgac aacacggacg cgatcgaccg catcaccatc 1740
cagttcggca cctccctggc ctacccccccc tccgccatgg gcgcccacct gtccgcggtg 1800
cccaacgccc agaccggccg cacggtgccc ttcaccttcc gcgcccacgt cgcgatgatg 1860
ggcggctcct tcggcctgga gctggacccc gccaccgtgg agggcgacga gatcgtgccc 1920
gagctgctgg cgctggccga gaaggtgaac cccatcatcc tgaacggcga cctgtaccgc 1980
ctgcgcctgc cccaggactc ccagtggccc gcggccctgt tcgtgtccca ggacggcgcc 2040
caggccgtcc tgttctactt ccaggtgcag cccaacgtca accacgccgt cccctgggtc 2100
cgcctgcagg gcctggaccc caaggccgac tacacggtgg acggcgacca gacgtactcc 2160
ggcgcgaccc tgatgaacct gggcctgcag tactccttcg acaccgagta cggctccaag 2220
gtggtgttcc tggagcgcca gtaa 2244

黑曲霉AlgC表达盒的相关序列

<210> 119
<211> 4011
<212> DNA
<213> 黑曲霉
<400> 119

gcggccgcgt ggacgagggc tacaaccccg cctacggcgc gcgcccgctg cgccgcgcca 60
tcatgcgcct gctggaggac gcgctggccg agcgcatgct cgccggcgac gtcaaggagg 120
gcgactcggt catcatggac gtcgatggcg atggcgccat cagcgtcctc aacggagacc 180
gcacccacac caccaccatc gactcctccc cggcgggcat ctcgtagacg cgagagggag 240
ggagggagga ttttcagagg gagatgagac gaggactggg ccgggggcct tggggcgctg 300
ctggagcgtg gtgagagcgc ggcggacgtg ccttttcttc ttccgtgcgc gcgctcttgg 360
ccattgatcc ccgattcgcg cccgcatccc cccactgccc ccatcatctt gcctgttgtc 420
gtggcactga cataaacccc ctgcgctgcg ctgctccgct actattgata taggtctcac 480

gcgccaatct tttttgctcc gggtaaccgt ctggacgcca gaattccttt cttgcgctat 540
gacacttcca gcaaaaggta gggcgggctg cgagacggct tcccggcgct gcatgcaaca 600
ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc atgggcgctc cgatgccgct 660
ccagggcgag cgctgtttaa atagccaggc ccccgattgc aaagacatta tagcgagcta 720
ccaaagccat attcaaacac ctagatcact accacttcta cacaggccac tcgagcttgt 780
gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt cagtcacaac ccgcaaacac 840
tagtatgatc ggctcctccc acgcggtcgt cgccctgggc ctgttcaccc tgtacggcca 900
ctccgccgcc gcgcccgcca tcggcgcctc caactcccag accatcgtca cgaacggcac 960
ctccttcgcc ctgaacggcg acaacgtgtc ctaccgcttc cacgtgaact cctccacggg 1020
cgacctgatc tccgaccact tcggcggcgt ggtgtccggc accatcccct cccccgtgga 1080
gcccgcggtc aacggctggg tgggcatgcc cggccgcatc cgccgcgagt tccccgacca 1140
gggccgcggc gacttccgca tccccgcggt gcgcatccgc gagtccgccg gctacaccgt 1200
ctccgacctg cagtacgtgt cccacgaggt gatcgagggc aagtacgcgc tgcccggcct 1260
gcccgccacg ttcggcgacg cccaggacgc caccaccctg gtggtgcacc tgtacgacaa 1320
ctactcctcc gtcgcggccg acctgtccta ctccatcttc cccaagtacg acgcgatcgt 1380
ccgctccgtg aacgtgacca accagggccc cggcaacatc accatcgagg cgctggcctc 1440
catctccatc gacttccgct acgaggacct ggacatggtg tccctgcgcg gcgactgggc 1500
ccgcgaggcg aacgtgcagc gctccaaggt ccagtacggc gtgcagggct tcggctcctc 1560
caccggctac tcctcccacc tgcacaaccc cttcctggcg atcgtcgacc ccgcgaccac 1620
cgagtcccag ggcgaggcct ggggcttcaa cctggtctac accggctcct tctccgccca 1680
ggtcgagaag ggctcccagg gcttcacgcg cgccctgctg ggcttcaacc ccgaccagct 1740
gtcctggaac ctgggccccg gcgagacgct gacgtccccc gagtgcgtcg ccgtctactc 1800
cgacaagggc ctgggctccg tctcccgcaa gttccaccgc ctgtaccgca accacctgat 1860
gaagtccaag ttcgccacgt ccgaccgccc cgtgctgctg aactcctggg agggcgtcta 1920
cttcgactac aaccagtcct ccatcgagac gctggcggag gagtccgccg ccctgggcgt 1980
gcacctgttc gtcatggacg acggctggtt cggcgacaag tacccccgcg tgtccgacaa 2040
cgccggcctg ggcgactgga tgcccaaccc cgcccgcttc cccgacggcc tgacgcccgt 2100
cgtgcaggac atcaccaacc tgaccgtcaa cggcaccgag tccaccaagc tgcgcttcgg 2160
catctgggtg gagcccgaga tggtgaaccc caactccacc ctgtaccacg agcaccccga 2220
gtgggcgctg cacgcgggcc cctacccccg caccgagcgc cgcaaccagc tggtcctgaa 2280
cctggccctg cccgcggtcc aggacttcat catcgacttc atgaccaacc tgctgcagga 2340
caccggcatc tcctacgtca agtgggacaa caaccgcggc atccacgaga cgccctcccc 2400
ctccacggac caccagtaca tgctgggcct gtaccgcgtg ttcgacacgc tgaccacgcg 2460
cttccccgac gtcctgtggg agggctgcgc gtccggcggc ggccggttcg acgccggcat 2520
gctgcagtac gtcccccaga tctggacgtc cgacaacacg gacgcgatcg accgcatcac 2580
catccagttc ggcacctccc tggcctaccc cccctccgcc atgggcgccc acctgtccgc 2640
ggtgcccaac gcccagaccg gccgcacggt gcccttcacc ttccgcgccc acgtcgcgat 2700
gatgggcggc tccttcggcc tggagctgga ccccgccacc gtggagggcg acgagatcgt 2760
gcccgagctg ctggcgctgg ccgagaaggt gaaccccatc atcctgaacg gcgacctgta 2820

ccgcctgcgc ctgccccagg actcccagtg gcccgcggcc ctgttcgtgt cccaggacgg 2880
cgcccaggcc gtcctgttct acttccaggt gcagcccaac gtcaaccacg ccgtcccctg 2940
ggtccgcctg cagggcctgg accccaaggc cgactacacg gtggacggcg accagacgta 3000
ctccggcgcg accctgatga acctgggcct gcagtactcc ttcgacaccg agtacggctc 3060
caaggtggtg ttcctggagc gccagtaaca attgaagcag cagcagctcg gatagtatcg 3120
acacactctg gacgctggtc gtgtgatgga ctgttgccgc cacacttgct gccttgacct 3180
gtgaatatcc ctgccgcttt tatcaaacag cctcagtgtg tttgatcttg tgtgtacgcg 3240
cttttgcgag ttgctagctg cttgtgctat ttgcgaatac cacccccagc atcccccttcc 3300
ctcgtttcat atcgcttgca tcccaaccgc aacttatcta cgctgtcctg ctatccctca 3360
gcgctgctcc tgctcctgct cactgcccct cgcacagcct tggtttgggc tccgcctgta 3420
ttctcctggt actgcaacct gtaaaccagc actgcaatgc tgatgcacgg gaagtagtgg 3480
gatgggaaca caaatggaaa gcttgagctc agaatagtat cgggtgatgc gaagtcagaa 3540
ccaggcaggg cctgtcgcct gaggtggcaa cgatgggaag caatcaatct gggtacagtc 3600
gtccgcacga tcccgtgatc tcccccaccg acacctatcc ccgcccatcc cggcccaccc 3660
tttcagtccc ctcagcatgc attgtgcacc gcgacaaagc atgtctgctc gtgcactggt 3720
tcaggccacg gcgcaccgag tcctcgccct tcgcagagtg atcaccctcc ccggaaccag 3780
ccacgctcgc tgctgcgggc cgatcagccg cgcgcactcc ctgcaactag ggacaactca 3840
ggcaaccacg cgcctcacaa gcatggccgc cgtggcatcc aacccgctcg tgacggtggg 3900
tgcgcaagtg ccaggggcct cgtcgtcacg gcgtgcatcc tcgagggatg cgatccggca 3960
actatatgtc gtttatctcc ccaccaatca caggatgagc ccctgtctag a 4011

瓜儿豆α-半乳糖苷酶氨基酸序列

<210> 120
<211> 411
<212> PRT
<213> 瓜儿豆
<400> 120

Met Ala Thr His Tyr Ser Ile Ile Gly Gly Met Ile Ile Val Val Leu
1 5 10 15
Leu Met Ile Ile Gly Ser Glu Gly Gly Arg Leu Leu Glu Lys Lys Asn
20 25 30
Arg Thr Ser Ala Glu Ala Glu His Tyr Asn Val Arg Arg Tyr Leu Ala
35 40 45
Glu Asn Gly Leu Gly Gln Thr Pro Pro Met Gly Trp Asn Ser Trp Asn
50 55 60
His Phe Gly Cys Asp Ile Asn Glu Asn Val Val Arg Glu Thr Ala Asp
65 70 75 80
Ala Met Val Ser Thr Gly Leu Ala Ala Leu Gly Tyr Gln Tyr Ile Asn
85 90 95
Leu Asp Asp Cys Trp Ala Glu Leu Asn Arg Asp Ser Glu Gly Asn Met

100 105 110

Val Pro Asn Ala Ala Ala Phe Pro Ser Gly Ile Lys Ala Leu Ala Asp
115 120 125

Tyr Val His Ser Lys Gly Leu Lys Leu Gly Val Tyr Ser Asp Ala Gly
130 135 140

Asn Gln Thr Cys Ser Lys Arg Met Pro Gly Ser Leu Gly His Glu Glu
145 150 155 160

Gln Asp Ala Lys Thr Phe Ala Ser Trp Gly Val Asp Tyr Leu Lys Tyr
165 170 175

Asp Asn Cys Glu Asn Leu Gly Ile Ser Val Lys Glu Arg Tyr Pro Pro
180 185 190

Met Gly Lys Ala Leu Leu Ser Ser Gly Arg Pro Ile Phe Phe Ser Met
195 200 205

Cys Glu Trp Gly Trp Glu Asp Pro Gln Ile Trp Ala Lys Ser Ile Gly
210 215 220

Asn Ser Trp Arg Thr Thr Gly Asp Ile Glu Asp Asn Trp Asn Ser Met
225 230 235 240

Thr Ser Ile Ala Asp Ser Asn Asp Lys Trp Ala Ser Tyr Ala Gly Pro
245 250 255

Gly Gly Trp Asn Asp Pro Asp Met Leu Glu Val Gly Asn Gly Gly Met
260 265 270

Thr Thr Glu Glu Tyr Arg Ser His Phe Ser Ile Trp Ala Leu Ala Lys
275 280 285

Ala Pro Leu Leu Val Gly Cys Asp Ile Arg Ala Met Asp Asp Thr Thr
290 295 300

His Glu Leu Ile Ser Asn Ala Glu Val Ile Ala Val Asn Gln Asp Lys
305 310 315 320

Leu Gly Val Gln Gly Lys Lys Val Lys Ser Thr Asn Asp Leu Glu Val
325 330 335

Trp Ala Gly Pro Leu Ser Asp Asn Lys Val Ala Val Ile Leu Trp Asn
340 345 350

Arg Ser Ser Ser Arg Ala Thr Val Thr Ala Ser Trp Ser Asp Ile Gly
355 360 365

Leu Gln Gln Gly Thr Thr Val Asp Ala Arg Asp Leu Trp Glu His Ser
370 375 380

Thr Gln Ser Leu Val Ser Gly Glu Ile Ser Ala Glu Ile Asp Ser His
385 390 395 400

Ala Cys Lys Met Tyr Val Leu Thr Pro Arg Ser
405 410

密码子优化的瓜儿豆α-半乳糖苷酶序列

<210> 121

<211> 1236

<212> DNA

<213> 瓜儿豆

<400> 121

atggccaccc actactccat catcggcggc atgatcatcg tcgtcctgct gatgatcatc 60
ggctccgagg gcggccggct gctggagaag aagaaccgca cctccgccga ggcggagcac 120
tacaacgtgc gccgctacct ggccgagaac ggcctgggcc agacccccc catgggctgg 180
aactcctgga accacttcgg ctgcgacatc aacgagaacg tcgtccgcga gacggcggac 240
gccatggtgt ccaccggcct ggccgccctg ggctaccagt acatcaacct ggacgactgc 300
tgggcggagc tgaaccgcga ctccgagggc aacatggtgc ccaacgccgc cgcgttcccc 360
tccggcatca aggcgctggc ggactacgtc cactccaagg gcctgaagct gggcgtctac 420
tccgacgcgg gcaaccagac ctgctccaag cgcatgcccg gctccctggg ccacgaggag 480
caggacgcga agacgttcgc ctcctggggc gtcgactacc tgaagtacga caactgcgag 540
aacctgggca tctccgtgaa ggagcgctac ccccccatgg gcaaggcgct gctgtcctcc 600
ggccgcccca tcttcttctc catgtgcgag tggggctggg aggaccccca gatctgggcc 660
aagtccatcg gcaactcctg gcgcaccacc ggcgacatcg aggacaactg gaactccatg 720
acctccatcg ccgactccaa cgacaagtgg gcctcctacg ccggccccgg cggctggaac 780
gaccccgaca tgctggaggt cggcaacggc ggcatgacga cggaggagta ccgctcccac 840
ttctccatct gggccctggc caaggcgccc ctgctggtgg gctgcgacat ccgcgcgatg 900
gacgacacga cgcacgagct gatctccaac gccgaggtga tcgcggtgaa ccaggacaag 960
ctgggcgtgc agggcaagaa ggtcaagtcc acgaacgacc tggaggtctg ggccggcccc 1020
ctgtccgaca acaaggtggc ggtgatcctg tggaaccgct cctcctcccg cgccaccgtc 1080
accgcgtcct ggtccgacat cggcctgcag cagggcacca ccgtcgacgc gcgcgacctg 1140
tgggagcact ccacgcagtc cctggtgtcc ggcgagatct ccgccgagat cgactcccac 1200
gcctgcaaga tgtacgtcct gacgccccgc tcctaa 1236

瓜儿豆α-半乳糖苷酶表达盒的相关序列

<210> 122

<211> 3003

<212> DNA

<213> 瓜儿豆

<400> 122

gcggccgcgt ggacgagggc tacaaccccg cctacggcgc gcgcccgctg cgccgcgcca 60
tcatgcgcct gctggaggac gcgctggccg agcgcatgct cgccggcgac gtcaaggagg 120
gcgactcggt catcatggac gtcgatggcg atggcgccat cagcgtcctc aacggagacc 180
gcacccacac caccaccatc gactcctccc cggcgggcat ctcgtagacg cgagagggag 240
ggagggagga ttttcagagg gagatgagac gaggactggg ccgggggcct tggggcgctg 300
ctggagcgtg gtgagagcgc ggcggacgtg ccttttcttc ttccgtgcgc gcgctcttgg 360

ccattgatcc ccgattcgcg cccgcatccc cccactgccc ccatcatctt gcctgttgtc 420
gtggcactga cataaacccc ctgcgctgcg ctgctccgct actattgata taggtctcac 480
gcgccaatct tttttgctcc gggtaaccgt ctggacgcca gaattccttt cttgcgctat 540
gacacttcca gcaaaaggta gggcgggctg cgagacggct tcccggcgct gcatgcaaca 600
ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc atgggcgctc cgatgccgct 660
ccagggcgag cgctgtttaa atagccaggc ccccgattgc aaagacatta tagcgagcta 720
ccaaagccat attcaaacac ctagatcact accacttcta cacaggccac tcgagcttgt 780
gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt cagtcacaac ccgcaaacac 840
tagtatggcc acccactact ccatcatcgg cggcatgatc atcgtcgtcc tgctgatgat 900
catcggctcc gagggcggcc ggctgctgga gaagaagaac cgcacctccg ccgaggcgga 960
gcactacaac gtgcgccgct acctggccga gaacggcctg ggccagaccc cccccatggg 1020
ctggaactcc tggaaccact tcggctgcga catcaacgag aacgtcgtcc gcgagacggc 1080
ggacgccatg gtgtccaccg gcctggccgc cctgggctac cagtacatca acctggacga 1140
ctgctgggcg gagctgaacc gcgactccga gggcaacatg gtgcccaacg ccgccgcgtt 1200
cccctccggc atcaaggcgc tggcggacta cgtccactcc aagggcctga agctgggcgt 1260
ctactccgac gcgggcaacc agacctgctc caagcgcatg cccggctccc tgggccacga 1320
ggagcaggac gcgaagacgt tcgcctcctg gggcgtcgac tacctgaagt acgacaactg 1380
cgagaacctg ggcatctccg tgaaggagcg ctacccccccc atgggcaagg cgctgctgtc 1440
ctccggccgc cccatcttct tctccatgtg cgagtggggc tgggaggacc cccagatctg 1500
ggccaagtcc atcggcaact cctggcgcac caccggcgac atcgaggaca actggaactc 1560
catgacctcc atcgccgact ccaacgacaa gtgggcctcc tacgccggcc ccggcggctg 1620
gaacgacccc gacatgctgg aggtcggcaa cggcggcatg acgacggagg agtaccgctc 1680
ccacttctcc atctgggccc tggccaaggc gcccctgctg gtgggctgcg acatccgcgc 1740
gatggacgac acgacgcacg agctgatctc caacgccgag gtgatcgcgg tgaaccagga 1800
caagctgggc gtgcagggca agaaggtcaa gtccacgaac gacctggagg tctgggccgg 1860
ccccctgtcc gacaacaagg tggcggtgat cctgtggaac cgctcctcct cccgcgccac 1920
cgtcaccgcg tcctggtccg acatcggcct gcagcagggc accaccgtcg acgcgcgcga 1980
cctgtgggag cactccacgc agtccctggt gtccggcgag atctccgccg agatcgactc 2040
ccacgcctgc aagatgtacg tcctgacgcc ccgctcctaa caattgaagc agcagcagct 2100
cggatagtat cgacacactc tggacgctgg tcgtgtgatg gactgttgcc gccacacttg 2160
ctgccttgac ctgtgaatat ccctgccgct tttatcaaac agcctcagtg tgtttgatct 2220
tgtgtgtacg cgcttttgcg agttgctagc tgcttgtgct atttgcgaat accacccccа 2280
gcatcccctt ccctcgtttc atatcgcttg catcccaacc gcaacttatc tacgctgtcc 2340
tgctatccct cagcgctgct cctgctcctg ctcactgccc ctcgcacagc cttggtttgg 2400
gctccgcctg tattctcctg gtactgcaac ctgtaaacca gcactgcaat gctgatgcac 2460
gggaagtagt gggatgggaa cacaaatgga aagcttgagc tcagaatagt atcgggtgat 2520
gcgaagtcag aaccaggcag ggcctgtcgc ctgaggtggc aacgatggga agcaatcaat 2580
ctgggtacag tcgtccgcac gatcccgtga tctcccccac cgacacctat ccccgcccat 2640
cccggcccac cctttcagtc ccctcagcat gcattgtgca ccgcgacaaa gcatgtctgc 2700

tcgtgcactg gttcaggcca cggcgcaccg agtcctcgcc cttcgcagag tgatcaccct 2760
ccccggaacc agccacgctc gctgctgcgg gccgatcagc cgcgcgcact ccctgcaact 2820
agggacaact caggcaacca cgcgcctcac aagcatggcc gccgtggcat ccaacccgct 2880
cgtgacggtg ggtgcgcaag tgccaggggc ctcgtcgtca cggcgtgcat cctcgaggga 2940
tgcgatccgg caactatatg tcgtttatct ccccaccaat cacaggatga gcccctgtct 3000
aga 3003

5'引物普通小球藻3'UTR:下游Clp序列

<210> 123
<211> 22
<212> DNA
<213> 普通小球藻
<400> 123

actgcaatgc tgatgcacgg ga 22

3'引物普通小球藻3'UTR:下游Clp序列

<210> 124
<211> 19
<212> DNA
<213> 普通小球藻
<400> 124

tccaggtcct tttcgcact 19

胶球藻属C-169 THIC氨基酸序列

<210> 125
<211> 632
<212> PRT
<213> 未知
<220>
<223> 胶球藻属
<400> 125

Met Thr Thr Asn Leu Ala Lys Leu Pro Leu Gly Gly Leu Ser Ser Arg
1 5 10 15
Ser Ser Ile Ala Gly Ala Pro Leu Arg Val Asn Ser His Ser Arg Glu
20 25 30
Arg His Leu Gly Ala Lys Thr Thr Ala Ile Ala Ala Pro Glu Arg Leu
35 40 45
Asp Tyr Leu Asp Asn Ala Glu Glu Ala Arg Leu Gln Gln Thr Asp Ala
50 55 60
Phe Ala Glu Leu Lys Ala Leu Ser Ser Arg Gln Ser Val Asn Arg Pro
65 70 75 80
Gln Lys Gly Glu Leu Ser Phe Arg Gln Ser Pro Thr Phe Gln Asp Cys

85 90 95

Phe Pro Gly Ser Glu Lys Cys Tyr Arg Glu Val Glu His Asp Gly Lys
100 105 110

Thr Leu Lys Val Pro Phe Arg Arg Val His Leu Gln Asp Asp Asn Gly
115 120 125

His Phe Asp Leu Tyr Asp Thr Ser Gly Pro Gln Gly Val Asn Pro Arg
130 135 140

Glu Gly Leu Pro Lys Ile Arg Ser Ser Trp Val Glu Pro Arg Glu Ala
145 150 155 160

Arg Gly Asp Lys Val Gln Thr Gln Gln Tyr Tyr Ala Lys Gln Gly Ile
165 170 175

Ile Thr Glu Glu Met Ala Phe Cys Ala Ala Arg Glu Arg Met Asp Pro
180 185 190

Glu Phe Ile Arg Ser Glu Val Ala Arg Gly Arg Ala Ile Ile Pro Ala
195 200 205

Asn Lys Arg His Leu Glu Leu Glu Pro Thr Val Val Gly Arg Asn Phe
210 215 220

Leu Val Lys Val Asn Ala Asn Ile Gly Asn Ser Ala Val Ser Ser Ser
225 230 235 240

Ile Glu Glu Glu Val Glu Lys Leu Gln Trp Ser Thr Ile Trp Gly Ala
245 250 255

Asp Thr Ala Met Asp Leu Ser Thr Gly Asn Asn Ile His Glu Thr Arg
260 265 270

Glu Trp Val Met Arg Asn Ser Pro Val Pro Val Gly Thr Val Pro Ile
275 280 285

Tyr Gln Cys Leu Glu Lys Ala Gly Gly Ile Val Glu Asn Ile Thr Trp
290 295 300

Glu Leu Phe Arg Glu Thr Leu Ile Glu Gln Ala Glu Gln Gly Val Asp
305 310 315 320

Tyr Phe Thr Ile His Ala Gly Val Leu Leu Arg Tyr Ile Pro Leu Thr
325 330 335

Ala Asn Arg Val Thr Gly Ile Val Ser Arg Gly Gly Ser Ile His Ala
340 345 350

Lys Leu Cys Leu Leu Asp His Thr Glu Asn Phe Ala Tyr Met His Trp
355 360 365

Asp Glu Ile Leu Asp Ile Cys Ala Gln Tyr Asp Ile Thr Leu Ser Ile
370 375 380

Gly Asp Gly Leu Arg Pro Gly Cys Ile Ala Asp Ala Asn Asp Ala Ala
385 390 395 400

Gln Phe Ala Glu Leu Lys Thr Gln Gly Glu Leu Thr Arg Arg Ala Trp
405 410 415
Ala Lys Asp Val Gln Val Met Asn Glu Gly Pro Gly His Val Pro Leu
420 425 430
His Lys Ile Pro Glu Asn Met Glu Lys Gln Leu Asp Trp Cys Ser Glu
435 440 445
Ala Pro Phe Tyr Thr Leu Gly Pro Leu Ala Thr Asp Ile Ala Pro Ala
450 455 460
Tyr Asp His Ile Thr Ser Ala Ile Gly Ala Ala Thr Ile Gly Ala Leu
465 470 475 480
Gly Thr Ala Leu Leu Cys Tyr Val Thr Pro Lys Glu His Leu Gly Leu
485 490 495
Pro Asp Arg Asp Asp Val Lys Ala Gly Val Ile Ala Tyr Lys Ile Ala
500 505 510
Ala His Ala Ala Asp Leu Ala Lys Gly His Pro Tyr Ala Gln Glu Trp
515 520 525
Asp Asn Ala Leu Ser Lys Ala Arg Phe Glu Phe Arg Trp Tyr Asp Gln
530 535 540
Phe His Leu Ser Leu Asp Pro Val Thr Ala Arg Leu Phe His Asp Ala
545 550 555 560
Thr Leu Pro Gln Glu Pro Ala Lys Thr Ala His Phe Cys Ser Met Cys
565 570 575
Gly Pro Lys Phe Cys Ser Met Gln Ile Thr Gln Asp Leu Arg Glu Tyr
580 585 590
Ala Lys Asn His Gln Met Glu Glu Asp Glu Ala Ile Gln Thr Gly Met
595 600 605
Ala Glu Met Ser Glu Gln Phe Lys Ser Ile Gly Ala Glu Val Tyr Leu
610 615 620
Asp Glu Ala Val Leu Glu His Ala
625 630

胶球藻属C-169 THIC天然转运肽氨基酸序列

<210> 126
<211> 40
<212> PRT
<213> 未知
<220>
<223> 胶球藻属
<400> 126

Met Thr Thr Asn Leu Ala Lys Leu Pro Leu Gly Gly Leu Ser Ser Arg

1 5 10 15

Ser Ser Ile Ala Gly Ala Pro Leu Arg Val Asn Ser His Ser Arg Glu

20 25 30

Arg His Leu Gly Ala Lys Thr Thr

35 40

胶球藻属C-169 THIC表达盒的相关序列

<210> 127

<211> 6649

<212> DNA

<213> 未知

<220>

<223> 胶球藻属

<400> 127

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60

ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120

tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180

ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240

gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300

ggaagacagg tgagggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360

atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420

cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480

gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540

ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600

ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660

cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720

ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 960

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 1020

cagtcacaac ccgcaaacgg cgcgccatgc tgctgcaggc cttcctgttc ctgctggccg 1080

gcttcgccgc caagatcagc gcctccatga cgaacgagac gtccgaccgc cccctggtgc 1140

acttcacccc caacaagggc tggatgaacg accccaacgg cctgtggtac gacgagaagg 1200

acgccaagtg gcacctgtac ttccagtaca acccgaacga caccgtctgg gggacgccct 1260

tgttctgggg ccacgccacg tccgacgacc tgaccaactg ggaggaccag cccatcgcca 1320

tcgccccgaa gcgcaacgac tccggcgcct tctccggctc catggtggtg gactacaaca 1380

acacctccgg cttcttcaac gacaccatcg acccgcgcca gcgctgcgtg gccatctgga 1440

cctacaacac cccggagtcc gaggagcagt acatctccta cagcctggac ggcggctaca 1500

ccttcaccga gtaccagaag aaccccgtgc tggccgccaa ctccacccag ttccgcgacc 1560

cgaaggtctt ctggtacgag ccctcccaga agtggatcat gaccgcggcc aagtcccagg 1620
actacaagat cgagatctac tcctccgacg acctgaagtc ctggaagctg gagtccgcgt 1680
tcgccaacga gggcttcctc ggctaccagt acgagtgccc cggcctgatc gaggtcccca 1740
ccgagcagga ccccagcaag tcctactggg tgatgttcat ctccatcaac cccggcgccc 1800
cggccggcgg ctccttcaac cagtacttcg tcggcagctt caacggcacc cacttcgagg 1860
ccttcgacaa ccagtcccgc gtggtggact tcggcaagga ctactacgcc ctgcagacct 1920
tcttcaacac cgacccgacc tacgggagcg ccctgggcat cgcgtgggcc tccaactggg 1980
agtactccgc cttcgtgccc accaacccct ggcgctcctc catgtccctc gtgcgcaagt 2040
tctccctcaa caccgagtac caggccaacc cggagacgga gctgatcaac ctgaaggccg 2100
agccgatcct gaacatcagc aacgccggcc cctggagccg gttcgccacc aacaccacgt 2160
tgacgaaggc caacagctac aacgtcgacc tgtccaacag caccggcacc ctggagttcg 2220
agctggtgta cgccgtcaac accacccaga cgatctccaa gtccgtgttc gcggacctct 2280
ccctctggtt caagggcctg gaggaccccg aggagtacct ccgcatgggc ttcgaggtgt 2340
ccgcgtcctc cttcttcctg gaccgcggga acagcaaggt gaagttcgtg aaggagaacc 2400
cctacttcac caaccgcatg agcgtgaaca accagccctt caagagcgag aacgacctgt 2460
cctactacaa ggtgtacggc ttgctggacc agaacatcct ggagctgtac ttcaacgacg 2520
gcgacgtcgt gtccaccaac acctacttca tgaccaccgg gaacgccctg ggctccgtga 2580
acatgacgac gggggtggac aacctgttct acatcgacaa gttccaggtg cgcgaggtca 2640
agtgacaatt ggcagcagca gctcggatag tatcgacaca ctctggacgc tggtcgtgtg 2700
atggactgtt gccgccacac ttgctgcctt gacctgtgaa tatccctgcc gcttttatca 2760
aacagcctca gtgtgtttga tcttgtgtgt acgcgctttt gcgagttgct agctgcttgt 2820
gctatttgcg aataccaccc ccagcatccc cttccctcgt ttcatatcgc ttgcatccca 2880
accgcaactt atctacgctg tcctgctatc cctcagcgct gctcctgctc ctgctcactg 2940
cccctcgcac agccttggtt tgggctccgc ctgtattctc ctggtactgc aacctgtaaa 3000
ccagcactgc aatgctgatg cacgggaagt agtgggatgg gaacacaaat ggaggatccc 3060
gcgtctcgaa cagagcgcgc agaggaacgc tgaaggtctc gcctctgtcg cacctcagcg 3120
cggcatacac cacaataacc acctgacgaa tgcgcttggt tcttcgtcca ttagcgaagc 3180
gtccggttca cacacgtgcc acgttggcga ggtggcaggt gacaatgatc ggtggagctg 3240
atggtcgaaa cgttcacagc ctagggatat cgaattcctt tcttgcgcta tgacacttcc 3300
agcaaaaggt agggcgggct gcgagacggc ttcccggcgc tgcatgcaac accgatgatg 3360
cttcgacccc ccgaagctcc ttcggggctg catgggcgct ccgatgccgc tccagggcga 3420
gcgctgttta aatagccagg cccccgattg caaagacatt atagcgagct accaaagcca 3480
tattcaaaca cctagatcac taccacttct acacaggcca ctcgagcttg tgatcgcact 3540
ccgctaaggg ggcgcctctt cctcttcgtt tcagtcacaa cccgcaaaca ctagtatgac 3600
caccaacctg gctaagctgc ctctcggcgg cctctcttct cggtcgtcca tcgccggggc 3660
tcctctgcgc gtgaacagcc acagccgcga gcgccacctg ggcgccaaga ccaccgcgat 3720
cgccgctccc gagcggctgg actacctcga caacgcggag gaggcgcggc tgcagcagac 3780
ggacgccttc gccgagctga aggcgctgtc gagccgccag tccgtgaacc ggccccaaaa 3840
gggcgagctg tcgttccggc agtcgccgac ctttcaggac tgcttccctg gctcggagaa 3900

gtgctaccgc gaggtggagc acgatggcaa gacgctcaag gtgcctttc gccgggtgca 3960
cctgcaggac gacaacggcc atttcgacct gtacgacacc tctggccccc agggcgtgaa 4020
cccccgcgag ggtctgccca agatccgctc ctcctgggtg gagccgcgcg aggcgcgggg 4080
cgacaaggtg caaacgcagc agtactacgc caagcaaggc atcatcaccg aggagatggc 4140
cttctgcgct gcgcgcgagc gcatggaccc cgagttcatc cggtctgagg tggcccgggg 4200
tcgggccatc atccctgcga acaagcgcca tctggagctg gagccgaccg tcgtgggccg 4260
caactttctg gtcaaggtga acgccaacat cggcaactct gcggtcagct cgagcatcga 4320
ggaggaggtg gagaagctgc agtggagcac catctggggt gccgacacgg ccatggacct 4380
gagcaccggc aacaacatcc acgagacgcg cgagtgggtg atgcgcaaca gcccggtccc 4440
tgtgggtacg gtccctatct accagtgcct ggagaaggcg ggcggcatcg tggagaacat 4500
cacctgggag ctgttccggg agacgctgat cgagcaggcc gagcagggcg tcgactactt 4560
taccatccac gcgggggtcc tgctccgcta catccccctg acggcgaacc gcgtcaccgg 4620
catcgtcagc cgcggcggct ctatccatgc caagctctgc ctgctggacc acaccgagaa 4680
cttcgcctac atgcactggg acgagatcct ggacatctgc gcccagtacg acatcaccct 4740
gtcgatcggc gacggcctgc gccccggctg catcgcggac gccaacgacg ctgcccagtt 4800
tgctgagctg aagacccagg gcgagctgac ccgccgcgcc tgggcgaagg acgtgcaggt 4860
gatgaacgag ggtcccggcc atgtgccgct gcacaagatc cccgagaaca tggagaagca 4920
gctggactgg tgcagcgagg cgcccttcta cacgctcggc cccctggcga cggacatcgc 4980
gcctgcgtac gaccacatca cctccgcgat cggcgccgcc acgatcggcg cgctggggac 5040
cgcgctgctg tgttacgtga ccccgaagga gcacctgggc ctccccgatc gcgacgacgt 5100
gaaggccggg gtgatcgcct acaagatcgc ggcgcacgcg gctgacctgg cgaagggtca 5160
cccctacgct caggagtggg ataacgccct ctctaaggcg cgcttcgagt tccggtggta 5220
cgaccagttc cacctgtcgc tggaccccgt caccgcccgc ctgttccatg acgcgaccct 5280
gccgcaggag cccgccaaga ccgcccactt ctgctccatg tgcggcccca agttctgctc 5340
catgcagatc acccaggacc tccgcgagta cgccaagaac catcaaatgg aggaggacga 5400
ggcgatccaa accggcatgg ccgagatgtc cgagcagttc aagagcatcg gcgcggaggt 5460
ctacctcgac gaggcggtgc tggagcacgc gtaagcagca gcagctcgga tagtatcgac 5520
acactctgga cgctggtcgt gtgatggact gttgccgcca cacttgctgc cttgacctgt 5580
gaatatccct gccgctttta tcaaacagcc tcagtgtgtt tgatcttgtg tgtacgcgct 5640
tttgcgagtt gctagctgct tgtgctattt gcgaatacca cccccagcat ccccttccct 5700
cgtttcatat cgcttgcatc ccaaccgcaa cttatctacg ctgtcctgct atccctcagc 5760
gctgctcctg ctcctgctca ctgcccctcg cacagccttg gtttgggctc cgcctgtatt 5820
ctcctggtac tgcaacctgt aaaccagcac tgcaatgctg atgcacggga agtagtggga 5880
tgggaacaca aatggaaagc ttgagctctt gttttccaga aggagttgct ccttgagcct 5940
ttcattctca gcctcgataa cctccaaagc cgctctaatt gtggaggggg ttcgaattta 6000
aaagcttgga atgttggttc gtgcgtctgg aacaagccca gacttgttgc tcactgggaa 6060
aaggaccatc agctccaaaa aacttgccgc tcaaaccgcg tacctctgct ttcgcgcaat 6120
ctgccctgtt gaaatcgcca ccacattcat attgtgacgc ttgagcagtc tgtaattgcc 6180
tcagaatgtg gaatcatctg ccccctgtgc gagcccatgc caggcatgtc gcgggcgagg 6240

acacccgcca ctcgtacagc agaccattat gctacctcac aatagttcat aacagtgacc 6300
atatttctcg aagctcccca acgagcacct ccatgctctg agtggccacc ccccggccct 6360
ggtgcttgcg gagggcaggt caaccggcat ggggctaccg aaatccccga ccggatccca 6420
ccacccccgc gatgggaaga atctctcccc gggatgtggg cccaccacca gcacaacctg 6480
ctggcccagg cgagcgtcaa accataccac acaaatatcc ttggcatcgg ccctgaattc 6540
cttctgccgc tctgctaccc ggtgcttctg tccgaagcag gggttgctag ggatcgctcc 6600
gagtccgcaa acccttgtcg cgtggcgggg cttgttcgag ctgaagagc 6649

密码子优化的胶球藻属C-169 THIC序列

<210> 128
<211> 1899
<212> DNA
<213> 未知
<220>
<223> 胶球藻属
<400> 128

atgaccacca acctggctaa gctgcctctc ggcggcctct cttctcggtc gtccatcgcc 60
ggggctcctc tgcgcgtgaa cagccacagc cgcgagcgcc acctgggcgc caagaccacc 120
gcgatcgccg ctcccgagcg gctggactac ctcgacaacg cggaggaggc gcggctgcag 180
cagacggacg ccttcgccga gctgaaggcg ctgtcgagcc gccagtccgt gaaccggccc 240
caaaagggcg agctgtcgtt ccggcagtcg ccgacctttc aggactgctt ccctggctcg 300
gagaagtgct accgcgaggt ggagcacgat ggcaagacgc tcaaggtgcc ttttcgccgg 360
gtgcacctgc aggacgacaa cggccatttc gacctgtacg acacctctgg cccccagggc 420
gtgaaccccc gcgagggtct gcccaagatc cgctcctcct gggtggagcc gcgcgaggcg 480
cggggcgaca aggtgcaaac gcagcagtac tacgccaagc aaggcatcat caccgaggag 540
atggccttct gcgctgcgcg cgagcgcatg gaccccgagt tcatccggtc tgaggtggcc 600
cggggtcggg ccatcatccc tgcgaacaag cgccatctgg agctggagcc gaccgtcgtg 660
ggccgcaact ttctggtcaa ggtgaacgcc aacatcggca actctgcggt cagctcgagc 720
atcgaggagg aggtggagaa gctgcagtgg agcaccatct ggggtgccga cacggccatg 780
gacctgagca ccggcaacaa catccacgag acgcgcgagt gggtgatgcg caacagcccg 840
gtccctgtgg gtacggtccc tatctaccag tgcctggaga aggcgggcgg catcgtggag 900
aacatcacct gggagctgtt ccgggagacg ctgatcgagc aggccgagca gggcgtcgac 960
tactttacca tccacgcggg ggtcctgctc cgctacatcc ccctgacggc gaaccgcgtc 1020
accggcatcg tcagccgcgg cggctctatc catgccaagc tctgcctgct ggaccacacc 1080
gagaacttcg cctacatgca ctgggacgag atcctggaca tctgcgccca gtacgacatc 1140
accctgtcga tcggcgacgg cctgcgcccc ggctgcatcg cggacgccaa cgacgctgcc 1200
cagtttgctg agctgaagac ccagggcgag ctgacccgcc gcgcctgggc gaaggacgtg 1260
caggtgatga acgagggtcc cggccatgtg ccgctgcaca agatccccga gaacatggag 1320
aagcagctgg actggtgcag cgaggcgccc ttctacacgc tcggcccccc ggcgacggac 1380
atcgcgcctg cgtacgacca catcacctcc gcgatcggcg ccgccacgat cggcgcgctg 1440

gggaccgcgc tgctgtgtta cgtgaccccg aaggagcacc tgggcctccc cgatcgcgac 1500
gacgtgaagg ccggggtgat cgcctacaag atcgcggcgc acgcggctga cctggcgaag 1560
ggtcacccct acgctcagga gtgggataac gccctctcta aggcgcgctt cgagttccgg 1620
tggtacgacc agttccacct gtcgctggac cccgtcaccg cccgcctgtt ccatgacgcg 1680
accctgccgc aggagcccgc caagaccgcc cacttctgct ccatgtgcgg ccccaagttc 1740
tgctccatgc agatcaccca ggacctccgc gagtacgcca agaaccatca aatggaggag 1800
gacgaggcga tccaaaccgg catggccgag atgtccgagc agttcaagag catcggcgcg 1860
gaggtctacc tcgacgaggc ggtgctggag cacgcgtaa 1899

密码子优化的酿酒酵母suc2序列

<210> 129
<211> 1613
<212> DNA
<213> 酿酒酵母
<400> 129

ggcgcgccat gctgctgcag gccttcctgt tcctgctggc cggcttcgcc gccaagatca 60
gcgcctccat gacgaacgag acgtccgacc gccccctggt gcacttcacc cccaacaagg 120
gctggatgaa cgaccccaac ggcctgtggt acgacgagaa ggacgccaag tggcacctgt 180
acttccagta caacccgaac gacaccgtct gggggacgcc cttgttctgg ggccacgcca 240
cgtccgacga cctgaccaac tgggaggacc agcccatcgc catcgccccg aagcgcaacg 300
actccggcgc cttctccggc tccatggtgg tggactacaa caacacctcc ggcttcttca 360
acgacaccat cgacccgcgc cagcgctgcg tggccatctg gacctacaac accccggagt 420
ccgaggagca gtacatctcc tacagcctgg acggcggcta caccttcacc gagtaccaga 480
agaaccccgt gctggccgcc aactccaccc agttccgcga cccgaaggtc ttctggtacg 540
agccctccca gaagtggatc atgaccgcgg ccaagtccca ggactacaag atcgagatct 600
actcctccga cgacctgaag tcctggaagc tggagtccgc gttcgccaac gagggcttcc 660
tcggctacca gtacgagtgc cccggcctga tcgaggtccc caccgagcag gaccccagca 720
agtcctactg ggtgatgttc atctccatca accccggcgc cccggccggc ggctccttca 780
accagtactt cgtcggcagc ttcaacggca cccacttcga ggccttcgac aaccagtccc 840
gcgtggtgga cttcggcaag gactactacg ccctgcagac cttcttcaac accgacccga 900
cctacgggag cgccctgggc atcgcgtggg cctccaactg ggagtactcc gccttcgtgc 960
ccaccaaccc ctggcgctcc tccatgtccc tcgtgcgcaa gttctccctc aacaccgagt 1020
accaggccaa cccggagacg gagctgatca acctgaaggc cgagccgatc ctgaacatca 1080
gcaacgccgg cccctggagc cggttcgcca ccaacaccac gttgacgaag gccaacagct 1140
acaacgtcga cctgtccaac agcaccggca ccctggagtt cgagctggtg tacgccgtca 1200
acaccaccca gacgatctcc aagtccgtgt tcgcggacct ctccctctgg ttcaagggcc 1260
tggaggaccc cgaggagtac ctccgcatgg gcttcgaggt gtccgcgtcc tccttcttcc 1320
tggaccgcgg gaacagcaag gtgaagttcg tgaaggagaa cccctacttc accaaccgca 1380
tgagcgtgaa caaccagccc ttcaagagcg agaacgacct gtcctactac aaggtgtacg 1440
gcttgctgga ccagaacatc ctggagctgt acttcaacga cggcgacgtc gtgtccacca 1500

acacctactt catgaccacc gggaacgccc tgggctccgt gaacatgacg acgggggtgg 1560
acaacctgtt ctacatcgac aagttccagg tgcgcgaggt caagtgacaa ttg 1613

具有原始小球藻转运肽的胶球藻属C-169 THIC表达构建体的相关序列

<210> 130
<211> 3660
<212> DNA
<213> 人工序列
<220>
<223> 合成的核苷酸
<400> 130

gcggccgcgt ggacgagggc tacaaccccg cctacggcgc gcgcccgctg cgccgcgcca 60
tcatgcgcct gctggaggac gcgctggccg agcgcatgct cgccggcgac gtcaaggagg 120
gcgactcggt catcatggac gtcgatggcg atggcgccat cagcgtcctc aacggagacc 180
gcacccacac caccaccatc gactcctccc cggcgggcat ctcgtagacg cgagagggag 240
ggagggagga ttttcagagg gagatgagac gaggactggg ccgggggcct tggggcgctg 300
ctggagcgtg gtgagagcgc ggcggacgtg ccttttcttc ttccgtgcgc gcgctcttgg 360
ccattgatcc ccgattcgcg cccgcatccc cccactgccc ccatcatctt gcctgttgtc 420
gtggcactga cataaacccc ctgcgctgcg ctgctccgct actattgata taggtctcac 480
gcgccaatct tttttgctcc gggtaaccgt ctggacgcca gaattccttt cttgcgctat 540
gacacttcca gcaaaaggta gggcgggctg cgagacggct tcccggcgct gcatgcaaca 600
ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc atgggcgctc cgatgccgct 660
ccagggcgag cgctgtttaa atagccaggc ccccgattgc aaagacatta tagcgagcta 720
ccaaagccat attcaaacac ctagatcact accacttcta cacaggccac tcgagcttgt 780
gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt cagtcacaac ccgcaaacac 840
tagtatggcc accgcatcca ctttctcggc gttcaatgcc cgctgcggcg acctgcgtcg 900
ctcggcgggc tccgggcccc ggcgcccagc gaggcccctc cccgtgcgcg ggcgcgccgc 960
gatcgccgct cccgagcggc tggactacct cgacaacgcg gaggaggcgc ggctgcagca 1020
gacggacgcc ttcgccgagc tgaaggcgct gtcgagccgc cagtccgtga accggcccca 1080
aaagggcgag ctgtcgttcc ggcagtcgcc gacctttcag gactgcttcc ctggctcgga 1140
gaagtgctac cgcgaggtgg agcacgatgg caagacgctc aaggtgcctt ttcgccgggt 1200
gcacctgcag gacgacaacg gccatttcga cctgtacgac acctctggcc cccagggcgt 1260
gaacccccgc gagggtctgc ccaagatccg ctcctcctgg gtggagccgc gcgaggcgcg 1320
gggcgacaag gtgcaaacgc agcagtacta cgccaagcaa ggcatcatca ccgaggagat 1380
ggccttctgc gctgcgcgcg agcgcatgga ccccgagttc atccggtctg aggtggcccg 1440
gggtcgggcc atcatccctg cgaacaagcg ccatctggag ctggagccga ccgtcgtggg 1500
ccgcaacttt ctggtcaagg tgaacgccaa catcggcaac tctgcggtca gctcgagcat 1560
cgaggaggag gtggagaagc tgcagtggag caccatctgg ggtgccgaca cggccatgga 1620
cctgagcacc ggcaacaaca tccacgagac gcgcgagtgg gtgatgcgca acagcccggt 1680
ccctgtgggt acggtcccta tctaccagtg cctggagaag gcgggcggca tcgtggagaa 1740

catcacctgg gagctgttcc gggagacgct gatcgagcag gccgagcagg gcgtcgacta 1800
ctttaccatc cacgcggggg tcctgctccg ctacatcccc ctgacggcga accgcgtcac 1860
cggcatcgtc agccgcggcg gctctatcca tgccaagctc tgcctgctgg accacaccga 1920
gaacttcgcc tacatgcact gggacgagat cctggacatc tgcgcccagt acgacatcac 1980
cctgtcgatc ggcgacggcc tgcgccccgg ctgcatcgcg gacgccaacg acgctgccca 2040
gtttgctgag ctgaagaccc agggcgagct gacccgccgc gcctgggcga aggacgtgca 2100
ggtgatgaac gagggtcccg gccatgtgcc gctgcacaag atccccgaga acatggagaa 2160
gcagctggac tggtgcagcg aggcgccctt ctacacgctc ggccccctgg cgacggacat 2220
cgcgcctgcg tacgaccaca tcacctccgc gatcggcgcc gccacgatcg gcgcgctggg 2280
gaccgcgctg ctgtgttacg tgaccccgaa ggagcacctg ggcctccccg atcgcgacga 2340
cgtgaaggcc ggggtgatcg cctacaagat cgcggcgcac gcggctgacc tggcgaaggg 2400
tcacccctac gctcaggagt gggataacgc cctctctaag gcgcgcttcg agttccggtg 2460
gtacgaccag ttccacctgt cgctggaccc cgtcaccgcc cgcctgttcc atgacgcgac 2520
cctgccgcag gagcccgcca agaccgccca cttctgctcc atgtgcggcc ccaagttctg 2580
ctccatgcag atcacccagg acctccgcga gtacgccaag aaccatcaaa tggaggagga 2640
cgaggcgatc caaaccggca tggccgagat gtccgagcag ttcaagagca tcggcgcgga 2700
ggtctacctc gacgaggcgg tgctggagca cgcgtaacaa ttgaagcagc agcagctcgg 2760
atagtatcga cacactctgg acgctggtcg tgtgatggac tgttgccgcc acacttgctg 2820
ccttgacctg tgaatatccc tgccgctttt atcaaacagc ctcagtgtgt ttgatcttgt 2880
gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt tgcgaatacc acccccagca 2940
tccccttccc tcgtttcata tcgcttgcat cccaaccgca acttatctac gctgtcctgc 3000
tatccctcag cgctgctcct gctcctgctc actgcccctc gcacagcctt ggtttgggct 3060
ccgcctgtat tctcctggta ctgcaacctg taaaccagca ctgcaatgct gatgcacggg 3120
aagtagtggg atgggaacac aaatggaaag cttgagctca gaatagtatc gggtgatgcg 3180
aagtcagaac caggcagggc ctgtcgcctg aggtggcaac gatgggaagc aatcaatctg 3240
ggtacagtcg tccgcacgat cccgtgatct cccccaccga cacctatccc cgcccatccc 3300
ggcccaccct ttcagtcccc tcagcatgca ttgtgcaccg cgacaaagca tgtctgctcg 3360
tgcactggtt caggccacgg cgcaccgagt cctcgccctt cgcagagtga tcaccctccc 3420
cggaaccagc cacgctcgct gctgcgggcc gatcagccgc gcgcactccc tgcaactagg 3480
gacaactcag gcaaccacgc gcctcacaag catggccgcc gtggcatcca acccgctcgt 3540
gacggtgggt gcgcaagtgc caggggcctc gtcgtcacgg cgtgcatcct cgagggatgc 3600
gatccggcaa ctatatgtcg tttatctccc caccaatcac aggatgagcc cctgtctaga 3660

具有原始小球藻转运肽的密码子优化的胶球藻属C-169 THIC序列

<210> 131
<211> 1905
<212> DNA
<213> 人工序列
<220>
<223> 合成的核苷酸

<400> 131

actagtatgg ccaccgcatc cactttctcg gcgttcaatg cccgctgcgg cgacctgcgt 60
cgctcggcgg gctccgggcc ccggcgccca gcgaggcccc tccccgtgcg cgggcgcgcc 120
gcgatcgccg ctcccgagcg gctggactac ctcgacaacg cggaggaggc gcggctgcag 180
cagacggacg ccttcgccga gctgaaggcg ctgtcgagcc gccagtccgt gaaccggccc 240
caaaagggcg agctgtcgtt ccggcagtcg ccgacctttc aggactgctt ccctggctcg 300
gagaagtgct accgcgaggt ggagcacgat ggcaagacgc tcaaggtgcc ttttcgccgg 360
gtgcacctgc aggacgacaa cggccatttc gacctgtacg acacctctgg cccccagggc 420
gtgaaccccc gcgagggtct gcccaagatc cgctcctcct gggtggagcc gcgcgaggcg 480
cggggcgaca aggtgcaaac gcagcagtac tacgccaagc aaggcatcat caccgaggag 540
atggccttct gcgctgcgcg cgagcgcatg gaccccgagt tcatccggtc tgaggtggcc 600
cggggtcggg ccatcatccc tgcgaacaag cgccatctgg agctggagcc gaccgtcgtg 660
ggccgcaact ttctggtcaa ggtgaacgcc aacatcggca actctgcggt cagctcgagc 720
atcgaggagg aggtggagaa gctgcagtgg agcaccatct ggggtgccga cacggccatg 780
gacctgagca ccggcaacaa catccacgag acgcgcgagt gggtgatgcg caacagcccg 840
gtccctgtgg gtacggtccc tatctaccag tgcctggaga aggcgggcgg catcgtggag 900
aacatcacct gggagctgtt ccgggagacg ctgatcgagc aggccgagca gggcgtcgac 960
tactttacca tccacgcggg ggtcctgctc cgctacatcc ccctgacggc gaaccgcgtc 1020
accggcatcg tcagccgcgg cggctctatc catgccaagc tctgcctgct ggaccacacc 1080
gagaacttcg cctacatgca ctgggacgag atcctggaca tctgcgccca gtacgacatc 1140
accctgtcga tcggcgacgg cctgcgcccc ggctgcatcg cggacgccaa cgacgctgcc 1200
cagtttgctg agctgaagac ccagggcgag ctgacccgcc gcgcctgggc gaaggacgtg 1260
caggtgatga acgagggtcc cggccatgtg ccgctgcaca agatccccga gaacatggag 1320
aagcagctgg actggtgcag cgaggcgccc ttctacacgc tcggcccccc ggcgacggac 1380
atcgcgcctg cgtacgacca catcacctcc gcgatcggcg ccgccacgat cggcgcgctg 1440
gggaccgcgc tgctgtgtta cgtgaccccg aaggagcacc tgggcctccc cgatcgcgac 1500
gacgtgaagg ccggggtgat cgcctacaag atcgcggcgc acgcggctga cctggcgaag 1560
ggtcacccct acgctcagga gtgggataac gccctctcta aggcgcgctt cgagttccgg 1620
tggtacgacc agttccacct gtcgctggac cccgtcaccg cccgcctgtt ccatgacgcg 1680
accctgccgc aggagcccgc caagaccgcc cacttctgct ccatgtgcgg ccccaagttc 1740
tgctccatgc agatcaccca ggacctccgc gagtacgcca agaaccatca aatggaggag 1800
gacgaggcga tccaaaccgg catggccgag atgtccgagc agttcaagag catcggcgcg 1860
gaggtctacc tcgacgaggc ggtgctggag cacgcgtaac aattg 1905

原始小球藻肌动蛋白启动子/5'UTR

<210> 132

<211> 1182

<212> DNA

<213> 原始小球藻

<400> 132

actagagagt ttaggtccag cgtccgtggg gggggacggg ctgggagctt gggccgggaa 60
gggcaagacg atgcagtccc tctggggagt cacagccgac tgtgtgtgtt gcactgtgcg 120
gcccgcagca ctcacacgca aaatgcctgg ccgacaggca ggccctgtcc agtgcaacat 180
ccacggtccc tctcatcagg ctcaccttgc tcattgacat aacggaatgc gtaccgctct 240
ttcagatctg tccatccaga gaggggagca ggctccccac cgacgctgtc aaacttgctt 300
cctgcccaac cgaaaacatt attgtttgag gggggggggg gggggcaga ttgcatggcg 360
ggatatctcg tgaggaacat cactgggaca ctgtggaaca cagtgagtgc agtatgcaga 420
gcatgtatgc taggggtcag cgcaggaagg gggcctttcc cagtctccca tgccactgca 480
ccgtatccac gactcaccag gaccagcttc ttgatcggct tccgctcccg tggacaccag 540
tgtgtagcct ctggactcca ggtatgcgtg caccgcaaag gccagccgat cgtgccgatt 600
cctggggtgg aggatatgag tcagccaact tggggctcag agtgcacact ggggcacgat 660
acgaaacaac atctacaccg tgtcctccat gctgacacac cacagcttcg ctccacctga 720
atgtgggcgc atgggcccga atcacagcca atgtcgctgc tgccataatg tgatccagac 780
cctctccgcc cagatgccga gcggatcgtg ggcgctgaat agattcctgt ttcgatcact 840
gtttgggtcc tttccttttc gtctcggatg cgcgtctcga aacaggctgc gtcgggcttt 900
cggatccctt ttgctccctc cgtcaccatc ctgcgcgcgg gcaagttgct tgaccctggg 960
ctggtaccag ggttggaggg tattaccgcg tcaggccatt cccagcccgg attcaattca 1020
aagtctgggc caccaccctc cgccgctctg tctgatcact ccacattcgt gcatacacta 1080
cgttcaagtc ctgatccagg cgtgtctcgg gacaaggtgt gcttgagttt gaatctcaag 1140
gacccactcc agcacagctg ctggttgacc ccgccctcgc aa 1182

原始小球藻EF1A 3'UTR

<210> 133
<211> 161
<212> DNA
<213> 原始小球藻
<400> 133

acggagcgtc gtgcgggagg gagtgtgccg agcggggagt cccggtctgt gcgaggcccg 60
gcagctgacg ctggcgagcc gtacgccccg agggtccccc tcccctgcac cctcttcccc 120
ttccctctga cggccgcgcc tgttcttgca tgttcagcga c 161

拟南芥THIC表达盒的相关序列

<210> 134
<211> 6306
<212> DNA
<213> 拟南芥
<400> 134

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60
ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120
tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180
ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240

gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300
ggaagacagg tgaggggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360
atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420
cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480
gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540
ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600
ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660
cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720
ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 960
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 1020
cagtcacaac ccgcaaactc tagaatatca atgatcgagc aggacggcct ccacgccggc 1080
tcccccgccg cctgggtgga gcgcctgttc ggctacgact gggcccagca gaccatcggc 1140
tgctccgacg ccgccgtgtt ccgcctgtcc gcccagggcc gccccgtgct gttcgtgaag 1200
accgacctgt ccggcgccct gaacgagctg caggacgagg ccgcccgcct gtcctggctg 1260
gccaccaccg gcgtgccctg cgccgccgtg ctggacgtgg tgaccgaggc cggccgcgac 1320
tggctgctgc tgggcgaggt gcccggccag gacctgctgt cctcccacct ggcccccgcc 1380
gagaaggtgt ccatcatggc cgacgccatg cgccgcctgc acaccctgga ccccgccacc 1440
tgcccccttcg accaccaggc caagcaccgc atcgagcgcg cccgcacccg catggaggcc 1500
ggcctggtgg accaggacga cctggacgag gagcaccagg gcctggcccc cgccgagctg 1560
ttcgcccgcc tgaaggcccg catgcccgac ggcgaggacc tggtggtgac ccacggcgac 1620
gcctgcctgc ccaacatcat ggtggagaac ggccgcttct ccggcttcat cgactgcggc 1680
cgcctgggcg tggccgaccg ctaccaggac atcgccctgg ccacccgcga catcgccgag 1740
gagctgggcg gcgagtgggc cgaccgcttc ctggtgctgt acggcatcgc cgcccccgac 1800
tcccagcgca tcgccttcta ccgcctgctg gacgagttct tctgacaatt ggcagcagca 1860
gctcggatag tatcgacaca ctctggacgc tggtcgtgtg atggactgtt gccgccacac 1920
ttgctgcctt gacctgtgaa tatccctgcc gcttttatca aacagcctca gtgtgtttga 1980
tcttgtgtgt acgcgctttt gcgagttgct agctgcttgt gctatttgcg aataccaccc 2040
ccagcatccc cttccctcgt ttcatatcgc ttgcatccca accgcaactt atctacgctg 2100
tcctgctatc cctcagcgct gctcctgctc ctgctcactg cccctcgcac agccttggtt 2160
tgggctccgc ctgtattctc ctggtactgc aacctgtaaa ccagcactgc aatgctgatg 2220
cacgggaagt agtgggatgg gaacacaaat ggaggatcca ctagagagtt taggtccagc 2280
gtccgtgggg ggggacgggc tgggagcttg ggccgggaag ggcaagacga tgcagtccct 2340
ctggggagtc acagccgact gtgtgtgttg cactgtgcgg cccgcagcac tcacacgcaa 2400
aatgcctggc cgacaggcag gccctgtcca gtgcaacatc cacggtccct ctcatcaggc 2460
tcaccttgct cattgacata acggaatgcg taccgctctt tcagatctgt ccatccagag 2520
aggggagcag gctccccacc gacgctgtca aacttgcttc ctgcccaacc gaaaacatta 2580

ttgtttgagg gggggggggg gggggcagat tgcatggcgg gatatctcgt gaggaacatc 2640
actgggacac tgtggaacac agtgagtgca gtatgcagag catgtatgct aggggtcagc 2700
gcaggaaggg ggcctttccc agtctcccat gccactgcac cgtatccacg actcaccagg 2760
accagcttct tgatcggctt ccgctcccgt ggacaccagt gtgtagcctc tggactccag 2820
gtatgcgtgc accgcaaagg ccagccgatc gtgccgattc ctggggtgga ggatatgagt 2880
cagccaactt ggggctcaga gtgcacactg gggcacgata cgaaacaaca tctacaccgt 2940
gtcctccatg ctgacacacc acagcttcgc tccacctgaa tgtgggcgca tgggcccgaa 3000
tcacagccaa tgtcgctgct gccataatgt gatccagacc ctctccgccc agatgccgag 3060
cggatcgtgg gcgctgaata gattcctgtt tcgatcactg tttgggtcct ttccttttcg 3120
tctcggatgc gcgtctcgaa acaggctgcg tcgggctttc ggatcccttt tgctccctcc 3180
gtcaccatcc tgcgcgcggg caagttgctt gaccctgggc tggtaccagg gttggagggt 3240
attaccgcgt caggccattc ccagcccgga ttcaattcaa agtctgggcc accaccctcc 3300
gccgctctgt ctgatcactc cacattcgtg catacactac gttcaagtcc tgatccaggc 3360
gtgtctcggg acaaggtgtg cttgagtttg aatctcaagg acccactcca gcacagctgc 3420
tggttgaccc cgccctcgca aactagtatg gccaccgcat ccactttctc ggcgttcaat 3480
gcccgctgcg gcgacctgcg tcgctcggcg ggctccgggc cccggcgccc agcgaggccc 3540
ctccccgtgc gcgggcgcgc cgtccaggcc gcggccaccc gcttcaagaa ggagacgacg 3600
accacccgcg ccacgctgac gttcgacccc cccacgacca actccgagcg cgccaagcag 3660
cgcaagcaca ccatcgaccc ctcctccccc gacttccagc ccatcccctc cttcgaggag 3720
tgcttcccca agtccacgaa ggagcacaag gaggtggtgc acgaggagtc cggccacgtc 3780
ctgaaggtgc ccttccgccg cgtgcacctg tccggcggcg agcccgcctt cgacaactac 3840
gacacgtccg gcccccagaa cgtcaacgcc cacatcggcc tggcgaagct gcgcaaggag 3900
tggatcgacc gccgcgagaa gctgggcacg ccccgctaca cgcagatgta ctacgcgaag 3960
cagggcatca tcacggagga gatgctgtac tgcgcgacgc gcgagaagct ggaccccgag 4020
ttcgtccgct ccgaggtcgc gcggggccgc gccatcatcc cctccaacaa gaagcacctg 4080
gagctggagc ccatgatcgt gggccgcaag ttcctggtga aggtgaacgc gaacatcggc 4140
aactccgccg tggcctcctc catcgaggag gaggtctaca aggtgcagtg ggccaccatg 4200
tggggcgccg acaccatcat ggacctgtcc acgggccgcc acatccacga gacgcgcgag 4260
tggatcctgc gcaactccgc ggtccccgtg ggcaccgtcc ccatctacca ggcgctggag 4320
aaggtggacg gcatcgcgga gaacctgaac tgggaggtgt tccgcgagac gctgatcgag 4380
caggccgagc agggcgtgga ctacttcacg atccacgcgg gcgtgctgct gcgctacatc 4440
cccctgaccg ccaagcgcct gacgggcatc gtgtcccgcg gcggctccat ccacgcgaag 4500
tggtgcctgg cctaccacaa ggagaacttc gcctacgagc actgggacga catcctggac 4560
atctgcaacc agtacgacgt cgccctgtcc atcggcgacg gcctgcgccc cggctccatc 4620
tacgacgcca acgacacggc ccagttcgcc gagctgctga cccagggcga gctgacgcgc 4680
cgcgcgtggg agaaggacgt gcaggtgatg aacgagggcc ccggccacgt gcccatgcac 4740
aagatccccg agaacatgca gaagcagctg gagtggtgca acgaggcgcc cttctacacc 4800
ctgggccccc tgacgaccga catcgcgccc ggctacgacc acatcacctc cgccatcggc 4860
gcggccaaca tcggcgccct gggcaccgcc ctgctgtgct acgtgacgcc caaggagcac 4920

ctgggcctgc ccaaccgcga cgacgtgaag gcgggcgtca tcgcctacaa gatcgccgcc 4980
cacgcggccg acctggccaa gcagcacccc cacgcccagg cgtgggacga cgcgctgtcc 5040
aaggcgcgct tcgagttccg ctggatggac cagttcgcgc tgtccctgga ccccatgacg 5100
gcgatgtcct tccacgacga gacgctgccc gcggacggcg cgaaggtcgc ccacttctgc 5160
tccatgtgcg gccccaagtt ctgctccatg aagatcacgg aggacatccg caagtacgcc 5220
gaggagaacg gctacggctc cgccgaggag gccatccgcc agggcatgga cgccatgtcc 5280
gaggagttca acatcgccaa gaagacgatc tccggcgagc agcacggcga ggtcggcggc 5340
gagatctacc tgcccgagtc ctacgtcaag gccgcgcaga agtgatacgt actcgagacg 5400
gagcgtcgtg cgggagggag tgtgccgagc ggggagtccc ggtctgtgcg aggcccggca 5460
gctgacgctg gcgagccgta cgccccgagg gtccccctcc cctgcaccct cttcccttc 5520
cctctgacgg ccgcgcctgt tcttgcatgt tcagcgacga gctcttgttt tccagaagga 5580
gttgctcctt gagcctttca ttctcagcct cgataacctc caaagccgct ctaattgtgg 5640
agggggttcg aatttaaaag cttggaatgt tggttcgtgc gtctggaaca agcccagact 5700
tgttgctcac tgggaaaagg accatcagct ccaaaaaact tgccgctcaa accgcgtacc 5760
tctgctttcg cgcaatctgc cctgttgaaa tcgccaccac attcatattg tgacgcttga 5820
gcagtctgta attgcctcag aatgtggaat catctgcccc ctgtgcgagc ccatgccagg 5880
catgtcgcgg gcgaggacac ccgccactcg tacagcagac cattatgcta cctcacaata 5940
gttcataaca gtgaccatat ttctcgaagc tccccaacga gcacctccat gctctgagtg 6000
gccacccccc ggccctggtg cttgcggagg gcaggtcaac cggcatgggg ctaccgaaat 6060
ccccgaccgg atcccaccac ccccgcgatg ggaagaatct ctccccggga tgtgggccca 6120
ccaccagcac aacctgctgg cccaggcgag cgtcaaacca taccacacaa atatccttgg 6180
catcggccct gaattccttc tgccgctctg ctacccggtg cttctgtccg aagcaggggt 6240
tgctagggat cgctccgagt ccgcaaaccc ttgtcgcgtg gcggggcttg ttcgagcttg 6300
aagagc 6306

具有原始小球藻转运肽的密码子优化的拟南芥THIC

<210> 135
<211> 1938
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 135

atggccaccg catccacttt ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg 60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgccgtccag 120
gccgcggcca cccgcttcaa gaaggagacg acgaccaccc gcgccacgct gacgttcgac 180
ccccccacga ccaactccga gcgcgccaag cagcgcaagc acaccatcga cccctcctcc 240
cccgacttcc agcccatccc ctccttcgag gagtgcttcc ccaagtccac gaaggagcac 300
aaggaggtgg tgcacgagga gtccggccac gtcctgaagg tgcccttccg ccgcgtgcac 360
ctgtccggcg gcgagcccgc cttcgacaac tacgacacgt ccggccccca gaacgtcaac 420

gcccacatcg gcctggcgaa gctgcgcaag gagtggatcg accgccgcga gaagctgggc 480
acgccccgct acacgcagat gtactacgcg aagcagggca tcatcacgga ggagatgctg 540
tactgcgcga cgcgcgagaa gctggacccc gagttcgtcc gctccgaggt cgcgcggggc 600
cgcgccatca tcccctccaa caagaagcac ctggagctgg agcccatgat cgtgggccgc 660
aagttcctgg tgaaggtgaa cgcgaacatc ggcaactccg ccgtggcctc ctccatcgag 720
gaggaggtct acaaggtgca gtgggccacc atgtggggcg ccgacaccat catggacctg 780
tccacgggcc gccacatcca cgagacgcgc gagtggatcc tgcgcaactc cgcggtcccc 840
gtgggcaccg tccccatcta ccaggcgctg gagaaggtgg acggcatcgc ggagaacctg 900
aactgggagg tgttccgcga gacgctgatc gagcaggccg agcagggcgt ggactacttc 960
acgatccacg cgggcgtgct gctgcgctac atccccctga ccgccaagcg cctgacgggc 1020
atcgtgtccc gcggcggctc catccacgcg aagtggtgcc tggcctacca caaggagaac 1080
ttcgcctacg agcactggga cgacatcctg gacatctgca accagtacga cgtcgccctg 1140
tccatcggcg acggcctgcg ccccggctcc atctacgacg ccaacgacac ggcccagttc 1200
gccgagctgc tgacccaggg cgagctgacg cgccgcgcgt gggagaagga cgtgcaggtg 1260
atgaacgagg gccccggcca cgtgcccatg cacaagatcc ccgagaacat gcagaagcag 1320
ctggagtggt gcaacgaggc gcccttctac accctgggcc ccctgacgac cgacatcgcg 1380
cccggctacg accacatcac ctccgccatc ggcgcggcca acatcggcgc cctgggcacc 1440
gccctgctgt gctacgtgac gcccaaggag cacctgggcc tgcccaaccg cgacgacgtg 1500
aaggcgggcg tcatcgccta caagatcgcc gcccacgcgg ccgacctggc caagcagcac 1560
ccccacgccc aggcgtggga cgacgcgctg tccaaggcgc gcttcgagtt ccgctggatg 1620
gaccagttcg cgctgtccct ggaccccatg acggcgatgt ccttccacga cgagacgctg 1680
cccgcggacg gcgcgaaggt cgcccacttc tgctccatgt gcggccccaa gttctgctcc 1740
atgaagatca cggaggacat ccgcaagtac gccgaggaga acggctacgg ctccgccgag 1800
gaggccatcc gccagggcat ggacgccatg tccgaggagt tcaacatcgc caagaagacg 1860
atctccggcg agcagcacgg cgaggtcggc ggcgagatct acctgcccga gtcctacgtc 1920
aaggccgcgc agaagtga 1938

拟南芥氨基酸序列(具有天然转运肽序列)

<210> 136
<211> 644
<212> PRT
<213> 拟南芥
<400> 136

Met Ala Ala Ser Val His Cys Thr Leu Met Ser Val Val Cys Asn Asn
1 5 10 15
Lys Asn His Ser Ala Arg Pro Lys Leu Pro Asn Ser Ser Leu Leu Pro
20 25 30
Gly Phe Asp Val Val Val Gln Ala Ala Ala Thr Arg Phe Lys Lys Glu
35 40 45
Thr Thr Thr Thr Arg Ala Thr Leu Thr Phe Asp Pro Pro Thr Thr Asn

50 55 60

Ser Glu Arg Ala Lys Gln Arg Lys His Thr Ile Asp Pro Ser Ser Pro
65 70 75 80

Asp Phe Gln Pro Ile Pro Ser Phe Glu Glu Cys Phe Pro Lys Ser Thr
85 90 95

Lys Glu His Lys Glu Val Val His Glu Glu Ser Gly His Val Leu Lys
100 105 110

Val Pro Phe Arg Arg Val His Leu Ser Gly Gly Glu Pro Ala Phe Asp
115 120 125

Asn Tyr Asp Thr Ser Gly Pro Gln Asn Val Asn Ala His Ile Gly Leu
130 135 140

Ala Lys Leu Arg Lys Glu Trp Ile Asp Arg Arg Glu Lys Leu Gly Thr
145 150 155 160

Pro Arg Tyr Thr Gln Met Tyr Tyr Ala Lys Gln Gly Ile Ile Thr Glu
165 170 175

Glu Met Leu Tyr Cys Ala Thr Arg Glu Lys Leu Asp Pro Glu Phe Val
180 185 190

Arg Ser Glu Val Ala Arg Gly Arg Ala Ile Ile Pro Ser Asn Lys Lys
195 200 205

His Leu Glu Leu Glu Pro Met Ile Val Gly Arg Lys Phe Leu Val Lys
210 215 220

Val Asn Ala Asn Ile Gly Asn Ser Ala Val Ala Ser Ser Ile Glu Glu
225 230 235 240

Glu Val Tyr Lys Val Gln Trp Ala Thr Met Trp Gly Ala Asp Thr Ile
245 250 255

Met Asp Leu Ser Thr Gly Arg His Ile His Glu Thr Arg Glu Trp Ile
260 265 270

Leu Arg Asn Ser Ala Val Pro Val Gly Thr Val Pro Ile Tyr Gln Ala
275 280 285

Leu Glu Lys Val Asp Gly Ile Ala Glu Asn Leu Asn Trp Glu Val Phe
290 295 300

Arg Glu Thr Leu Ile Glu Gln Ala Glu Gln Gly Val Asp Tyr Phe Thr
305 310 315 320

Ile His Ala Gly Val Leu Leu Arg Tyr Ile Pro Leu Thr Ala Lys Arg
325 330 335

Leu Thr Gly Ile Val Ser Arg Gly Gly Ser Ile His Ala Lys Trp Cys
340 345 350

Leu Ala Tyr His Lys Glu Asn Phe Ala Tyr Glu His Trp Asp Asp Ile
355 360 365

Leu Asp Ile Cys Asn Gln Tyr Asp Val Ala Leu Ser Ile Gly Asp Gly
370 375 380
Leu Arg Pro Gly Ser Ile Tyr Asp Ala Asn Asp Thr Ala Gln Phe Ala
385 390 395 400
Glu Leu Leu Thr Gln Gly Glu Leu Thr Arg Arg Ala Trp Glu Lys Asp
405 410 415
Val Gln Val Met Asn Glu Gly Pro Gly His Val Pro Met His Lys Ile
420 425 430
Pro Glu Asn Met Gln Lys Gln Leu Glu Trp Cys Asn Glu Ala Pro Phe
435 440 445
Tyr Thr Leu Gly Pro Leu Thr Thr Asp Ile Ala Pro Gly Tyr Asp His
450 455 460
Ile Thr Ser Ala Ile Gly Ala Ala Asn Ile Gly Ala Leu Gly Thr Ala
465 470 475 480
Leu Leu Cys Tyr Val Thr Pro Lys Glu His Leu Gly Leu Pro Asn Arg
485 490 495
Asp Asp Val Lys Ala Gly Val Ile Ala Tyr Lys Ile Ala Ala His Ala
500 505 510
Ala Asp Leu Ala Lys Gln His Pro His Ala Gln Ala Trp Asp Asp Ala
515 520 525
Leu Ser Lys Ala Arg Phe Glu Phe Arg Trp Met Asp Gln Phe Ala Leu
530 535 540
Ser Leu Asp Pro Met Thr Ala Met Ser Phe His Asp Glu Thr Leu Pro
545 550 555 560
Ala Asp Gly Ala Lys Val Ala His Phe Cys Ser Met Cys Gly Pro Lys
565 570 575
Phe Cys Ser Met Lys Ile Thr Glu Asp Ile Arg Lys Tyr Ala Glu Glu
580 585 590
Asn Gly Tyr Gly Ser Ala Glu Glu Ala Ile Arg Gln Gly Met Asp Ala
595 600 605
Met Ser Glu Glu Phe Asn Ile Ala Lys Lys Thr Ile Ser Gly Glu Gln
610 615 620
His Gly Glu Val Gly Gly Glu Ile Tyr Leu Pro Glu Ser Tyr Val Lys
625 630 635 640
Ala Ala Gln Lys

集胞藻种类thiC表达盒的相关序列

〈210〉 137

〈211〉 6130

〈212〉 DNA

<213> 未知

<220>

<223> 集胞藻属

<400> 137

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60
ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120
tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180
ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240
gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300
ggaagacagg tgaggggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360
atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420
cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480
gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540
ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600
ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660
cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720
ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 960
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 1020
cagtcacaac ccgcaaactc tagaatatca atgatcgagc aggacggcct ccacgccggc 1080
tcccccgccg cctgggtgga gcgcctgttc ggctacgact gggcccagca gaccatcggc 1140
tgctccgacg ccgccgtgtt ccgcctgtcc gcccagggcc gccccgtgct gttcgtgaag 1200
accgacctgt ccggcgccct gaacgagctg caggacgagg ccgcccgcct gtcctggctg 1260
gccaccaccg gcgtgccctg cgccgccgtg ctggacgtgg tgaccgaggc cggccgcgac 1320
tggctgctgc tgggcgaggt gcccggccag gacctgctgt cctcccacct ggcccccgcc 1380
gagaaggtgt ccatcatggc cgacgccatg cgccgcctgc acaccctgga ccccgccacc 1440
tgcccccttcg accaccaggc caagcaccgc atcgagcgcg cccgcacccg catggaggcc 1500
ggcctggtgg accaggacga cctggacgag gagcaccagg gcctggcccc cgccgagctg 1560
ttcgcccgcc tgaaggcccg catgcccgac ggcgaggacc tggtggtgac ccacggcgac 1620
gcctgcctgc ccaacatcat ggtggagaac ggccgcttct ccggcttcat cgactgcggc 1680
cgcctgggcg tggccgaccg ctaccaggac atcgccctgg ccacccgcga catcgccgag 1740
gagctgggcg gcgagtgggc cgaccgcttc ctggtgctgt acggcatcgc cgcccccgac 1800
tcccagcgca tcgccttcta ccgcctgctg gacgagttct tctgacaatt ggcagcagca 1860
gctcggatag tatcgacaca ctctggacgc tggtcgtgtg atggactgtt gccgccacac 1920
ttgctgcctt gacctgtgaa tatccctgcc gcttttatca aacagcctca gtgtgtttga 1980
tcttgtgtgt acgcgctttt gcgagttgct agctgcttgt gctatttgcg aataccaccc 2040
ccagcatccc cttccctcgt ttcatatcgc ttgcatccca accgcaactt atctacgctg 2100

tcctgctatc cctcagcgct gctcctgctc ctgctcactg cccctcgcac agccttggtt 2160
tgggctccgc ctgtattctc ctggtactgc aacctgtaaa ccagcactgc aatgctgatg 2220
cacgggaagt agtgggatgg gaacacaaat ggaggatcca ctagagagtt taggtccagc 2280
gtccgtgggg ggggacgggc tgggagcttg ggccgggaag ggcaagacga tgcagtccct 2340
ctggggagtc acagccgact gtgtgtgttg cactgtgcgg cccgcagcac tcacacgcaa 2400
aatgcctggc cgacaggcag gccctgtcca gtgcaacatc cacggtccct ctcatcaggc 2460
tcaccttgct cattgacata acggaatgcg taccgctctt tcagatctgt ccatccagag 2520
aggggagcag gctccccacc gacgctgtca aacttgcttc ctgcccaacc gaaaacatta 2580
ttgtttgagg gggggggggg gggggcagat tgcatggcgg gatatctcgt gaggaacatc 2640
actgggacac tgtggaacac agtgagtgca gtatgcagag catgtatgct aggggtcagc 2700
gcaggaaggg ggcctttccc agtctcccat gccactgcac cgtatccacg actcaccagg 2760
accagcttct tgatcggctt ccgctcccgt ggacaccagt gtgtagcctc tggactccag 2820
gtatgcgtgc accgcaaagg ccagccgatc gtgccgattc ctggggtgga ggatatgagt 2880
cagccaactt ggggctcaga gtgcacactg gggcacgata cgaaacaaca tctacaccgt 2940
gtcctccatg ctgacacacc acagcttcgc tccacctgaa tgtgggcgca tgggcccgaa 3000
tcacagccaa tgtcgctgct gccataatgt gatccagacc ctctccgccc agatgccgag 3060
cggatcgtgg gcgctgaata gattcctgtt tcgatcactg tttgggtcct ttccttttcg 3120
tctcggatgc gcgtctcgaa acaggctgcg tcgggctttc ggatcccttt tgctccctcc 3180
gtcaccatcc tgcgcgcggg caagttgctt gaccctgggc tggtaccagg gttggagggt 3240
attaccgcgt caggccattc ccagcccgga ttcaattcaa agtctgggcc accaccctcc 3300
gccgctctgt ctgatcactc cacattcgtg catacactac gttcaagtcc tgatccaggc 3360
gtgtctcggg acaaggtgtg cttgagtttg aatctcaagg acccactcca gcacagctgc 3420
tggttgaccc cgccctcgca aactagtatg gccaccgcat ccactttctc ggcgttcaat 3480
gcccgctgcg gcgacctgcg tcgctcggcg ggctccgggc cccggcgccc agcgaggccc 3540
ctccccgtgc gcgggcgcgc catgcgcacc gcgtgggtcg ccaagcgcca gggccagacg 3600
aacgtctccc agatgcacta cgcgcgcaag ggcgtgatca ccgaggagat ggactacgtg 3660
gcgaagcgcg agaacctgcc cgtggagctg atcaaggacg aggtggcgcg gggccgcatg 3720
atcatccccg ccaacatcaa ccacacgaac ctggagccca tggccatcgg catcgcctcc 3780
aagtgcaagg tcaacgcgaa catcggcgcg tcccccaact cctccaacat cgacgaggag 3840
gtggagaagc tgctgctgtc cgtgaagtac ggcgcggaca cggtgatgga cctgtccacc 3900
ggcggcggcg acctggacgt gatccgcacc gccatcatca acgcctcccc cgtgcccatc 3960
ggcacggtgc ccatctacca ggccctggag tccgtgcacg gctccatcga gaacctgacg 4020
cccgacgact tcctgcacat catcgagaag cacgcccagc agggcgtgga ctacatgacg 4080
atccacgcgg gcctgctgat cgagtacctg cccctggtga agtcccgcat cacgggcatc 4140
gtgtcccgcg gcggcggcat catcgccaag tggatgctgc accaccacaa gcagaacccc 4200
ctgtacacgc acttcgacga gatcatcgag atcttcaaga agtacgacgt gtccttctcc 4260
ctgggcgact ccctgcgccc cggctgcacc cacgacgcct ccgacgacgc ccagctgtcc 4320
gagctgaaga cgctgggcca gctgacgcgc cgcgcctggg agcacgacgt ccaggtcatg 4380
gtggagggcc ccggccacgt ccccatcgac cagatcgagt tcaacgtcaa gaagcagatg 4440

gaggagtgct ccgaggcccc cttctacgtg ctgggccccc tggtcaccga catcgcgccc 4500
ggctacgacc acatcacctc cgcgatcggc gccgccatcg cgggctggca cggcaccgcg 4560
atgctgtgct acgtgacgcc caaggagcac ctgggcctgc ccaacgcgga ggacgtgcgc 4620
aacggcctga tcgcctacaa gatcgccgcg cacgccgcgg acatcgcccg ccaccgcccc 4680
ggcgcgcgcg accgcgacga cgagctgtcc aaggcgcgct acaacttcga ctggaaccgc 4740
cagttcgagc tgtccctgga ccccgagcgc gcgaaggagt accacgacga gacgctgccc 4800
gcggacatct acaagaccgc ggagttctgc tccatgtgcg gccccaagtt ctgccccatg 4860
cagaccaagg tggacgccga gatgctggag gagctggagg tcttcctggc caaggacaag 4920
gagatggtgt cccagcgctg atacgtactc gagacggagc gtcgtgcggg agggagtgtg 4980
ccgagcgggg agtcccggtc tgtgcgaggc ccggcagctg acgctggcga gccgtacgcc 5040
ccgagggtcc ccctcccctg caccctcttc cccttccctc tgacggccgc gcctgttctt 5100
gcatgttcag cgacgagctc ttgttttcca gaaggagttg ctccttgagc ctttcattct 5160
cagcctcgat aacctccaaa gccgctctaa ttgtggaggg ggttcgaatt taaaagcttg 5220
gaatgttggt tcgtgcgtct ggaacaagcc cagacttgtt gctcactggg aaaaggacca 5280
tcagctccaa aaaacttgcc gctcaaaccg cgtacctctg ctttcgcgca atctgccctg 5340
ttgaaatcgc caccacattc atattgtgac gcttgagcag tctgtaattg cctcagaatg 5400
tggaatcatc tgccccctgt gcgagcccat gccaggcatg tcgcgggcga ggacacccgc 5460
cactcgtaca gcagaccatt atgctacctc acaatagttc ataacagtga ccatatttct 5520
cgaagctccc caacgagcac ctccatgctc tgagtggcca ccccccggcc ctggtgcttg 5580
cggagggcag gtcaaccggc atggggctac cgaaatcccc gaccggatcc caccaccccc 5640
gcgatgggaa gaatctctcc ccgggatgtg ggcccaccac cagcacaacc tgctggccca 5700
ggcgagcgtc aaaccatacc acacaaatat ccttggcatc ggccctgaat tccttctgcc 5760
gctctgctac ccggtgcttc tgtccgaagc aggggttgct agggatcgct ccgagtccgc 5820
aaacccttgt cgcgtggcgg ggcttgttcg agcttgaaga gcctctagag tcgacctgca 5880
ggcatgcaag cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc 5940
tcacaattcc acacaacata cgagccggaa gcataaagtg taaagcctgg ggtgcctaat 6000
gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc 6060
tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg 6120
ggcgctcttc 6130

具有原始小球藻转运肽的密码子优化的集胞藻种类thiC序列

<210> 138
<211> 1494
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 138

atggccaccg catccacttt ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg 60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgccatgcgc 120

accgcgtggg tcgccaagcg ccagggccag acgaacgtct cccagatgca ctacgcgcgc 180
aagggcgtga tcaccgagga gatggactac gtggcgaagc gcgagaacct gcccgtggag 240
ctgatcaagg acgaggtggc gcggggccgc atgatcatcc ccgccaacat caaccacacg 300
aacctggagc ccatggccat cggcatcgcc tccaagtgca aggtcaacgc gaacatcggc 360
gcgtccccca actcctccaa catcgacgag gaggtggaga agctgctgct gtccgtgaag 420
tacggcgcgg acacggtgat ggacctgtcc accggcggcg gcgacctgga cgtgatccgc 480
accgccatca tcaacgcctc ccccgtgccc atcggcacgg tgcccatcta ccaggccctg 540
gagtccgtgc acggctccat cgagaacctg acgcccgacg acttcctgca catcatcgag 600
aagcacgccc agcagggcgt ggactacatg acgatccacg cgggcctgct gatcgagtac 660
ctgcccctgg tgaagtcccg catcacgggc atcgtgtccc gcggcggcgg catcatcgcc 720
aagtggatgc tgcaccacca caagcagaac cccctgtaca cgcacttcga cgagatcatc 780
gagatcttca agaagtacga cgtgtccttc tccctgggcg actccctgcg ccccggctgc 840
acccacgacg cctccgacga cgcccagctg tccgagctga agacgctggg ccagctgacg 900
cgccgcgcct gggagcacga cgtccaggtc atggtggagg gccccggcca cgtccccatc 960
gaccagatcg agttcaacgt caagaagcag atggaggagt gctccgaggc ccccttctac 1020
gtgctgggcc ccctggtcac cgacatcgcg cccggctacg accacatcac ctccgcgatc 1080
ggcgccgcca tcgcgggctg gcacggcacc gcgatgctgt gctacgtgac gcccaaggag 1140
cacctgggcc tgcccaacgc ggaggacgtg cgcaacggcc tgatcgccta caagatcgcc 1200
gcgcacgccg cggacatcgc ccgccaccgc cccggcgcgc gcgaccgcga cgacgagctg 1260
tccaaggcgc gctacaactt cgactggaac cgccagttcg agctgtccct ggaccccgag 1320
cgcgcgaagg agtaccacga cgagacgctg cccgcggaca tctacaagac cgcggagttc 1380
tgctccatgt gcggccccaa gttctgcccc atgcagacca aggtggacgc cgagatgctg 1440
gaggagctgg aggtcttcct ggccaaggac aaggagatgg tgtcccagcg ctga 1494

集胞藻种类thiC氨基酸序列(具有天然转运肽序列)

<210> 139
<211> 459
<212> PRT
<213> 未知
<220>
<223> 集胞藻属
<400> 139

Met Arg Thr Ala Trp Val Ala Lys Arg Gln Gly Gln Thr Asn Val Ser
1 5 10 15
Gln Met His Tyr Ala Arg Lys Gly Val Ile Thr Glu Glu Met Asp Tyr
20 25 30
Val Ala Lys Arg Glu Asn Leu Pro Val Glu Leu Ile Lys Asp Glu Val
35 40 45
Ala Arg Gly Arg Met Ile Ile Pro Ala Asn Ile Asn His Thr Asn Leu
50 55 60

Glu Pro Met Ala Ile Gly Ile Ala Ser Lys Cys Lys Val Asn Ala Asn
65 70 75 80
Ile Gly Ala Ser Pro Asn Ser Ser Asn Ile Asp Glu Glu Val Glu Lys
85 90 95
Leu Leu Leu Ser Val Lys Tyr Gly Ala Asp Thr Val Met Asp Leu Ser
100 105 110
Thr Gly Gly Gly Asp Leu Asp Val Ile Arg Thr Ala Ile Ile Asn Ala
115 120 125
Ser Pro Val Pro Ile Gly Thr Val Pro Ile Tyr Gln Ala Leu Glu Ser
130 135 140
Val His Gly Ser Ile Glu Asn Leu Thr Pro Asp Asp Phe Leu His Ile
145 150 155 160
Ile Glu Lys His Ala Gln Gln Gly Val Asp Tyr Met Thr Ile His Ala
165 170 175
Gly Leu Leu Ile Glu Tyr Leu Pro Leu Val Lys Ser Arg Ile Thr Gly
180 185 190
Ile Val Ser Arg Gly Gly Gly Ile Ile Ala Lys Trp Met Leu His His
195 200 205
His Lys Gln Asn Pro Leu Tyr Thr His Phe Asp Glu Ile Ile Glu Ile
210 215 220
Phe Lys Lys Tyr Asp Val Ser Phe Ser Leu Gly Asp Ser Leu Arg Pro
225 230 235 240
Gly Cys Thr His Asp Ala Ser Asp Asp Ala Gln Leu Ser Glu Leu Lys
245 250 255
Thr Leu Gly Gln Leu Thr Arg Arg Ala Trp Glu His Asp Val Gln Val
260 265 270
Met Val Glu Gly Pro Gly His Val Pro Ile Asp Gln Ile Glu Phe Asn
275 280 285
Val Lys Lys Gln Met Glu Glu Cys Ser Glu Ala Pro Phe Tyr Val Leu
290 295 300
Gly Pro Leu Val Thr Asp Ile Ala Pro Gly Tyr Asp His Ile Thr Ser
305 310 315 320
Ala Ile Gly Ala Ala Ile Ala Gly Trp His Gly Thr Ala Met Leu Cys
325 330 335
Tyr Val Thr Pro Lys Glu His Leu Gly Leu Pro Asn Ala Glu Asp Val
340 345 350
Arg Asn Gly Leu Ile Ala Tyr Lys Ile Ala Ala His Ala Ala Asp Ile
355 360 365
Ala Arg His Arg Pro Gly Ala Arg Asp Arg Asp Asp Glu Leu Ser Lys

370 375 380

Ala Arg Tyr Asn Phe Asp Trp Asn Arg Gln Phe Glu Leu Ser Leu Asp
385 390 395 400
Pro Glu Arg Ala Lys Glu Tyr His Asp Glu Thr Leu Pro Ala Asp Ile
405 410 415
Tyr Lys Thr Ala Glu Phe Cys Ser Met Cys Gly Pro Lys Phe Cys Pro
420 425 430
Met Gln Thr Lys Val Asp Ala Glu Met Leu Glu Glu Leu Glu Val Phe
435 440 445
Leu Ala Lys Asp Lys Glu Met Val Ser Gln Arg
450 455

密码子优化的neoR基因

<210> 140
<211> 312
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 140

ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60
cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc 120
gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc aggcccccga ttgcaaagac 180
attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240
ccactcgagc ttgtgatcgc actccgctaa gggggcgcct cttcctcttc gtttcagtca 300
caacccgcaa ac 312

5’THIC胶球藻属C-169确认引物

<210> 141
<211> 18
<212> DNA
<213> 未知
<220>
<223> 胶球藻属
<400> 141

gggtgatcgc ctacaaga 18

3’THIC确认引物

<210> 142
<211> 20
<212> DNA
<213> 人工序列

<220>
<223> 合成的引物
<400> 142
acgtcgcgac ccatgcttcc 20
5’THIC拟南芥确认引物
<210> 143
<211> 18
<212> DNA
<213> 拟南芥
<400> 143
gcgtcatcgc ctacaaga 18
5’thiC集胞藻种类确认引物
<210> 144
<211> 18
<212> DNA
<213> 未知
<220>
<223> 集胞藻属
<400> 144
cgatgctgtg ctacgtga 18
来自pSZ1318的肉豆蔻广泛基础C14: 0-C18:0硫酯酶的密码子优化的编码区
<210> 145
<211> 1260
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 145
actagtatgg ctatcaagac gaacaggcag cctgtggaga agcctccgtt cacgatcggg 60
acgctgcgca aggccatccc cgcgcactgt ttcgagcgct cggcgcttcg tgggcgcgcc 120
gccaacgccc acaccgtgcc caagatcaac ggcaacaagg ccggcctgct gacccccatg 180
gagagcacca aggacgagga catcgtcgcg gcccccaccg tggcgcccaa gcgcaccttc 240
atcaaccagc tgcccgactg gtcgatgctg ctggccgcga tcaccaccat cttcctggcg 300
gccgagaagc agtggaccaa cctggactgg aagccccgcc gccccgacat gctggtcgac 360
ttcgacccct tctccctggg ccgcttcgtg caggacggcc tgatcttccg ccagaacttc 420
agcatccgct cgtacgagat cggcgcggac cgcaccgcct ccatcgagac cctgatgaac 480
cacctgcagg agaccgcgct gaaccacgtc cgctgcatcg gcctgctgga cgacggcttc 540
ggcagcaccc ccgagatgac ccgccgcgac ctgatctggg tggtcacccg catgcaggtc 600
ctggtggacc gctacccctc gtggggcgac gtgatcgagg tcgactcctg ggtgacccca 660

agcggcaaga acggcatgaa gcgcgagtgg ttcctgcgcg actgcaagac cggcgagatc 720
ctgacccgcg ccacctcggt ctgggtgatg atgaacaagc gcacccgccg cctgtccaag 780
atccccgagg aggtccgcgt ggagatcgag ccctacttcg tcgagcacgg cgtgctggac 840
gaggactcgc gcaagctgcc caagctgaac gacaacaccg ccaactacat ccgccgcggc 900
ctggcgcccc gctggtccga cctggacgtc aaccagcacg tgaacaacgt caagtacatc 960
ggctggatcc tggagagcgt gcccagcagc ctgctggagt cgcacgagct gtacggcatg 1020
accctggagt accgcaagga gtgcggcaag gacggcctgc tgcagtccct gaccgccgtc 1080
gccagcgact acggcggcgg ctcgctggag gccggcgtgg agtgcgacca cctgctgcgc 1140
ctggaggacg gctccgagat catgcgcggc aagaccgagt ggcgccccaa gcgcgccgcg 1200
aacaccacct acttcggcag cgtcgacgac atccccccg ccaacaacgc gtgactcgag 1260

肉豆蔻广泛基础C14: 0-C18:0硫酯酶的氨基酸序列(登录号AAB71729)

<210> 146
<211> 429
<212> PRT
<213> 肉豆蔻
<400> 146

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

180 185 190
Gly Ser Thr Pro Glu Met Thr Arg Arg Asp Leu Ile Trp Val Val Thr
195 200 205
Arg Met Gln Val Leu Val Asp Arg Tyr Pro Ser Trp Gly Asp Val Ile
210 215 220
Glu Val Asp Ser Trp Val Thr Pro Ser Gly Lys Asn Gly Met Lys Arg
225 230 235 240
Glu Trp Phe Leu Arg Asp Cys Lys Thr Gly Glu Ile Leu Thr Arg Ala
245 250 255
Thr Ser Val Trp Val Met Met Asn Lys Arg Thr Arg Arg Leu Ser Lys
260 265 270
Ile Pro Glu Glu Val Arg Val Glu Ile Glu Pro Tyr Phe Val Glu His
275 280 285
Gly Val Leu Asp Glu Asp Ser Arg Lys Leu Pro Lys Leu Asn Asp Asn
290 295 300
Thr Ala Asn Tyr Ile Arg Arg Gly Leu Ala Pro Arg Trp Ser Asp Leu
305 310 315 320
Asp Val Asn Gln His Val Asn Asn Val Lys Tyr Ile Gly Trp Ile Leu
325 330 335
Glu Ser Val Pro Ser Ser Leu Leu Glu Ser His Glu Leu Tyr Gly Met
340 345 350
Thr Leu Glu Tyr Arg Lys Glu Cys Gly Lys Asp Gly Leu Leu Gln Ser
355 360 365
Leu Thr Ala Val Ala Ser Asp Tyr Gly Gly Gly Ser Leu Glu Ala Gly
370 375 380
Val Glu Cys Asp His Leu Leu Arg Leu Glu Asp Gly Ser Glu Ile Met
385 390 395 400
Arg Gly Lys Thr Glu Trp Arg Pro Lys Arg Ala Ala Asn Thr Thr Tyr
405 410 415
Phe Gly Ser Val Asp Asp Ile Pro Pro Ala Asn Asn Ala
420 425

来自pSZ1452的山竹C16:0-偏好型硫酯酶的密码子优化的编码区

<210> 147
<211> 1269
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 147

actagtatgg ccaccgcatc cactttctcg gcgttcaatg cccgctgcgg cgacctgcgt 60
cgctcggcgg gctccgggcc ccggcgccca gcgaggcccc tccccgtgcg cgggcgcgcc 120
gccaacgccc aggcccccac caagatcaac ggctccaccg acgacgccca gctgcccgcc 180
ccccgcacct tcatcaacca gctgcccgac tggtccatgc tgctggccgc catcaccacc 240
gtgttcctgg ccgccgagaa gcagtggatg atgctggact ggaagccccg ccgccccgac 300
atgctgatcg acaccttcgg cctgggccgc atcgtgcagg acggcctggt gttccgccag 360
aacttctcca tccgctccta cgagatcggc gccgaccgca ccgcctccat cgagaccgtg 420
atgaaccacc tgcaggagac cgccctgaac cacgtgaaga ccgccggcct gctgggcgac 480
ggcttcggct ccacccccga gatgtccaag cgcaacctga tctgggtggt gaccaagatg 540
caggtggagg tggaccgcta ccccacctgg ggcgacgtgg tgcaggtgga cacctgggtg 600
tccgcctccg gcaagaacgg catgcgccgc gactggctgc tgcgcgacgg caacaccggc 660
gagaccctga cccgcgcctc ctccgtgtgg gtgatgatga acaagctgac ccgccgcctg 720
tccaagatcc ccgaggaggt gcgcgaggag atcggctcct acttcgtgaa ctccgacccc 780
gtggtggagg aggacggccg caaggtgacc aagctggacg acaacaccgc cgacttcgtg 840
cgcaagggcc tgacccccaa gtggaacgac ctggacatca accagcacgt gaacaacgtg 900
aagtacatcg gctggatcct ggagtccgcc ccccagccca tcctggagac ccgcgagctg 960
tccgccgtga ccctggagta ccgccgcgag tgcggccgcg actccgtgct gcgctccctg 1020
accgccgtgt ccggcggcgg cgtgggcgac ctgggccacg ccggcaacgt ggagtgccag 1080
cacgtgctgc gcctggagga cggcgccgag atcgtgcgcg gccgcaccga gtggcgcccc 1140
aagtacatca acaacttctc catcatgggc cagatcccca ccgacgcctc catggactac 1200
aaggaccacg acggcgacta caaggaccac gacatcgact acaaggacga cgacgacaag 1260
tgaatcgat 1269

山竹C16:0-偏好型硫酯酶的氨基酸序列(登录号AAB51525.1)

<210> 148
<211> 406
<212> PRT
<213> 山竹
<400> 148

Met Val Ala Thr Ala Ala Thr Ser Ser Phe Phe Pro Leu Thr Ser Pro
1 5 10 15
Ser Gly Asp Ala Lys Ser Gly Asn Pro Gly Lys Gly Ser Val Ser Phe
20 25 30
Gly Ser Met Lys Ser Lys Ser Ala Ala Ser Ser Arg Gly Leu Gln Val
35 40 45
Lys Ala Asn Ala Gln Ala Pro Thr Lys Ile Asn Gly Ser Thr Asp Asp
50 55 60
Ala Gln Leu Pro Ala Pro Arg Thr Phe Ile Asn Gln Leu Pro Asp Trp
65 70 75 80
Ser Met Leu Leu Ala Ala Ile Thr Thr Val Phe Leu Ala Ala Glu Lys

85 90 95

Gln Trp Met Met Leu Asp Trp Lys Pro Arg Arg Pro Asp Met Leu Ile
100 105 110

Asp Thr Phe Gly Leu Gly Arg Ile Val Gln Asp Gly Leu Val Phe Arg
115 120 125

Gln Asn Phe Ser Ile Arg Ser Tyr Glu Ile Gly Ala Asp Arg Thr Ala
130 135 140

Ser Ile Glu Thr Val Met Asn His Leu Gln Glu Thr Ala Leu Asn His
145 150 155 160

Val Lys Thr Ala Gly Leu Leu Gly Asp Gly Phe Gly Ser Thr Pro Glu
165 170 175

Met Ser Lys Arg Asn Leu Ile Trp Val Val Thr Lys Met Gln Val Glu
180 185 190

Val Asp Arg Tyr Pro Thr Trp Gly Asp Val Val Gln Val Asp Thr Trp
195 200 205

Val Ser Ala Ser Gly Lys Asn Gly Met Arg Arg Asp Trp Leu Leu Arg
210 215 220

Asp Gly Asn Thr Gly Glu Thr Leu Thr Arg Ala Ser Ser Val Trp Val
225 230 235 240

Met Met Asn Lys Leu Thr Arg Arg Leu Ser Lys Ile Pro Glu Glu Val
245 250 255

Arg Glu Glu Ile Gly Ser Tyr Phe Val Asn Ser Asp Pro Val Val Glu
260 265 270

Glu Asp Gly Arg Lys Val Thr Lys Leu Asp Asp Asn Thr Ala Asp Phe
275 280 285

Val Arg Lys Gly Leu Thr Pro Lys Trp Asn Asp Leu Asp Ile Asn Gln
290 295 300

His Val Asn Asn Val Lys Tyr Ile Gly Trp Ile Leu Glu Ser Ala Pro
305 310 315 320

Gln Pro Ile Leu Glu Thr Arg Glu Leu Ser Ala Val Thr Leu Glu Tyr
325 330 335

Arg Arg Glu Cys Gly Arg Asp Ser Val Leu Arg Ser Leu Thr Ala Val
340 345 350

Ser Gly Gly Gly Val Gly Asp Leu Gly His Ala Gly Asn Val Glu Cys
355 360 365

Gln His Val Leu Arg Leu Glu Asp Gly Ala Glu Ile Val Arg Gly Arg
370 375 380

Thr Glu Trp Arg Pro Lys Tyr Ile Asn Asn Phe Ser Ile Met Gly Gln
385 390 395 400

Ile Pro Thr Asp Ala Ser
405

来自pSZ1417的萼距花C16:0-偏好型硫酯酶的密码子优化的编码区

<210> 149
<211> 1338
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 149

actagtatgg tggccaccgc cgcctcctcc gccttcttcc ccctgccctc cgccgacacc 60
tcctcccgcc ccggcaagct gggcaacaag ccctcctccc tgtccccccc gaagcccaag 120
tccacccccа acggcggcct gcaggtgaag gccaacgcct ccgccccccc caagatcaac 180
ggctcccccg tgggcctgaa gtccggcggc ctgaaggggc gcgccaccca ggaggacgcc 240
cactccgccc ccccccccсg caccttcatc aaccagctgc ccgactggtc catgctgctg 300
gccgccatca ccaccgtgtt cctggccgcc gagaagcagt ggatgatgct ggactggaag 360
cccaagcgcc ccgacatgct ggtggacccc ttcggcctgg gctccatcgt gcaggacggc 420
ctggtgttcc gccagaactt ctccatccgc tcctacgaga tcggcgccga ccgcaccgcc 480
tccatcgaga ccgtgatgaa ccacctgcag gagaccgccc tgaaccacgt gaagatcgcc 540
ggcctgtcca acgacggctt cggccgcacc cccgagatgt acaagcgcga cctgatctgg 600
gtggtggcca agatgcaggt gatggtgaac cgctacccca cctggggcga caccgtggag 660
gtgaacacct gggtggccaa gtccggcaag aacggcatgc gccgcgactg gctgatctcc 720
gactgcaaca ccggcgagat cctgacccgc gcctcctccg tgtgggtgat gatgaaccag 780
aagacccgcc gcctgtccaa gatccccgac gaggtgcgca acgagatcga gccccacttc 840
gtggactccc cccccgtgat cgaggacgac gaccgcaagc tgcccaagct ggacgagaag 900
accgccgact ccatccgcaa gggcctgacc ccccgctgga acgacctgga cgtgaaccag 960
cacgtgaaca acgtgaagta catcggctgg atcctggagt ccacccccсc cgaggtgctg 1020
gagacccagg agctgtgctc cctgaccctg gagtaccgcc gcgagtgcgg ccgcgagtcc 1080
gtgctggagt ccctgaccgc catggacccc tccggcggcg gctacggctc ccagttccag 1140
cacctgctgc gcctggagga cggcggcgag atcgtgaagg gccgcaccga gtggcgcccc 1200
aagaacggcg tgatcaacgg cgtggtgccc accggcgagt cctcccccgg cgactactcc 1260
atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 1320
gacgacaagt gaatcgat 1338

萼距花C16:0-偏好型硫酯酶的氨基酸序列(登录号Q39513.1)

<210> 150
<211> 415
<212> PRT
<213> 萼距花
<400> 150

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

305 310 315 320
Lys Tyr Ile Gly Trp Ile Leu Glu Ser Thr Pro Pro Glu Val Leu Glu
325 330 335
Thr Gln Glu Leu Cys Ser Leu Thr Leu Glu Tyr Arg Arg Glu Cys Gly
340 345 350
Arg Glu Ser Val Leu Glu Ser Leu Thr Ala Met Asp Pro Ser Gly Gly
355 360 365
Gly Tyr Gly Ser Gln Phe Gln His Leu Leu Arg Leu Glu Asp Gly Gly
370 375 380
Glu Ile Val Lys Gly Arg Thr Glu Trp Arg Pro Lys Asn Gly Val Ile
385 390 395 400
Asn Gly Val Val Pro Thr Gly Glu Ser Ser Pro Gly Asp Tyr Ser
405 410 415

来自pSZ1437的油棕 C16:0-偏好型硫酯酶的密码子优化的编码区

<210> 151
<211> 1344
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 151

actagtatgg tggcctccat cgtggcctgg gccttcttcc ccacccccctc cttctccccc 60
accgggcgcg ccgcctccgc caaggcctcc aagaccatcg gcgagggctc cgagaacctg 120
aacgtgcgcg gcatcatcgc caagcccacc tcctcctccg ccgccaagca gggcaaggtg 180
atggcccagg ccgtgcccaa gatcaacggc gccaaggtgg gcctgaaggc cgagtcccag 240
aaggccgagg aggacgccgc cccctcctcc gccccccgca ccttctacaa ccagctgccc 300
gactggtccg tgctgctggc cgccgtgacc accatcttcc tggccgccga gaagcagtgg 360
accctgctgg actggaagcc ccgccgcccc gacatgctga ccggcgcctt ctccctgggc 420
aagatcgtgc aggacggcct ggtgttccgc cagaacttct ccatccgctc ctacgagatc 480
ggcgccgacc gcaccgcctc catcgagacc ctgatgaacc acctgcagga gaccgccctg 540
aaccacgtgc gcaacgccgg cctgctgggc gacggcttcg gcgccacccc cgagatgtcc 600
aagcgcaacc tgatctgggt ggtgaccaag atgcaggtgc tgatcgagca ctacccctcc 660
tggggcgacg tggtggaggt ggacacctgg gtgggcgcct ccggcaagaa cggcatgcgc 720
cgcgactggc acgtgcgcga ctaccgcacc ggccagacca tcctgcgcgc cacctccatc 780
tgggtgatga tggacaagca cacccgcaag ctgtccaaga tgcccgagga ggtgcgcgcc 840
gagatcggcc cctacttcat ggagcacgcc gccatcgtgg acgaggactc ccgcaagctg 900
cccaagctgg acgacgacac cgccgactac atcaagtggg gcctgacccc ccgctggtcc 960
gacctggacg tgaaccagca cgtgaacaac gtgaagtaca tcggctggat cctggagtcc 1020
gcccccatct ccatcctgga gaaccacgag ctggcctcca tgaccctgga gtaccgccgc 1080

gagtgcggcc gcgactccgt gctgcagtcc ctgaccgccg tggccaacga ctgcaccggc 1140
ggcctgcccg aggcctccat cgagtgccag cacctgctgc agctggagtg cggcgccgag 1200
atcgtgcgcg gccgcaccca gtggcgcccc cgccgcgcct ccggccccac ctccgccggc 1260
tccgccatgg actacaagga ccacgacggc gactacaagg accacgacat cgactacaag 1320
gacgacgacg acaagtgaat cgat 1344

油棕C16:0-偏好型硫酯酶的氨基酸序列(登录号AAD42220.2)

<210> 152
<211> 417
<212> PRT
<213> 油棕
<400> 152

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

Trp Val Gly Ala Ser Gly Lys Asn Gly Met Arg Arg Asp Trp His Val
225 230 235 240
Arg Asp Tyr Arg Thr Gly Gln Thr Ile Leu Arg Ala Thr Ser Ile Trp
245 250 255
Val Met Met Asp Lys His Thr Arg Lys Leu Ser Lys Met Pro Glu Glu
260 265 270
Val Arg Ala Glu Ile Gly Pro Tyr Phe Met Glu His Ala Ala Ile Val
275 280 285
Asp Glu Asp Ser Arg Lys Leu Pro Lys Leu Asp Asp Asp Thr Ala Asp
290 295 300
Tyr Ile Lys Trp Gly Leu Thr Pro Arg Trp Ser Asp Leu Asp Val Asn
305 310 315 320
Gln His Val Asn Asn Val Lys Tyr Ile Gly Trp Ile Leu Glu Ser Ala
325 330 335
Pro Ile Ser Ile Leu Glu Asn His Glu Leu Ala Ser Met Thr Leu Glu
340 345 350
Tyr Arg Arg Glu Cys Gly Arg Asp Ser Val Leu Gln Ser Leu Thr Ala
355 360 365
Val Ala Asn Asp Cys Thr Gly Gly Leu Pro Glu Ala Ser Ile Glu Cys
370 375 380
Gln His Leu Leu Gln Leu Glu Cys Gly Ala Glu Ile Val Arg Gly Arg
385 390 395 400
Thr Gln Trp Arg Pro Arg Arg Ala Ser Gly Pro Thr Ser Ala Gly Ser
405 410 415
Ala

来自pSZ1358的欧洲油菜C18:0-偏好型硫酯酶的密码子优化的编码区

<210> 153
<211> 1191
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 153

actagtatgc tgaagctgtc ctgcaacgtg accaacaacc tgcacacctt ctccttcttc 60
tccgactcct ccctgttcat ccccgtgaac cgccgcacca tcgccgtgtc ctccgggcgc 120
gcctcccagc tgcgcaagcc cgccctggac cccctgcgcg ccgtgatctc cgccgaccag 180
ggctccatct cccccgtgaa ctcctgcacc cccgccgacc gcctgcgcgc cggccgcctg 240
atggaggacg gctactccta caaggagaag ttcatcgtgc gctcctacga ggtgggcatc 300
aacaagaccg ccaccgtgga gaccatcgcc aacctgctgc aggaggtggc ctgcaaccac 360

gtgcagaagt gcggcttctc caccgacggc ttcgccacca ccctgaccat gcgcaagctg 420
cacctgatct gggtgaccgc ccgcatgcac atcgagatct acaagtaccc cgcctggtcc 480
gacgtggtgg agatcgagac ctggtgccag tccgagggcc gcatcggcac ccgccgcgac 540
tggatcctgc gcgactccgc caccaacgag gtgatcggcc gcgccacctc caagtgggtg 600
atgatgaacc aggacacccg ccgcctgcag cgcgtgaccg acgaggtgcg cgacgagtac 660
ctggtgttct gcccccgcga gccccgcctg gccttccccg aggagaacaa ctcctccctg 720
aagaagatcc ccaagctgga ggaccccgcc cagtactcca tgctggagct gaagccccgc 780
cgcgccgacc tggacatgaa ccagcacgtg aacaacgtga cctacatcgg ctgggtgctg 840
gagtccatcc cccaggagat catcgacacc cacgagctgc aggtgatcac cctggactac 900
cgccgcgagt gccagcagga cgacatcgtg gactccctga ccacctccga gatccccgac 960
gaccccatct ccaagttcac cggcaccaac ggctccgcca tgtcctccat ccagggccac 1020
aacgagtccc agttcctgca catgctgcgc ctgtccgaga acggccagga gatcaaccgc 1080
ggccgcaccc agtggcgcaa gaagtcctcc cgcatggact acaaggacca cgacggcgac 1140
tacaaggacc acgacatcga ctacaaggac gacgacgaca agtgaatcga t 1191

欧洲油菜C18:0-偏好型硫酯酶的氨基酸序列(登录号CAA52070.1)

<210> 154
<211> 366
<212> PRT
<213> 欧洲油菜
<400> 154

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

145 150 155 160
Glu Thr Trp Cys Gln Ser Glu Gly Arg Ile Gly Thr Arg Arg Asp Trp
165 170 175
Ile Leu Arg Asp Ser Ala Thr Asn Glu Val Ile Gly Arg Ala Thr Ser
180 185 190
Lys Trp Val Met Met Asn Gln Asp Thr Arg Arg Leu Gln Arg Val Thr
195 200 205
Asp Glu Val Arg Asp Glu Tyr Leu Val Phe Cys Pro Arg Glu Pro Arg
210 215 220
Leu Ala Phe Pro Glu Glu Asn Asn Ser Ser Leu Lys Lys Ile Pro Lys
225 230 235 240
Leu Glu Asp Pro Ala Gln Tyr Ser Met Leu Glu Leu Lys Pro Arg Arg
245 250 255
Ala Asp Leu Asp Met Asn Gln His Val Asn Asn Val Thr Tyr Ile Gly
260 265 270
Trp Val Leu Glu Ser Ile Pro Gln Glu Ile Ile Asp Thr His Glu Leu
275 280 285
Gln Val Ile Thr Leu Asp Tyr Arg Arg Glu Cys Gln Gln Asp Asp Ile
290 295 300
Val Asp Ser Leu Thr Thr Ser Glu Ile Pro Asp Asp Pro Ile Ser Lys
305 310 315 320
Phe Thr Gly Thr Asn Gly Ser Ala Met Ser Ser Ile Gln Gly His Asn
325 330 335
Glu Ser Gln Phe Leu His Met Leu Arg Leu Ser Glu Asn Gly Gln Glu
340 345 350
Ile Asn Arg Gly Arg Thr Gln Trp Arg Lys Lys Ser Ser Arg
355 360 365

来自pSZ1375的具有原始小球藻UTEX250硬脂酰-ACP去饱和酶转运肽的蓖麻ACP-硫酯酶的密码子优化的编码序列

<210> 155
<211> 1038
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 155

actagtatgg ccaccgcatc cactttctcg gcgttcaatg cccgctgcgg cgacctgcgt 60
cgctcggcgg gctccgggcc ccggcgccca gcgaggcccc tccccgtgcg cgggcgcgcc 120
cgctcctacg aggtgggcat caacaagacc gccaccgtgg agaccatcgc caacctgctg 180

caggaggtgg gctgcaacca cgcccagtcc gtgggcttct ccaccgacgg cttcgccacc 240
accacctcca tgcgcaagat gcacctgatc tgggtgaccg cccgcatgca catcgagatc 300
tacaagtacc ccgcctggtc cgacgtggtg gaggtggaga cctggtgcca gtccgagggc 360
cgcatcggca cccgccgcga ctggatcctg accgactacg ccaccggcca gatcatcggc 420
cgcgccacct ccaagtgggt gatgatgaac caggacaccc gccgcctgca gaaggtgacc 480
gacgacgtgc gcgaggagta cctggtgttc tgcccccgcg agctgcgcct ggccttcccc 540
gaggagaaca accgctcctc caagaagatc tccaagctgg aggaccccgc ccagtactcc 600
aagctgggcc tggtgccccg ccgcgccgac ctggacatga accagcacgt gaacaacgtg 660
acctacatcg gctgggtgct ggagtccatc ccccaggaga tcatcgacac ccacgagctg 720
cagaccatca ccctggacta ccgccgcgag tgccagcacg acgacatcgt ggactccctg 780
acctccgtgg agccctccga gaacctggag gccgtgtccg agctgcgcgg caccaacggc 840
tccgccacca ccaccgccgg cgacgaggac tgccgcaact tcctgcacct gctgcgcctg 900
tccggcgacg gcctggagat caaccgcggc cgcaccgagt ggcgcaagaa gtccgcccgc 960
atggactaca aggaccacga cggcgactac aaggaccacg acatcgacta caaggacgac 1020
gacgacaagt gaatcgat 1038

蓖麻ACP-硫酯酶的氨基酸序列(登录号ABS30422.1)

<210> 156
<211> 371
<212> PRT
<213> 蓖麻
<400> 156

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

Ala Arg Met His Ile Glu Ile Tyr Lys Tyr Pro Ala Trp Ser Asp Val
145 150 155 160
Val Glu Val Glu Thr Trp Cys Gln Ser Glu Gly Arg Ile Gly Thr Arg
165 170 175
Arg Asp Trp Ile Leu Thr Asp Tyr Ala Thr Gly Gln Ile Ile Gly Arg
180 185 190
Ala Thr Ser Lys Trp Val Met Met Asn Gln Asp Thr Arg Arg Leu Gln
195 200 205
Lys Val Thr Asp Asp Val Arg Glu Glu Tyr Leu Val Phe Cys Pro Arg
210 215 220
Glu Leu Arg Leu Ala Phe Pro Glu Glu Asn Asn Arg Ser Ser Lys Lys
225 230 235 240
Ile Ser Lys Leu Glu Asp Pro Ala Gln Tyr Ser Lys Leu Gly Leu Val
245 250 255
Pro Arg Arg Ala Asp Leu Asp Met Asn Gln His Val Asn Asn Val Thr
260 265 270
Tyr Ile Gly Trp Val Leu Glu Ser Ile Pro Gln Glu Ile Ile Asp Thr
275 280 285
His Glu Leu Gln Thr Ile Thr Leu Asp Tyr Arg Arg Glu Cys Gln His
290 295 300
Asp Asp Ile Val Asp Ser Leu Thr Ser Val Glu Pro Ser Glu Asn Leu
305 310 315 320
Glu Ala Val Ser Glu Leu Arg Gly Thr Asn Gly Ser Ala Thr Thr Thr
325 330 335
Ala Gly Asp Glu Asp Cys Arg Asn Phe Leu His Leu Leu Arg Leu Ser
340 345 350
Gly Asp Gly Leu Glu Ile Asn Arg Gly Arg Thr Glu Trp Arg Lys Lys
355 360 365
Ser Ala Arg
370

来自pSZ1317的具有原始小球藻UTEX250硬脂酰-ACP转运肽的肉豆蔻广泛基础硫酯酶的密码子优化的编码区

<210> 157
<211> 1248
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 157

atggccaccg catccacttt ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg 60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgccgccaac 120
gcccacaccg tgcccaagat caacggcaac aaggccggcc tgctgacccc catggagagc 180
accaaggacg aggacatcgt cgcggccccc accgtggcgc ccaagcgcac cttcatcaac 240
cagctgcccg actggtcgat gctgctggcc gcgatcacca ccatcttcct ggcggccgag 300
aagcagtgga ccaacctgga ctggaagccc cgccgccccg acatgctggt cgacttcgac 360
cccttctccc tgggccgctt cgtgcaggac ggcctgatct tccgccagaa cttcagcatc 420
cgctcgtacg agatcggcgc ggaccgcacc gcctccatcg agaccctgat gaaccacctg 480
caggagaccg cgctgaacca cgtccgctgc atcggcctgc tggacgacgg cttcggcagc 540
acccccgaga tgacccgccg cgacctgatc tgggtggtca cccgcatgca ggtcctggtg 600
gaccgctacc cctcgtgggg cgacgtgatc gaggtcgact cctgggtgac ccccagcggc 660
aagaacggca tgaagcgcga gtggttcctg cgcgactgca agaccggcga gatcctgacc 720
cgcgccacct cggtctgggt gatgatgaac aagcgcaccc gccgcctgtc caagatcccc 780
gaggaggtcc gcgtggagat cgagccctac ttcgtcgagc acggcgtgct ggacgaggac 840
tcgcgcaagc tgcccaagct gaacgacaac accgccaact acatccgccg cggcctggcg 900
ccccgctggt ccgacctgga cgtcaaccag cacgtgaaca acgtcaagta catcggctgg 960
atcctggaga gcgtgcccag cagcctgctg gagtcgcacg agctgtacgg catgaccctg 1020
gagtaccgca aggagtgcgg caaggacggc ctgctgcagt ccctgaccgc cgtcgccagc 1080
gactacggcg gcggctcgct ggaggccggc gtggagtgcg accacctgct gcgcctggag 1140
gacggctccg agatcatgcg cggcaagacc gagtggcgcc ccaagcgcgc cgcgaacacc 1200
acctacttcg gcagcgtcga cgacatcccc cccgccaaca acgcgtga 1248

具有原始小球藻UTEX250硬脂酰-ACP转运肽序列的肉豆蔻广泛基础硫酯酶的氨基酸序列

<210> 158
<211> 415
<212> PRT
<213> 人工序列
<220>
<223> 合成的构建体
<400> 158

Met Ala Thr Ala Ser Thr Phe Ser Ala Phe Asn Ala Arg Cys Gly Asp
1 5 10 15
Leu Arg Arg Ser Ala Gly Ser Gly Pro Arg Arg Pro Ala Arg Pro Leu
20 25 30
Pro Val Arg Gly Arg Ala Ala Asn Ala His Thr Val Pro Lys Ile Asn
35 40 45
Gly Asn Lys Ala Gly Leu Leu Thr Pro Met Glu Ser Thr Lys Asp Glu
50 55 60
Asp Ile Val Ala Ala Pro Thr Val Ala Pro Lys Arg Thr Phe Ile Asn

65 70 75 80
Gln Leu Pro Asp Trp Ser Met Leu Leu Ala Ala Ile Thr Thr Ile Phe
85 90 95
Leu Ala Ala Glu Lys Gln Trp Thr Asn Leu Asp Trp Lys Pro Arg Arg
100 105 110
Pro Asp Met Leu Val Asp Phe Asp Pro Phe Ser Leu Gly Arg Phe Val
115 120 125
Gln Asp Gly Leu Ile Phe Arg Gln Asn Phe Ser Ile Arg Ser Tyr Glu
130 135 140
Ile Gly Ala Asp Arg Thr Ala Ser Ile Glu Thr Leu Met Asn His Leu
145 150 155 160
Gln Glu Thr Ala Leu Asn His Val Arg Cys Ile Gly Leu Leu Asp Asp
165 170 175
Gly Phe Gly Ser Thr Pro Glu Met Thr Arg Arg Asp Leu Ile Trp Val
180 185 190
Val Thr Arg Met Gln Val Leu Val Asp Arg Tyr Pro Ser Trp Gly Asp
195 200 205
Val Ile Glu Val Asp Ser Trp Val Thr Pro Ser Gly Lys Asn Gly Met
210 215 220
Lys Arg Glu Trp Phe Leu Arg Asp Cys Lys Thr Gly Glu Ile Leu Thr
225 230 235 240
Arg Ala Thr Ser Val Trp Val Met Met Asn Lys Arg Thr Arg Arg Leu
245 250 255
Ser Lys Ile Pro Glu Glu Val Arg Val Glu Ile Glu Pro Tyr Phe Val
260 265 270
Glu His Gly Val Leu Asp Glu Asp Ser Arg Lys Leu Pro Lys Leu Asn
275 280 285
Asp Asn Thr Ala Asn Tyr Ile Arg Arg Gly Leu Ala Pro Arg Trp Ser
290 295 300
Asp Leu Asp Val Asn Gln His Val Asn Asn Val Lys Tyr Ile Gly Trp
305 310 315 320
Ile Leu Glu Ser Val Pro Ser Ser Leu Leu Glu Ser His Glu Leu Tyr
325 330 335
Gly Met Thr Leu Glu Tyr Arg Lys Glu Cys Gly Lys Asp Gly Leu Leu
340 345 350
Gln Ser Leu Thr Ala Val Ala Ser Asp Tyr Gly Gly Gly Ser Leu Glu
355 360 365
Ala Gly Val Glu Cys Asp His Leu Leu Arg Leu Glu Asp Gly Ser Glu
370 375 380

Ile Met Arg Gly Lys Thr Glu Trp Arg Pro Lys Arg Ala Ala Asn Thr
385 390 395 400
Thr Tyr Phe Gly Ser Val Asp Asp Ile Pro Pro Ala Asn Asn Ala
405 410 415

具有莱茵衣藻β-微管蛋白启动子/5'UTR和普通小球藻硝酸盐还原酶3'UTR的酿酒酵母suc2基因的密码子优化的表达盒的核苷酸序列

<210> 159
<211> 2333
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 159

ctttcttgcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60
cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc 120
gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc aggcccccga ttgcaaagac 180
attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240
ccactcgagc ttgtgatcgc actccgctaa gggggcgcct cttcctcttc gtttcagtca 300
caacccgcaa acggcgcgcc atgctgctgc aggccttcct gttcctgctg gccggcttcg 360
ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg gtgcacttca 420
cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag aaggacgcca 480
agtggcacct gtacttccag tacaacccga acgacaccgt ctgggggacg cccttgttct 540
ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc gccatcgccc 600
cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac aacaacacct 660
ccggcttctt caacgacacc atcgacccgc gccagcgctg cgtggccatc tggacctaca 720
acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc tacaccttca 780
ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc gacccgaagg 840
tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc caggactaca 900
agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc gcgttcgcca 960
acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc cccaccgagc 1020
aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc gccccggccg 1080
gcggctcctt caaccagtac ttcgtcggca gcttcaacgg cacccacttc gaggccttcg 1140
acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag accttcttca 1200
acaccgaccc gacctacggg agcgccctgg gcatcgcgtg ggcctccaac tgggagtact 1260
ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc aagttctccc 1320
tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag gccgagccga 1380
tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc acgttgacga 1440
aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag ttcgagctgg 1500
tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac ctctccctct 1560

ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag gtgtccgcgt 1620
cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag aacccctact 1680
tcaccaaccg catgagcgtg aacaaccagc ccttcaagag cgagaacgac ctgtcctact 1740
acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac gacggcgacg 1800
tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc gtgaacatga 1860
cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag gtcaagtgac 1920
aattggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg tgtgatggac 1980
tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt atcaaacagc 2040
ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc ttgtgctatt 2100
tgcgaatacc acccccagca tcccccttccc tcgtttcata tcgcttgcat cccaaccgca 2160
acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc actgcccctc 2220
gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg taaaccagca 2280
ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga tcc 2333

来自pSZ1462的具有原始小球藻UTEX250硬脂酰-ACP去饱和酶转运肽的萼距花C16:0-偏好型硫酯酶的密码子优化的编码区

<210> 160
<211> 1233
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 160

actagtatgg ccaccgcatc cactttctcg gcgttcaatg cccgctgcgg cgacctgcgt 60
cgctcggcgg gctccgggcc ccggcgccca gcgaggcccc tccccgtgcg cgggcgcgcc 120
acccaggagg acgcccactc cgccccccccc ccccgcacct tcatcaacca gctgcccgac 180
tggtccatgc tgctggccgc catcaccacc gtgttcctgg ccgccgagaa gcagtggatg 240
atgctggact ggaagcccaa gcgccccgac atgctggtgg acccccttcgg cctgggctcc 300
atcgtgcagg acggcctggt gttccgccag aacttctcca tccgctccta cgagatcggc 360
gccgaccgca ccgcctccat cgagaccgtg atgaaccacc tgcaggagac cgccctgaac 420
cacgtgaaga tcgccggcct gtccaacgac ggcttcggcc gcacccccga gatgtacaag 480
cgcgacctga tctgggtggt ggccaagatg caggtgatgg tgaaccgcta ccccacctgg 540
ggcgacaccg tggaggtgaa cacctgggtg gccaagtccg gcaagaacgg catgcgccgc 600
gactggctga tctccgactg caacaccggc gagatcctga cccgcgcctc ctccgtgtgg 660
gtgatgatga accagaagac ccgccgcctg tccaagatcc ccgacgaggt gcgcaacgag 720
atcgagcccc acttcgtgga ctccccccccc gtgatcgagg acgacgaccg caagctgccc 780
aagctggacg agaagaccgc cgactccatc cgcaagggcc tgacccccccg ctggaacgac 840
ctggacgtga accagcacgt gaacaacgtg aagtacatcg gctggatcct ggagtccacc 900
ccccccgagg tgctggagac ccaggagctg tgctccctga ccctggagta ccgccgcgag 960
tgcggccgcg agtccgtgct ggagtccctg accgccatgg acccctccgg cggcggctac 1020

ggctcccagt tccagcacct gctgcgcctg gaggacggcg gcgagatcgt gaagggccgc 1080
accgagtggc gccccaagaa cggcgtgatc aacggcgtgg tgcccaccgg cgagtcctcc 1140
cccggcgact actccatgga ctacaaggac cacgacggcg actacaagga ccacgacatc 1200
gactacaagg acgacgacga caagtgaatc gat 1233

来自pSZ1436的油棕C16:0-偏好型硫酯酶PATE的密码子优化的编码区

<210> 161
<211> 1344
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 161

actagtatgg tggcctccat cgccgcctcc gccttcttcc ccacccccctc cttctccccc 60
accgggcgcg ccgcctccgc caaggcctcc aagaccatcg gcgagggctc cgagaacctg 120
gacgcccgcg gcatcatcgc caagcccacc tcctcctccg ccgccatgca gggcaaggtg 180
atggcccagg ccgtgtccaa gatcaacggc gccaaggtgg gcctgaaggc cgagtcccag 240
aaggccgagg aggacgccgc cccctcctcc gccccccgca ccttctacaa ccagctgccc 300
gactggtccg tgctgctggc cgccgtgacc accatcttcc tggccgccga gaagcagtgg 360
accctgctgg actggaagcc ccgccgcccc gacatgctga ccgacgcctt ctccctgggc 420
aagatcgtgc aggacggcct ggtgttcaag cagaacttct ccatccgctc ctacgagatc 480
ggcgccgacc gcaccgcctc catcgagacc ctgatgaacc acctgcagga gaccgccctg 540
aaccacgtgc gctccgccgg cctgatgggc gacggcttcg gcgccacccc cgagatgtcc 600
aagcgcaacc tgatctgggt ggtgaccaag gtgcgcgtgc tgatcgagca ctacccctcc 660
tggggcgacg tggtggaggt ggacacctgg gtgggccccg ccggcaagaa cggcatgcgc 720
cgcgactggc acgtgcgcga ccaccgcacc ggccagacca tcctgcgcgc cacccgcgtg 780
tgggtgatga tgaacaagaa cacccgcaag ctgtccaagg tgcccgagga ggtgcgcgcc 840
gagatcggcc cctacttcgt ggagcgcgcc gccatcgtgg acgaggactc ccgcaagctg 900
cccaagctgg acgaggacac caccgactac gtgaagaagg gcctgacccc ccgctggtcc 960
gacctggacg tgaaccagca cgtgaacaac gtgaagtaca tcggctggat cctggagtcc 1020
gcccccatct ccatcctgga gaaccacgag ctggcctcca tgaccctgga gtaccgccgc 1080
gagtgcggcc gcgactccgt gctgcagtcc ctgaccgccg tggccaacga ctgcaccggc 1140
ggcctgcccg aggcctccat cgagtgccag cacctgctgc agctggagtg cggcgccgag 1200
atcgtgcgcg gccgcaccca gtggcgcccc cgccgcgcct ccggccccac ctccgccggc 1260
tccgccatgg actacaagga ccacgacggc gactacaagg accacgacat cgactacaag 1320
gacgacgacg acaagtgaat cgat 1344

油棕C16:0-偏好型硫酯酶PATE的氨基酸序列(Genbank登录号ABD83939)

<210> 162
<211> 417
<212> PRT

<213> 油棕

<400> 162

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

290 295 300

Tyr Val Lys Lys Gly Leu Thr Pro Arg Trp Ser Asp Leu Asp Val Asn
305 310 315 320
Gln His Val Asn Asn Val Lys Tyr Ile Gly Trp Ile Leu Glu Ser Ala
325 330 335
Pro Ile Ser Ile Leu Glu Asn His Glu Leu Ala Ser Met Thr Leu Glu
340 345 350
Tyr Arg Arg Glu Cys Gly Arg Asp Ser Val Leu Gln Ser Leu Thr Ala
355 360 365
Val Ala Asn Asp Cys Thr Gly Gly Leu Pro Glu Ala Ser Ile Glu Cys
370 375 380
Gln His Leu Leu Gln Leu Glu Cys Gly Ala Glu Ile Val Arg Gly Arg
385 390 395 400
Thr Gln Trp Arg Pro Arg Arg Ala Ser Gly Pro Thr Ser Ala Gly Ser
405 410 415
Ala

来自pSZ1468的RNAi表达盒的FADc部分

<210> 163
<211> 595
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 163

actagtatgg ctatcaagac gaacaggcag cctgtggaga agcctccgtt cacgatcggg 60
acgctgcgca aggccatccc cgcgcactgt ttcgagcgct cggcgcttcg tagcagcatg 120
tacctggcct ttgacatcgc ggtcatgtcc ctgctctacg tcgcgtcgac gtacatcgac 180
cctgcaccgg tgcctacgtg ggtcaagtac ggcatcatgt ggccgctcta ctggttcttc 240
caggtgtgtt tgagggtttt ggttgcccgt attgaggtcc tggtggcgcg catggaggag 300
aaggcgcctg tcccgctgac ccccccggct accctcccgg caccttccag ggcgcgtacg 360
ggaagaacca gtagagcggc cacatgatgc cgtacttgac ccacgtaggc accggtgcag 420
ggtcgatgta cgtcgacgcg acgtagagca gggacatgac cgcgatgtca aaggccaggt 480
acatgctgct acgaagcgcc gagcgctcga aacagtgcgc ggggatggcc ttgcgcagcg 540
tcccgatcgt gaacggaggc ttctccacag gctgcctgtt cgtcttgata gccat 595

来自pSZ1468的FADc RNAi表达盒的相关部分

<210> 164
<211> 6101
<212> DNA
<213> 人工序列

<220>

<223> 合成的构建体

<400> 164

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60
ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120
tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180
ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240
gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300
ggaagacagg tgagggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360
atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420
cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480
gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540
ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600
ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660
cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720
ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 960
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 1020
cagtcacaac ccgcaaacgg cgcgccatgc tgctgcaggc cttcctgttc ctgctggccg 1080
gcttcgccgc caagatcagc gcctccatga cgaacgagac gtccgaccgc cccctggtgc 1140
acttcacccc caacaagggc tggatgaacg accccaacgg cctgtggtac gacgagaagg 1200
acgccaagtg gcacctgtac ttccagtaca acccgaacga caccgtctgg gggacgccct 1260
tgttctgggg ccacgccacg tccgacgacc tgaccaactg ggaggaccag cccatcgcca 1320
tcgccccgaa gcgcaacgac tccggcgcct tctccggctc catggtggtg gactacaaca 1380
acacctccgg cttcttcaac gacaccatcg acccgcgcca gcgctgcgtg gccatctgga 1440
cctacaacac cccggagtcc gaggagcagt acatctccta cagcctggac ggcggctaca 1500
ccttcaccga gtaccagaag aaccccgtgc tggccgccaa ctccacccag ttccgcgacc 1560
cgaaggtctt ctggtacgag ccctcccaga agtggatcat gaccgcggcc aagtcccagg 1620
actacaagat cgagatctac tcctccgacg acctgaagtc ctggaagctg gagtccgcgt 1680
tcgccaacga gggcttcctc ggctaccagt acgagtgccc cggcctgatc gaggtcccca 1740
ccgagcagga ccccagcaag tcctactggg tgatgttcat ctccatcaac cccggcgccc 1800
cggccggcgg ctccttcaac cagtacttcg tcggcagctt caacggcacc cacttcgagg 1860
ccttcgacaa ccagtcccgc gtggtggact tcggcaagga ctactacgcc ctgcagacct 1920
tcttcaacac cgacccgacc tacgggagcg ccctgggcat cgcgtgggcc tccaactggg 1980
agtactccgc cttcgtgccc accaaccct ggcgctcctc catgtccctc gtgcgcaagt 2040
tctccctcaa caccgagtac caggccaacc cggagacgga gctgatcaac ctgaaggccg 2100
agccgatcct gaacatcagc aacgccggcc cctggagccg gttcgccacc aacaccacgt 2160

tgacgaaggc caacagctac aacgtcgacc tgtccaacag caccggcacc ctggagttcg 2220
agctggtgta cgccgtcaac accacccaga cgatctccaa gtccgtgttc gcggacctct 2280
ccctctggtt caagggcctg gaggaccccg aggagtacct ccgcatgggc ttcgaggtgt 2340
ccgcgtcctc cttcttcctg gaccgcggga acagcaaggt gaagttcgtg aaggagaacc 2400
cctacttcac caaccgcatg agcgtgaaca accagcccct caagagcgag aacgacctgt 2460
cctactacaa ggtgtacggc ttgctggacc agaacatcct ggagctgtac ttcaacgacg 2520
gcgacgtcgt gtccaccaac acctacttca tgaccaccgg gaacgccctg ggctccgtga 2580
acatgacgac gggggtggac aacctgttct acatcgacaa gttccaggtg cgcgaggtca 2640
agtgacaatt ggcagcagca gctcggatag tatcgacaca ctctggacgc tggtcgtgtg 2700
atggactgtt gccgccacac ttgctgcctt gacctgtgaa tatccctgcc gcttttatca 2760
aacagcctca gtgtgtttga tcttgtgtgt acgcgctttt gcgagttgct agctgcttgt 2820
gctatttgcg aataccaccc ccagcatccc cttccctcgt ttcatatcgc ttgcatccca 2880
accgcaactt atctacgctg tcctgctatc cctcagcgct gctcctgctc ctgctcactg 2940
cccctcgcac agccttggtt tgggctccgc ctgtattctc ctggtactgc aacctgtaaa 3000
ccagcactgc aatgctgatg cacgggaagt agtgggatgg gaacacaaat ggaggatccc 3060
gcgtctcgaa cagagcgcgc agaggaacgc tgaaggtctc gcctctgtcg cacctcagcg 3120
cggcatacac cacaataacc acctgacgaa tgcgcttggt tcttcgtcca ttagcgaagc 3180
gtccggttca cacacgtgcc acgttggcga ggtggcaggt gacaatgatc ggtggagctg 3240
atggtcgaaa cgttcacagc ctagggatat cgaattcggc cgacaggacg cgcgtcaaag 3300
gtgctggtcg tgtatgccct ggccggcagg tcgttgctgc tgctggttag tgattccgca 3360
accctgattt tggcgtctta ttttggcgtg gcaaacgctg gcgcccgcga gccgggccgg 3420
cggcgatgcg gtgccccacg gctgccggaa tccaagggag gcaagagcgc ccgggtcagt 3480
tgaagggctt tacgcgcaag gtacagccgc tcctgcaagg ctgcgtggtg gaattggacg 3540
tgcaggtcct gctgaagttc ctccaccgcc tcaccagcgg acaaagcacc ggtgtatcag 3600
gtccgtgtca tccactctaa agagctcgac tacgacctac tgatggccct agattcttca 3660
tcaaaaacgc ctgagacact tgcccaggat tgaaactccc tgaagggacc accaggggcc 3720
ctgagttgtt ccttcccccc gtggcgagct gccagccagg ctgtacctgt gatcgaggct 3780
ggcgggaaaa taggcttcgt gtgctcaggt catgggaggt gcaggacagc tcatgaaacg 3840
ccaacaatcg cacaattcat gtcaagctaa tcagctattt cctcttcacg agctgtaatt 3900
gtcccaaaat tctggtctac cgggggtgat ccttcgtgta cgggcccttc cctcaaccct 3960
aggtatgcgc gcatgcggtc gccgcgcaac tcgcgcgagg gccgagggtt tgggacgggc 4020
cgtcccgaaa tgcagttgca cccggatgcg tggcaccttt tttgcgataa tttatgcaat 4080
ggactgctct gcaaaattct ggctctgtcg ccaaccctag gatcagcggc gtaggatttc 4140
gtaatcattc gtcctgatgg ggagctaccg actaccctaa tatcagcccg actgcctgac 4200
gccagcgtcc acttttgtgc acacattcca ttcgtgccca agacatttca ttgtggtgcg 4260
aagcgtcccc agttacgctc acctgtttcc cgacctcctt actgttctgt cgacagagcg 4320
ggcccacagg ccggtcgcag ccactagtat ggctatcaag acgaacaggc agcctgtgga 4380
gaagcctccg ttcacgatcg ggacgctgcg caaggccatc cccgcgcact gtttcgagcg 4440
ctcggcgctt cgtagcagca tgtacctggc ctttgacatc gcggtcatgt ccctgctcta 4500

cgtcgcgtcg acgtacatcg accctgcacc ggtgcctacg tgggtcaagt acggcatcat 4560
gtggccgctc tactggttct tccaggtgtg tttgagggtt ttggttgccc gtattgaggt 4620
cctggtggcg cgcatggagg agaaggcgcc tgtcccgctg acccccccgg ctaccctccc 4680
ggcaccttcc agggcgcgta cgggaagaac cagtagagcg gccacatgat gccgtacttg 4740
acccacgtag gcaccggtgc agggtcgatg tacgtcgacg cgacgtagag cagggacatg 4800
accgcgatgt caaaggccag gtacatgctg ctacgaagcg ccgagcgctc gaaacagtgc 4860
gcggggatgg ccttgcgcag cgtcccgatc gtgaacggag gcttctccac aggctgcctg 4920
ttcgtcttga tagccatctc gaggcagcag cagctcggat agtatcgaca cactctggac 4980
gctggtcgtg tgatggactg ttgccgccac acttgctgcc ttgacctgtg aatatccctg 5040
ccgcttttat caaacagcct cagtgtgttt gatcttgtgt gtacgcgctt ttgcgagttg 5100
ctagctgctt gtgctatttg cgaataccac ccccagcatc cccttccctc gtttcatatc 5160
gcttgcatcc caaccgcaac ttatctacgc tgtcctgcta tccctcagcg ctgctcctgc 5220
tcctgctcac tgcccctcgc acagccttgg tttgggctcc gcctgtattc tcctggtact 5280
gcaacctgta aaccagcact gcaatgctga tgcacgggaa gtagtgggat gggaacacaa 5340
atggaaagct gtagagctct tgttttccag aaggagttgc tccttgagcc tttcattctc 5400
agcctcgata acctccaaag ccgctctaat tgtggagggg gttcgaattt aaaagcttgg 5460
aatgttggtt cgtgcgtctg gaacaagccc agacttgttg ctcactggga aaaggaccat 5520
cagctccaaa aaacttgccg ctcaaaccgc gtacctctgc tttcgcgcaa tctgccctgt 5580
tgaaatcgcc accacattca tattgtgacg cttgagcagt ctgtaattgc ctcagaatgt 5640
ggaatcatct gccccctgtg cgagcccatg ccaggcatgt cgcgggcgag gacacccgcc 5700
actcgtacag cagaccatta tgctacctca caatagttca taacagtgac catatttctc 5760
gaagctcccc aacgagcacc tccatgctct gagtggccac cccccggccc tggtgcttgc 5820
ggagggcagg tcaaccggca tggggctacc gaaatccccg accggatccc accacccccg 5880
cgatgggaag aatctctccc cgggatgtgg gcccaccacc agcacaacct gctggcccag 5940
gcgagcgtca aaccatacca cacaaatatc cttggcatcg gccctgaatt ccttctgccg 6000
ctctgctacc cggtgcttct gtccgaagca ggggttgcta gggatcgctc cgagtccgca 6060
aacccttgtc gcgtggcggg gcttgttcga gcttgaagag c 6101

来自pSZ1469的RNAi表达盒的FADc部分

<210> 165
<211> 595
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 165

actagtatgg ctatcaagac gaacaggcag cctgtggaga agcctccgtt cacgatcggg 60
acgctgcgca aggccatccc cgcgcactgt ttcgagcgct cggcgcttcg tagcagcatg 120
tacctggcct ttgacatcgc ggtcatgtcc ctgctctacg tcgcgtcgac gtacatcgac 180
cctgcaccgg tgcctacgtg ggtcaagtac ggcatcatgt ggccgctcta ctggttcttc 240

caggtgtgtt tgagggtttt ggttgcccgt attgaggtcc tggtggcgcg catggaggag 300
aaggcgcctg tcccgctgac ccccccggct accctcccgg caccttccag ggcgcgtacg 360
ggaagaacca gtagagcggc cacatgatgc cgtacttgac ccacgtaggc accggtgcag 420
ggtcgatgta cgtcgacgcg acgtagagca gggacatgac cgcgatgtca aaggccaggt 480
acatgctgct acgaagcgcc gagcgctcga aacagtgcgc ggggatggcc ttgcgcagcg 540
tcccgatcgt gaacggaggc ttctccacag gctgcctgtt cgtcttgata gccat 595

来自pSZ1469的FADc RNAi表达盒的相关部分

<210> 166
<211> 5348
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 166

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60
ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120
tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180
ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240
gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300
ggaagacagg tgagggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360
atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420
cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480
gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540
ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600
ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660
cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720
ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 960
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 1020
cagtcacaac ccgcaaacgg cgcgccatgc tgctgcaggc cttcctgttc ctgctggccg 1080
gcttcgccgc caagatcagc gcctccatga cgaacgagac gtccgaccgc cccctggtgc 1140
acttcacccc caacaagggc tggatgaacg accccaacgg cctgtggtac gacgagaagg 1200
acgccaagtg gcacctgtac ttccagtaca acccgaacga caccgtctgg gggacgccct 1260
tgttctgggg ccacgccacg tccgacgacc tgaccaactg ggaggaccag cccatcgcca 1320
tcgccccgaa gcgcaacgac tccggcgcct tctccggctc catggtggtg gactacaaca 1380
acacctccgg cttcttcaac gacaccatcg acccgcgcca gcgctgcgtg gccatctgga 1440
cctacaacac cccggagtcc gaggagcagt acatctccta cagcctggac ggcggctaca 1500

ccttcaccga gtaccagaag aaccccgtgc tggccgccaa ctccacccag ttccgcgacc 1560
cgaaggtctt ctggtacgag ccctcccaga agtggatcat gaccgcggcc aagtcccagg 1620
actacaagat cgagatctac tcctccgacg acctgaagtc ctggaagctg gagtccgcgt 1680
tcgccaacga gggcttcctc ggctaccagt acgagtgccc cggcctgatc gaggtcccca 1740
ccgagcagga ccccagcaag tcctactggg tgatgttcat ctccatcaac cccggcgccc 1800
cggccggcgg ctccttcaac cagtacttcg tcggcagctt caacggcacc cacttcgagg 1860
ccttcgacaa ccagtcccgc gtggtggact tcggcaagga ctactacgcc ctgcagacct 1920
tcttcaacac cgacccgacc tacgggagcg ccctgggcat cgcgtgggcc tccaactggg 1980
agtactccgc cttcgtgccc accaacccct ggcgctcctc catgtccctc gtgcgcaagt 2040
tctccctcaa caccgagtac caggccaacc cggagacgga gctgatcaac ctgaaggccg 2100
agccgatcct gaacatcagc aacgccggcc cctggagccg gttcgccacc aacaccacgt 2160
tgacgaaggc caacagctac aacgtcgacc tgtccaacag caccggcacc ctggagttcg 2220
agctggtgta cgccgtcaac accacccaga cgatctccaa gtccgtgttc gcggacctct 2280
ccctctggtt caagggcctg gaggaccccg aggagtacct ccgcatgggc ttcgaggtgt 2340
ccgcgtcctc cttcttcctg gaccgcggga acagcaaggt gaagttcgtg aaggagaacc 2400
cctacttcac caaccgcatg agcgtgaaca accagccctt caagagcgag aacgacctgt 2460
cctactacaa ggtgtacggc ttgctggacc agaacatcct ggagctgtac ttcaacgacg 2520
gcgacgtcgt gtccaccaac acctacttca tgaccaccgg gaacgccctg ggctccgtga 2580
acatgacgac gggggtggac aacctgttct acatcgacaa gttccaggtg cgcgaggtca 2640
agtgacaatt ggcagcagca gctcggatag tatcgacaca ctctggacgc tggtcgtgtg 2700
atggactgtt gccgccacac ttgctgcctt gacctgtgaa tatccctgcc gcttttatca 2760
aacagcctca gtgtgtttga tcttgtgtgt acgcgctttt gcgagttgct agctgcttgt 2820
gctatttgcg aataccaccc ccagcatccc cttccctcgt ttcatatcgc ttgcatccca 2880
accgcaactt atctacgctg tcctgctatc cctcagcgct gctcctgctc ctgctcactg 2940
cccctcgcac agccttggtt tgggctccgc ctgtattctc ctggtactgc aacctgtaaa 3000
ccagcactgc aatgctgatg cacgggaagt agtgggatgg gaacacaaat ggaggatccc 3060
gcgtctcgaa cagagcgcgc agaggaacgc tgaaggtctc gcctctgtcg cacctcagcg 3120
cggcatacac cacaataacc acctgacgaa tgcgcttggt tcttcgtcca ttagcgaagc 3180
gtccggttca cacacgtgcc acgttggcga ggtggcaggt gacaatgatc ggtggagctg 3240
atggtcgaaa cgttcacagc ctagggatat cgaattcctt tcttgcgcta tgacacttcc 3300
agcaaaaggt agggcgggct gcgagacggc ttcccggcgc tgcatgcaac accgatgatg 3360
cttcgacccc ccgaagctcc ttcggggctg catgggcgct ccgatgccgc tccagggcga 3420
gcgctgttta aatagccagg cccccgattg caaagacatt atagcgagct accaaagcca 3480
tattcaaaca cctagatcac taccacttct acacaggcca ctcgagcttg tgatcgcact 3540
ccgctaaggg ggcgcctctt cctcttcgtt tcagtcacaa cccgcaaaca ctagtatggc 3600
tatcaagacg aacaggcagc ctgtggagaa gcctccgttc acgatcggga cgctgcgcaa 3660
ggccatcccc gcgcactgtt tcgagcgctc ggcgcttcgt agcagcatgt acctggcctt 3720
tgacatcgcg gtcatgtccc tgctctacgt cgcgtcgacg tacatcgacc ctgcaccggt 3780
gcctacgtgg gtcaagtacg gcatcatgtg gccgctctac tggttcttcc aggtgtgttt 3840

gagggttttg gttgcccgta ttgaggtcct ggtggcgcgc atggaggaga aggcgcctgt 3900
cccgctgacc cccccggcta ccctcccggc accttccagg gcgcgtacgg gaagaaccag 3960
tagagcggcc acatgatgcc gtacttgacc cacgtaggca ccggtgcagg gtcgatgtac 4020
gtcgacgcga cgtagagcag ggacatgacc gcgatgtcaa aggccaggta catgctgcta 4080
cgaagcgccg agcgctcgaa acagtgcgcg gggatggcct tgcgcagcgt cccgatcgtg 4140
aacggaggct tctccacagg ctgcctgttc gtcttgatag ccatctcgag gcagcagcag 4200
ctcggatagt atcgacacac tctggacgct ggtcgtgtga tggactgttg ccgccacact 4260
tgctgccttg acctgtgaat atccctgccg cttttatcaa acagcctcag tgtgtttgat 4320
cttgtgtgta cgcgcttttg cgagttgcta gctgcttgtg ctatttgcga ataccacccc 4380
cagcatcccc ttccctcgtt tcatatcgct tgcatcccaa ccgcaactta tctacgctgt 4440
cctgctatcc ctcagcgctg ctcctgctcc tgctcactgc ccctcgcaca gccttggttt 4500
gggctccgcc tgtattctcc tggtactgca acctgtaaac cagcactgca atgctgatgc 4560
acgggaagta gtgggatggg aacacaaatg gaaagctgta gagctcttgt tttccagaag 4620
gagttgctcc ttgagccttt cattctcagc ctcgataacc tccaaagccg ctctaattgt 4680
ggaggggttt cgaatttaaa agcttggaat gttggttcgt gcgtctggaa caagcccaga 4740
cttgttgctc actgggaaaa ggaccatcag ctccaaaaaa cttgccgctc aaaccgcgta 4800
cctctgcttt cgcgcaatct gccctgttga aatcgccacc acattcatat tgtgacgctt 4860
gagcagtctg taattgcctc agaatgtgga atcatctgcc ccctgtgcga gcccatgcca 4920
ggcatgtcgc gggcgaggac acccgccact cgtacagcag accattatgc tacctcacaa 4980
tagttcataa cagtgaccat atttctcgaa gctccccaac gagcacctcc atgctctgag 5040
tggccacccc ccggccctgg tgcttgcgga gggcaggtca accggcatgg ggctaccgaa 5100
atccccgacc ggatcccacc acccccgcga tgggaagaat ctctccccgg gatgtgggcc 5160
caccaccagc acaacctgct ggcccaggcg agcgtcaaac cataccacac aaatatcctt 5220
ggcatcggcc ctgaattcct tctgccgctc tgctacccgg tgcttctgtc cgaagcaggg 5280
gttgctaggg atcgctccga gtccgcaaac ccttgtcgcg tggcggggct tgttcgagct 5340
tgaagagc 5348

来自pSZ1470的RNAi表达盒的FADc部分

<210> 167
<211> 1220
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 167

actagttcac ttgtggaacc agagcgcgga gtcgtcctcg ggcgcgtccg ggacgacgta 60
gcggcagtcg cgccagtcct cccacagggc gcggccgacc cagcggctgt cggactggta 120
gtacttgccc aggatgggcc tgatggcggc ggaggcctcc tcggcgtggt agtgcgggat 180
ggtgctgaag aggtggtgca gcacgtgggt gtcggagatg tggtgcagga tgttgtccat 240
gaagggcggg cccatggagc ggtccacggt ggccatggcg ccgcgcagcc agtcccagtc 300

cttctcgaag tagtgcggca gcgccgggtg cgtgtgctgg agcagcgtga tgagcacgag 360
ccacatgttc acgatcaggt agggcaccac gtaggtcttg accagccagg cccagcccat 420
ggtgcggccc agcacgctga gcccgctgag caccgccacc agcgccaggt cggagatgac 480
cacctcgatg cgctcgcgct tgctgaagat gggcgaccac gggtcaaagt ggttggcgaa 540
gcgcgggtac ggccgcgagg cgacgttgaa catgaggtac agcggccagc ccagggtcag 600
ggtgaccagc accttgccca tgcggatggg cagccactcc tcccactcca ggccctcgtg 660
cgccactgcg cggtgcggcg gcacaaacac ctcgtccttg tccaggcacc ccgtgttgga 720
gtggtggcgg cggtgcgagt gcttccagga gtagtagggc accagcagca ggctgtggaa 780
caccaggccc acgccgtcgt tgatggcctg gctggaggaa aaggcctggt ggccgcactc 840
gtgcgcgcac acccagacac ccgtgccgaa ggcgccctgg aaggtgccgg gagggtagcc 900
gggggggtca gcgggacagg cgccttctcc tccatgcgcg ccaccaggac ctcaatacgg 960
gcaaccaaaa ccctcaaaca cacctggaag aaccagtaga gcggccacat gatgccgtac 1020
ttgacccacg taggcaccgg tgcagggtcg atgtacgtcg acgcgacgta gagcagggac 1080
atgaccgcga tgtcaaaggc caggtacatg ctgctacgaa gcgccgagcg ctcgaaacag 1140
tgcgcgggga tggccttgcg cagcgtcccg atcgtgaacg gaggcttctc cacaggctgc 1200
ctgttcgtct tgatagccat 1220

来自pSZ1470的FADc RNAi表达盒的相关部分

<210> 168
<211> 6726
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 168

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60
ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120
tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180
ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240
gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300
ggaagacagg tgaggggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360
atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420
cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480
gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540
ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600
ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660
cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720
ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 960
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 1020
cagtcacaac ccgcaaacgg cgcgccatgc tgctgcaggc cttcctgttc ctgctggccg 1080
gcttcgccgc caagatcagc gcctccatga cgaacgagac gtccgaccgc cccctggtgc 1140
acttcacccc caacaagggc tggatgaacg accccaacgg cctgtggtac gacgagaagg 1200
acgccaagtg gcacctgtac ttccagtaca acccgaacga caccgtctgg gggacgccct 1260
tgttctgggg ccacgccacg tccgacgacc tgaccaactg ggaggaccag cccatcgcca 1320
tcgccccgaa gcgcaacgac tccggcgcct tctccggctc catggtggtg gactacaaca 1380
acacctccgg cttcttcaac gacaccatcg acccgcgcca gcgctgcgtg gccatctgga 1440
cctacaacac cccggagtcc gaggagcagt acatctccta cagcctggac ggcggctaca 1500
ccttcaccga gtaccagaag aaccccgtgc tggccgccaa ctccacccag ttccgcgacc 1560
cgaaggtctt ctggtacgag ccctcccaga agtggatcat gaccgcggcc aagtcccagg 1620
actacaagat cgagatctac tcctccgacg acctgaagtc ctggaagctg gagtccgcgt 1680
tcgccaacga gggcttcctc ggctaccagt acgagtgccc cggcctgatc gaggtcccca 1740
ccgagcagga ccccagcaag tcctactggg tgatgttcat ctccatcaac cccggcgccc 1800
cggccggcgg ctccttcaac cagtacttcg tcggcagctt caacggcacc cacttcgagg 1860
ccttcgacaa ccagtcccgc gtggtggact tcggcaagga ctactacgcc ctgcagacct 1920
tcttcaacac cgacccgacc tacgggagcg ccctgggcat cgcgtgggcc tccaactggg 1980
agtactccgc cttcgtgccc accaacccct ggcgctcctc catgtccctc gtgcgcaagt 2040
tctccctcaa caccgagtac caggccaacc cggagacgga gctgatcaac ctgaaggccg 2100
agccgatcct gaacatcagc aacgccggcc cctggagccg gttcgccacc aacaccacgt 2160
tgacgaaggc caacagctac aacgtcgacc tgtccaacag caccggcacc ctggagttcg 2220
agctggtgta cgccgtcaac accacccaga cgatctccaa gtccgtgttc gcggacctct 2280
ccctctggtt caagggcctg gaggaccccg aggagtacct ccgcatgggc ttcgaggtgt 2340
ccgcgtcctc cttcttcctg gaccgcggga acagcaaggt gaagttcgtg aaggagaacc 2400
cctacttcac caaccgcatg agcgtgaaca accagccctt caagagcgag aacgacctgt 2460
cctactacaa ggtgtacggc ttgctggacc agaacatcct ggagctgtac ttcaacgacg 2520
gcgacgtcgt gtccaccaac acctacttca tgaccaccgg gaacgccctg ggctccgtga 2580
acatgacgac gggggtggac aacctgttct acatcgacaa gttccaggtg cgcgaggtca 2640
agtgacaatt ggcagcagca gctcggatag tatcgacaca ctctggacgc tggtcgtgtg 2700
atggactgtt gccgccacac ttgctgcctt gacctgtgaa tatccctgcc gcttttatca 2760
aacagcctca gtgtgtttga tcttgtgtgt acgcgctttt gcgagttgct agctgcttgt 2820
gctatttgcg aataccaccc ccagcatccc cttccctcgt ttcatatcgc ttgcatccca 2880
accgcaactt atctacgctg tcctgctatc cctcagcgct gctcctgctc ctgctcactg 2940
cccctcgcac agccttggtt tgggctccgc ctgtattctc ctggtactgc aacctgtaaa 3000
ccagcactgc aatgctgatg cacgggaagt agtgggatgg gaacacaaat ggaggatccc 3060
gcgtctcgaa cagagcgcgc agaggaacgc tgaaggtctc gcctctgtcg cacctcagcg 3120
cggcatacac cacaataacc acctgacgaa tgcgcttggt tcttcgtcca ttagcgaagc 3180
gtccggttca cacacgtgcc acgttggcga ggtggcaggt gacaatgatc ggtggagctg 3240

atggtcgaaa cgttcacagc ctagggatat cgaattcggc cgacaggacg cgcgtcaaag 3300
gtgctggtcg tgtatgccct ggccggcagg tcgttgctgc tgctggttag tgattccgca 3360
accctgattt tggcgtctta ttttggcgtg gcaaacgctg gcgcccgcga gccgggccgg 3420
cggcgatgcg gtgccccacg gctgccggaa tccaagggag gcaagagcgc ccgggtcagt 3480
tgaagggctt tacgcgcaag gtacagccgc tcctgcaagg ctgcgtggtg gaattggacg 3540
tgcaggtcct gctgaagttc ctccaccgcc tcaccagcgg acaaagcacc ggtgtatcag 3600
gtccgtgtca tccactctaa agagctcgac tacgacctac tgatggccct agattcttca 3660
tcaaaaacgc ctgagacact tgcccaggat tgaaactccc tgaagggacc accaggggcc 3720
ctgagttgtt ccttcccccc gtggcgagct gccagccagg ctgtacctgt gatcgaggct 3780
ggcgggaaaa taggcttcgt gtgctcaggt catgggaggt gcaggacagc tcatgaaacg 3840
ccaacaatcg cacaattcat gtcaagctaa tcagctattt cctcttcacg agctgtaatt 3900
gtcccaaaat tctggtctac cgggggtgat ccttcgtgta cgggcccttc cctcaaccct 3960
aggtatgcgc gcatgcggtc gccgcgcaac tcgcgcgagg gccgagggtt tgggacgggc 4020
cgtcccgaaa tgcagttgca cccggatgcg tggcaccttt tttgcgataa tttatgcaat 4080
ggactgctct gcaaaattct ggctctgtcg ccaaccctag gatcagcggc gtaggatttc 4140
gtaatcattc gtcctgatgg ggagctaccg actaccctaa tatcagcccg actgcctgac 4200
gccagcgtcc acttttgtgc acacattcca ttcgtgccca agacatttca ttgtggtgcg 4260
aagcgtcccc agttacgctc acctgtttcc cgacctcctt actgttctgt cgacagagcg 4320
ggcccacagg ccggtcgcag ccactagttc acttgtggaa ccagagcgcg gagtcgtcct 4380
cgggcgcgtc cgggacgacg tagcggcagt cgcgccagtc ctcccacagg gcgcggccga 4440
cccagcggct gtcggactgg tagtacttgc ccaggatggg cctgatggcg gcggaggcct 4500
cctcggcgtg gtagtgcggg atggtgctga agaggtggtg cagcacgtgg gtgtcggaga 4560
tgtggtgcag gatgttgtcc atgaagggcg ggcccatgga gcggtccacg gtggccatgg 4620
cgccgcgcag ccagtcccag tccttctcga agtagtgcgg cagcgccggg tgcgtgtgct 4680
ggagcagcgt gatgagcacg agccacatgt tcacgatcag gtagggcacc acgtaggtct 4740
tgaccagcca ggcccagccc atggtgcggc ccagcacgct gagcccgctg agcaccgcca 4800
ccagcgccag gtcggagatg accacctcga tgcgctcgcg cttgctgaag atgggcgacc 4860
acgggtcaaa gtggttggcg aagcgcgggt acggccgcga ggcgacgttg aacatgaggt 4920
acagcggcca gcccagggtc agggtgacca gcaccttgcc catgcggatg ggcagccact 4980
cctcccactc caggccctcg tgcgccactg cgcggtgcgg cggcacaaac acctcgtcct 5040
tgtccaggca ccccgtgttg gagtggtggc ggcggtgcga gtgcttccag gagtagtagg 5100
gcaccagcag caggctgtgg aacaccaggc ccacgccgtc gttgatggcc tggctggagg 5160
aaaaggcctg gtggccgcac tcgtgcgcgc acacccagac acccgtgccg aaggcgccct 5220
ggaaggtgcc gggagggtag ccgggggggt cagcgggaca ggcgccttct cctccatgcg 5280
cgccaccagg acctcaatac gggcaaccaa aaccctcaaa cacacctgga agaaccagta 5340
gagcggccac atgatgccgt acttgaccca cgtaggcacc ggtgcagggt cgatgtacgt 5400
cgacgcgacg tagagcaggg acatgaccgc gatgtcaaag gccaggtaca tgctgctacg 5460
aagcgccgag cgctcgaaac agtgcgcggg gatggccttg cgcagcgtcc cgatcgtgaa 5520
cggaggcttc tccacaggct gcctgttcgt cttgatagcc atctcgaggc agcagcagct 5580

cggatagtat cgacacactc tggacgctgg tcgtgtgatg gactgttgcc gccacacttg 5640
ctgccttgac ctgtgaatat ccctgccgct tttatcaaac agcctcagtg tgtttgatct 5700
tgtgtgtacg cgcttttgcg agttgctagc tgcttgtgct atttgcgaat accaccccca 5760
gcatcccctt ccctcgtttc atatcgcttg catcccaacc gcaacttatc tacgctgtcc 5820
tgctatccct cagcgctgct cctgctcctg ctcactgccc ctcgcacagc cttggtttgg 5880
gctccgcctg tattctcctg gtactgcaac ctgtaaacca gcactgcaat gctgatgcac 5940
gggaagtagt gggatgggaa cacaaatgga aagctgtaga gctcttgttt tccagaagga 6000
gttgctcctt gagcctttca ttctcagcct cgataacctc caaagccgct ctaattgtgg 6060
agggggttcg aatttaaaag cttggaatgt tggttcgtgc gtctggaaca agcccagact 6120
tgttgctcac tgggaaaagg accatcagct ccaaaaaact tgccgctcaa accgcgtacc 6180
tctgctttcg cgcaatctgc cctgttgaaa tcgccaccac attcatattg tgacgcttga 6240
gcagtctgta attgcctcag aatgtggaat catctgcccc ctgtgcgagc ccatgccagg 6300
catgtcgcgg gcgaggacac ccgccactcg tacagcagac cattatgcta cctcacaata 6360
gttcataaca gtgaccatat ttctcgaagc tccccaacga gcacctccat gctctgagtg 6420
gccacccccc ggccctggtg cttgcggagg gcaggtcaac cggcatgggg ctaccgaaat 6480
ccccgaccgg atcccaccac ccccgcgatg ggaagaatct ctccccggga tgtgggccca 6540
ccaccagcac aacctgctgg cccaggcgag cgtcaaacca taccacacaa atatccttgg 6600
catcggccct gaattccttc tgccgctctg ctacccggtg cttctgtccg aagcaggggt 6660
tgctagggat cgctccgagt ccgcaaaccc ttgtcgcgtg gcggggcttg ttcgagcttg 6720
aagagc 6726

来自pSZ1471的RNAi表达盒的FADc部分

<210> 169
<211> 1220
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 169

actagttcac ttgtggaacc agagcgcgga gtcgtcctcg ggcgcgtccg ggacgacgta 60
gcggcagtcg cgccagtcct cccacagggc gcggccgacc cagcggctgt cggactggta 120
gtacttgccc aggatgggcc tgatggcggc ggaggcctcc tcggcgtggt agtgcgggat 180
ggtgctgaag aggtggtgca gcacgtgggt gtcggagatg tggtgcagga tgttgtccat 240
gaagggcggg cccatggagc ggtccacggt ggccatggcg ccgcgcagcc agtcccagtc 300
cttctcgaag tagtgcggca gcgccgggtg cgtgtgctgg agcagcgtga tgagcacgag 360
ccacatgttc acgatcaggt agggcaccac gtaggtcttg accagccagg cccagcccat 420
ggtgcggccc agcacgctga gcccgctgag caccgccacc agcgccaggt cggagatgac 480
cacctcgatg cgctcgcgct tgctgaagat gggcgaccac gggtcaaagt ggttggcgaa 540
gcgcgggtac ggccgcgagg cgacgttgaa catgaggtac agcggccagc ccagggtcag 600
ggtgaccagc accttgccca tgcggatggg cagccactcc tcccactcca ggccctcgtg 660

cgccactgcg cggtgcggcg gcacaaacac ctcgtccttg tccaggcacc ccgtgttgga 720
gtggtggcgg cggtgcgagt gcttccagga gtagtagggc accagcagca ggctgtggaa 780
caccaggccc acgccgtcgt tgatggcctg gctggaggaa aaggcctggt ggccgcactc 840
gtgcgcgcac acccagacac ccgtgccgaa ggcgccctgg aaggtgccgg gagggtagcc 900
gggggggtca gcgggacagg cgccttctcc tccatgcgcg ccaccaggac ctcaatacgg 960
gcaaccaaaa ccctcaaaca cacctggaag aaccagtaga gcggccacat gatgccgtac 1020
ttgacccacg taggcaccgg tgcagggtcg atgtacgtcg acgcgacgta gagcagggac 1080
atgaccgcga tgtcaaaggc caggtacatg ctgctacgaa gcgccgagcg ctcgaaacag 1140
tgcgcgggga tggccttgcg cagcgtcccg atcgtgaacg gaggcttctc cacaggctgc 1200
ctgttcgtct tgatagccat 1220

来自pSZ1471的FADc RNAi表达盒的相关部分

<210> 170
<211> 5973
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 170

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60
ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120
tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180
ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240
gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300
ggaagacagg tgagggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360
atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420
cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480
gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540
ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600
ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660
cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720
ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 960
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 1020
cagtcacaac ccgcaaacgg cgcgccatgc tgctgcaggc cttcctgttc ctgctggccg 1080
gcttcgccgc caagatcagc gcctccatga cgaacgagac gtccgaccgc cccctggtgc 1140
acttcacccc caacaagggc tggatgaacg accccaacgg cctgtggtac gacgagaagg 1200
acgccaagtg gcacctgtac ttccagtaca acccgaacga caccgtctgg gggacgccct 1260

tgttctgggg ccacgccacg tccgacgacc tgaccaactg ggaggaccag cccatcgcca 1320
tcgccccgaa gcgcaacgac tccggcgcct tctccggctc catggtggtg gactacaaca 1380
acacctccgg cttcttcaac gacaccatcg acccgcgcca gcgctgcgtg gccatctgga 1440
cctacaacac cccggagtcc gaggagcagt acatctccta cagcctggac ggcggctaca 1500
ccttcaccga gtaccagaag aaccccgtgc tggccgccaa ctccacccag ttccgcgacc 1560
cgaaggtctt ctggtacgag ccctcccaga agtggatcat gaccgcggcc aagtcccagg 1620
actacaagat cgagatctac tcctccgacg acctgaagtc ctggaagctg gagtccgcgt 1680
tcgccaacga gggcttcctc ggctaccagt acgagtgccc cggcctgatc gaggtcccca 1740
ccgagcagga ccccagcaag tcctactggg tgatgttcat ctccatcaac cccggcgccc 1800
cggccggcgg ctccttcaac cagtacttcg tcggcagctt caacggcacc cacttcgagg 1860
ccttcgacaa ccagtcccgc gtggtggact tcggcaagga ctactacgcc ctgcagacct 1920
tcttcaacac cgacccgacc tacgggagcg ccctgggcat cgcgtgggcc tccaactggg 1980
agtactccgc cttcgtgccc accaacccct ggcgctcctc catgtccctc gtgcgcaagt 2040
tctccctcaa caccgagtac caggccaacc cggagacgga gctgatcaac ctgaaggccg 2100
agccgatcct gaacatcagc aacgccggcc cctggagccg gttcgccacc aacaccacgt 2160
tgacgaaggc caacagctac aacgtcgacc tgtccaacag caccggcacc ctggagttcg 2220
agctggtgta cgccgtcaac accacccaga cgatctccaa gtccgtgttc gcggacctct 2280
ccctctggtt caagggcctg gaggaccccg aggagtacct ccgcatgggc ttcgaggtgt 2340
ccgcgtcctc cttcttcctg gaccgcggga acagcaaggt gaagttcgtg aaggagaacc 2400
cctacttcac caaccgcatg agcgtgaaca accagccctt caagagcgag aacgacctgt 2460
cctactacaa ggtgtacggc ttgctggacc agaacatcct ggagctgtac ttcaacgacg 2520
gcgacgtcgt gtccaccaac acctacttca tgaccaccgg gaacgccctg ggctccgtga 2580
acatgacgac gggggtggac aacctgttct acatcgacaa gttccaggtg cgcgaggtca 2640
agtgacaatt ggcagcagca gctcggatag tatcgacaca ctctggacgc tggtcgtgtg 2700
atggactgtt gccgccacac ttgctgcctt gacctgtgaa tatccctgcc gcttttatca 2760
aacagcctca gtgtgtttga tcttgtgtgt acgcgctttt gcgagttgct agctgcttgt 2820
gctatttgcg aataccaccc ccagcatccc cttccctcgt ttcatatcgc ttgcatccca 2880
accgcaactt atctacgctg tcctgctatc cctcagcgct gctcctgctc ctgctcactg 2940
cccctcgcac agccttggtt tgggctccgc ctgtattctc ctggtactgc aacctgtaaa 3000
ccagcactgc aatgctgatg cacgggaagt agtgggatgg gaacacaaat ggaggatccc 3060
gcgtctcgaa cagagcgcgc agaggaacgc tgaaggtctc gcctctgtcg cacctcagcg 3120
cggcatacac cacaataacc acctgacgaa tgcgcttggt tcttcgtcca ttagcgaagc 3180
gtccggttca cacacgtgcc acgttggcga ggtggcaggt gacaatgatc ggtggagctg 3240
atggtcgaaa cgttcacagc ctagggatat cgaattcctt tcttgcgcta tgacacttcc 3300
agcaaaaggt agggcgggct gcgagacggc ttcccggcgc tgcatgcaac accgatgatg 3360
cttcgacccc ccgaagctcc ttcggggctg catgggcgct ccgatgccgc tccagggcga 3420
gcgctgttta aatagccagg cccccgattg caaagacatt atagcgagct accaaagcca 3480
tattcaaaca cctagatcac taccacttct acacaggcca ctcgagcttg tgatcgcact 3540
ccgctaaggg ggcgcctctt cctcttcgtt tcagtcacaa cccgcaaaca ctagttcact 3600

tgtggaacca gagcgcggag tcgtcctcgg gcgcgtccgg gacgacgtag cggcagtcgc 3660
gccagtcctc ccacagggcg cggccgaccc agcggctgtc ggactggtag tacttgccca 3720
ggatgggcct gatggcggcg gaggcctcct cggcgtggta gtgcgggatg gtgctgaaga 3780
ggtggtgcag cacgtgggtg tcggagatgt ggtgcaggat gttgtccatg aagggcgggc 3840
ccatggagcg gtccacggtg gccatggcgc cgcgcagcca gtcccagtcc ttctcgaagt 3900
agtgcggcag cgccgggtgc gtgtgctgga gcagcgtgat gagcacgagc cacatgttca 3960
cgatcaggta gggcaccacg taggtcttga ccagccaggc ccagcccatg gtgcggccca 4020
gcacgctgag cccgctgagc accgccacca gcgccaggtc ggagatgacc acctcgatgc 4080
gctcgcgctt gctgaagatg ggcgaccacg ggtcaaagtg gttggcgaag cgcgggtacg 4140
gccgcgaggc gacgttgaac atgaggtaca gcggccagcc cagggtcagg gtgaccagca 4200
ccttgcccat gcggatgggc agccactcct cccactccag gccctcgtgc gccactgcgc 4260
ggtgcggcgg cacaaacacc tcgtccttgt ccaggcaccc cgtgttggag tggtggcggc 4320
ggtgcgagtg cttccaggag tagtagggca ccagcagcag gctgtggaac accaggccca 4380
cgccgtcgtt gatggcctgg ctggaggaaa aggcctggtg gccgcactcg tgcgcgcaca 4440
cccagacacc cgtgccgaag gcgccctgga aggtgccggg agggtagccg ggggggtcag 4500
cgggacaggc gccttctcct ccatgcgcgc caccaggacc tcaatacggg caaccaaaac 4560
cctcaaacac acctggaaga accagtagag cggccacatg atgccgtact tgacccacgt 4620
aggcaccggt gcagggtcga tgtacgtcga cgcgacgtag agcagggaca tgaccgcgat 4680
gtcaaaggcc aggtacatgc tgctacgaag cgccgagcgc tcgaaacagt gcgcggggat 4740
ggccttgcgc agcgtcccga tcgtgaacgg aggcttctcc acaggctgcc tgttcgtctt 4800
gatagccatc tcgaggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg 4860
tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt 4920
atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc 4980
ttgtgctatt tgcgaatacc acccccagca tccccttccc tcgtttcata tcgcttgcat 5040
cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc 5100
actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg 5160
taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggaaag 5220
ctgtagagct cttgttttcc agaaggagtt gctccttgag cctttcattc tcagcctcga 5280
taacctccaa agccgctcta attgtggagg gggttcgaat ttaaaagctt ggaatgttgg 5340
ttcgtgcgtc tggaacaagc ccagacttgt tgctcactgg gaaaaggacc atcagctcca 5400
aaaaacttgc cgctcaaacc gcgtacctct gctttcgcgc aatctgccct gttgaaatcg 5460
ccaccacatt catattgtga cgcttgagca gtctgtaatt gcctcagaat gtggaatcat 5520
ctgccccctg tgcgagccca tgccaggcat gtcgcgggcg aggacacccg ccactcgtac 5580
agcagaccat tatgctacct cacaatagtt cataacagtg accatatttc tcgaagctcc 5640
ccaacgagca cctccatgct ctgagtggcc acccccggc cctggtgctt gcggagggca 5700
ggtcaaccgg catggggcta ccgaaatccc cgaccggatc ccaccacccc cgcgatggga 5760
agaatctctc cccgggatgt gggcccacca ccagcacaac ctgctggccc aggcgagcgt 5820
caaaccatac cacacaaata tccttggcat cggccctgaa ttccttctgc cgctctgcta 5880
cccggtgctt ctgtccgaag caggggttgc tagggatcgc tccgagtccg caaacccttg 5940

tcgcgtggcg gggcttgttc gagcttgaag agc 5973

来自pSZ1377的具有原始小球藻UTEX250硬脂酰-ACP转运肽的油橄榄硬脂酰-ACP去饱和酶的密码子优化的编码序列

<210> 171

<211> 1221

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 171

actagtatgg ccaccgcatc cactttctcg gcgttcaatg cccgctgcgg cgacctgcgt 60

cgctcggcgg gctccgggcc ccggcgccca gcgaggcccc tccccgtgcg cgggcgcgcc 120

gaggtgcacg tgcaggtgac ccactccctg gcccccgaga agcgcgagat cttcaactcc 180

ctgaacaact gggcccagga gaacatcctg gtgctgctga aggacgtgga caagtgctgg 240

cagccctccg acttcctgcc cgactccgcc tccgagggct tcgacgagca ggtgatggag 300

ctgcgcaagc gctgcaagga gatccccgac gactacttca tcgtgctggt gggcgacatg 360

atcaccgagg aggccctgcc cacctaccag accatgctga acaccctgga cggcgtgcgc 420

gacgagaccg gcgcctccct gacccccctgg gccatctgga cccgcgcctg gaccgccgag 480

gagaaccgcc acggcgacct gctgaacaag tacctgtacc tgtccggccg cgtggacatg 540

aagcagatcg agaagaccat ccagtacctg atcggctccg gcatggaccc ccgcaccgag 600

aacaacccct acctgggctt catctacacc tccttccagg agcgcgccac cttcatctcc 660

cacggcaaca ccgcccgcct ggccaaggag cacggcgacc tgaagctggc ccagatctgc 720

ggcatcatcg ccgccgacga gaagcgccac gagaccgcct acaccaagat cgtggagaag 780

ctgttcgaga tcgaccccga cggcaccgtg ctggccctgg ccgacatgat gcgcaagaag 840

gtgtccatgc ccgcccacct gatgtacgac ggccaggacg acaacctgtt cgagaacttc 900

tcctccgtgg cccagcgcct gggcgtgtac accgccaagg actacgccga catcctggag 960

ttcctggtgg gccgctggga catcgagaag ctgaccggcc tgtccggcga gggccgcaag 1020

gcccaggact acgtgtgcac cctgcccccc cgcatccgcc gcctggagga gcgcgcccag 1080

tcccgcgtga agaaggcctc cgccaccccc ttctcctgga tcttcggccg cgagatcaac 1140

ctgatggact acaaggacca cgacggcgac tacaaggacc acgacatcga ctacaaggac 1200

gacgacgaca agtgaatcga t 1221

油橄榄硬脂酰-ACP去饱和酶的氨基酸序列(登录号AAB67840.1)

<210> 172

<211> 390

<212> PRT

<213> 油橄榄

<400> 172

Met Ala Leu Lys Leu Cys Phe Pro Pro His Lys Met Pro Ser Phe Pro

1 5 10 15

Asp Ala Arg Ile Arg Ser His Arg Val Phe Met Ala Ser Thr Ile His
20 25 30
Ser Pro Ser Met Glu Val Gly Lys Val Lys Lys Pro Phe Thr Pro Pro
35 40 45
Arg Glu Val His Val Gln Val Thr His Ser Leu Ala Pro Glu Lys Arg
50 55 60
Glu Ile Phe Asn Ser Leu Asn Asn Trp Ala Gln Glu Asn Ile Leu Val
65 70 75 80
Leu Leu Lys Asp Val Asp Lys Cys Trp Gln Pro Ser Asp Phe Leu Pro
85 90 95
Asp Ser Ala Ser Glu Gly Phe Asp Glu Gln Val Met Glu Leu Arg Lys
100 105 110
Arg Cys Lys Glu Ile Pro Asp Asp Tyr Phe Ile Val Leu Val Gly Asp
115 120 125
Met Ile Thr Glu Glu Ala Leu Pro Thr Tyr Gln Thr Met Leu Asn Thr
130 135 140
Leu Asp Gly Val Arg Asp Glu Thr Gly Ala Ser Leu Thr Pro Trp Ala
145 150 155 160
Ile Trp Thr Arg Ala Trp Thr Ala Glu Glu Asn Arg His Gly Asp Leu
165 170 175
Leu Asn Lys Tyr Leu Tyr Leu Ser Gly Arg Val Asp Met Lys Gln Ile
180 185 190
Glu Lys Thr Ile Gln Tyr Leu Ile Gly Ser Gly Met Asp Pro Arg Thr
195 200 205
Glu Asn Asn Pro Tyr Leu Gly Phe Ile Tyr Thr Ser Phe Gln Glu Arg
210 215 220
Ala Thr Phe Ile Ser His Gly Asn Thr Ala Arg Leu Ala Lys Glu His
225 230 235 240
Gly Asp Leu Lys Leu Ala Gln Ile Cys Gly Ile Ile Ala Ala Asp Glu
245 250 255
Lys Arg His Glu Thr Ala Tyr Thr Lys Ile Val Glu Lys Leu Phe Glu
260 265 270
Ile Asp Pro Asp Gly Thr Val Leu Ala Leu Ala Asp Met Met Arg Lys
275 280 285
Lys Val Ser Met Pro Ala His Leu Met Tyr Asp Gly Gln Asp Asp Asn
290 295 300
Leu Phe Glu Asn Phe Ser Ser Val Ala Gln Arg Leu Gly Val Tyr Thr
305 310 315 320
Ala Lys Asp Tyr Ala Asp Ile Leu Glu Phe Leu Val Gly Arg Trp Asp

325 330 335
Ile Glu Lys Leu Thr Gly Leu Ser Gly Glu Gly Arg Lys Ala Gln Asp
340 345 350
Tyr Val Cys Thr Leu Pro Pro Arg Ile Arg Arg Leu Glu Glu Arg Ala
355 360 365
Gln Ser Arg Val Lys Lys Ala Ser Ala Thr Pro Phe Ser Trp Ile Phe
370 375 380
Gly Arg Glu Ile Asn Leu
385 390

来自pSZ1419的蓖麻油酸12-羟化酶的密码子优化的编码序列

<210> 173
<211> 1254
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 173

actagtatgg gcggcggcgg ccgcatgtcc accgtgatca cctccaacaa ctccgagaag 60
aagggcggct cctcccacct gaagcgcgcc ccccacacca agcccccctt caccctgggc 120
gacctgaagc gcgccatccc cccccactgc ttcgagcgct ccttcgtgcg cgggcgcgcc 180
tccttctcct acgtggccta cgacgtgtgc ctgtccttcc tgttctactc catcgccacc 240
aacttcttcc cctacatctc ctcccccctg tcctacgtgg cctggctggt gtactggctg 300
ttccagggct gcatcctgac cggcctgtgg gtgatcggcc acgagtgcgg ccaccacgcc 360
ttctccgagt accagctggc cgacgacatc gtgggcctga tcgtgcactc cgccctgctg 420
gtgccctact tctcctggaa gtactcccac cgccgccacc actccaacat cggctccctg 480
gagcgcgacg aggtgttcgt gcccaagtcc aagtccaaga tctcctggta ctccaagtac 540
tccaacaacc ccccggccg cgtgctgacc ctggccgcca ccctgctgct gggctggccc 600
ctgtacctgg ccttcaacgt gtccggccgc ccctacgacc gcttcgcctg ccactacgac 660
ccctacggcc ccatcttctc cgagcgcgag cgcctgcaga tctacatcgc cgacctgggc 720
atcttcgcca ccaccttcgt gctgtaccag gccaccatgg ccaagggcct ggcctgggtg 780
atgcgcatct acggcgtgcc cctgctgatc gtgaactgct tcctggtgat gatcacctac 840
ctgcagcaca cccaccccgc catcccccgc tacggctcct ccgagtggga ctggctgcgc 900
ggcgccatgg tgaccgtgga ccgcgactac ggcgtgctga acaaggtgtt ccacaacatc 960
gccgacaccc acgtggccca ccacctgttc gccaccgtgc cccactacca cgccatggag 1020
gccaccaagg ccatcaagcc catcatgggc gagtactacc gctacgacgg caccccccttc 1080
tacaaggccc tgtggcgcga ggccaaggag tgcctgttcg tggagcccga cgagggcgcc 1140
cccacccagg gcgtgttctg gtaccgcaac aagtacatgg actacaagga ccacgacggc 1200
gactacaagg accacgacat cgactacaag gacgacgacg acaagtgaat cgat 1254

蓖麻油酸12-羟化酶的氨基酸序列(Genbank登录号AAC49010.1)

<210> 174
<211> 387
<212> PRT
<213> 蓖麻
<400> 174

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

Ile Thr Tyr Leu Gln His Thr His Pro Ala Ile Pro Arg Tyr Gly Ser
275 280 285
Ser Glu Trp Asp Trp Leu Arg Gly Ala Met Val Thr Val Asp Arg Asp
290 295 300
Tyr Gly Val Leu Asn Lys Val Phe His Asn Ile Ala Asp Thr His Val
305 310 315 320
Ala His His Leu Phe Ala Thr Val Pro His Tyr His Ala Met Glu Ala
325 330 335
Thr Lys Ala Ile Lys Pro Ile Met Gly Glu Tyr Tyr Arg Tyr Asp Gly
340 345 350
Thr Pro Phe Tyr Lys Ala Leu Trp Arg Glu Ala Lys Glu Cys Leu Phe
355 360 365
Val Glu Pro Asp Glu Gly Ala Pro Thr Gln Gly Val Phe Trp Tyr Arg
370 375 380
Asn Lys Tyr
385

蓖麻

<210> 175
<211> 535
<212> PRT
<213> 蓖麻
<400> 175

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

130 135 140

Tyr Asn Asn Leu Leu Asp Gly Ser Ser Gly Ile Ser Gln Ile Asp Ser
145 150 155 160
Phe Asp Cys Ala Gln Phe Pro Thr Arg Ile Ala Gly Glu Ile Lys Ser
165 170 175
Phe Ser Thr Asp Gly Trp Val Ala Pro Lys Leu Ser Lys Arg Met Asp
180 185 190
Lys Phe Met Leu Tyr Met Leu Thr Ala Gly Lys Lys Ala Leu Ala Asp
195 200 205
Gly Gly Ile Thr Glu Asp Met Met Asp Glu Leu Asp Lys Ala Arg Cys
210 215 220
Gly Val Leu Ile Gly Ser Ala Met Gly Gly Met Lys Val Phe Asn Asp
225 230 235 240
Ala Ile Glu Ala Leu Arg Ile Ser Tyr Arg Lys Met Asn Pro Phe Cys
245 250 255
Val Pro Phe Ala Thr Thr Asn Met Gly Ser Ala Met Leu Ala Met Asp
260 265 270
Leu Gly Trp Met Gly Pro Asn Tyr Ser Ile Ser Thr Ala Cys Ala Thr
275 280 285
Ser Asn Phe Cys Ile Leu Asn Ala Ala Asn His Ile Ile Arg Gly Glu
290 295 300
Ala Asp Ile Met Leu Cys Gly Gly Ser Asp Ala Ala Ile Ile Pro Ile
305 310 315 320
Gly Leu Gly Gly Phe Val Ala Cys Arg Ala Leu Ser Gln Arg Asn Asp
325 330 335
Asp Pro Thr Lys Ala Ser Arg Pro Trp Asp Met Asn Arg Asp Gly Phe
340 345 350
Val Met Gly Glu Gly Ala Gly Val Leu Leu Leu Glu Glu Leu Glu His
355 360 365
Ala Lys Lys Arg Gly Ala Asn Ile Tyr Ala Glu Phe Leu Gly Gly Ser
370 375 380
Phe Thr Cys Asp Ala Tyr His Met Thr Glu Pro Arg Pro Asp Gly Val
385 390 395 400
Gly Val Ile Leu Cys Ile Glu Lys Ala Leu Ala Arg Ser Gly Val Ser
405 410 415
Lys Glu Glu Val Asn Tyr Ile Asn Ala His Ala Thr Ser Thr Pro Ala
420 425 430
Gly Asp Leu Lys Glu Tyr Glu Ala Leu Met Arg Cys Phe Ser Gln Asn
435 440 445

Pro Asp Leu Arg Val Asn Ser Thr Lys Ser Met Ile Gly His Leu Leu
450 455 460
Gly Ala Ala Gly Ala Val Glu Ala Ile Ala Thr Ile Gln Ala Ile Arg
465 470 475 480
Thr Gly Trp Val His Pro Asn Ile Asn Leu Glu Asn Pro Glu Glu Gly
485 490 495
Val Asp Thr Lys Val Leu Val Gly Pro Lys Lys Glu Arg Leu Asp Ile
500 505 510
Lys Val Ala Leu Ser Asn Ser Phe Gly Phe Gly Gly His Asn Ser Ser
515 520 525
Ile Ile Phe Ala Pro Tyr Lys
530 535

向日葵

<210> 176
<211> 491
<212> PRT
<213> 向日葵
<400> 176

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

165 170 175
Lys Ser Arg Cys Gly Val Leu Ile Gly Ser Ala Met Gly Gly Met Lys
180 185 190
Val Phe Asn Asp Ala Ile Glu Ala Leu Arg Val Ser Tyr Arg Lys Met
195 200 205
Asn Pro Phe Cys Val Pro Phe Ala Thr Thr Asn Met Gly Ser Ala Met
210 215 220
Leu Ala Met Asp Leu Gly Trp Met Gly Pro Asn Tyr Ser Ile Ser Thr
225 230 235 240
Ala Cys Ala Thr Ser Asn Phe Cys Ile Leu Asn Ala Ala Asn His Ile
245 250 255
Ile Arg Gly Glu Ala Asp Met Met Leu Cys Gly Gly Ser Asp Ala Val
260 265 270
Ile Ile Pro Ile Gly Leu Gly Gly Phe Val Ala Cys Arg Ala Leu Ser
275 280 285
Glu Arg Asn Thr Asp Pro Ala Lys Ala Ser Arg Pro Trp Asp Ser Gly
290 295 300
Arg Asp Gly Phe Val Met Gly Glu Gly Ala Gly Val Leu Leu Leu Glu
305 310 315 320
Glu Leu Glu His Ala Lys Lys Arg Gly Ala Lys Ile Tyr Ala Glu Phe
325 330 335
Leu Gly Gly Ser Phe Thr Cys Asp Ala Tyr His Met Thr Glu Pro His
340 345 350
Pro Glu Gly Ala Gly Val Ile Leu Cys Ile Glu Lys Ala Leu Ser Gln
355 360 365
Ala Gly Val Arg Arg Glu Asp Val Asn Tyr Ile Asn Ala His Ala Thr
370 375 380
Ser Thr Pro Ala Gly Asp Leu Lys Glu Tyr His Ala Leu Leu His Cys
385 390 395 400
Phe Gly Asn Asn Gln Glu Leu Arg Val Asn Ser Thr Lys Ser Met Ile
405 410 415
Gly His Leu Leu Gly Ala Ala Gly Ala Val Glu Ala Val Ala Thr Val
420 425 430
Gln Ala Ile Arg Thr Gly Trp Ile His Pro Asn Ile Asn Leu Glu Asn
435 440 445
Pro Asp Gln Gly Val Asp Thr Lys Val Leu Val Gly Ser Lys Lys Glu
450 455 460
Arg Leu Asn Val Lys Val Gly Leu Ser Asn Ser Phe Gly Phe Gly Gly
465 470 475 480

His Asn Ser Ser Ile Leu Phe Ala Pro Phe Gln
485 490

欧洲油菜

<210> 177

<211> 433

<212> PRT

<213> 欧洲油菜

<400> 177

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

245 250 255
Gly Val Leu Leu Leu Glu Glu Leu Glu His Ala Lys Lys Arg Gly Ala
260 265 270
Thr Ile Tyr Ala Glu Phe Leu Gly Gly Ser Phe Thr Cys Asp Ala Tyr
275 280 285
His Met Thr Glu Pro Arg Pro Asp Gly Ala Gly Val Ile Leu Cys Ile
290 295 300
Glu Arg Ala Leu Ala Asp Ala Gly Ile Ser Lys Glu Gln Ile Asn Tyr
305 310 315 320
Ile Asn Ala His Ala Thr Ser Thr Pro Ala Gly Asp Leu Lys Glu Tyr
325 330 335
Gln Ala Leu Ala His Cys Phe Gly Gln Asn Pro Glu Ile Lys Val Asn
340 345 350
Ser Thr Lys Ser Met Ile Gly His Leu Leu Gly Ala Ala Gly Ala Val
355 360 365
Glu Ala Val Ala Thr Val Gln Ala Ile Arg Thr Gly Trp Val His Pro
370 375 380
Asn Ile Asn Leu Glu Ser Pro Asp Asn Gly Val Asp Thr Asn Leu Leu
385 390 395 400
Val Gly Pro Glu Lys Glu Arg Leu Asp Ile Lys Ala Ala Leu Ser Asn
405 410 415
Ser Phe Gly Phe Gly Gly His Asn Ser Ser Ile Ile Phe Ala Pro Tyr
420 425 430
Lys

分别具有GenBank登录号ABR01158.1、ABI18155.1、AAF61739.1以及AAW88763.1的大豆

<210> 178
<211> 488
<212> PRT
<213> 大豆
<400> 178
Met Ala Ser Thr Thr Thr Ser Ser Leu Cys Thr Trp Leu Val Ala Ala
1 5 10 15
Cys Met Ser Val Thr Cys His Ala Asp Arg Thr Lys Thr Pro His Ala
20 25 30
Met Phe Arg Ser Ser Lys Lys Ser Arg Tyr Ser Gln Phe Asn Val Cys
35 40 45
Arg Ser Thr His Ser Gly Lys Thr Met Ala Val Ala Leu Gln Pro Thr
50 55 60
Gln Glu Ile Thr Thr Ile Lys Lys Pro Pro Thr Lys Gln Arg Arg Val

65 70 75 80
Val Val Thr Gly Leu Gly Val Val Thr Pro Leu Gly His Glu Pro Asp
85 90 95
Ile Phe Tyr Asn Asn Leu Leu Asp Gly Ala Ser Gly Ile Ser Glu Ile
100 105 110
Glu Thr Phe Asp Cys Ala Glu Tyr Pro Thr Arg Ile Ala Gly Glu Ile
115 120 125
Lys Ser Phe Ser Thr Asp Gly Trp Val Ala Pro Lys Leu Ser Lys Arg
130 135 140
Met Asp Lys Phe Met Leu Tyr Met Leu Thr Ala Gly Lys Lys Ala Leu
145 150 155 160
Val Asp Gly Gly Ile Thr Asp Asp Val Met Asp Glu Leu Asn Lys Glu
165 170 175
Lys Cys Gly Val Leu Ile Gly Ser Ala Met Gly Gly Met Lys Val Phe
180 185 190
Asn Asp Ala Ile Glu Ala Leu Arg Ile Ser Tyr Lys Lys Met Asn Pro
195 200 205
Phe Cys Val Pro Phe Ala Thr Thr Asn Met Gly Ser Ala Met Leu Ala
210 215 220
Met Asp Leu Gly Trp Met Gly Pro Asn Tyr Ser Ile Ser Thr Ala Cys
225 230 235 240
Ala Thr Ser Asn Phe Cys Ile Leu Asn Ala Ala Asn His Ile Ile Arg
245 250 255
Gly Glu Ala Asp Val Met Leu Cys Gly Gly Ser Asp Ala Ala Ile Ile
260 265 270
Pro Ile Gly Leu Gly Gly Phe Val Ala Cys Arg Ala Leu Ser Gln Arg
275 280 285
Asn Thr Asp Pro Thr Lys Ala Ser Arg Pro Trp Asp Ile Asn Arg Asp
290 295 300
Gly Phe Val Met Gly Glu Gly Ala Gly Val Leu Leu Leu Glu Glu Leu
305 310 315 320
Glu His Ala Lys Glu Arg Gly Ala Thr Ile Tyr Ala Glu Phe Leu Gly
325 330 335
Gly Ser Phe Thr Cys Asp Ala Tyr His Val Thr Glu Pro Arg Pro Asp
340 345 350
Gly Ala Gly Val Ile Leu Cys Ile Glu Lys Ala Leu Ala Gln Ser Gly
355 360 365
Val Ser Lys Glu Asp Val Asn Tyr Ile Asn Ala His Ala Thr Ser Thr
370 375 380

Pro Ala Gly Asp Leu Lys Glu Tyr Gln Ala Leu Met His Cys Phe Gly
385 390 395 400
Gln Asn Pro Glu Leu Arg Val Asn Ser Thr Lys Ser Met Ile Gly His
405 410 415
Leu Leu Gly Ala Ala Gly Gly Val Glu Ala Val Ala Thr Ile Gln Ala
420 425 430
Ile Arg Thr Gly Trp Val His Pro Asn Ile Asn Leu Glu Asn Pro Asp
435 440 445
Asn Gly Val Asp Ala Lys Val Leu Val Gly Ser Lys Lys Glu Arg Leu
450 455 460
Asp Val Lys Ala Ala Leu Ser Asn Ser Phe Gly Phe Gly Gly His Asn
465 470 475 480
Ser Ser Ile Ile Phe Ala Pro Tyr
485

P.moriformis
<210> 179
<211> 467
<212> PRT
<213> Prototheca moriformis
<400> 179
Met Gln Thr Ala His Gln Arg Pro Pro Thr Glu Gly His Cys Phe Gly
1 5 10 15
Ala Arg Leu Pro Thr Ala Ser Arg Arg Ala Val Arg Arg Ala Trp Ser
20 25 30
Arg Ile Ala Arg Ala Ala Ala Ala Ala Asp Ala Asn Pro Ala Arg Pro
35 40 45
Glu Arg Arg Val Val Ile Thr Gly Gln Gly Val Val Thr Ser Leu Gly
50 55 60
Gln Thr Ile Glu Gln Phe Tyr Ser Ser Leu Leu Glu Gly Val Ser Gly
65 70 75 80
Ile Ser Gln Ile Gln Lys Phe Asp Thr Thr Gly Tyr Thr Thr Thr Ile
85 90 95
Ala Gly Glu Ile Lys Ser Leu Gln Leu Asp Pro Tyr Val Pro Lys Arg
100 105 110
Trp Ala Lys Arg Val Asp Asp Val Ile Lys Tyr Val Tyr Ile Ala Gly
115 120 125
Lys Gln Ala Leu Glu Ser Ala Gly Leu Pro Ile Glu Ala Ala Gly Leu
130 135 140
Ala Gly Ala Gly Leu Asp Pro Ala Leu Cys Gly Val Leu Ile Gly Thr

145 150 155 160
Ala Met Ala Gly Met Thr Ser Phe Ala Ala Gly Val Glu Ala Leu Thr
165 170 175
Arg Gly Gly Val Arg Lys Met Asn Pro Phe Cys Ile Pro Phe Ser Ile
180 185 190
Ser Asn Met Gly Gly Ala Met Leu Ala Met Asp Ile Gly Phe Met Gly
195 200 205
Pro Asn Tyr Ser Ile Ser Thr Ala Cys Ala Thr Gly Asn Tyr Cys Ile
210 215 220
Leu Gly Ala Ala Asp His Ile Arg Arg Gly Asp Ala Asn Val Met Leu
225 230 235 240
Ala Gly Gly Ala Asp Ala Ala Ile Ile Pro Ser Gly Ile Gly Gly Phe
245 250 255
Ile Ala Cys Lys Ala Leu Ser Lys Arg Asn Asp Glu Pro Glu Arg Ala
260 265 270
Ser Arg Pro Trp Asp Ala Asp Arg Asp Gly Phe Val Met Gly Glu Gly
275 280 285
Ala Gly Val Leu Val Leu Glu Glu Leu Glu His Ala Lys Arg Arg Gly
290 295 300
Ala Thr Ile Leu Ala Glu Leu Val Gly Gly Ala Ala Thr Ser Asp Ala
305 310 315 320
His His Met Thr Glu Pro Asp Pro Gln Gly Arg Gly Val Arg Leu Cys
325 330 335
Leu Glu Arg Ala Leu Glu Arg Ala Arg Leu Ala Pro Glu Arg Val Gly
340 345 350
Tyr Val Asn Ala His Gly Thr Ser Thr Pro Ala Gly Asp Val Ala Glu
355 360 365
Tyr Arg Ala Ile Arg Ala Val Ile Pro Gln Asp Ser Leu Arg Ile Asn
370 375 380
Ser Thr Lys Ser Met Ile Gly His Leu Leu Gly Gly Ala Gly Ala Val
385 390 395 400
Glu Ala Val Ala Ala Ile Gln Ala Leu Arg Thr Gly Trp Leu His Pro
405 410 415
Asn Leu Asn Leu Glu Asn Pro Ala Pro Gly Val Asp Pro Val Val Leu
420 425 430
Val Gly Pro Arg Lys Glu Arg Ala Glu Asp Leu Asp Val Val Leu Ser
435 440 445
Asn Ser Phe Gly Phe Gly Gly His Asn Ser Cys Val Ile Phe Arg Lys
450 455 460

Tyr Asp Glu
465

P.moriformis硬脂酰ACP硫酯酶

<210> 180
<211> 400
<212> PRT
<213> Prototheca moriformis
<400> 180

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

245 250 255
Ser Asp Glu Gly Arg His Glu Ile Ala Tyr Thr Arg Ile Val Asp Glu
260 265 270
Phe Phe Arg Leu Asp Pro Glu Gly Ala Val Ala Ala Tyr Ala Asn Met
275 280 285
Met Arg Lys Gln Ile Thr Met Pro Ala His Leu Met Asp Asp Met Gly
290 295 300
His Gly Glu Ala Asn Pro Gly Arg Asn Leu Phe Ala Asp Phe Ser Ala
305 310 315 320
Val Ala Glu Lys Ile Asp Val Tyr Asp Ala Glu Asp Tyr Cys Arg Ile
325 330 335
Leu Glu His Leu Asn Ala Arg Trp Lys Val Asp Glu Arg Gln Val Ser
340 345 350
Gly Gln Ala Ala Ala Asp Gln Glu Tyr Val Leu Gly Leu Pro Gln Arg
355 360 365
Phe Arg Lys Leu Ala Glu Lys Thr Ala Ala Lys Arg Lys Arg Val Ala
370 375 380
Arg Arg Pro Val Ala Phe Ser Trp Ile Ser Gly Arg Glu Ile Met Val
385 390 395 400

亚麻12去饱和酶

<210> 181
<211> 378
<212> PRT
<213> 亚麻
<400> 181

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

Leu Asp Asp Met Asp Gly Phe Val Leu His Ser Ser Leu Leu Val Pro
115 120 125
Tyr Phe Ser Trp Lys His Ser His Arg Arg His His Ser Asn Thr Gly
130 135 140
Ser Leu Asp Arg Asp Glu Val Phe Val Pro Lys Gln Lys Ala Glu Ile
145 150 155 160
Gly Trp Tyr Ser Lys Tyr Leu Asn Asn Pro Pro Gly Arg Val Ile Thr
165 170 175
Leu Ala Val Thr Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Phe Asn
180 185 190
Val Ser Gly Arg Pro Tyr Asp Arg Phe Ala Cys His Phe Asp Pro His
195 200 205
Gly Pro Ile Tyr Asn Asp Arg Glu Arg Met Glu Ile Tyr Leu Ser Asp
210 215 220
Ala Gly Ile Phe Thr Val Cys Tyr Ile Leu Tyr Arg Leu Val Leu Thr
225 230 235 240
Lys Gly Leu Val Trp Val Val Ser Ile Tyr Gly Val Pro Leu Leu Ile
245 250 255
Val Asn Gly Phe Leu Val Leu Ile Thr Phe Leu Gln His Thr His Pro
260 265 270
Ser Leu Pro His Tyr Lys Ser Ser Glu Trp Asp Trp Met Arg Gly Ala
275 280 285
Leu Ser Thr Val Asp Arg Asp Tyr Gly Leu Leu Asn Thr Val Phe His
290 295 300
Asn Ile Thr Asp Thr His Val Ala His His Leu Phe Ser Thr Met Pro
305 310 315 320
His Tyr His Ala Met Glu Ala Thr Lys Ala Ile Lys Pro Val Leu Gly
325 330 335
Glu Tyr Tyr Gln Phe Asp Gly Thr Pro Phe Val Lys Ala Met Trp Arg
340 345 350
Glu Ala Lys Glu Cys Ile Tyr Val Glu Pro Asp Glu Gly Asp Pro Ser
355 360 365
Gln Gly Val Phe Trp Tyr Asn Asn Lys Leu
370 375

亚麻15去饱和酶

<210> 182

<211> 391

<212> PRT

<213> 亚麻

<400> 182

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

His Val Ile His His Leu Phe Pro Gln Met Pro His Tyr His Leu Val
305 310 315 320
Glu Ala Thr Gln Ala Ala Lys His Val Leu Gly Lys Tyr Tyr Arg Glu
325 330 335
Pro Lys Lys Ser Gly Pro Phe Pro Phe His Leu Phe Gly Tyr Leu Val
340 345 350
Arg Ser Leu Gly Glu Asp His Tyr Val Ser Asp Thr Gly Asp Val Val
355 360 365
Phe Tyr Gln Ser Asp Pro His Ile Pro Lys Phe Arg Thr Ser Ser Ala
370 375 380
Thr Thr Lys Ser Lys Ser Ser
385 390

亚麻15去饱和酶

<210> 183
<211> 392
<212> PRT
<213> 亚麻
<400> 183

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

165 170 175
Thr Ile Pro Leu Pro Met Phe Ala Tyr Pro Ile Tyr Leu Trp Arg Arg
180 185 190
Ser Pro Gly Lys Lys Gly Ser His Phe Asn Pro Tyr Ser Asp Leu Phe
195 200 205
Ala Pro Asn Glu Arg Thr Ser Val Met Ile Ser Thr Leu Cys Trp Thr
210 215 220
Ala Met Ala Leu Leu Leu Cys Tyr Ser Ser Phe Ile Tyr Gly Phe Leu
225 230 235 240
Pro Val Phe Lys Ile Tyr Gly Val Pro Tyr Leu Ile Phe Val Ala Trp
245 250 255
Leu Asp Met Val Thr Tyr Leu His His His Gly Tyr Glu Gln Lys Leu
260 265 270
Pro Trp Tyr Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr
275 280 285
Thr Val Asp Arg Asp Tyr Gly Val Ile Asn Asn Ile His His Asp Ile
290 295 300
Gly Thr His Val Ile His His Leu Phe Pro Gln Met Pro His Tyr His
305 310 315 320
Leu Val Glu Ala Thr Gln Ala Ala Lys His Val Leu Gly Lys Tyr Tyr
325 330 335
Arg Glu Pro Lys Lys Ser Gly Pro Phe Pro Phe His Leu Phe Gly Tyr
340 345 350
Leu Val Arg Ser Leu Gly Glu Asp His Tyr Val Ser Asp Thr Gly Asp
355 360 365
Val Val Phe Tyr Gln Ser Asp Pro His Ile Pro Lys Phe Pro Thr Ser
370 375 380
Ala Thr Thr Lys Ser Lys Ser Ser
385 390

红花12去饱和酶

<210> 184
<211> 383
<212> PRT
<213> 红花
<400> 184

Met Gly Ala Gly Gly Arg Met Ser Asn Pro Ser Glu Gly Glu Lys Lys
1 5 10 15
Thr Glu Leu Glu Gly Ile Gln Arg Val Pro Tyr Gln Lys Pro Pro Phe
20 25 30

Thr Val Gly Asp Val Lys Lys Ala Ile Pro Pro His Cys Phe Asn Arg
35 40 45
Ser Val Ile Arg Ser Phe Ser Tyr Val Val Tyr Asp Leu Thr Ile Ala
50 55 60
Ser Ile Leu Tyr Tyr Ile Ala Thr Thr Phe Ile Pro Leu Leu Pro His
65 70 75 80
Pro Leu Ala Tyr Val Ala Trp Pro Ile Tyr Trp Ala Val Gln Gly Cys
85 90 95
Val Met Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His Ala
100 105 110
Phe Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Leu His
115 120 125
Ser Val Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg
130 135 140
His His Ser Asn Thr Gly Ser Ile Glu His Asp Glu Val Phe Val Pro
145 150 155 160
Lys Leu Lys Ser Gly Val Arg Ser Thr Ala Lys Tyr Leu Asn Asn Pro
165 170 175
Pro Gly Arg Ile Leu Thr Leu Leu Val Thr Leu Thr Leu Gly Trp Pro
180 185 190
Leu Tyr Leu Met Phe Asn Val Ser Gly Arg Tyr Tyr Asp Arg Phe Ala
195 200 205
Cys His Phe Asp Pro Asn Ser Pro Ile Tyr Ser Asn Arg Glu Arg Ala
210 215 220
Gln Ile Phe Ile Ser Asp Ala Gly Ile Phe Ala Val Leu Tyr Gly Leu
225 230 235 240
Tyr Arg Leu Ala Ala Val Lys Gly Leu Val Trp Val Leu Thr Val Tyr
245 250 255
Ala Gly Pro Leu Leu Val Val Asn Gly Phe Leu Val Leu Ile Thr Phe
260 265 270
Leu Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Thr Glu Trp
275 280 285
Asp Trp Leu Arg Gly Ala Leu Ala Thr Ile Asp Arg Asp Tyr Gly Ile
290 295 300
Leu Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Thr His His
305 310 315 320
Leu Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr Lys Ala
325 330 335
Ile Ile Pro Ile Leu Gly Asp Tyr Tyr Gln Phe Asp Gly Thr Ser Val

340 345 350

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

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

分别具有GenBank登录号ABB05230.1、ABA02173.1、ABA02172.1、ADM48789.1以及AAL68983.1的向日葵12去饱和酶

<210> 185

<211> 382

<212> PRT

<213> 向日葵

<400> 185

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

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

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

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

Ile Phe Tyr Tyr Leu Ala Asn Asn Tyr Ile Pro Leu Leu Pro Asn Ser
65 70 75 80

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

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

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

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

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

Leu Lys Ser Ser Val Arg Ser Thr Ala Lys Tyr Leu Asn Asn Pro Pro
165 170 175

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

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

His Phe Asp Pro Asn Ser Pro Ile Tyr Ser Asn Arg Glu Arg Ala Gln

210 215 220
Ile Phe Ile Ser Asp Ala Gly Ile Leu Thr Val Phe Tyr Ile Leu Phe
225 230 235 240
Arg Leu Ala Ser Thr Lys Gly Leu Val Trp Val Leu Thr Met Tyr Gly
245 250 255
Gly Pro Leu Leu Val Val Asn Gly Phe Leu Val Leu Ile Thr Phe Leu
260 265 270
Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Thr Glu Trp Asp
275 280 285
Trp Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asp Tyr Gly Ile Leu
290 295 300
Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Thr His His Leu
305 310 315 320
Phe Ser Thr Met Pro His Tyr His Ala Met Glu Ala Thr Lys Ala Ile
325 330 335
Lys Pro Ile Leu Gly Asp Tyr Tyr Gln Phe Asp Gly Thr Ser Ile Phe
340 345 350
Lys Ala Met Tyr Arg Glu Thr Lys Glu Cys Ile Tyr Val Asp Lys Asp
355 360 365
Glu Asp Val Lys Asp Gly Val Tyr Trp Tyr Arg Asn Lys Ile
370 375 380
Cuphea wrightii FatB1硫酯酶;GenBank登录号U56103
<210> 186
<211> 397
<212> PRT
<213> Cuphea wrightii
<400> 186
Met Val Ala Ala Ala Ala Ser Ser Ala Phe Phe Ser Val Pro Thr Pro
1 5 10 15
Gly Thr Ser Pro Lys Pro Gly Lys Phe Gly Asn Trp Pro Ser Ser Leu
20 25 30
Ser Val Pro Phe Lys Pro Asp Asn Gly Gly Phe Val Lys Ala Asn Ala
35 40 45
Ser Ala His Pro Lys Ala Asn Gly Ser Ala Val Asn Leu Lys Ser Gly
50 55 60
Ser Leu Glu Thr Pro Pro Arg Ser Phe Ile Asn Gln Leu Pro Asp Leu
65 70 75 80
Ser Met Leu Leu Ser Lys Ile Thr Thr Val Phe Gly Ala Ala Glu Lys
85 90 95

Gln Trp Lys Arg Pro Gly Met Leu Val Glu Pro Phe Gly Val Asp Arg
100 105 110
Ile Phe Gln Asp Gly Val Phe Phe Arg Gln Ser Phe Ser Ile Arg Ser
115 120 125
Tyr Glu Ile Gly Val Asp Arg Thr Ala Ser Ile Glu Thr Leu Met Asn
130 135 140
Ile Phe Gln Glu Thr Ser Leu Asn His Cys Lys Ser Ile Gly Leu Leu
145 150 155 160
Asn Asp Gly Phe Gly Arg Thr Pro Glu Met Cys Lys Arg Asp Leu Ile
165 170 175
Trp Val Val Thr Lys Ile Gln Val Glu Val Asn Arg Tyr Pro Thr Trp
180 185 190
Gly Asp Thr Ile Glu Val Asn Thr Trp Val Ser Glu Ser Gly Lys Asn
195 200 205
Gly Met Gly Arg Asp Trp Leu Ile Ser Asp Cys Arg Thr Gly Glu Ile
210 215 220
Leu Ile Arg Ala Thr Ser Val Trp Ala Met Met Asn Gln Asn Thr Arg
225 230 235 240
Arg Leu Ser Lys Phe Pro Tyr Glu Val Arg Gln Glu Ile Ala Pro His
245 250 255
Phe Val Asp Ser Ala Pro Val Ile Glu Asp Asp Arg Lys Leu His Lys
260 265 270
Leu Asp Val Lys Thr Gly Asp Ser Ile Arg Asp Gly Leu Thr Pro Arg
275 280 285
Trp Asn Asp Leu Asp Val Asn Gln His Val Asn Asn Val Lys Tyr Ile
290 295 300
Gly Trp Ile Leu Lys Ser Val Pro Ile Glu Val Phe Glu Thr Gln Glu
305 310 315 320
Leu Cys Gly Val Thr Leu Glu Tyr Arg Arg Glu Cys Gly Arg Asp Ser
325 330 335
Val Leu Glu Ser Val Thr Thr Met Asp Pro Ala Lys Glu Gly Asp Arg
340 345 350
Cys Val Tyr Gln His Leu Leu Arg Leu Glu Asp Gly Ala Asp Ile Thr
355 360 365
Ile Gly Arg Thr Glu Trp Arg Pro Lys Asn Ala Gly Ala Asn Gly Ala
370 375 380
Ile Ser Ser Gly Lys Thr Ser Asn Gly Asn Ser Val Ser
385 390 395

Cuphea wrightii FatB2硫酯酶;GenBank登录号U56104

<210> 187
<211> 408
<212> PRT
<213> Cuphea wrightii
<400> 187

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

Ile Ala Pro His Phe Val Asp Ala Pro Pro Val Ile Glu Asp Asn Asp
275 280 285
Arg Lys Leu His Lys Phe Asp Val Lys Thr Gly Asp Ser Ile Cys Lys
290 295 300
Gly Leu Thr Pro Gly Trp Asn Asp Phe Asp Val Asn Gln His Val Ser
305 310 315 320
Asn Val Lys Tyr Ile Gly Trp Ile Leu Glu Ser Met Pro Thr Glu Val
325 330 335
Leu Glu Thr Gln Glu Leu Cys Ser Leu Thr Leu Glu Tyr Arg Arg Glu
340 345 350
Cys Gly Arg Glu Ser Val Val Glu Ser Val Thr Ser Met Asn Pro Ser
355 360 365
Lys Val Gly Asp Arg Ser Gln Tyr Gln His Leu Leu Arg Leu Glu Asp
370 375 380
Gly Ala Asp Ile Met Lys Gly Arg Thr Glu Trp Arg Pro Lys Asn Ala
385 390 395 400
Gly Thr Asn Arg Ala Ile Ser Thr
405

转化构建体,6S-CrbTub_suc2_nr::CrbTub_CwFatBl_nr-6S,指定为构建体1

<210> 188
<211> 5975
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 188

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60
ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120
tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180
ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240
gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300
ggaagacagg tgagggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360
atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420
cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480
gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540
ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600
ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660
cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720
ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900
aaagacatta tagcgagcta ccaaagcata ttcaaacacc tagatcacta ccacttctac 960
acaggccact cgagcttgtg atcgcactcc gctaaggggg cgcctcttcc cttcgtttca 1020
gtcacaaccc gcaaacggcg cgccatatca atgctgctgc aggccttcct gttcctgctg 1080
gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg 1140
gtgcacttca cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag 1200
aaggacgcca agtggcacct gtacttccag tacaacccga acgacaccgt ctgggggacg 1260
cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc 1320
gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac 1380
aacaacacct ccggcttctt caacgacacc atcgacccgc gccagcgctg cgtggccatc 1440
tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc 1500
tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc 1560
gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc 1620
caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc 1680
gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc 1740
cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc 1800
gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg cacccacttc 1860
gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag 1920
accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcgtg ggcctccaac 1980
tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc 2040
aagttctccc tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag 2100
gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc 2160
acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag 2220
ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac 2280
ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag 2340
gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag 2400
aacccctact tcaccaaccg catgagcgtg aacaaccagc ccttcaagag cgagaacgac 2460
ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac 2520
gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc 2580
gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag 2640
gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg 2700
tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt 2760
atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc 2820
ttgtgctatt tgcgaatacc acccccagca tccccttccc tcgtttcata tcgcttgcat 2880
cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc 2940
actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg 3000
taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga 3060
tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgcctct gtcgcacctc 3120

agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg 3180
aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga 3240
gctgatggtc gaaacgttca cagcctaggg atatcgaatt cctttcttgc gctatgacac 3300
ttccagcaaa aggtagggcg ggctgcgaga cggcttcccg gcgctgcatg caacaccgat 3360
gatgcttcga ccccccgaag ctccttcggg gctgcatggg cgctccgatg ccgctccagg 3420
gcgagcgctg tttaaatagc caggcccccg attgcaaaga cattatagcg agctaccaaa 3480
gccatattca aacacctaga tcactaccac ttctacacag gccactcgag cttgtgatcg 3540
cactccgcta agggggcgcc tcttcctctt cgtttcagtc acaacccgca aacactagta 3600
tggtggccgc cgccgccagc agcgccttct tcagcgtgcc caccccggc accagcccca 3660
agcccggcaa gttcggcaac tggcccagca gcctgagcgt gcccttcaag cccgacaacg 3720
gcggcttcca cgtgaaggcc aacgccagcg cccacgggcg cgcccccaag gccaacggca 3780
gcgccgtgaa cctgaagtcc ggcagcctgg agacccccccc ccgcagcttc atcaaccagc 3840
tgcccgacct gagcatgctg ctgagcaaga tcaccaccgt gttcggcgcc gccgagaagc 3900
agtggaagcg ccccggcatg ctggtggagc ccttcggcgt ggaccgcatc ttccaggacg 3960
gcgtgttctt ccgccagagc ttcagcatcc gcagctacga gatcggcgtg gaccgcaccg 4020
ccagcatcga gaccctgatg aacatcttcc aggagaccag cctgaaccac tgcaagagca 4080
tcggcctgct gaacgacggc ttcggccgca cccccgagat gtgcaagcgc gacctgatct 4140
gggtggtgac caagatccag gtggaggtga accgctaccc cacctggggc gacaccatcg 4200
aggtgaacac ctgggtgagc gagagcggca agaacggcat gggccgcgac tggctgatca 4260
gcgactgccg caccggcgag atcctgatcc gcgccaccag cgtgtgggcc atgatgaacc 4320
agaacacccg ccgcctgagc aagttcccct acgaggtgcg ccaggagatc gccccccact 4380
tcgtggacag cgcccccgtg atcgaggacg accgcaagct gcacaagctg gacgtgaaga 4440
ccggcgacag catccgcgac ggcctgaccc cccgctggaa cgacctggac gtgaaccagc 4500
acgtgaacaa cgtgaagtac atcggctgga ttctgaagtc cgtgcccatc gaggtgttcg 4560
agacccagga gctgtgcggc gtgaccctgg agtaccgccg cgagtgcggc cgcgacagcg 4620
tgctggagag cgtgaccacc atggaccccg ccaaggaggg cgaccgctgc gtgtaccagc 4680
acctgctgcg cctggaggac ggcgccgaca tcaccatcgg ccgcaccgag tggcgcccca 4740
agaacgccgg cgccaacggc gccatcagca gcggcaagac cagcaacggc aacagcgtga 4800
gctgattaat taactcgagg cagcagcagc tcggatagta tcgacacact ctggacgctg 4860
gtcgtgtgat ggactgttgc cgccacactt gctgccttga cctgtgaata tccctgccgc 4920
ttttatcaaa cagcctcagt gtgtttgatc ttgtgtgtac gcgcttttgc gagttgctag 4980
ctgcttgtgc tatttgcgaa taccaccccc agcatcccct tccctcgttt catatcgctt 5040
gcatcccaac cgcaacttat ctacgctgtc ctgctatccc tcagcgctgc tcctgctcct 5100
gctcactgcc cctcgcacag ccttggtttg ggctccgcct gtattctcct ggtactgcaa 5160
cctgtaaacc agcactgcaa tgctgatgca cgggaagtag tgggatggga acacaaatgg 5220
aaagcttgag ctcttgtttt ccagaaggag ttgctccttg agcctttcat tctcagcctc 5280
gataacctcc aaagccgctc taattgtgga gggggttcga atttaaaagc ttggaatgtt 5340
ggttcgtgcg tctggaacaa gcccagactt gttgctcact gggaaaagga ccatcagctc 5400
caaaaaactt gccgctcaaa ccgcgtacct ctgctttcgc gcaatctgcc ctgttgaaat 5460

cgccaccaca ttcatattgt gacgcttgag cagtctgtaa ttgcctcaga atgtggaatc 5520
atctgccccc tgtgcgagcc catgccaggc atgtcgcggg cgaggacacc cgccactcgt 5580
acagcagacc attatgctac ctcacaatag ttcataacag tgaccatatt tctcgaagct 5640
ccccaacgag cacctccatg ctctgagtgg ccacccccg gccctggtgc ttgcggaggg 5700
caggtcaacc ggcatggggc taccgaaatc cccgaccgga tcccaccacc cccgcgatgg 5760
gaagaatctc tccccgggat gtgggcccac caccagcaca acctgctggc ccaggcgagc 5820
gtcaaaccat accacacaaa tatccttggc atcggccctg aattccttct gccgctctgc 5880
tacccggtgc ttctgtccga agcaggggtt gctagggatc gctccgagtc cgcaaaccct 5940
tgtcgcgtgg cggggcttgt tcgagcttga agagc 5975

转化构建体,6S-CrbTub_suc2_nr::CrbTub_CwFatB2_nr-6S,指定为构建体2

<210> 189

<211> 1256

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 189

actagtatgg tggtggccgc cgccgccagc agcgccttct tccccgtgcc cgcccccccgc 60
cccacccccа agcccggcaa gttcggcaac tggcccagca gcctgagcca gcccttcaag 120
cccaagagca accccaacgg ccgcttccag gtgaaggcca acgtgagccc ccacgggcgc 180
gcccccaagg ccaacggcag cgccgtgagc ctgaagtccg gcagcctgaa caccctggag 240
gacccccccа gcagcccccc cccccgcacc ttcctgaacc agctgcccga ctggagccgc 300
ctgcgcaccg ccatcaccac cgtgttcgtg gccgccgaga agcagttcac ccgcctggac 360
cgcaagagca agcgccccga catgctggtg gactggttcg gcagcgagac catcgtgcag 420
gacggcctgg tgttccgcga gcgcttcagc atccgcagct acgagatcgg cgccgaccgc 480
accgccagca tcgagaccct gatgaaccac ctgcaggaca ccagcctgaa ccactgcaag 540
agcgtgggcc tgctgaacga cggcttcggc cgcacccccg agatgtgcac ccgcgacctg 600
atctgggtgc tgaccaagat gcagatcgtg gtgaaccgct accccacctg gggcgacacc 660
gtggagatca acagctggtt cagccagagc ggcaagatcg gcatgggccg cgagtggctg 720
atcagcgact gcaacaccgg cgagatcctg gtgcgcgcca ccagcgcctg ggccatgatg 780
aaccagaaga cccgccgctt cagcaagctg ccctgcgagg tgcgccagga gatcgccccc 840
cacttcgtgg acgccccccc cgtgatcgag gacaacgacc gcaagctgca caagttcgac 900
gtgaagaccg gcgacagcat ctgcaagggc ctgacccccg gctggaacga cttcgacgtg 960
aaccagcacg tgagcaacgt gaagtacatc ggctggattc tggagagcat gcccaccgag 1020
gtgctggaga cccaggagct gtgcagcctg accctggagt accgccgcga gtgcggccgc 1080
gagagcgtgg tggagagcgt gaccagcatg aaccccagca aggtgggcga ccgcagccag 1140
taccagcacc tgctgcgcct ggaggacggc gccgacatca tgaagggccg caccgagtgg 1200
cgccccaaga acgccggcac caaccgcgcc atcagcacct gattaattaa ctcgag 1256

转化构建体6S-CrbTub_suc2_nr::Amt3_CwFatBl_nr-6S,指定为构建体3,以及6S-

CrbTub_suc2_nr::Amt3_CwFatB2_nr-6S,指定为构建体4

<210> 190

<211> 1370

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 190

gaattcggcg ctatgacact tccagcaaaa ggtagggcgg gctgcgagac ggcttcccgg 60
cgctgcatgc aacaccgatg atgcttcgac cccccgaagc tccttcgggg ctgcatgggc 120
gctccgatgc cgctccaggg cgagcgctgt ttaaatagcc aggccccccga ttgcaaagac 180
attatagcga gctaccaaag ccatattcaa acacctagat cactaccact tctacacagg 240
ccactcgagc ttgtgatcgc actccgctaa gggggcgcct cttcctcttc gtttcagtcg 300
gccgacagga cgcgcgtcaa aggtgctggt cgtgtatgcc ctggccggca ggtcgttgct 360
gctgctggtt agtgattccg caaccctgat tttggcgtct tattttggcg tggcaaacgc 420
tggcgcccgc gagccgggcc ggcggcgatg cggtgcccca cggctgccgg aatccaaggg 480
aggcaagagc gcccgggtca gttgaagggc tttacgcgca aggtacagcc gctcctgcaa 540
ggctgcgtgg tggaattgga cgtgcaggtc ctgctgaagt tcctccaccg cctcaccagc 600
ggacaaagca ccggtgtatc aggtccgtgt catccactct aaagagctcg actacgacct 660
actgatggcc ctagattctt catcaaaaac gcctgagaca cttgcccagg attgaaactc 720
cctgaaggga ccaccagggg ccctgagttg ttccttcccc ccgtggcgag ctgccagcca 780
ggctgtacct gtgatcgagg ctggcgggaa aataggcttc gtgtgctcag gtcatgggag 840
gtgcaggaca gctcatgaaa cgccaacaat cgcacaattc atgtcaagct aatcagctat 900
ttcctcttca cgagctgtaa ttgtcccaaa attctggtct accgggggtg atccttcgtg 960
tacgggccct tccctcaacc ctaggtatgc gcgcatgcgg tcgccgcgca actcgcgcga 1020
gggccgaggg tttgggacgg gccgtcccga aatgcagttg cacccggatg cgtggcacct 1080
tttttgcgat aatttatgca atggactgct ctgcaaaatt ctggctctgt cgccaaccct 1140
aggatcagcg gcgtaggatt tcgtaatcat tcgtcctgat ggggagctac cgactaccct 1200
aatatcagcc cgactgcctg acgccagcgt ccacttttgt gcacacattc cattcgtgcc 1260
caagacattt cattgtggtg cgaagcgtcc ccagttacgc tcacctgttt cccgacctcc 1320
ttactgttct gtcgacagag cgggcccaca ggccggtcgc agccactagt 1370

TP1(UTEX250 硬脂酰ACP去饱和酶转运肽序列)

<210> 191

<211> 120

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 191

actagtatgg ccaccgcatc cactttctcg gcgttcaatg cccgctgcgg cgacctgcgt 60
cgctcggcgg gctccgggcc ccggcgccca gcgaggcccc tccccgtgcg cgggcgcgcc 120
TP2(UTEX1435 硬脂酰ACP去饱和酶转运肽序列)
<210> 192
<211> 114
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 192
actagtatgg cttccgcggc attcaccatg tcggcgtgcc ccgcgatgac tggcagggcc 60
cctggggcac gtcgctccgg acggccagtc gccacccgcc tgaggggcg cgcc 114
TP3(UTEX1435 Δ12脂肪酸去饱和酶转运肽序列)
<210> 193
<211> 120
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 193
actagtatgg ctatcaagac gaacaggcag cctgtggaga agcctccgtt cacgatcggg 60
acgctgcgca aggccatccc cgcgcactgt ttcgagcgct cggcgcttcg tgggcgcgcc 120
TP4 (UTEX1435 异戊烯二磷酸合成酶转运肽序列)
<210> 194
<211> 117
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 194
actagtatga cgttcggggt cgccctcccg gccatgggcc gcggtgtctc ccttccccgg 60
cccagggtcg cggtgcgcgc ccagtcggcg agtcaggttt tggagagcgg gcgcgcc 117
红花油酰基-酰基载体蛋白硫酯酶,GenBank登录号AAA33019.1
<210> 195
<211> 385
<212> PRT
<213> 红花
<400> 195
Met Leu Ser Lys Gly Ala Pro Ala Ala Pro Ala Val Ala Ala Met Tyr

1 5 10 15
Asn Ala Ser Ala Lys Asp Thr Thr Phe Ala Leu Thr His Ser Arg Ser
20 25 30
Ile Gly Ser Val Ser Ile Arg Arg Arg Tyr Asn Val Phe Leu Cys Asn
35 40 45
Ser Ser Ser Ser Ser Arg Lys Val Ser Pro Leu Leu Ala Val Ala Thr
50 55 60
Gly Glu Gln Pro Ser Gly Val Ala Ser Leu Arg Glu Ala Asp Lys Glu
65 70 75 80
Lys Ser Leu Gly Asn Arg Leu Arg Leu Gly Ser Leu Thr Glu Asp Gly
85 90 95
Leu Ser Tyr Lys Glu Lys Phe Val Ile Arg Cys Tyr Glu Val Gly Ile
100 105 110
Asn Lys Thr Ala Thr Ile Glu Thr Ile Ala Asn Leu Leu Gln Glu Val
115 120 125
Gly Gly Asn His Ala Gln Gly Val Gly Phe Ser Thr Asp Gly Phe Ala
130 135 140
Thr Thr Thr Thr Met Arg Lys Leu His Leu Ile Trp Val Thr Ala Arg
145 150 155 160
Met His Ile Glu Ile Tyr Arg Tyr Pro Ala Trp Ser Asp Val Ile Glu
165 170 175
Ile Glu Thr Trp Val Gln Gly Glu Gly Lys Val Gly Thr Arg Arg Asp
180 185 190
Trp Ile Leu Lys Asp Tyr Ala Asn Gly Glu Val Ile Gly Arg Ala Thr
195 200 205
Ser Lys Trp Val Met Met Asn Glu Asp Thr Arg Arg Leu Gln Lys Val
210 215 220
Ser Asp Asp Val Arg Glu Glu Tyr Leu Val Phe Cys Pro Arg Thr Leu
225 230 235 240
Arg Leu Ala Phe Pro Glu Glu Asn Asn Asn Ser Met Lys Lys Ile Pro
245 250 255
Lys Leu Glu Asp Pro Ala Glu Tyr Ser Arg Leu Gly Leu Val Pro Arg
260 265 270
Arg Ser Asp Leu Asp Met Asn Lys His Val Asn Asn Val Thr Tyr Ile
275 280 285
Gly Trp Ala Leu Glu Ser Ile Pro Pro Glu Ile Ile Asp Thr His Glu
290 295 300
Leu Gln Ala Ile Thr Leu Asp Tyr Arg Arg Glu Cys Gln Arg Asp Asp
305 310 315 320

Ile Val Asp Ser Leu Thr Ser Arg Glu Pro Leu Gly Asn Ala Ala Gly
325 330 335
Val Lys Phe Lys Glu Ile Asn Gly Ser Val Ser Pro Lys Lys Asp Glu
340 345 350
Gln Asp Leu Ser Arg Phe Met His Leu Leu Arg Ser Ala Gly Ser Gly
355 360 365
Leu Glu Ile Asn Arg Cys Arg Thr Glu Trp Arg Lys Lys Pro Ala Lys
370 375 380
Arg
385

蓖麻硬脂酰-酰基载体蛋白去饱和酶,GenBank登录号AACG59946.1

<210> 196
<211> 395
<212> PRT
<213> 蓖麻
<400> 196

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

180 185 190
Asp Met Lys His Ile Glu Lys Thr Ile Gln Tyr Leu Ile Gly Ser Gly
195 200 205
Met Asp Pro Lys Thr Glu Asn Asn Pro Tyr Leu Gly Phe Ile Tyr Thr
210 215 220
Ser Phe Gln Glu Arg Ala Thr Phe Ile Ser His Gly Asn Thr Ala Arg
225 230 235 240
Leu Ala Lys Glu His Gly Asp Met Lys Leu Ala Gln Ile Cys Gly Ile
245 250 255
Ile Ala Ala Asp Glu Lys Arg His Glu Thr Ala Tyr Thr Lys Ile Ala
260 265 270
Glu Lys Leu Phe Glu Ile Asp Pro Asp Gly Thr Val Met Ala Leu Ala
275 280 285
Asp Met Met Arg Lys Lys Ile Ser Met Pro Ala His Leu Met Phe Asp
290 295 300
Gly Gln Asp Asp Asn Leu Phe Asp His Tyr Ser Ser Val Ala Gln Arg
305 310 315 320
Ile Gly Val Tyr Thr Ala Lys Asp Tyr Ala Asp Ile Leu Glu Phe Leu
325 330 335
Val Gly Arg Trp Lys Val Glu Asn Leu Ile Gly Leu Ser Gly Glu Gly
340 345 350
Arg Arg Ala Gln Asp Phe Val Cys Gly Leu Pro Gly Arg Ile Arg Arg
355 360 365
Leu Glu Glu Arg Ala Gln Gly Arg Ala Lys Gln Ala Ala Thr Val Pro
370 375 380
Phe Ser Trp Ile Phe Gly Arg Glu Leu Lys Leu
385 390 395

向日葵硬脂酰-ACP去饱和酶,Genbank登录号AAB65145.1

<210> 197
<211> 396
<212> PRT
<213> 向日葵
<400> 197

Met Ala Ile Arg Ile Asn Thr Ala Thr Phe Gln Ser Asp Leu Tyr Arg
1 5 10 15
Ser Phe Ala Phe Pro Gln Pro Lys Pro Leu Arg Ser Pro Lys Phe Ala
20 25 30
Met Ala Ser Thr Ile Gly Ser Ala Thr Thr Lys Val Glu Ser Thr Lys
35 40 45

Lys Pro Phe Thr Pro Pro Arg Glu Val His Gln Gln Val Leu His Ser
50 55 60
Met Pro Pro Gln Lys Ile Glu Ile Phe Lys Ser Met Glu Gly Trp Ala
65 70 75 80
Glu Asp Asn Ile Leu Val His Leu Lys Pro Val Glu Lys Cys Trp Gln
85 90 95
Ala Gln Asp Phe Leu Pro Asp Pro Ala Ser Asp Gly Phe Met Glu Gln
100 105 110
Val Glu Glu Leu Arg Ala Arg Ala Lys Glu Ile Pro Asp Asp Tyr Phe
115 120 125
Val Val Leu Val Gly Asp Met Ile Thr Glu Glu Ala Leu Pro Thr Tyr
130 135 140
Gln Thr Met Leu Asn Thr Leu Asp Gly Val Arg Asp Glu Thr Gly Ala
145 150 155 160
Ser Pro Thr Ser Trp Ala Ile Trp Thr Arg Ala Trp Thr Ala Glu Glu
165 170 175
Asn Arg His Gly Asp Leu Leu His Gln Tyr Leu Tyr Leu Ser Gly Arg
180 185 190
Val Asp Met Arg Gln Ile Gln Lys Thr Ile Gln Tyr Leu Ile Gly Ser
195 200 205
Gly Met Asp Pro Arg Thr Glu Asn Ser Pro Tyr Leu Gly Phe Ile Tyr
210 215 220
Thr Ser Phe Gln Glu Arg Ala Thr Phe Ile Ser His Gly Asn Thr Ala
225 230 235 240
Arg His Ala Lys Glu His Gly Asp Val Lys Leu Ala Gln Met Cys Gly
245 250 255
Ile Ile Ala Ala Asp Glu Lys Arg His Glu Thr Ala Tyr Thr Lys Ile
260 265 270
Val Glu Lys Leu Phe Glu Ile Asp Pro Asp Gly Thr Val Leu Ala Phe
275 280 285
Ala Asp Met Met Arg Lys Lys Ile Ser Met Pro Ala His Leu Met Tyr
290 295 300
Asp Gly Arg Asp Asp Asn Leu Phe Glu Asn Phe Ser Ala Val Ala Gln
305 310 315 320
Arg Leu Gly Val Tyr Thr Ala Lys Asp Tyr Ala Asp Ile Leu Glu Phe
325 330 335
Leu Val Gly Arg Trp Lys Val Ala Asp Leu Thr Gly Leu Ser Gly Glu
340 345 350
Gly Arg Lys Ala Gln Asp Tyr Val Cys Gly Leu Ala Pro Arg Ile Arg

355 360 365
Arg Leu Glu Glu Arg Asn Ser Ala Arg Ala Lys Glu Ser Val Asn Val
370 375 380
Pro Phe Ser Trp Ile Phe Asp Arg Glu Val Lys Leu
385 390 395

芥菜质体Δ-9-硬脂酰-酰基载体蛋白去饱和酶,GenBank登录号AAD40245.1

<210> 198
<211> 401
<212> PRT
<213> 芥菜
<400> 198

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

Leu Gly Phe Ile Tyr Thr Ser Phe Gln Glu Arg Ala Thr Phe Val Ser
225 230 235 240
His Gly Asn Thr Ala Arg Gln Ala Lys Glu His Gly Asp Leu Lys Leu
245 250 255
Ala Gln Ile Cys Gly Thr Ile Ala Ala Asp Glu Lys Arg His Glu Thr
260 265 270
Ala Tyr Thr Lys Ile Val Glu Lys Leu Leu Glu Ile Asp Pro Asp Gly
275 280 285
Thr Val Val Ala Phe Ala Asp Met Met Arg Lys Lys Ile Ser Met Pro
290 295 300
Ala His Leu Met Tyr Asp Gly Arg Asp Asp Asn Leu Phe Asp Asn Phe
305 310 315 320
Ser Ser Val Ala Gln Arg Leu Gly Val Tyr Thr Ala Lys Asp Tyr Ala
325 330 335
Asp Ile Leu Glu Phe Leu Ala Gly Arg Trp Lys Ile Glu Ser Leu Ser
340 345 350
Gly Leu Ser Gly Glu Gly Asn Lys Ala Gln Glu Tyr Leu Cys Gly Leu
355 360 365
Thr Pro Arg Ile Arg Arg Leu Asp Glu Arg Ala Gln Ala Arg Ala Lys
370 375 380
Lys Gly Pro Lys Ile Pro Phe Ser Trp Ile His Asp Arg Glu Val Gln
385 390 395 400
Leu

大豆硬脂酰-ACP去饱和酶,Genbank登录号ACJ39209.1

<210> 199
<211> 391
<212> PRT
<213> 大豆
<400> 199

Met Ala Leu Arg Leu Asn Pro Ile Pro Thr Gln Thr Phe Ser Leu Pro
1 5 10 15
Gln Met Ala Ser Leu Arg Ser Pro Arg Phe Arg Met Ala Ser Thr Leu
20 25 30
Arg Ser Gly Ser Lys Glu Val Glu Asn Ile Lys Lys Pro Phe Thr Pro
35 40 45
Pro Arg Glu Val His Val Gln Val Thr His Ser Met Pro Pro Gln Lys
50 55 60
Ile Glu Ile Phe Lys Ser Leu Glu Asp Trp Ala Glu Gln Asn Ile Leu
65 70 75 80

Thr His Leu Lys Pro Val Glu Lys Cys Trp Gln Pro Gln Asp Phe Leu
85 90 95
Pro Asp Pro Ser Ser Asp Gly Phe Glu Glu Gln Val Lys Glu Leu Arg
100 105 110
Glu Arg Ala Lys Glu Leu Pro Asp Asp Tyr Phe Val Val Leu Val Gly
115 120 125
Asp Met Ile Thr Glu Glu Ala Leu Pro Thr Tyr Gln Thr Met Leu Asn
130 135 140
Thr Leu Asp Gly Val Arg Asp Glu Thr Gly Ala Ser Leu Thr Ser Trp
145 150 155 160
Ala Ile Trp Thr Arg Ala Trp Thr Ala Glu Glu Asn Arg His Gly Asp
165 170 175
Leu Leu Asn Lys Tyr Leu Tyr Leu Ser Gly Arg Val Asp Met Lys Gln
180 185 190
Ile Glu Lys Thr Ile Gln Tyr Leu Ile Gly Ser Gly Met Asp Pro Arg
195 200 205
Thr Glu Asn Ser Pro Tyr Leu Gly Phe Ile Tyr Thr Ser Phe Gln Glu
210 215 220
Arg Ala Thr Phe Ile Ser His Gly Asn Thr Ala Arg Leu Ala Lys Glu
225 230 235 240
His Gly Asp Ile Lys Leu Ala Gln Ile Cys Gly Met Ile Ala Ser Asp
245 250 255
Glu Lys Arg His Glu Thr Ala Tyr Thr Lys Ile Val Glu Lys Leu Phe
260 265 270
Glu Val Asp Pro Asp Gly Thr Val Met Ala Phe Ala Asp Met Met Arg
275 280 285
Lys Lys Ile Ala Met Pro Ala His Leu Met Tyr Asp Gly Arg Asp Asp
290 295 300
Asn Leu Phe Asp Ser Tyr Ser Ser Val Ala Gln Arg Ile Gly Val Tyr
305 310 315 320
Thr Ala Lys Asp Tyr Ala Asp Ile Leu Glu Phe Leu Val Gly Arg Trp
325 330 335
Lys Val Glu Gln Leu Thr Gly Leu Ser Gly Glu Gly Arg Lys Ala Gln
340 345 350
Glu Tyr Ile Cys Gly Leu Pro Pro Arg Ile Arg Arg Leu Glu Glu Arg
355 360 365
Ala Gln Ala Arg Val Lys Glu Ser Ser Thr Leu Lys Phe Ser Trp Ile
370 375 380
His Asp Arg Glu Val Leu Leu

385 390

油橄榄硬脂酰-ACP去饱和酶,Genbank登录号AAB67840.1

<210> 200

<211> 390

<212> PRT

<213> 油橄榄

<400> 200

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

Lys Arg His Glu Thr Ala Tyr Thr Lys Ile Val Glu Lys Leu Phe Glu
260 265 270
Ile Asp Pro Asp Gly Thr Val Leu Ala Leu Ala Asp Met Met Arg Lys
275 280 285
Lys Val Ser Met Pro Ala His Leu Met Tyr Asp Gly Gln Asp Asp Asn
290 295 300
Leu Phe Glu Asn Phe Ser Ser Val Ala Gln Arg Leu Gly Val Tyr Thr
305 310 315 320
Ala Lys Asp Tyr Ala Asp Ile Leu Glu Phe Leu Val Gly Arg Trp Asp
325 330 335
Ile Glu Lys Leu Thr Gly Leu Ser Gly Glu Gly Arg Lys Ala Gln Asp
340 345 350
Tyr Val Cys Thr Leu Pro Pro Arg Ile Arg Arg Leu Glu Glu Arg Ala
355 360 365
Gln Ser Arg Val Lys Lys Ala Ser Ala Thr Pro Phe Ser Trp Ile Phe
370 375 380
Gly Arg Glu Ile Asn Leu
385 390

油桐硬脂酰-酰基载体蛋白去饱和酶,GenBank登录号ADC32803.1

<210> 201
<211> 392
<212> PRT
<213> 油桐
<400> 201

Met Ala Leu Lys Leu Asn Pro Phe Ile Ser Gln Ser His Lys Phe Pro
1 5 10 15
Thr Phe Ala Leu Pro Pro Met Ala Asn Leu Arg Ser Pro Lys Phe Tyr
20 25 30
Met Ala Ser Thr Leu Arg Glu Ile Glu His Leu Lys Lys Pro Phe Met
35 40 45
Pro Pro Arg Glu Val His Val Gln Val Thr His Ser Met Pro Pro Gln
50 55 60
Lys Ile Glu Ile Phe Lys Ser Leu Glu Asp Trp Ala Glu Gln Asn Ile
65 70 75 80
Leu Val His Leu Lys Pro Val Glu Lys Cys Trp Gln Pro Gln Asp Phe
85 90 95
Leu Pro Asp Pro Ala Ser Asp Asp Phe Asp Glu Gln Val Lys Glu Leu
100 105 110
Arg Glu Arg Ala Lys Glu Ile Pro Asp Asp Tyr Phe Val Val Leu Val

115 120 125

Gly Asp Met Ile Thr Glu Glu Ala Leu Pro Thr Tyr Gln Thr Met Leu
130 135 140

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

Trp Ala Ile Trp Thr Arg Ala Trp Thr Ala Glu Glu Asn Arg His Gly
165 170 175

Asp Leu Leu Asn Lys Tyr Leu Tyr Leu Ser Gly Arg Val Asp Met Arg
180 185 190

Gln Ile Glu Lys Thr Ile Gln Tyr Leu Ile Gly Ser Gly Met Asp Pro
195 200 205

Arg Thr Glu Asn Ser Pro Tyr Leu Gly Phe Ile Tyr Thr Ser Phe Gln
210 215 220

Glu Arg Ala Thr Phe Ile Ser His Gly Asn Thr Ala Arg His Ala Lys
225 230 235 240

Glu His Gly Asp Ile Lys Leu Ala Gln Ile Cys Gly Thr Ile Ala Ser
245 250 255

Asp Glu Lys Arg His Glu Thr Ala Tyr Thr Lys Ile Val Glu Lys Leu
260 265 270

Phe Glu Ile Asp Pro Asn Gly Thr Val Leu Ala Phe Ala Asp Met Met
275 280 285

Arg Lys Lys Ile Ser Met Pro Ala His Leu Met Tyr Asp Gly Cys Asn
290 295 300

Asp Asn Leu Phe Asp His Phe Ser Ala Val Ala Gln Arg Leu Gly Val
305 310 315 320

Tyr Thr Ala Lys Asp Tyr Ala Asp Ile Leu Glu Phe Leu Val Gly Arg
325 330 335

Trp Lys Val Asp Lys Leu Met Gly Leu Ser Ala Glu Gly Gln Lys Ala
340 345 350

Gln Asp Tyr Val Cys Arg Leu Pro Pro Arg Ile Arg Arg Leu Glu Glu
355 360 365

Arg Ala Gln Gly Arg Ala Lys Glu Ala Thr Ile Val Pro Phe Ser Trp
370 375 380

Ile Phe Glu Arg Glu Val Gln Leu
385 390

向日葵酰基-ACP硫酯酶FATA1,GenBank登录号AAL79361.1

<210> 202

<211> 365

<212> PRT

<213> 向日葵

<400> 202

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

290 295 300

Tyr Arg Arg Glu Cys Gln His Asp Asp Val Val Asp Ser Leu Thr Ser
305 310 315 320

Ser Glu Ser Pro Ala Val Asn Gly Cys Ala Ser Gly Glu Asn Leu Ser
325 330 335

Lys Phe Leu His Leu Leu Arg Ser Ser Gly Glu Gly Leu Glu Leu Asn
340 345 350

Arg Gly Arg Thr Glu Trp Arg Lys Lys Pro Ala Lys Lys
355 360 365

白菜型油菜FATA1,GenBank登录号AAC49002.1

<210> 203

<211> 362

<212> PRT

<213> 白菜型油菜

<400> 203

Met Leu Lys Leu Ser Cys Asn Ala Thr Asp Lys Leu Gln Thr Leu Phe
1 5 10 15

Ser His Ser His Gln Pro Asp Pro Ala His Arg Arg Thr Val Ser Ser
20 25 30

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

Ile Val Ser Ala Asp Gln Gly Ser Val Ile Arg Ala Glu Gln Gly Leu
50 55 60

Gly Ser Leu Ala Asp Gln Leu Arg Leu Gly Ser Leu Thr Glu Asp Gly
65 70 75 80

Leu Ser Tyr Lys Glu Lys Phe Ile Val Arg Ser Tyr Glu Val Gly Ser
85 90 95

Asn Lys Thr Ala Thr Val Glu Thr Val Ala Asn Leu Leu Gln Glu Val
100 105 110

Gly Cys Asn His Ala Gln Ser Val Gly Phe Ser Thr Asp Gly Phe Ala
115 120 125

Thr Thr Pro Thr Met Arg Lys Leu His Leu Ile Trp Val Thr Ala Arg
130 135 140

Met His Ile Glu Ile Tyr Lys Tyr Pro Ala Trp Gly Asp Val Val Glu
145 150 155 160

Ile Glu Thr Trp Cys Gln Ser Glu Gly Arg Ile Gly Thr Arg Arg Asp
165 170 175

Trp Ile Leu Lys Asp Val Ala Thr Gly Glu Val Thr Gly Arg Ala Thr
180 185 190

Ser Lys Trp Val Met Met Asn Gln Asp Thr Arg Arg Leu Gln Lys Val
195 200 205
Ser Asp Asp Val Arg Asp Glu Tyr Leu Val Phe Cys Pro Lys Glu Leu
210 215 220
Arg Leu Ala Phe Pro Glu Glu Asn Asn Arg Ser Leu Lys Lys Ile Pro
225 230 235 240
Lys Leu Glu Asp Pro Ala Gln Tyr Ser Met Ile Gly Leu Lys Pro Arg
245 250 255
Arg Ala Asp Leu Asp Met Asn Gln His Val Asn Asn Val Thr Tyr Ile
260 265 270
Gly Trp Val Leu Glu Ser Ile Pro Gln Glu Ile Val Asp Thr His Glu
275 280 285
Leu Gln Val Ile Thr Leu Asp Tyr Arg Arg Glu Cys Gln Gln Asp Asp
290 295 300
Val Val Asp Ser Leu Thr Thr Thr Thr Ser Glu Ile Gly Gly Thr Asn
305 310 315 320
Gly Ser Ala Ser Ser Gly Thr Gln Gly Gln Asn Asp Ser Gln Phe Leu
325 330 335
His Leu Leu Arg Leu Ser Gly Asp Gly Gln Glu Ile Asn Arg Gly Thr
340 345 350
Thr Leu Trp Arg Lys Lys Pro Ser Asn Leu
355 360

麻疯树酰基-ACP硫酯酶,GenBank登录号ABX82799.3

<210> 204
<211> 369
<212> PRT
<213> 麻疯树
<400> 204

Met Met Leu Lys Val Pro Cys Thr Ala Pro Asp Gln Ile Gln Ser Leu
1 5 10 15
Ala Pro Gln Cys Arg Phe Leu Asn Gln His Arg Ser Ser Ser His Phe
20 25 30
Thr Arg Arg Arg Ala Asn Ala Val Val Tyr Ser Ser Pro Val Ala Lys
35 40 45
Ile Ser Thr Gln Pro Val Val Ser Asp Gly Val Thr Lys Val Glu Ser
50 55 60
Gly Cys Glu Ser Leu Ala Asp Arg Leu Arg Leu Gly His Leu Thr Glu
65 70 75 80
Asp Gly Phe Ser Tyr Lys Glu Lys Phe Ile Val Arg Ser Tyr Glu Val

85 90 95

Gly Ile Asn Lys Thr Ala Thr Val Glu Thr Ile Ala Asn Phe Leu Gln
100 105 110

Glu Val Gly Cys Asn His Ala Gln Ser Val Gly Phe Ser Thr Asp Gly
115 120 125

Phe Ala Thr Thr Pro Thr Met Arg Lys Leu His Leu Ile Trp Val Thr
130 135 140

Ala Arg Met His Ile Glu Ile Tyr Lys Tyr Pro Ala Trp Ser Asp Val
145 150 155 160

Val Glu Ile Glu Thr Trp Cys Gln Ser Glu Gly Arg Ile Gly Thr Arg
165 170 175

Arg Asp Trp Ile Leu Lys Asp Tyr Ala Thr Gly Gln Val Ile Gly Arg
180 185 190

Ala Thr Ser Lys Trp Val Met Met Asn Gln Asp Thr Arg Arg Leu Gln
195 200 205

Lys Val Thr Asp Asp Val Arg Asp Glu Tyr Leu Val Phe Cys Pro Arg
210 215 220

Glu Leu Arg Leu Ala Phe Pro Glu Glu Asn Asn Arg Ser Leu Lys Lys
225 230 235 240

Ile Ser Lys Leu Glu Asp Thr Thr Gln Tyr Ser Lys Leu Gly Leu Val
245 250 255

Pro Arg Arg Ala Asp Leu Asp Met Asn Gln His Val Asn Asn Val Thr
260 265 270

Tyr Ile Gly Trp Val Leu Glu Ser Met Pro Gln Glu Ile Ile Asp Thr
275 280 285

His Glu Leu Gln Thr Ile Thr Leu Asp Tyr Arg Arg Glu Cys Gln His
290 295 300

Asp Asp Val Val Asp Ser Leu Thr Ser Ala Glu Ser Leu Glu Gly Ala
305 310 315 320

Gly Leu Asp Leu His Ala Thr Asn Gly Ser Ala Thr Ala Ile Ala Gly
325 330 335

Glu Gln Asp Ser Arg Asn Phe Leu His Leu Leu Arg Phe Ser Ser Asp
340 345 350

Gly Leu Glu Ile Asn Arg Gly Arg Thr Glu Trp Arg Lys Lys Pro Ser
355 360 365

Arg

玉米油酰基-酰基载体蛋白硫酯酶,GenBank登录号ACG40089.1

〈210〉 205

〈211〉 229

<212> PRT
<213> 玉米
<400> 205

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

玉米油酰基-酰基载体蛋白硫酯酶,GenBank登录号ACG42559.1

<210> 206
<211> 366
<212> PRT
<213> 玉米
<400> 206

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

305 310 315 320
Glu Lys Ser Met Asn Gly Ser Ala Ser Ala Ala Ala Pro His Lys Glu
325 330 335
Glu Arg Gln Gln Phe Leu His Cys Leu Arg Phe Ala Ala Asn Gly His
340 345 350
Glu Ile Asn Arg Gly Arg Thr Val Trp Arg Lys Leu Ala Arg
355 360 365

红花Δ12脂肪酸CtFAD2-3,GenBank登录号ADM48790.1

<210> 207
<211> 381
<212> PRT
<213> 红花
<400> 207

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

Phe Tyr Pro Tyr Ser Pro Ile Phe Asn Asp Ser Glu Arg Phe Gln Ile
210 215 220
Phe Leu Thr Asp Leu Gly Val Phe Gly Thr Leu Tyr Gly Val Tyr Arg
225 230 235 240
Leu Ala Leu Ile Lys Gly Thr Glu Trp Val Ile Asn Phe Tyr Gly Met
245 250 255
Pro Ile Leu Phe Met Ser Gly Ser Phe Ile Leu Leu Thr Tyr Leu His
260 265 270
His Thr His Pro Ser Ile Pro His Tyr Asp Ser Thr Glu Trp Asp Trp
275 280 285
Leu Arg Gly Ala Leu Ala Thr Val Asp Arg Asn Phe Gly Phe Leu Asn
290 295 300
His Ala Phe His Asp Ala Thr Arg Thr His Ala Val His His Leu Phe
305 310 315 320
Pro Thr Ile Pro His Tyr His Thr Phe Glu Ala Arg Gln Ala Val Met
325 330 335
Pro Ile Leu Gly Asp Tyr Tyr Lys Tyr Asp Asp Thr Pro Ile Leu Gln
340 345 350
Ala Ile Trp Arg Glu Thr Lys Asp Cys Thr Phe Ile Glu Pro Glu Glu
355 360 365
Val Asn Gly Glu Lys Lys Gly Ile Tyr Trp Phe Tyr Lys
370 375 380

陆地棉ω-6去饱和酶,GenBank登录号CAA71199.1

<210> 208
<211> 383
<212> PRT
<213> 陆地棉
<400> 208

Met Gly Ala Gly Gly Arg Met Ser Val Pro Thr Ser Pro Lys Lys Pro
1 5 10 15
Glu Phe Asn Ser Leu Lys Arg Val Pro Tyr Ser Lys Pro Pro Phe Thr
20 25 30
Leu Ser Glu Ile Lys Lys Ala Ile Pro Pro His Cys Phe Gln Arg Ser
35 40 45
Val Leu Arg Ser Phe Ser Tyr Leu Leu Tyr Asp Phe Ile Leu Ala Ser
50 55 60
Leu Phe Tyr His Val Ala Thr Asn Tyr Phe Pro Asn Leu Pro Gln Ala
65 70 75 80
Leu Ser Asn Val Ala Trp Pro Leu Tyr Trp Ala Met Gln Gly Cys Ile

85 90 95
Leu Thr Gly Val Trp Val Ile Ala His Glu Cys Gly His His Ala Phe
100 105 110
Ser Asp Tyr Gln Trp Leu Asp Asp Thr Val Gly Leu Ile Leu His Ser
115 120 125
Ser Leu Leu Val Pro Tyr Phe Ser Trp Lys Tyr Ser His Arg Arg His
130 135 140
His Ser Asn Thr Gly Ser Leu Glu Arg Asp Glu Val Phe Val Pro Lys
145 150 155 160
Lys Lys Ser Gly Leu Arg Trp Trp Ala Lys His Phe Asn Asn Pro Pro
165 170 175
Gly Arg Phe Leu Ser Ile Thr Ile Gln Leu Thr Leu Gly Trp Pro Leu
180 185 190
Tyr Leu Ala Phe Asn Val Ala Gly Arg Pro Tyr Asp Arg Phe Ala Cys
195 200 205
His Tyr Asp Pro Tyr Gly Pro Ile Phe Ser Asp Arg Glu Arg Leu Gln
210 215 220
Ile Tyr Ile Ser Asp Ala Gly Val Leu Ala Val Ala Tyr Ala Leu Tyr
225 230 235 240
Arg Leu Val Leu Ala Lys Gly Val Gly Trp Val Ile Ser Val Tyr Gly
245 250 255
Val Pro Leu Leu Val Val Asn Ala Phe Leu Val Met Ile Thr Tyr Leu
260 265 270
Gln His Thr His Pro Ser Leu Pro His Tyr Asp Ser Ser Glu Trp Asp
275 280 285
Trp Met Arg Gly Ala Leu Ser Thr Val Asp Arg Asp Tyr Gly Ile Leu
290 295 300
Asn Lys Val Phe His Asn Ile Thr Asp Thr His Val Ala His His Leu
305 310 315 320
Phe Ser Thr Met Pro His Tyr His Ala Met Val Ala Thr Lys Ala Ile
325 330 335
Lys Pro Ile Leu Gly Glu Tyr Tyr Gln Phe Asp Gly Met Pro Val Tyr
340 345 350
Lys Ala Ile Trp Arg Glu Ala Lys Glu Cys Leu Tyr Val Glu Pro Asp
355 360 365
Glu Gly Asp Lys Asp Lys Gly Val Phe Trp Phe Arg Asn Lys Leu
370 375 380

大豆微粒体去饱和酶GmFAD2-2a,GenBank登录号BAD89862.1

<210> 209

<211> 383
<212> PRT
<213> 大豆
<400> 209

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

275 280 285

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

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

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

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

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

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

玉米脂肪酸去饱和酶FAD2,GenBank登录号ABF50053.1

<210> 210

<211> 387

<212> PRT

<213> 玉米

<400> 210

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

Gln Leu Ala Arg Ala Thr Gly Gly Ala Ala Met Gln Arg Ser Pro Val
20 25 30

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

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

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

Pro Ala Leu Pro Ser Pro Leu Gly Tyr Ala Ala Trp Pro Leu Tyr Trp
85 90 95

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

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

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

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

Glu Val Phe Val Pro Lys Lys Lys Glu Ala Leu Pro Trp Tyr Thr Pro
165 170 175
Tyr Val Tyr Asn Asn Pro Val Gly Arg Leu Ala His Ile Val Val Gln
180 185 190
Leu Thr Leu Gly Trp Pro Leu Tyr Leu Ala Thr Asn Ala Ser Gly Arg
195 200 205
Pro Tyr Pro Arg Phe Ala Cys His Phe Asp Pro Tyr Gly Pro Ile Tyr
210 215 220
Asn Asp Arg Glu Arg Ala Gln Ile Phe Val Ser Asp Ala Gly Val Ala
225 230 235 240
Ala Val Ala Phe Gly Leu Tyr Lys Leu Ala Ala Ala Phe Gly Val Trp
245 250 255
Trp Val Val Arg Val Tyr Ala Val Pro Leu Leu Ile Val Asn Ala Trp
260 265 270
Leu Val Leu Ile Thr Tyr Leu Gln His Thr His Pro Ser Leu Pro His
275 280 285
Tyr Asp Ser Ser Glu Trp Asp Trp Leu Arg Gly Ala Leu Ala Thr Met
290 295 300
Asp Arg Asp Tyr Gly Ile Leu Asn Arg Val Phe His Asn Ile Thr Asp
305 310 315 320
Thr His Val Ala His His Leu Phe Ser Thr Met Pro His Tyr His Ala
325 330 335
Met Glu Ala Thr Lys Ala Ile Arg Pro Ile Leu Gly Asp Tyr Tyr His
340 345 350
Phe Asp Pro Thr Pro Val Ala Lys Ala Thr Trp Arg Glu Ala Gly Glu
355 360 365
Cys Ile Tyr Val Glu Pro Glu Asp Arg Lys Gly Val Phe Trp Tyr Asn
370 375 380
Lys Lys Phe
385

Prototheca moriformis d12去饱和酶等位基因1 CHP-与Riyaz一致

<210> 211
<211> 368
<212> PRT
<213> Prototheca moriformis
<400> 211

Met Ala Ile Lys Thr Asn Arg Gln Pro Val Glu Lys Pro Pro Phe Thr
1 5 10 15
Ile Gly Thr Leu Arg Lys Ala Ile Pro Ala His Cys Phe Glu Arg Ser

20 25 30

Ala Leu Arg Ser Ser Met Tyr Leu Ala Phe Asp Ile Ala Val Met Ser
35 40 45

Leu Leu Tyr Val Ala Ser Thr Tyr Ile Asp Pro Ala Pro Val Pro Thr
50 55 60

Trp Val Lys Tyr Gly Val Met Trp Pro Leu Tyr Trp Phe Phe Gln Gly
65 70 75 80

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

Ala Phe Ser Ser Ser Gln Ala Ile Asn Asp Gly Val Gly Leu Val Phe
100 105 110

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

Arg His His Ser Asn Thr Gly Cys Leu Asp Lys Asp Glu Val Phe Val
130 135 140

Pro Pro His Arg Ala Val Ala His Glu Gly Leu Glu Trp Glu Glu Trp
145 150 155 160

Leu Pro Ile Arg Met Gly Lys Val Leu Val Thr Leu Thr Leu Gly Trp
165 170 175

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

Ala Asn His Phe Asp Pro Trp Ser Pro Ile Phe Ser Lys Arg Glu Arg
195 200 205

Ile Glu Val Val Ile Ser Asp Leu Ala Leu Val Ala Val Leu Ser Gly
210 215 220

Leu Ser Val Leu Gly Arg Thr Met Gly Trp Ala Trp Leu Val Lys Thr
225 230 235 240

Tyr Val Val Pro Tyr Leu Ile Val Asn Met Trp Leu Val Leu Ile Thr
245 250 255

Leu Leu Gln His Thr His Pro Ala Leu Pro His Tyr Phe Glu Lys Asp
260 265 270

Trp Asp Trp Leu Arg Gly Ala Met Ala Thr Val Asp Arg Ser Met Gly
275 280 285

Pro Pro Phe Met Asp Asn Ile Leu His His Ile Ser Asp Thr His Val
290 295 300

Leu His His Leu Phe Ser Thr Ile Pro His Tyr His Ala Glu Glu Ala
305 310 315 320

Ser Ala Ala Ile Arg Pro Ile Leu Gly Lys Tyr Tyr Gln Ser Asp Ser
325 330 335

Arg Trp Val Gly Arg Ala Leu Trp Glu Asp Trp Arg Asp Cys Arg Tyr
340 345 350
Val Val Pro Asp Ala Pro Glu Asp Asp Ser Ala Leu Trp Phe His Lys
355 360 365

Prototheca moriformis Δ12去饱和酶等位基因2

<210> 212
<211> 368
<212> PRT
<213> Prototheca moriformis
<400> 212

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

225 230 235 240
Tyr Val Val Pro Tyr Met Ile Val Asn Met Trp Leu Val Leu Ile Thr
245 250 255
Leu Leu Gln His Thr His Pro Ala Leu Pro His Tyr Phe Glu Lys Asp
260 265 270
Trp Asp Trp Leu Arg Gly Ala Met Ala Thr Val Asp Arg Ser Met Gly
275 280 285
Pro Pro Phe Met Asp Ser Ile Leu His His Ile Ser Asp Thr His Val
290 295 300
Leu His His Leu Phe Ser Thr Ile Pro His Tyr His Ala Glu Glu Ala
305 310 315 320
Ser Ala Ala Ile Arg Pro Ile Leu Gly Lys Tyr Tyr Gln Ser Asp Ser
325 330 335
Arg Trp Val Gly Arg Ala Leu Trp Glu Asp Trp Arg Asp Cys Arg Tyr
340 345 350
Val Val Pro Asp Ala Pro Glu Asp Asp Ser Ala Leu Trp Phe His Lys
355 360 365

甘蓝型油菜亚油酸去饱和酶FAD3,GenBank登录号AAA32994.1

<210> 213
<211> 383
<212> PRT
<213> 甘蓝型油菜
<400> 213

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

Gly Trp Arg Ile Ser His Arg Thr His His Gln Asn His Gly His Val
130 135 140
Glu Asn Asp Glu Ser Trp Val Pro Leu Pro Glu Lys Leu Tyr Lys Asn
145 150 155 160
Leu Pro His Ser Thr Arg Met Leu Arg Tyr Thr Val Pro Leu Pro Met
165 170 175
Leu Ala Tyr Pro Ile Tyr Leu Trp Tyr Arg Ser Pro Gly Lys Glu Gly
180 185 190
Ser His Phe Asn Pro Tyr Ser Ser Leu Phe Ala Pro Ser Glu Arg Lys
195 200 205
Leu Ile Ala Thr Ser Thr Thr Cys Trp Ser Ile Met Leu Ala Thr Leu
210 215 220
Val Tyr Leu Ser Phe Leu Val Asp Pro Val Thr Val Leu Lys Val Tyr
225 230 235 240
Gly Val Pro Tyr Ile Ile Phe Val Met Trp Leu Asp Ala Val Thr Tyr
245 250 255
Leu His His His Gly His Asp Glu Lys Leu Pro Trp Tyr Arg Gly Lys
260 265 270
Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr Thr Ile Asp Arg Asp Tyr
275 280 285
Gly Ile Phe Asn Asn Ile His His Asp Ile Gly Thr His Val Ile His
290 295 300
His Leu Phe Pro Gln Ile Pro His Tyr His Leu Val Asp Ala Thr Arg
305 310 315 320
Ala Ala Lys His Val Leu Gly Arg Tyr Tyr Arg Glu Pro Lys Thr Ser
325 330 335
Gly Ala Ile Pro Ile His Leu Val Glu Ser Leu Val Ala Ser Ile Lys
340 345 350
Lys Asp His Tyr Val Ser Asp Thr Gly Asp Ile Val Phe Tyr Glu Thr
355 360 365
Asp Pro Asp Leu Tyr Val Tyr Ala Ser Asp Lys Ser Lys Ile Asn
370 375 380

亚麻荠ω-3 FAD7-2

<210> 214

<211> 448

<212> PRT

<213> 亚麻荠

<400> 214

Met Ala Asn Leu Val Leu Ser Glu Cys Gly Ile Arg Pro Leu Pro Arg

1 5 10 15
Ile Tyr Thr Thr Pro Arg Ser Asn Phe Val Ser Asn Asn Asn Lys Pro
20 25 30
Ile Phe Lys Phe Arg Pro Phe Thr Ser Tyr Lys Thr Ser Ser Ser Pro
35 40 45
Leu Ala Cys Ser Arg Asp Gly Phe Gly Lys Asn Trp Ser Leu Asn Val
50 55 60
Ser Val Pro Leu Thr Thr Thr Thr Pro Ile Val Asp Glu Ser Pro Leu
65 70 75 80
Lys Glu Glu Glu Glu Glu Lys Gln Arg Phe Asp Pro Gly Ala Pro Pro
85 90 95
Pro Phe Asn Leu Ala Asp Ile Arg Ala Ala Ile Pro Lys His Cys Trp
100 105 110
Val Lys Asn Pro Trp Lys Ser Met Ser Tyr Val Leu Arg Asp Val Ala
115 120 125
Ile Val Phe Ala Leu Ala Ala Gly Ala Ser Tyr Leu Asn Asn Trp Ile
130 135 140
Val Trp Pro Leu Tyr Trp Leu Ala Gln Gly Thr Met Phe Trp Ala Leu
145 150 155 160
Phe Val Leu Gly His Asp Cys Gly His Gly Ser Phe Ser Asn Asn Pro
165 170 175
Arg Leu Asn Asn Val Val Gly His Leu Leu His Ser Ser Ile Leu Val
180 185 190
Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr His His Gln Asn His
195 200 205
Gly His Val Glu Asn Asp Glu Ser Trp His Pro Met Ser Glu Lys Ile
210 215 220
Tyr Gln Ser Leu Asp Lys Pro Thr Arg Phe Phe Arg Phe Thr Leu Pro
225 230 235 240
Leu Val Met Leu Ala Tyr Pro Phe Tyr Leu Trp Ala Arg Ser Pro Gly
245 250 255
Lys Lys Gly Ser His Tyr His Pro Glu Ser Asp Leu Phe Leu Pro Lys
260 265 270
Glu Lys Thr Asp Val Leu Thr Ser Thr Ala Cys Trp Thr Ala Met Ala
275 280 285
Ala Leu Leu Ile Cys Leu Asn Phe Val Val Gly Pro Val Gln Met Leu
290 295 300
Lys Leu Tyr Gly Ile Pro Tyr Trp Ile Asn Val Met Trp Leu Asp Phe
305 310 315 320

Val Thr Tyr Leu His His His Gly His Glu Asp Lys Leu Pro Trp Tyr
325 330 335
Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr Thr Leu Asp
340 345 350
Arg Asp Tyr Gly Val Ile Asn Asn Ile His His Asp Ile Gly Thr His
355 360 365
Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr His Leu Val Glu
370 375 380
Ala Thr Glu Ala Val Lys Pro Val Leu Gly Lys Tyr Tyr Arg Glu Pro
385 390 395 400
Asp Lys Ser Gly Pro Leu Pro Leu His Leu Leu Gly Ile Leu Ala Lys
405 410 415
Ser Ile Lys Glu Asp His Tyr Val Ser Asp Glu Gly Asp Val Val Tyr
420 425 430
Tyr Lys Ala Asp Pro Asn Met Tyr Gly Glu Ile Lys Val Gly Ala Asp
435 440 445

亚麻荠ω-3 FAD7-1

<210> 215
<211> 448
<212> PRT
<213> 亚麻荠
<400> 215

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

130 135 140

Val Trp Pro Leu Tyr Trp Leu Ala Gln Gly Thr Met Phe Trp Ala Leu
145 150 155 160

Phe Val Leu Gly His Asp Cys Gly His Gly Ser Phe Ser Asn Asn Pro
165 170 175

Arg Leu Asn Asn Val Val Gly His Leu Leu His Ser Ser Ile Leu Val
180 185 190

Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr His His Gln Asn His
195 200 205

Gly His Val Glu Asn Asp Glu Ser Trp His Pro Met Ser Glu Lys Ile
210 215 220

Tyr Gln Ser Leu Asp Lys Pro Thr Arg Phe Phe Arg Phe Thr Leu Pro
225 230 235 240

Leu Val Met Leu Ala Tyr Pro Phe Tyr Leu Trp Ala Arg Ser Pro Gly
245 250 255

Lys Lys Gly Ser His Tyr His Pro Glu Ser Asp Leu Phe Leu Pro Lys
260 265 270

Glu Lys Thr Asp Val Leu Thr Ser Thr Ala Cys Trp Thr Ala Met Ala
275 280 285

Ala Leu Leu Ile Cys Leu Asn Phe Val Val Gly Pro Val Gln Met Leu
290 295 300

Lys Leu Tyr Gly Ile Pro Tyr Trp Ile Asn Val Met Trp Leu Asp Phe
305 310 315 320

Val Thr Tyr Leu His His His Gly His Glu Asp Lys Leu Pro Trp Tyr
325 330 335

Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu Thr Thr Leu Asp
340 345 350

Arg Asp Tyr Gly Val Ile Asn Asn Ile His His Asp Ile Gly Thr His
355 360 365

Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr His Leu Val Glu
370 375 380

Ala Thr Glu Ala Val Lys Pro Val Leu Gly Lys Tyr Tyr Arg Glu Pro
385 390 395 400

Asp Lys Ser Gly Pro Leu Pro Leu His Leu Leu Gly Ile Leu Ala Lys
405 410 415

Ser Ile Lys Glu Asp His Tyr Val Ser Asp Glu Gly Asp Val Val Tyr
420 425 430

Tyr Lys Ala Asp Pro Asn Met Tyr Gly Glu Ile Lys Val Gly Ala Asp
435 440 445

大豆叶绿体ω 3脂肪酸去饱和酶亚型2,GenBank登录号ACF19424.1

<210> 216

<211> 453

<212> PRT

<213> 大豆

<400> 216

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

260 265 270

Ser Asp Leu Phe Val Pro Asn Glu Arg Lys Asp Val Ile Thr Ser Thr

275 280 285

Ala Cys Trp Ala Ala Met Leu Gly Leu Leu Val Gly Leu Gly Phe Val

290 295 300

Met Gly Pro Ile Gln Leu Leu Lys Leu Tyr Gly Val Pro Tyr Val Ile

305 310 315 320

Phe Val Met Trp Leu Asp Leu Val Thr Tyr Leu His His His Gly His

325 330 335

Glu Asp Lys Leu Pro Trp Tyr Arg Gly Lys Glu Trp Ser Tyr Leu Arg

340 345 350

Gly Gly Leu Thr Thr Leu Asp Arg Asp Tyr Gly Leu Ile Asn Asn Ile

355 360 365

His His Asp Ile Gly Thr His Val Ile His His Leu Phe Pro Gln Ile

370 375 380

Pro His Tyr His Leu Val Glu Ala Thr Glu Ala Ala Lys Pro Val Phe

385 390 395 400

Gly Lys Tyr Tyr Arg Glu Pro Lys Lys Ser Ser Pro Leu Pro Phe His

405 410 415

Leu Ile Gly Glu Leu Ile Lys Ser Phe Lys Thr Asp His Phe Val Ser

420 425 430

Asp Lys Gly Asp Val Val Tyr Tyr Gln Thr Asp Ser Glu Ile Asn Gly

435 440 445

Ser Ser Lys Ser Glu

450

油桐ω-3脂肪酸去饱和酶,GenBank登录号AAF12821.1

<210> 217

<211> 451

<212> PRT

<213> 油桐

<400> 217

Met Ala Val Trp Ala Leu Ser Asp Cys Gly Ile Arg Pro Leu Pro Lys

1 5 10 15

Ile Tyr Ser Lys Pro Arg Leu Ala Phe Thr Ser Asn Asn Pro Gln Pro

20 25 30

Thr Lys Pro Pro Ile Pro Arg Pro Glu Leu Arg Asn Gly Ser Ser Phe

35 40 45

Lys Leu Ser Ser Ala Gly Leu Arg Asp Lys Trp Val Val Asn Leu Ser

50 55 60

Ala Pro Leu Arg Gly Gly Leu Val Glu Glu Asp Glu Asp Asn Phe Glu
65 70 75 80
Gly Asn Arg Val Ile Ser Ile Asp Glu Ser Gly Gly Glu Asp Phe Asp
85 90 95
Pro Asp Ala Pro Pro Pro Asp Thr Leu Ala Asp Ile Arg Ala Ala Ile
100 105 110
Pro Lys His Arg Trp Thr Lys Asn Pro Glu Ile Ser Leu Ser Tyr Val
115 120 125
Val Arg Asn Val Ala Val Val Ser Gly Leu Ala Ala Val Ala Ala Tyr
130 135 140
Phe Asn Asp Trp Ala Phe Trp Ser Pro Tyr Trp Ile Cys Gln Gly Thr
145 150 155 160
Met Leu Ser Ala Leu Phe Val Leu Gly His Asp Cys Gly His Gly Ser
165 170 175
Phe Ser Asn Asn Pro Asn Leu Asp Ser Val Val Gly His Leu Leu His
180 185 190
Tyr Ser Ile Leu Val Pro Tyr His Gly Trp Arg Ile Ser Tyr Arg Thr
195 200 205
His Tyr Arg His His Gln His Ala Glu Asn Asp Glu Ser Trp His Pro
210 215 220
Leu Ser Glu Lys Ile Tyr Lys Gly Leu Asn Asn Val Thr Arg Thr Leu
225 230 235 240
Arg Leu Ser Leu Pro Phe Ser Leu Leu Ala Tyr Pro Ser Tyr Leu Trp
245 250 255
Ser Arg Ser Pro Gly Lys Lys Gly Phe His Phe His Pro Ser Ser Asp
260 265 270
Leu Phe Val Pro Ser Glu Lys Lys Asp Val Ile Thr Ser Thr Val Cys
275 280 285
Trp Thr Ala Met Leu Ala Leu Leu Val Gly Ser Ser Phe Val Met Gly
290 295 300
Leu Val Gln Leu Leu Lys Pro Tyr Gly Ile Pro Tyr Leu Ile Phe Val
305 310 315 320
Met Trp Leu Asp Leu Val Thr Tyr Leu His His His Gly His Glu Asn
325 330 335
Lys Ile Pro Trp Tyr Arg Gly Glu Glu Trp Ser Tyr Leu Arg Gly Arg
340 345 350
Leu Thr Thr Val Gly Arg Asp Tyr Gly Trp Phe Asn Lys Ile His His
355 360 365
Asp Ile Gly Ile His Val Ile His His Leu Phe Pro Gln Ile Pro His

370 375 380
Tyr His Leu Ile Glu Ala Thr Glu Ala Ala Lys Pro Val Leu Gly Lys
385 390 395 400
Tyr Tyr Arg Glu Pro Lys Lys Ser Gly Pro Leu Pro Phe His Leu Leu
405 410 415
Gly Ser Leu Ile Arg Ser Met Lys Glu Asp His Tyr Val Ser Asp Ala
420 425 430
Gly Asp Val Val Tyr Tyr Gln Lys Glu Pro Lys Leu Ser Gly Ala Arg
435 440 445
Cys Leu Ser
450

蓖麻 ω-3脂肪酸去饱和酶,内质网,假定的,GenBank登录号EEF36775.1

<210> 218
<211> 506
<212> PRT
<213> 蓖麻
<400> 218

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

Ser Ile Leu Val Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr His
180 185 190
His Gln Asn His Gly Asn Val Asp Lys Asp Glu Ser Trp Val Pro Leu
195 200 205
Pro Glu Lys Ile Tyr Asn Ser Leu Asp Val Ser Thr Arg Ile Met Arg
210 215 220
Phe Thr Val Pro Leu Pro Leu Leu Ala Tyr Pro Ile Tyr Leu Trp Thr
225 230 235 240
Arg Gly Pro Gly Lys Asp Gly Ser His Phe Ser Pro Tyr Ser Asp Leu
245 250 255
Phe Ala Pro Asn Glu Arg Arg Ala Val Leu Thr Ser Thr Val Cys Trp
260 265 270
Ile Leu Met Ile Leu Leu Leu Leu Tyr Ser Cys Phe Val Phe Gly Ser
275 280 285
Val Gln Val Leu Lys Leu Tyr Gly Val Pro Tyr Ile Ile Phe Ile Met
290 295 300
Trp Leu Asp Phe Val Thr Tyr Leu His His His Gly His Glu Gln Lys
305 310 315 320
Leu Pro Trp Tyr Arg Gly Lys Glu Trp Ser Tyr Leu Arg Gly Gly Leu
325 330 335
Thr Thr Val Asp Arg Asp Tyr Gly Trp Ile Asn Asn Ile His His Asp
340 345 350
Ile Gly Thr His Val Ile His His Leu Phe Pro Gln Ile Pro His Tyr
355 360 365
His Leu Val Glu Ala Thr Lys Ala Ala Lys Pro Val Leu Gly Lys Tyr
370 375 380
Tyr Arg Glu Pro Arg Lys Ser Gly Pro Phe Pro Leu His Leu Phe Glu
385 390 395 400
Asn Leu Val Lys Ser Ile Ser Gln Asp His Tyr Val Ser Asp Ile Gly
405 410 415
Asp Val Val Tyr Tyr Gln Thr Asp Pro Glu Leu Tyr Lys Lys Leu Ala
420 425 430
Lys Thr Lys Asp Val Leu Glu Asp Leu Glu Ser Cys Asp Ala Leu Gly
435 440 445
Ile Leu Lys Cys Lys Asn Cys Ser Ala Lys Ser Gly Ser Ser Tyr Ser
450 455 460
Ala Leu Gln Phe Leu Glu Leu Lys Gly Arg Ala Ala Gly Ala Glu Lys
465 470 475 480
Gly Arg Glu Trp Lys Ile Leu Val Phe Glu Ile Leu Thr Leu His His

485 490 495

Arg Ile Ile Ala Gln Met Ser Phe Met Gly

500 505

亚麻脂肪酸去饱和酶3C,GenBank登录号ADV92272.1

<210> 219

<211> 392

<212> PRT

<213> 亚麻

<400> 219

Met Ser Pro Pro Asn Ser Met Gly Met Glu Ala Ala His Pro Thr Gly
1 5 10 15

Asn Gly Asn Gly Val Ala Val Met Asn Gly Ala Ser Ala Asn Lys Pro
20 25 30

Asp Phe Asp Pro Ser Glu Ala Pro Pro Phe Lys Ile Ala Asp Ile Arg
35 40 45

Ala Ala Ile Pro Pro His Cys Trp Val Lys Asn Pro Trp Arg Ser Leu
50 55 60

Ser Tyr Val Leu Arg Asp Ala Val Val Ile Leu Ala Phe Ala Ala Ala
65 70 75 80

Ala Leu Lys Leu Asp Leu Trp Ala Val Trp Pro Leu Tyr Trp Ile Ala
85 90 95

Gln Gly Thr Met Phe Trp Ala Val Phe Val Leu Gly His Asp Cys Gly
100 105 110

His Gly Ser Phe Ser Asp Ser Trp Trp Leu Asn Asn Val Val Gly His
115 120 125

Ile Leu His Ser Ala Ile Leu Val Pro Tyr His Gly Trp Arg Ile Ser
130 135 140

His Lys Thr His His Gln Asn His Gly Asn Val Glu Lys Asp Glu Ser
145 150 155 160

Trp Val Pro Leu Pro Glu Lys Val Tyr Lys Thr Leu Asp Thr Ser Thr
165 170 175

Lys Phe Met Arg Phe Thr Ile Pro Leu Pro Met Phe Ala Tyr Pro Ile
180 185 190

Tyr Leu Trp Thr Arg Ser Pro Gly Lys Lys Gly Ser His Phe Asn Pro
195 200 205

Tyr Ser Asp Leu Phe Ala Pro Gln Glu Arg Lys Ser Val Leu Ile Ser
210 215 220

Thr Ile Ser Trp Ile Ser Met Val Leu Ile Leu Leu Tyr Ala Ser Phe
225 230 235 240

Leu Phe Gly Phe Leu Thr Val Phe Lys Val Tyr Thr Val Pro Tyr Leu
245 250 255
Ile Phe Val Ala Trp Leu Asp Met Val Thr Tyr Leu His His His Gly
260 265 270
His Glu Glu Lys Leu Pro Trp Tyr Arg Gly Gln Glu Trp Ser Tyr Leu
275 280 285
Arg Gly Gly Leu Thr Thr Val Asp Arg Asp Tyr Gly Ile Ile Asn Asn
290 295 300
Ile His His Asp Ile Gly Thr His Val Ile His His Leu Phe Pro Gln
305 310 315 320
Ile Pro His Tyr His Leu Val Glu Ala Thr Lys Ala Ala Lys Ser Val
325 330 335
Leu Gly Lys Tyr Tyr Arg Glu Pro Lys Lys Ser Gly Pro Phe Pro Phe
340 345 350
His Leu Phe Asp Asn Leu Val Arg Ser Leu Gly Glu Asp His Tyr Val
355 360 365
Ser Asp Ala Gly Asp Val Val Phe Tyr Gln Ser Asp Pro Glu Ile Phe
370 375 380
Lys Phe Ser Lys Ser Lys Ser Ala
385 390

Prototheca moriformis FAD-A亚油酸去饱和酶

<210> 220
<211> 315
<212> PRT
<213> Prototheca moriformis
<400> 220

Met Phe Trp Ala Leu Phe Val Val Gly His Asp Cys Gly His Gln Ser
1 5 10 15
Phe Ser Thr Ser Lys Arg Leu Asn Asp Ala Val Gly Leu Phe Val His
20 25 30
Ser Ile Ile Gly Val Pro Tyr His Gly Trp Arg Ile Ser His Arg Thr
35 40 45
His His Asn Asn His Gly His Val Glu Asn Asp Glu Ser Trp Tyr Pro
50 55 60
Pro Thr Glu Ser Gly Leu Lys Ala Met Thr Asp Met Gly Arg Gln Gly
65 70 75 80
Arg Phe His Phe Pro Ser Met Leu Phe Val Tyr Pro Phe Tyr Leu Phe
85 90 95
Trp Arg Ser Pro Gly Lys Thr Gly Ser His Phe Ser Pro Ala Thr Asp

100 105 110
Leu Phe Ala Leu Trp Glu Ala Pro Leu Ile Arg Thr Ser Asn Ala Cys
115 120 125
Gln Leu Ala Trp Leu Gly Ala Leu Ala Ala Gly Thr Trp Ala Leu Gly
130 135 140
Val Leu Pro Met Leu Asn Leu Tyr Leu Ala Pro Tyr Val Ile Ser Val
145 150 155 160
Ala Trp Leu Asp Leu Val Thr Tyr Leu His His His Gly Pro Ser Asp
165 170 175
Pro Arg Glu Glu Met Pro Trp Tyr Arg Gly Arg Glu Trp Ser Tyr Met
180 185 190
Arg Gly Gly Leu Thr Thr Ile Asp Arg Asp Tyr Gly Leu Phe Asn Lys
195 200 205
Val His His Asp Ile Gly Thr His Val Val His His Leu Phe Pro Gln
210 215 220
Ile Pro His Tyr Asn Leu Cys Arg Ala Thr Lys Ala Ala Lys Lys Val
225 230 235 240
Leu Gly Pro Tyr Tyr Arg Glu Pro Glu Arg Cys Pro Leu Gly Leu Leu
245 250 255
Pro Val His Leu Leu Ala Pro Leu Leu Arg Ser Leu Gly Gln Asp His
260 265 270
Phe Val Asp Asp Ala Gly Ser Val Leu Phe Tyr Arg Arg Ala Glu Gly
275 280 285
Ile Asn Pro Trp Ile Gln Lys Leu Leu Pro Trp Leu Gly Gly Ala Arg
290 295 300
Arg Gly Ala Asp Ala Gln Arg Asp Ala Ala Gln
305 310 315

Prototheca moriformis FAD-D ω 3去饱和酶

<210> 221
<211> 452
<212> PRT
<213> Prototheca moriformis
<400> 221

Met Ser Ile Gln Phe Ala Leu Arg Ala Ala Tyr Ile Lys Gly Thr Cys
1 5 10 15
Gln Arg Leu Ser Gly Arg Gly Ala Ala Leu Gly Leu Ser Arg Asp Trp
20 25 30
Thr Pro Gly Trp Thr Leu Pro Arg Cys Trp Pro Ala Ser Ala Ala Ala
35 40 45

Thr Ala Pro Pro Arg Ala Arg His Gln Glu Arg Ala Ile His Leu Thr
50 55 60
Ser Gly Arg Arg Arg His Ser Ala Leu Ala Ser Asp Ala Asp Glu Arg
65 70 75 80
Ala Leu Pro Ser Asn Ala Pro Gly Leu Val Met Ala Ser Gln Ala Asn
85 90 95
Tyr Phe Arg Val Arg Leu Leu Pro Glu Gln Glu Glu Gly Glu Leu Glu
100 105 110
Ser Trp Ser Pro Asn Val Arg His Thr Thr Leu Leu Cys Lys Pro Arg
115 120 125
Ala Met Leu Ser Lys Leu Gln Met Arg Val Met Val Gly Asp Arg Val
130 135 140
Ile Val Thr Ala Ile Asp Pro Val Asn Met Thr Val His Ala Pro Pro
145 150 155 160
Phe Asp Pro Leu Pro Ala Thr Arg Phe Leu Val Ala Gly Glu Ala Ala
165 170 175
Asp Met Asp Ile Thr Val Val Leu Asn Lys Ala Asp Leu Val Pro Glu
180 185 190
Glu Glu Ser Ala Ala Leu Ala Gln Glu Val Ala Ser Trp Gly Pro Val
195 200 205
Val Leu Thr Ser Thr Leu Thr Gly Arg Gly Leu Gln Glu Leu Glu Arg
210 215 220
Gln Leu Gly Ser Thr Thr Ala Val Leu Ala Gly Pro Ser Gly Ala Gly
225 230 235 240
Lys Ser Ser Ile Ile Asn Ala Leu Ala Arg Ala Ala Arg Glu Arg Pro
245 250 255
Ser Asp Ala Ser Val Ser Asn Val Pro Glu Glu Gln Val Val Gly Glu
260 265 270
Asp Gly Arg Ala Leu Ala Asn Pro Pro Pro Phe Thr Leu Ala Asp Ile
275 280 285
Arg Asn Ala Ile Pro Lys Asp Cys Phe Arg Lys Ser Ala Ala Lys Ser
290 295 300
Leu Ala Tyr Leu Gly Asp Leu Ser Ile Thr Gly Met Ala Val Leu Ala
305 310 315 320
Tyr Lys Ile Asn Ser Pro Trp Leu Trp Pro Leu Tyr Trp Phe Ala Gln
325 330 335
Gly Thr Met Phe Trp Ala Leu Phe Val Val Gly His Asp Cys Gly His
340 345 350
Gln Ser Phe Ser Thr Ser Lys Arg Leu Asn Asp Ala Leu Ala Trp Leu

355 360 365

Gly Ala Leu Ala Ala Gly Thr Trp Thr Trp Ala Leu Gly Val Leu Pro
370 375 380

Met Leu Asn Leu Tyr Leu Ala Pro Tyr Val Trp Leu Leu Val Thr Tyr
385 390 395 400

Leu His His His Gly Pro Ser Asp Pro Arg Glu Glu Met Pro Trp Tyr
405 410 415

Arg Gly Arg Glu Trp Ser Tyr Met Arg Gly Gly Leu Thr Thr Ile Asp
420 425 430

Arg Asp Tyr Gly Leu Phe Asn Lys Val His His Asp Ile Gly Thr His
435 440 445

Val Val His His
450

针对在P.moriformis中表达密码子优化的针对Bn FAD3-3xFLAG的pSZ2124转化序列的部分

<210> 222

<211> 6756

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 222

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60

ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120

tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180

ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240

gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300

ggaagacagg tgagggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360

atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420

cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480

gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540

ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600

ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660

cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720

ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840

atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900

aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 960

cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 1020

cagtcacaac ccgcaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg 1080
gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg 1140
gtgcacttca cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag 1200
aaggacgcca agtggcacct gtacttccag tacaacccga acgacaccgt ctgggggacg 1260
cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc 1320
gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac 1380
aacaacacct ccggcttctt caacgacacc atcgacccgc gccagcgctg cgtggccatc 1440
tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc 1500
tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc 1560
gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc 1620
caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc 1680
gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc 1740
cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc 1800
gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg cacccacttc 1860
gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag 1920
accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcgtg ggcctccaac 1980
tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc 2040
aagttctccc tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag 2100
gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc 2160
acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag 2220
ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac 2280
ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag 2340
gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag 2400
aacccctact tcaccaaccg catgagcgtg aacaaccagc ccttcaagag cgagaacgac 2460
ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac 2520
gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc 2580
gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag 2640
gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg 2700
tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt 2760
atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc 2820
ttgtgctatt tgcgaatacc acccccagca tcccccttccc tcgtttcata tcgcttgcat 2880
cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc 2940
actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg 3000
taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga 3060
tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgcctct gtcgcacctc 3120
agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg 3180
aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga 3240
gctgatggtc gaaacgttca cagcctaggg atatcgaatt cggccgacag gacgcgcgtc 3300
aaaggtgctg gtcgtgtatg ccctggccgg caggtcgttg ctgctgctgg ttagtgattc 3360

cgcaaccctg attttggcgt cttattttgg cgtggcaaac gctggcgccc gcgagccggg 3420
ccggcggcga tgcggtgccc cacggctgcc ggaatccaag ggaggcaaga gcgcccgggt 3480
cagttgaagg gctttacgcg caaggtacag ccgctcctgc aaggctgcgt ggtggaattg 3540
gacgtgcagg tcctgctgaa gttcctccac cgcctcacca gcggacaaag caccggtgta 3600
tcaggtccgt gtcatccact ctaaagaact cgactacgac ctactgatgg ccctagattc 3660
ttcatcaaaa acgcctgaga cacttgccca ggattgaaac tccctgaagg gaccaccagg 3720
ggccctgagt tgttccttcc ccccgtggcg agctgccagc caggctgtac ctgtgatcga 3780
ggctggcggg aaaataggct tcgtgtgctc aggtcatggg aggtgcagga cagctcatga 3840
aacgccaaca atcgcacaat tcatgtcaag ctaatcagct atttcctctt cacgagctgt 3900
aattgtccca aaattctggt ctaccggggg tgatccttcg tgtacgggcc cttccctcaa 3960
ccctaggtat gcgcgcatgc ggtcgccgcg caactcgcgc gagggccgag ggtttgggac 4020
gggccgtccc gaaatgcagt tgcacccgga tgcgtggcac ctttttgcg ataatttatg 4080
caatggactg ctctgcaaaa ttctggctct gtcgccaacc ctaggatcag cggcgtagga 4140
tttcgtaatc attcgtcctg atggggagct accgactacc ctaatatcag cccgactgcc 4200
tgacgccagc gtccactttt gtgcacacat tccattcgtg cccaagacat ttcattgtgg 4260
tgcgaagcgt ccccagttac gctcacctgt ttcccgacct ccttactgtt ctgtcgacag 4320
agcgggccca caggccggtc gcagccacta gtatggtggt ggccatggac cagcgctcca 4380
acgtgaacgg cgactccggc gcccgcaagg aggagggcgg gcgcgccatg gactacaagg 4440
accacgacgg cgactacaag gaccacgaca tcgactacaa ggacgacgac gacaagttcg 4500
acccctccgc ccagcccccc ttcaagatcg gcgacatccg cgccgccatc cccaagcact 4560
gctgggtgaa gtccccccctg cgctccatgt cctacgtgac ccgcgacatc ttcgccgtgg 4620
ccgccctggc catggccgcc gtgtacttcg actcctggtt cctgtggccc ctgtactggg 4680
tggcccaggg caccctgttc tgggccatct tcgtgctggg ccacgactgc ggccacggct 4740
ccttctccga catcccccctg ctgaactccg tggtgggcca catcctgcac tccttcatcc 4800
tggtgcccta ccacggctgg cgcatctccc accgcaccca ccaccagaac cacggccacg 4860
tggagaacga cgagtcctgg gtgcccctgc ccgagaagct gtacaagaac ctgccccact 4920
ccacccgcat gctgcgctac accgtgcccc tgcccatgct ggcctacccc atctacctgt 4980
ggtaccgctc ccccggcaag gagggctccc acttcaaccc ctactcctcc ctgttcgccc 5040
cctccgagcg caagctgatc gccacctcca ccacctgctg gtccatcatg ctggccaccc 5100
tggtgtacct gtccttcctg gtggaccccg tgaccgtgct gaaggtgtac ggcgtgccct 5160
acatcatctt cgtgatgtgg ctggacgccg tgacctacct gcaccaccac ggccacgacg 5220
agaagctgcc ctggtaccgc ggcaaggagt ggtcctacct gcgcggcggc ctgaccacca 5280
tcgaccgcga ctacggcatc ttcaacaaca tccaccacga catcggcacc cacgtgatcc 5340
accacctgtt cccccagatc ccccactacc acctggtgga cgccacccgc gccgccaagc 5400
acgtgctggg ccgctactac cgcgagccca agacctccgg cgccatcccc atccacctgg 5460
tggagtccct ggtggcctcc atcaagaagg accactacgt gtccgacacc ggcgacatcg 5520
tgttctacga gaccgacccc gacctgtacg tgtacgcctc cgacaagtcc aagatcaact 5580
agatcgatct taaggcagca gcagctcgga tagtatcgac acactctgga cgctggtcgt 5640
gtgatggact gttgccgcca cacttgctgc cttgacctgt gaatatccct gccgctttta 5700

tcaaacagcc tcagtgtgtt tgatcttgtg tgtacgcgct tttgcgagtt gctagctgct 5760
tgtgctattt gcgaatacca cccccagcat ccccttccct cgtttcatat cgcttgcatc 5820
ccaaccgcaa cttatctacg ctgtcctgct atccctcagc gctgctcctg ctcctgctca 5880
ctgcccctcg cacagccttg gtttgggctc cgcctgtatt ctcctggtac tgcaacctgt 5940
aaaccagcac tgcaatgctg atgcacggga agtagtggga tgggaacaca aatggaaagc 6000
ttaattaaga gctcttgttt tccagaagga gttgctcctt gagcctttca ttctcagcct 6060
cgataacctc caaagccgct ctaattgtgg agggggttcg aatttaaaag cttggaatgt 6120
tggttcgtgc gtctggaaca agcccagact tgttgctcac tgggaaaagg accatcagct 6180
ccaaaaaact tgccgctcaa accgcgtacc tctgctttcg cgcaatctgc cctgttgaaa 6240
tcgccaccac attcatattg tgacgcttga gcagtctgta attgcctcag aatgtggaat 6300
catctgcccc ctgtgcgagc ccatgccagg catgtcgcgg gcgaggacac ccgccactcg 6360
tacagcagac cattatgcta cctcacaata gttcataaca gtgaccatat ttctcgaagc 6420
tccccaacga gcacctccat gctctgagtg gccacccccc ggccctggtg cttgcggagg 6480
gcaggtcaac cggcatgggg ctaccgaaat ccccgaccgg atcccaccac ccccgcgatg 6540
ggaagaatct ctccccggga tgtgggccca ccaccagcac aacctgctgg cccaggcgag 6600
cgtcaaacca taccacacaa atatccttgg catcggccct gaattccttc tgccgctctg 6660
ctacccggtg cttctgtccg aagcaggggt tgctagggat cgctccgagt ccgcaaaccc 6720
ttgtcgcgtg gcggggcttg ttcgagcttg aagagc 6756

pSZ2125的部分,针对在P.moriformis中表达的密码子优化的亚麻荠FAD-7-3XFLAG的编码序列

<210> 223
<211> 1437
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 223

actagtatgg ccaacctggt gctgtccgag tgcggcatcc gccccctgcc ccgcatctac 60
accacccccc gctccaactt cgtgtccaac aacaacaagc ccatcttcaa gttccgcccc 120
ctgacctcct acaagacctc ctccccccctg ttctgctccc gcgacggctt cggccgcaac 180
tggtccctga acgtgtccgt gcccctggcc accaccaccc ccatcgtgga cgagtccccc 240
ctggaggagg aggaggagga ggagaagcag cgcgggcgcg ccatggacta caaggaccac 300
gacggcgact acaaggacca cgacatcgac tacaaggacg acgacgacaa gttcgacccc 360
ggcgcccccc ccccccttcaa cctggccgac atccgcgccg ccatccccaa gcactgctgg 420
gtgaagaacc cctggaagtc catgtcctac gtgctgcgcg acgtggccat cgtgttcgcc 480
ctggccgccg gcgccgccta cctgaacaac tggatcgtgt ggcccctgta ctggctggcc 540
cagggcacca tgttctgggc cctgttcgtg ctgggccacg actgcggcca cggctccttc 600
tccaacaacc cccgcctgaa caacgtggtg ggccacctgc tgcactcctc catcctggtg 660
ccctaccacg gctggcgcat ctcccaccgc acccaccacc agaaccacgg ccacgtggag 720

aacgacgagt cctggcaccc catgtccgag aagatctacc agtccctgga caagcccacc 780
cgcttcttcc gcttcaccct gcccctggtg atgctggcct acccccttcta cctgtgggcc 840
cgctcccccg gcaagaaggg ctcccactac caccccgagt ccgacctgtt cctgcccaag 900
gagaagaccg acgtgctgac ctccaccgcc tgctggaccg ccatggccgc cctgctgatc 960
tgcctgaact tcgtggtggg ccccgtgcag atgctgaagc tgtacggcat cccctactgg 1020
atcaacgtga tgtggctgga cttcgtgacc tacctgcacc accacggcca cgaggacaag 1080
ctgccctggt accgcggcaa ggagtggtcc tacctgcgcg gcggcctgac caccctggac 1140
cgcgactacg gcgtgatcaa caacatccac cacgacatcg gcacccacgt gatccaccac 1200
ctgttccccc agatccccca ctaccacctg gtggaggcca ccgaggccgt gaagcccgtg 1260
ctgggcaagt actaccgcga gcccgacaag tccggccccc tgcccctgca cctgctgggc 1320
atcctggcca agtccatcaa ggaggaccac tacgtgtccg acgagggcga cgtggtgtac 1380
tacaaggccg accccaacat gtacggcgag atcaaggtgg gcgccgacta gatcgat 1437

pSZ2126的部分,针对在P.moriformis中表达的密码子优化的亚麻FAD3A-3XFLAG的编码序列

<210> 224
<211> 1269
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 224

actagtatgt ccccccccaa ctccatgtcc cccgccacca acggctccac caacggcgtg 60
gccatcaacg gcgccaagaa gctgctggac gggcgcgcca tggactacaa ggaccacgac 120
ggcgactaca aggaccacga catcgactac aaggacgacg acgacaagtt cgacccctcc 180
gccgcccccc ccttcaagat cgccgacatc cgcgccgcca tccccccccca ctgctgggtg 240
aagaacccct ggcgctccct gtcctacgtg ctgcgcgacc tgctggtgat cctgtccttc 300
gccgtggccg ccaccaagct ggactcctgg accgtgtggc ccctgtactg gatcgcccag 360
ggcaccatgt tctgggccgt gttcgtgctg ggccacgact gcggccacgg ctccttctcc 420
gactcctggc tgctgaacaa cgtgatgggc cacatcctgc actcctccat cctggtgccc 480
taccacggct ggcgcatctc ccacaagacc caccaccaga accacggcaa cgtggagaag 540
gacgagtcct gggtgcccct gcccgagaag gtgtacaagt ccctggacac cggcaccaag 600
ttcatgcgct tcaccatccc cctgcccatg ttcgcctacc ccatctacct gtggcgccgc 660
tcccccggca agaagggctc ccacttcaac ccctactccg acctgttcgc ccccaacgag 720
cgcacctccg tgatgatctc caccctgtgc tggaccgcca tggccctgct gctgtgctac 780
tcctccttca tctacggctt cctgcccgtg ttcaagatct acggcgtgcc ctacctgatc 840
ttcgtggcct ggctggacat ggtgacctac ctgcaccacc acggctacga gcagaagctg 900
ccctggtacc gcggcaagga gtggtcctac ctgcgcggcg gcctgaccac cgtggaccgc 960
gactacggcg tgatcaacaa catccaccac gacatcggca cccacgtgat ccaccacctg 1020
ttcccccaga tgccccacta ccacctggtg gaggccaccc aggccgccaa gcacgtgctg 1080

ggcaagtact accgcgagcc caagaagtcc ggccccttcc ccttccacct gttcggctac 1140
ctggtgcgct ccctgggcga ggaccactac gtgtccgaca ccggcgacgt ggtgttctac 1200
cagtccgacc cccacatccc caagttcccc acctccgcca ccaccaagtc caagtcctcc 1260
tagatcgat 1269

pSZ2127的部分,针对在P.moriformis中表达的密码子优化的亚麻FAD3B-3XFLAG的编码序列

<210> 225
<211> 1266
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 225

actagtatgt ccccccccaa ctccatgtcc cccaccacca acggcaacgg cgtggccatg 60
aacggcgcca agaagcagct ggacgggcgc gccatggact acaaggacca cgacggcgac 120
tacaaggacc acgacatcga ctacaaggac gacgacgaca agttcgaccc ctccgccgcc 180
ccccccttca agatcgccga catccgcgcc gccatccccc cccactgctg ggtgaagaac 240
ccctggcgct ccctgtccta cgtgctgcgc gacctgctgg tgatcctgtc cttcgccgtg 300
gccgccgcca agctggactc ctggaccttc tggcccctgt actgggtggc ccagggcacc 360
atgttctggg ccgtgttcgt gctgggccac gactgcggcc acggctcctt ctccgacatc 420
tggctgctga acaacgtgat gggccacatc ctgcactcct ccatcctggt gccctaccac 480
ggctggcgca tctcccacaa gacccaccac cagaaccacg gcaacgtgga gaaggacgag 540
tcctgggtgc ccctgcccga gaaggtgtac aagtccctgg acacctccac caagttcatg 600
cgcttcacca tccccctgcc catgttcgcc taccccatct acctgtggac ccgctccccc 660
ggcaagaagg gctcccactt caacccctac tccgacctgt tcgcccccaa cgagcgcgcc 720
gccgtgctga tctccaccct gtgctggacc gccatggccc tgctgctgtg ctactcctcc 780
ttcatctacg gcttcgcccc cgtgctgaag atctacggcg tgccctacct gatcttcgtg 840
gcctggctgg acatggtgac ctacctgcac caccacggct acgagcagaa gctgccctgg 900
taccgcggca aggagtggtc ctacctgcgc ggcggcctga ccaccgtgga ccgcgactac 960
ggcgtgatca acaacatcca ccacgacatc ggcacccacg tgatccacca cctgttcccc 1020
cagatgcccc actaccacct ggtggaggcc acccaggccg ccaagcacgt gctgggcaag 1080
tactaccgcg agcccaagaa gtccggcccc ttccccttcc acctgttcgg ctacctggtg 1140
cgctccctgg gcgaggacca ctacgtgtcc gacaccggcg acgtggtgtt ctaccagtcc 1200
gacccccaca tccccaagtt ccgcacctcc tccgccacca ccaagtccaa gtcctcctag 1260
atcgat 1266

pSZ2139的部分,Prototheca moriformis SAD2-发夹A

<210> 226
<211> 5246
<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 226

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60
ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120
tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180
ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240
gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300
ggaagacagg tgaggggtgt atgaattgta cagaacaacc acgagccttg tctaggcaga 360
atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420
cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480
gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540
ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600
ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660
cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720
ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 960
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 1020
cagtcacaac ccgcaaacgg cgcgccatgc tgctgcaggc cttcctgttc ctgctggccg 1080
gcttcgccgc caagatcagc gcctccatga cgaacgagac gtccgaccgc cccctggtgc 1140
acttcacccc caacaagggc tggatgaacg accccaacgg cctgtggtac gacgagaagg 1200
acgccaagtg gcacctgtac ttccagtaca acccgaacga caccgtctgg gggacgccct 1260
tgttctgggg ccacgccacg tccgacgacc tgaccaactg ggaggaccag cccatcgcca 1320
tcgccccgaa gcgcaacgac tccggcgcct tctccggctc catggtggtg gactacaaca 1380
acacctccgg cttcttcaac gacaccatcg acccgcgcca gcgctgcgtg gccatctgga 1440
cctacaacac cccggagtcc gaggagcagt acatctccta cagcctggac ggcggctaca 1500
ccttcaccga gtaccagaag aaccccgtgc tggccgccaa ctccacccag ttccgcgacc 1560
cgaaggtctt ctggtacgag ccctcccaga agtggatcat gaccgcggcc aagtcccagg 1620
actacaagat cgagatctac tcctccgacg acctgaagtc ctggaagctg gagtccgcgt 1680
tcgccaacga gggcttcctc ggctaccagt acgagtgccc cggcctgatc gaggtcccca 1740
ccgagcagga ccccagcaag tcctactggg tgatgttcat ctccatcaac cccggcgccc 1800
cggccggcgg ctccttcaac cagtacttcg tcggcagctt caacggcacc cacttcgagg 1860
ccttcgacaa ccagtcccgc gtggtggact tcggcaagga ctactacgcc ctgcagacct 1920
tcttcaacac cgacccgacc tacgggagcg ccctgggcat cgcgtgggcc tccaactggg 1980
agtactccgc cttcgtgccc accaacccct ggcgctcctc catgtccctc gtgcgcaagt 2040
tctccctcaa caccgagtac caggccaacc cggagacgga gctgatcaac ctgaaggccg 2100

agccgatcct gaacatcagc aacgccggcc cctggagccg gttcgccacc aacaccacgt 2160
tgacgaaggc caacagctac aacgtcgacc tgtccaacag caccggcacc ctggagttcg 2220
agctggtgta cgccgtcaac accacccaga cgatctccaa gtccgtgttc gcggacctct 2280
ccctctggtt caagggcctg gaggaccccg aggagtacct ccgcatgggc ttcgaggtgt 2340
ccgcgtcctc cttcttcctg gaccgcggga acagcaaggt gaagttcgtg aaggagaacc 2400
cctacttcac caaccgcatg agcgtgaaca accagccctt caagagcgag aacgacctgt 2460
cctactacaa ggtgtacggc ttgctggacc agaacatcct ggagctgtac ttcaacgacg 2520
gcgacgtcgt gtccaccaac acctacttca tgaccaccgg gaacgccctg ggctccgtga 2580
acatgacgac gggggtggac aacctgttct acatcgacaa gttccaggtg cgcgaggtca 2640
agtgacaatt ggcagcagca gctcggatag tatcgacaca ctctggacgc tggtcgtgtg 2700
atggactgtt gccgccacac ttgctgcctt gacctgtgaa tatccctgcc gcttttatca 2760
aacagcctca gtgtgtttga tcttgtgtgt acgcgctttt gcgagttgct agctgcttgt 2820
gctatttgcg aataccaccc ccagcatccc cttccctcgt ttcatatcgc ttgcatccca 2880
accgcaactt atctacgctg tcctgctatc cctcagcgct gctcctgctc ctgctcactg 2940
cccctcgcac agccttggtt tgggctccgc ctgtattctc ctggtactgc aacctgtaaa 3000
ccagcactgc aatgctgatg cacgggaagt agtgggatgg gaacacaaat ggaggatccc 3060
gcgtctcgaa cagagcgcgc agaggaacgc tgaaggtctc gcctctgtcg cacctcagcg 3120
cggcatacac cacaataacc acctgacgaa tgcgcttggt tcttcgtcca ttagcgaagc 3180
gtccggttca cacacgtgcc acgttggcga ggtggcaggt gacaatgatc ggtggagctg 3240
atggtcgaaa cgttcacagc ctagggatat cgaattcctt tcttgcgcta tgacacttcc 3300
agcaaaaggt agggcgggct gcgagacggc ttcccggcgc tgcatgcaac accgatgatg 3360
cttcgacccc ccgaagctcc ttcggggctg catgggcgct ccgatgccgc tccagggcga 3420
gcgctgttta aatagccagg cccccgattg caaagacatt atagcgagct accaaagcca 3480
tattcaaaca cctagatcac taccacttct acacaggcca ctcgagcttg tgatcgcact 3540
ccgctaaggg ggcgcctctt cctcttcgtt tcagtcacaa cccgcaaaca ctagtgggtc 3600
gccgcgctgc ctcgcgtccc ctggtggtgc acgccgtcgc cagcgaggcc ccgctgggcg 3660
tgccgccctc ggtgcagcgc ccctcccccg tggtctactc caagctggac aagcaacacc 3720
gcctgacgcc cgagcgcctg gagctggtgc agagcatggg gcagtttgcg gaggagaggt 3780
gtgtttgagg gttttggttg cccgtatcga ggtcctggtg gcgcgcatgg gggagaaggc 3840
gcctgtcccg ctgacccccc cggctaccct cccggcacct tccagggcgc gtacgggatc 3900
cctctcctcc acaaactgcc ccatgctctg caccagctcc aggcgctcgg gcgtcaggcg 3960
gtgctgcttg tccagcttgg agtagaccac gggggagggg cgctgcaccg agggcggaac 4020
gcccagcggg gcctcgctgg cgaccgcgcg caccaccagg ggacgcgagg cagcgcggcg 4080
acccaagctt gcagcagcag ctcggatagt atcgacacac tctggacgct ggtcgtgtga 4140
tggactgttg ccgccacact tgctgccttg acctgtgaat atccctgccg cttttatcaa 4200
acagcctcag tgtgtttgat cttgtgtgta cgcgcttttg cgagttgcta gctgcttgtg 4260
ctatttgcga ataccaccccc cagcatcccc ttccctcgtt tcatatcgct tgcatcccaa 4320
ccgcaactta tctacgctgt cctgctatcc ctcagcgctg ctcctgctcc tgctcactgc 4380
ccctcgcaca gccttggttt gggctccgcc tgtattctcc tggtactgca acctgtaaac 4440

cagcactgca atgctgatgc acgggaagta gtgggatggg aacacaaatg gaaagctgga 4500
gctcttgttt tccagaagga gttgctcctt gagcctttca ttctcagcct cgataacctc 4560
caaagccgct ctaattgtgg agggggttcg aatttaaaag cttggaatgt tggttcgtgc 4620
gtctggaaca agcccagact tgttgctcac tgggaaaagg accatcagct ccaaaaaact 4680
tgccgctcaa accgcgtacc tctgctttcg cgcaatctgc cctgttgaaa tcgccaccac 4740
attcatattg tgacgcttga gcagtctgta attgcctcag aatgtggaat catctgcccc 4800
ctgtgcgagc ccatgccagg catgtcgcgg gcgaggacac ccgccactcg tacagcagac 4860
cattatgcta cctcacaata gttcataaca gtgaccatat ttctcgaagc tccccaacga 4920
gcacctccat gctctgagtg gccacccccc ggccctggtg cttgcggagg gcaggtcaac 4980
cggcatgggg ctaccgaaat ccccgaccgg atcccaccac ccccgcgatg ggaagaatct 5040
ctccccggga tgtgggccca ccaccagcac aacctgctgg cccaggcgag cgtcaaacca 5100
taccacacaa atatccttgg catcggccct gaattccttc tgccgctctg ctacccggtg 5160
cttctgtccg aagcaggggt tgctagggat cgctccgagt ccgcaaaccc ttgtcgcgtg 5220
gcggggcttg ttcgagcttg aagagc 5246

pSZ2140的部分,Prototheca moriformis SAD2-发夹B

<210> 227
<211> 615
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 227

actagttgcc cgtgctgcac cccgtggaca agctgtggca gccgcaggac ttcctgcccg 60
accccgagtc gcccgacttc gaggaccagg tggcggagct gcgcgcgcgc gccaaggacc 120
tgcccgacga gtactttgtg gtgctggtgg gcgacatgat cacggaggag gcgctgccga 180
cctacatggc catgctcaac accttggacg gtgtgcgcga cgacacgggc gcggccgacc 240
acccgtgtgt gtttgagggt tttggttgcc cgtattgagg tcctggtggc gcgcatggag 300
gagaaggcgc ctgtcccgct gacccccccg gctaccctcc cggcaccttc cagggcgcgt 360
acgggatcca cgggtggtcg gccgcgcccg tgtcgtcgcg cacaccgtcc aaggtgttga 420
gcatggccat gtaggtcggc agcgcctcct ccgtgatcat gtcgcccacc agcaccacaa 480
agtactcgtc gggcaggtcc ttggcgcgcg cgcgcagctc cgccacctgg tcctcgaagt 540
cgggcgactc ggggtcgggc aggaagtcct gcggctgcca cagcttgtcc acggggtgca 600
gcacgggcaa agctt 615

pSZ2141的部分,Prototheca moriformis SAD2-发夹C

<210> 228
<211> 615
<212> DNA
<213> 人工序列
<220>

<223> 合成的构建体

<400> 228

actagtgcgc tggacgcggc agtgggtggc cgaggagaac cggcacggcg acctgctgaa 60
caagtactgt tggctgacgg ggcgcgtcaa catgcgggcc gtggaggtga ccatcaacaa 120
cctgatcaag agcggcatga acccgcagac ggacaacaac ccttacttgg gcttcgtcta 180
cacctccttc caggagcgcg cgaccaagta cagccacggc aacaccgcgc gccttgcggc 240
cgagcagtgt gtttgagggt tttggttgcc cgtatcgagg tcctggtggc gcgcatgggg 300
gagaaggcgc ctgtcccgct gacccccccg gctaccctcc cggcaccttc cagggcgcgt 360
acgggatcct gctcggccgc aaggcgcgcg gtgttgccgt ggctgtactt ggtcgcgcgc 420
tcctggaagg aggtgtagac gaagcccaag taagggttgt tgtccgtctg cgggttcatg 480
ccgctcttga tcaggttgtt gatggtcacc tccacggccc gcatgttgac gcgccccgtc 540
agccaacagt acttgttcag caggtcgccg tgccggttct cctcggccac ccactgccgc 600
gtccagcgca agctt 615

pSZ2142的部分,Prototheca moriformis SAD2-发夹D

<210> 229

<211> 613

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 229

actagtgctg atcgccagcg acgagggccg gcacgagatc gcctacacgc gcatcgtgga 60
cgagttcttc cgcctcgacc ccgagggcgc cgtcgccgcc tacgccaaca tgatgcgcaa 120
gcagatcacc atgcccgcgc acctcatgga cgacatgggc cacggcgagg ccaacccggg 180
ccgcaacctc ttcgccgact tctccgcggt cgccgagaag atcgacgtct acgacgccga 240
ggactgtgtg tttgagggtt ttggttgccc gtatcgaggt cctggtggcg cgcatggggg 300
agaaggcgcc tgtcccgctg acccccccgg ctaccctccc ggcaccttcc agggcgcgta 360
cgggatccag tcctcggcgt cgtagacgtc gatcttctcg gcgaccgcgg agaagtcggc 420
gaagaggttg cggcccgggt tggcctcgcc gtggcccatg tcgtccatga ggtgcgcggg 480
catggtgatc tgcttgcgca tcatgttggc gtaggcggcg acggcgccct cggggtcgag 540
gcggaagaac tcgtccacga tgcgcgtgta ggcgatctcg tgccggccct cgtcgctggc 600
gatcagcaag ctt 613

包括密码子优化的大豆KASII 的pSZ1747的部分

<210> 230

<211> 1503

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 230

atggccaccg catccacttt ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg 60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgcctccacc 120
cactccggca agaccatggc cgtggccctg cagcccaccc aggagatcac caccatcaag 180
aagcccccca ccaagcagcg ccgcgtggtg gtgaccggcc tgggcgtggt gacccccctg 240
ggccacgagc ccgacatctt ctacaacaac ctgctggacg gcgcctccgg catctccgag 300
atcgagacct tcgactgcgc cgagtacccc acccgcatcg ccggcgagat caagtccttc 360
tccaccgacg gctgggtggc ccccaagctg tccaagcgca tggacaagtt catgctgtac 420
atgctgaccg ccggcaagaa ggccctggtg gacggcggca tcaccgacga cgtgatggac 480
gagctgaaca aggagaagtg cggcgtgctg atcggctccg ccatgggcgg catgaaggtg 540
ttcaacgacg ccatcgaggc cctgcgcatc tcctacaaga agatgaaccc cttctgcgtg 600
cccttcgcca ccaccaacat gggctccgcc atgctggcca tggacctggg ctggatgggc 660
cccaactact ccatctccac cgcctgcgcc acctccaact tctgcatcct gaacgccgcc 720
aaccacatca tccgcggcga ggccgacgtg atgctgtgcg gcggctccga cgccgccatc 780
atccccatcg gcctgggcgg cttcgtggcc tgccgcgccc tgtcccagcg caacaccgac 840
cccaccaagg cctcccgccc ctgggacatc aaccgcgacg gcttcgtgat gggcgagggc 900
gccggcgtgc tgctgctgga ggagctggag cacgccaagg agcgcggcgc caccatctac 960
gccgagttcc tgggcggctc cttcacctgc gacgcctacc acgtgaccga gccccgcccc 1020
gacggcgccg gcgtgatcct gtgcatcgag aaggccctgg cccagtccgg cgtgtccaag 1080
gaggacgtga actacatcaa cgcccacgcc acctccaccc ccgccggcga cctgaaggag 1140
taccaggccc tgatgcactg cttcggccag aaccccgagc tgcgcgtgaa ctccaccaag 1200
tccatgatcg gccacctgct gggcgccgcc ggcggcgtgg aggccgtggc caccatccag 1260
gccatccgca ccggctgggt gcaccccaac atcaacctgg agaaccccga caacggcgtg 1320
gacgccaagg tgctggtggg ctccaagaag gagcgcctgg acgtgaaggc cgccctgtcc 1380
aactccttcg gcttcggcgg ccacaactcc tccatcatct tcgcccccta catggactac 1440
aaggaccacg acggcgacta caaggaccac gacatcgact acaaggacga cgacgacaag 1500
tga 1503

包括密码子优化的向日葵KASII 的pSZ1750的相关部分

<210> 231
<211> 1530
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 231

atggccaccg catccacttt ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg 60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgcctgccgc 120
cgcggcggcc gcgtggccat ggccatcgcc atccagccct cctccatcga gatggaggag 180
gagaccaccc tgaccaagcg caagcagccc cccaccaagc agcgccgcgt ggtggtgacc 240

ggcatgggcg tggagacccc catcggcaac aaccccgacc agttctacaa caacctgctg 300
cagggcgtgt ccggcatcac ccagatcgag gccttcgact gctcctccta ccccacccgc 360
atcgccggcg agatcaagaa cttctccacc gacggctggg tggcccccaa gctgtccaag 420
cgcatggacc gcttcatgct gtacatgctg accgccggca agaaggccct ggccgacgcc 480
ggcatctccc cctccgactc cgacgagatc gacaagtccc gctgcggcgt gctgatcggc 540
tccgccatgg gcggcatgaa ggtgttcaac gacgccatcg aggccctgcg cgtgtcctac 600
cgcaagatga acccttctg cgtgcccttc gccaccacca acatgggctc cgccatgctg 660
gccatggacc tgggctggat gggccccaac tactccatct ccaccgcctg cgccacctcc 720
aacttctgca tcctgaacgc cgccaaccac atcatccgcg gcgaggccga catgatgctg 780
tgcggcggct ccgacgccgt gatcatcccc atcggcctgg gcggcttcgt ggcctgccgc 840
gccctgtccg agcgcaacac cgaccccgcc aaggcctccc gcccctggga ctccggccgc 900
gacggcttcg tgatgggcga gggcgccggc gtgctgctgc tggaggagct ggagcacgcc 960
aagaagcgcg gcgccaagat ctacgccgag ttcctgggcg gctccttcac ctgcgacgcc 1020
taccacatga ccgagcccca ccccgagggc gccggcgtga tcctgtgcat cgagaaggcc 1080
ctgtcccagg ccggcgtgcg ccgcgaggac gtgaactaca tcaacgccca cgccacctcc 1140
acccccgccg gcgacctgaa ggagtaccac gccctgctgc actgcttcgg caacaaccag 1200
gagctgcgcg tgaactccac caagtccatg atcggccacc tgctgggcgc cgccggcgcc 1260
gtggaggccg tggccaccgt gcaggccatc cgcaccggct ggatccaccc caacatcaac 1320
ctggagaacc ccgaccaggg cgtggacacc aaggtgctgg tgggctccaa gaaggagcgc 1380
ctgaacgtga aggtgggcct gtccaactcc ttcggcttcg gcggccacaa ctcctccatc 1440
ctgttcgccc ccttccagat ggactacaag gaccacgacg gcgactacaa ggaccacgac 1500
atcgactaca aggacgacga cgacaagtga 1530

包括密码子优化的蓖麻KASII 的pSZ1754的部分

<210> 232
<211> 1584
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 232

atggccaccg catccacttt ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg 60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgcctcccac 120
tactactcct ccaacggcct gttccccaac acccccctgc tgcccaagcg ccacccccgc 180
ctgcaccacc gcctgccccg ctccggcgag gccatggccg tggccgtgca gcccgagaag 240
gaggtggcca ccaacaagaa gcccctgatg aagcagcgcc gcgtggtggt gaccggcatg 300
ggcgtggtgt ccccccctggg ccacgacatc gacgtgtact acaacaacct gctggacggc 360
tcctccggca tctcccagat cgactccttc gactgcgccc agttccccac ccgcatcgcc 420
ggcgagatca agtccttctc caccgacggc tgggtggccc ccaagctgtc caagcgcatg 480
gacaagttca tgctgtacat gctgaccgcc ggcaagaagg ccctggccga cggcggcatc 540

accgaggaca tgatggacga gctggacaag gcccgctgcg gcgtgctgat cggctccgcc 600
atgggcggca tgaaggtgtt caacgacgcc atcgaggccc tgcgcatctc ctaccgcaag 660
atgaacccct tctgcgtgcc cttcgccacc accaacatgg gctccgccat gctggccatg 720
gacctgggct ggatgggccc caactactcc atctccaccg cctgcgccac ctccaacttc 780
tgcatcctga acgccgccaa ccacatcatc cgcggcgagg ccgacatcat gctgtgcggc 840
ggctccgacg ccgccatcat ccccatcggc ctgggcggct tcgtggcctg ccgcgccctg 900
tcccagcgca acgacgaccc caccaaggcc tcccgccccc gggacatgaa ccgcgacggc 960
ttcgtgatgg gcgagggcgc cggcgtgctg ctgctggagg agctggagca cgccaagaag 1020
cgcggcgcca acatctacgc cgagttcctg ggcggctcct tcacctgcga cgcctaccac 1080
atgaccgagc cccgccccga cggcgtgggc gtgatcctgt gcatcgagaa ggccctggcc 1140
cgctccggcg tgtccaagga ggaggtgaac tacatcaacg cccacgccac ctccaccccc 1200
gccggcgacc tgaaggagta cgaggccctg atgcgctgct tctcccagaa ccccgacctg 1260
cgcgtgaact ccaccaagtc catgatcggc cacctgctgg gcgccgccgg cgccgtggag 1320
gccatcgcca ccatccaggc catccgcacc ggctgggtgc accccaacat caacctggag 1380
aaccccgagg agggcgtgga caccaaggtg ctggtgggcc ccaagaagga gcgcctggac 1440
atcaaggtgg ccctgtccaa ctccttcggc ttcggcggcc acaactcctc catcatcttc 1500
gccccctaca agatggacta caaggaccac gacggcgact acaaggacca cgacatcgac 1560
tacaaggacg acgacgacaa gtga 1584

包括密码子优化的Prototheca moriformis KASII 的pSZ2041转化载体的相关部分

<210> 233
<211> 7008
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 233

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60
ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120
tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180
ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240
gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300
ggaagacagg tgagggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360
atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420
cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480
gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540
ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600
ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660
cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720
ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780

tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 960
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 1020
cagtcacaac ccgcaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg 1080
gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg 1140
gtgcacttca cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag 1200
aaggacgcca agtggcacct gtacttccag tacaacccga acgacaccgt ctgggggacg 1260
cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc 1320
gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac 1380
aacaacacct ccggcttctt caacgacacc atcgacccgc gccagcgctg cgtggccatc 1440
tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc 1500
tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc 1560
gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc 1620
caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc 1680
gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc 1740
cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc 1800
gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg cacccacttc 1860
gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag 1920
accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcgtg ggcctccaac 1980
tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc 2040
aagttctccc tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag 2100
gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc 2160
acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag 2220
ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac 2280
ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag 2340
gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag 2400
aacccctact tcaccaaccg catgagcgtg aacaaccagc ccttcaagag cgagaacgac 2460
ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac 2520
gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc 2580
gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag 2640
gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg 2700
tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt 2760
atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc 2820
ttgtgctatt tgcgaatacc acccccagca tccccttccc tcgtttcata tcgcttgcat 2880
cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc 2940
actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg 3000
taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga 3060
tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgcctct gtcgcacctc 3120

agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg 3180
aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga 3240
gctgatggtc gaaacgttca cagcctaggg atatcgaatt cggccgacag gacgcgcgtc 3300
aaaggtgctg gtcgtgtatg ccctggccgg caggtcgttg ctgctgctgg ttagtgattc 3360
cgcaaccctg attttggcgt cttattttgg cgtggcaaac gctggcgccc gcgagccggg 3420
ccggcggcga tgcggtgccc cacggctgcc ggaatccaag ggaggcaaga gcgcccgggt 3480
cagttgaagg gctttacgcg caaggtacag ccgctcctgc aaggctgcgt ggtggaattg 3540
gacgtgcagg tcctgctgaa gttcctccac cgcctcacca gcggacaaag caccggtgta 3600
tcaggtccgt gtcatccact ctaaagaact cgactacgac ctactgatgg ccctagattc 3660
ttcatcaaaa acgcctgaga cacttgccca ggattgaaac tccctgaagg gaccaccagg 3720
ggccctgagt tgttccttcc ccccgtggcg agctgccagc caggctgtac ctgtgatcga 3780
ggctggcggg aaaataggct tcgtgtgctc aggtcatggg aggtgcagga cagctcatga 3840
aacgccaaca atcgcacaat tcatgtcaag ctaatcagct atttcctctt cacgagctgt 3900
aattgtccca aaattctggt ctaccggggg tgatccttcg tgtacgggcc cttccctcaa 3960
ccctaggtat gcgcgcatgc ggtcgccgcg caactcgcgc gagggccgag ggtttgggac 4020
gggccgtccc gaaatgcagt tgcacccgga tgcgtggcac ctttttgcg ataatttatg 4080
caatggactg ctctgcaaaa ttctggctct gtcgccaacc ctaggatcag cggcgtagga 4140
tttcgtaatc attcgtcctg atggggagct accgactacc ctaatatcag cccgactgcc 4200
tgacgccagc gtccactttt gtgcacacat tccattcgtg cccaagacat ttcattgtgg 4260
tgcgaagcgt ccccagttac gctcacctgt ttcccgacct ccttactgtt ctgtcgacag 4320
agcgggccca caggccggtc gcagccacta gtatggccac cgcatccact ttctcggcgt 4380
tcaatgcccg ctgcggcgac ctgcgtcgct cggcgggctc cgggccccgg cgcccagcga 4440
ggcccctccc cgtgcgcggg cgcgccgccg ccgccgccga cgccaacccc gcccgccccg 4500
agcgccgcgt ggtgatcacc ggccagggcg tggtgacctc cctgggccag accatcgagc 4560
agttctactc ctccctgctg gagggcgtgt ccggcatctc ccagatccag aagttcgaca 4620
ccaccggcta caccaccacc atcgccggcg agatcaagtc cctgcagctg gacccctacg 4680
tgcccaagcg ctgggccaag cgcgtggacg acgtgatcaa gtacgtgtac atcgccggca 4740
agcaggccct ggagtccgcc ggcctgccca tcgaggccgc cggcctggcc ggcgccggcc 4800
tggaccccgc cctgtgcggc gtgctgatcg gcaccgccat ggccggcatg acctccttcg 4860
ccgccggcgt ggaggccctg acccgcggcg gcgtgcgcaa gatgaacccc ttctgcatcc 4920
ccttctccat ctccaacatg ggcggcgcca tgctggccat ggacatcggc ttcatgggcc 4980
ccaactactc catctccacc gcctgcgcca ccggcaacta ctgcatcctg ggcgccgccg 5040
accacatccg ccgcggcgac gccaacgtga tgctggccgg cggcgccgac gccgccatca 5100
tcccctccgg catcggcggc ttcatcgcct gcaaggccct gtccaagcgc aacgacgagc 5160
ccgagcgcgc ctcccgcccc tgggacgccg accgcgacgg cttcgtgatg ggcgagggcg 5220
ccggcgtgct ggtgctggag gagctggagc acgccaagcg ccgcggcgcc accatcctgg 5280
ccgagctggt gggcggcgcc gccacctccg acgcccacca catgaccgag cccgaccccc 5340
agggccgcgg cgtgcgcctg tgcctggagc gcgccctgga gcgcgcccgc ctggcccccg 5400
agcgcgtggg ctacgtgaac gcccacggca cctccacccc cgccggcgac gtggccgagt 5460

accgcgccat ccgcgccgtg atcccccagg actccctgcg catcaactcc accaagtcca 5520
tgatcggcca cctgctgggc ggcgccggcg ccgtggaggc cgtggccgcc atccaggccc 5580
tgcgcaccgg ctggctgcac cccaacctga acctggagaa ccccgccccc ggcgtggacc 5640
ccgtggtgct ggtgggcccc cgcaaggagc gcgccgagga cctggacgtg gtgctgtcca 5700
actccttcgg cttcggcggc cacaactcct gcgtgatctt ccgcaagtac gacgagatgg 5760
actacaagga ccacgacggc gactacaagg accacgacat cgactacaag gacgacgacg 5820
acaagtgaat cgatagatct cttaaggcag cagcagctcg gatagtatcg acacactctg 5880
gacgctggtc gtgtgatgga ctgttgccgc cacacttgct gccttgacct gtgaatatcc 5940
ctgccgcttt tatcaaacag cctcagtgtg tttgatcttg tgtgtacgcg cttttgcgag 6000
ttgctagctg cttgtgctat ttgcgaatac cacccccagc atccccttcc ctcgtttcat 6060
atcgcttgca tcccaaccgc aacttatcta cgctgtcctg ctatccctca gcgctgctcc 6120
tgctcctgct cactgccccc cgcacagcct tggtttgggc tccgcctgta ttctcctggt 6180
actgcaacct gtaaaccagc actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca 6240
caaatggaaa gcttaattaa gagctcttgt tttccagaag gagttgctcc ttgagccttt 6300
cattctcagc ctcgataacc tccaaagccg ctctaattgt ggaggggttt cgaatttaaa 6360
agcttggaat gttggttcgt gcgtctggaa caagcccaga cttgttgctc actgggaaaa 6420
ggaccatcag ctccaaaaaa cttgccgctc aaaccgcgta cctctgcttt cgcgcaatct 6480
gccctgttga aatcgccacc acattcatat tgtgacgctt gagcagtctg taattgcctc 6540
agaatgtgga atcatctgcc ccctgtgcga gcccatgcca ggcatgtcgc gggcgaggac 6600
acccgccact cgtacagcag accattatgc tacctcacaa tagttcataa cagtgaccat 6660
atttctcgaa gctccccaac gagcacctcc atgctctgag tggccacccc ccggccctgg 6720
tgcttgcgga gggcaggtca accggcatgg ggctaccgaa atccccgacc ggatcccacc 6780
acccccgcga tgggaagaat ctctccccgg gatgtgggcc caccaccagc acaacctgct 6840
ggcccaggcg agcgtcaaac cataccacac aaatatcctt ggcatcggcc ctgaattcct 6900
tctgccgctc tgctacccgg tgcttctgtc cgaagcaggg gttgctaggg atcgctccga 6960
gtccgcaaac ccttgtcgcg tggcggggct tgttcgagct tgaagagc 7008

<210> 234
<211> 7008
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 234

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60
ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120
tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180
ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240
gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300
ggaagacagg tgaggggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360

atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420
cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480
gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540
ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600
ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660
cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720
ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 960
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 1020
cagtcacaac ccgcaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg 1080
gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg 1140
gtgcacttca cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag 1200
aaggacgcca agtggcacct gtacttccag tacaacccga acgacaccgt ctgggggacg 1260
cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc 1320
gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac 1380
aacaacacct ccggcttctt caacgacacc atcgacccgc gccagcgctg cgtggccatc 1440
tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc 1500
tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc 1560
gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc 1620
caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc 1680
gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc 1740
cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc 1800
gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg cacccacttc 1860
gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag 1920
accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcgtg ggcctccaac 1980
tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc 2040
aagttctccc tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag 2100
gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc 2160
acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag 2220
ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac 2280
ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag 2340
gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag 2400
aacccctact tcaccaaccg catgagcgtg aacaaccagc ccttcaagag cgagaacgac 2460
ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac 2520
gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc 2580
gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag 2640
gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg 2700

tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt 2760
atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc 2820
ttgtgctatt tgcgaatacc acccccagca tcccccttccc tcgtttcata tcgcttgcat 2880
cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc 2940
actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg 3000
taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga 3060
tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgcctct gtcgcacctc 3120
agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg 3180
aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga 3240
gctgatggtc gaaacgttca cagcctaggg atatcgaatt cggccgacag gacgcgcgtc 3300
aaaggtgctg gtcgtgtatg ccctggccgg caggtcgttg ctgctgctgg ttagtgattc 3360
cgcaaccctg attttggcgt cttattttgg cgtggcaaac gctggcgccc gcgagccggg 3420
ccggcggcga tgcggtgccc cacggctgcc ggaatccaag ggaggcaaga gcgcccgggt 3480
cagttgaagg gctttacgcg caaggtacag ccgctcctgc aaggctgcgt ggtggaattg 3540
gacgtgcagg tcctgctgaa gttcctccac cgcctcacca gcggacaaag caccggtgta 3600
tcaggtccgt gtcatccact ctaaagaact cgactacgac ctactgatgg ccctagattc 3660
ttcatcaaaa acgcctgaga cacttgccca ggattgaaac tccctgaagg gaccaccagg 3720
ggccctgagt tgttccttcc ccccgtggcg agctgccagc caggctgtac ctgtgatcga 3780
ggctggcggg aaaataggct tcgtgtgctc aggtcatggg aggtgcagga cagctcatga 3840
aacgccaaca atcgcacaat tcatgtcaag ctaatcagct atttcctctt cacgagctgt 3900
aattgtccca aaattctggt ctaccggggg tgatccttcg tgtacgggcc cttccctcaa 3960
ccctaggtat gcgcgcatgc ggtcgccgcg caactcgcgc gagggccgag ggtttgggac 4020
gggccgtccc gaaatgcagt tgcacccgga tgcgtggcac ctttttgcg ataatttatg 4080
caatggactg ctctgcaaaa ttctggctct gtcgccaacc ctaggatcag cggcgtagga 4140
tttcgtaatc attcgtcctg atggggagct accgactacc ctaatatcag cccgactgcc 4200
tgacgccagc gtccactttt gtgcacacat tccattcgtg cccaagacat ttcattgtgg 4260
tgcgaagcgt ccccagttac gctcacctgt ttcccgacct ccttactgtt ctgtcgacag 4320
agcgggccca caggccggtc gcagccacta gtatggccac cgcatccact ttctcggcgt 4380
tcaatgcccg ctgcggcgac ctgcgtcgct cggcgggctc cgggccccgg cgcccagcga 4440
ggcccctccc cgtgcgcggg cgcgccgccg ccgccgccga cgccaacccc gcccgccccg 4500
agcgccgcgt ggtgatcacc ggccagggcg tggtgacctc cctgggccag accatcgagc 4560
agttctactc ctccctgctg gagggcgtgt ccggcatctc ccagatccag aagttcgaca 4620
ccaccggcta caccaccacc atcgccggcg agatcaagtc cctgcagctg gacccctacg 4680
tgcccaagcg ctgggccaag cgcgtggacg acgtgatcaa gtacgtgtac atcgccggca 4740
agcaggccct ggagtccgcc ggcctgccca tcgaggccgc cggcctggcc ggcgccggcc 4800
tggaccccgc cctgtgcggc gtgctgatcg gcaccgccat ggccggcatg acctccttcg 4860
ccgccggcgt ggaggccctg acccgcggcg gcgtgcgcaa gatgaacccc ttctgcatcc 4920
ccttctccat ctccaacatg ggcggcgcca tgctggccat ggacatcggc ttcatgggcc 4980
ccaactactc catctccacc gcctgcgcca ccggcaacta ctgcatcctg ggcgccgccg 5040

accacatccg ccgcggcgac gccaacgtga tgctggccgg cggcgccgac gccgccatca 5100
tcccctccgg catcggcggc ttcatcgcct gcaaggccct gtccaagcgc aacgacgagc 5160
ccgagcgcgc ctcccgcccc tgggacgccg accgcgacgg cttcgtgatg ggcgagggcg 5220
ccggcgtgct ggtgctggag gagctggagc acgccaagcg ccgcggcgcc accatcctgg 5280
ccgagctggt gggcggcgcc gccacctccg acgcccacca catgaccgag cccgaccccc 5340
agggccgcgg cgtgcgcctg tgcctggagc gcgccctgga gcgcgcccgc ctggcccccg 5400
agcgcgtggg ctacgtgaac gcccacggca cctccacccc cgccggcgac gtggccgagt 5460
accgcgccat ccgcgccgtg atcccccagg actccctgcg catcaactcc accaagtcca 5520
tgatcggcca cctgctgggc ggcgccggcg ccgtggaggc cgtggccgcc atccaggccc 5580
tgcgcaccgg ctggctgcac cccaacctga acctggagaa ccccgccccc ggcgtggacc 5640
ccgtggtgct ggtgggcccc cgcaaggagc gcgccgagga cctggacgtg gtgctgtcca 5700
actccttcgg cttcggcggc cacaactcct gcgtgatctt ccgcaagtac gacgagatgg 5760
actacaagga ccacgacggc gactacaagg accacgacat cgactacaag gacgacgacg 5820
acaagtgaat cgatagatct cttaaggcag cagcagctcg gatagtatcg acacactctg 5880
gacgctggtc gtgtgatgga ctgttgccgc cacacttgct gccttgacct gtgaatatcc 5940
ctgccgcttt tatcaaacag cctcagtgtg tttgatcttg tgtgtacgcg cttttgcgag 6000
ttgctagctg cttgtgctat ttgcgaatac cacccccagc atcccccttcc ctcgtttcat 6060
atcgcttgca tcccaaccgc aacttatcta cgctgtcctg ctatccctca gcgctgctcc 6120
tgctcctgct cactgccccct cgcacagcct tggtttgggc tccgcctgta ttctcctggt 6180
actgcaacct gtaaaccagc actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca 6240
caaatggaaa gcttaattaa gagctcttgt tttccagaag gagttgctcc ttgagccttt 6300
cattctcagc ctcgataacc tccaaagccg ctctaattgt ggaggggttt cgaatttaaa 6360
agcttggaat gttggttcgt gcgtctggaa caagcccaga cttgttgctc actgggaaaa 6420
ggaccatcag ctccaaaaaa cttgccgctc aaaccgcgta cctctgcttt cgcgcaatct 6480
gccctgttga aatcgccacc acattcatat tgtgacgctt gagcagtctg taattgcctc 6540
agaatgtgga atcatctgcc ccctgtgcga gcccatgcca ggcatgtcgc gggcgaggac 6600
acccgccact cgtacagcag accattatgc tacctcacaa tagttcataa cagtgaccat 6660
atttctcgaa gctccccaac gagcacctcc atgctctgag tggccacccc ccggccctgg 6720
tgcttgcgga gggcaggtca accggcatgg ggctaccgaa atccccgacc ggatcccacc 6780
acccccgcga tgggaagaat ctctccccgg gatgtgggcc caccaccagc acaacctgct 6840
ggcccaggcg agcgtcaaac cataccacac aaatatcctt ggcatcggcc ctgaattcct 6900
tctgccgctc tgctacccgg tgcttctgtc cgaagcaggg gttgctaggg atcgctccga 6960
gtccgcaaac ccttgtcgcg tggcggggct tgttcgagct tgaagagc 7008

<210> 235
<211> 1476
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体

<400> 235

atggccaccg catccacttt ctcggcgttc aatgcccgct gcggcgacct gcgtcgctcg 60
gcgggctccg ggccccggcg cccagcgagg cccctccccg tgcgcgggcg cgccgccgcc 120
gccgccgacg ccaaccccgc ccgccccgag cgccgcgtgg tgatcaccgg ccagggcgtg 180
gtgacctccc tgggccagac catcgagcag ttctactcct ccctgctgga gggcgtgtcc 240
ggcatctccc agatccagaa gttcgacacc accggctaca ccaccaccat cgccggcgag 300
atcaagtccc tgcagctgga cccctacgtg cccaagcgct gggccaagcg cgtggacgac 360
gtgatcaagt acgtgtacat cgccggcaag caggccctgg agtccgccgg cctgcccatc 420
gaggccgccg gcctggccgg cgccggcctg gaccccgccc tgtgcggcgt gctgatcggc 480
accgccatgg ccggcatgac ctccttcgcc gccggcgtgg aggccctgac ccgcggcggc 540
gtgcgcaaga tgaacccctt ctgcatcccc ttctccatct ccaacatggg cggcgccatg 600
ctggccatgg acatcggctt catgggcccc aactactcca tctccaccgc ctgcgccacc 660
ggcaactact gcatcctggg cgccgccgac cacatccgcc gcggcgacgc caacgtgatg 720
ctggccggcg gcgccgacgc cgccatcatc ccctccggca tcggcggctt catcgcctgc 780
aaggccctgt ccaagcgcaa cgacgagccc gagcgcgcct cccgcccctg ggacgccgac 840
cgcgacggct tcgtgatggg cgagggcgcc ggcgtgctgg tgctggagga gctggagcac 900
gccaagcgcc gcggcgccac catcctggcc gagctggtgg gcggcgccgc cacctccgac 960
gcccaccaca tgaccgagcc cgacccccag ggccgcggcg tgcgcctgtg cctggagcgc 1020
gccctggagc gcgcccgcct ggcccccgag cgcgtgggct acgtgaacgc ccacggcacc 1080
tccacccccg ccggcgacgt ggccgagtac cgcgccatcc gcgccgtgat cccccaggac 1140
tccctgcgca tcaactccac caagtccatg atcggccacc tgctgggcgg cgccggcgcc 1200
gtggaggccg tggccgccat ccaggccctg cgcaccggct ggctgcaccc caacctgaac 1260
ctggagaacc ccgccccccgg cgtggacccc gtggtgctgg tgggcccccg caaggagcgc 1320
gccgaggacc tggacgtggt gctgtccaac tccttcggct tcggcggcca caactcctgc 1380
gtgatcttcc gcaagtacga cgagatggac tacaaggacc acgacggcga ctacaaggac 1440
cacgacatcg actacaagga cgacgacgac aagtga 1476

<210> 236
<211> 491
<212> PRT
<213> 人工序列
<220>
<223> 合成的构建体
<400> 236

Met Ala Thr Ala Ser Thr Phe Ser Ala Phe Asn Ala Arg Cys Gly Asp
1 5 10 15
Leu Arg Arg Ser Ala Gly Ser Gly Pro Arg Arg Pro Ala Arg Pro Leu
20 25 30
Pro Val Arg Gly Arg Ala Ala Ala Ala Ala Asp Ala Asn Pro Ala Arg
35 40 45

Pro Glu Arg Arg Val Val Ile Thr Gly Gln Gly Val Val Thr Ser Leu
50 55 60
Gly Gln Thr Ile Glu Gln Phe Tyr Ser Ser Leu Leu Glu Gly Val Ser
65 70 75 80
Gly Ile Ser Gln Ile Gln Lys Phe Asp Thr Thr Gly Tyr Thr Thr Thr
85 90 95
Ile Ala Gly Glu Ile Lys Ser Leu Gln Leu Asp Pro Tyr Val Pro Lys
100 105 110
Arg Trp Ala Lys Arg Val Asp Asp Val Ile Lys Tyr Val Tyr Ile Ala
115 120 125
Gly Lys Gln Ala Leu Glu Ser Ala Gly Leu Pro Ile Glu Ala Ala Gly
130 135 140
Leu Ala Gly Ala Gly Leu Asp Pro Ala Leu Cys Gly Val Leu Ile Gly
145 150 155 160
Thr Ala Met Ala Gly Met Thr Ser Phe Ala Ala Gly Val Glu Ala Leu
165 170 175
Thr Arg Gly Gly Val Arg Lys Met Asn Pro Phe Cys Ile Pro Phe Ser
180 185 190
Ile Ser Asn Met Gly Gly Ala Met Leu Ala Met Asp Ile Gly Phe Met
195 200 205
Gly Pro Asn Tyr Ser Ile Ser Thr Ala Cys Ala Thr Gly Asn Tyr Cys
210 215 220
Ile Leu Gly Ala Ala Asp His Ile Arg Arg Gly Asp Ala Asn Val Met
225 230 235 240
Leu Ala Gly Gly Ala Asp Ala Ala Ile Ile Pro Ser Gly Ile Gly Gly
245 250 255
Phe Ile Ala Cys Lys Ala Leu Ser Lys Arg Asn Asp Glu Pro Glu Arg
260 265 270
Ala Ser Arg Pro Trp Asp Ala Asp Arg Asp Gly Phe Val Met Gly Glu
275 280 285
Gly Ala Gly Val Leu Val Leu Glu Glu Leu Glu His Ala Lys Arg Arg
290 295 300
Gly Ala Thr Ile Leu Ala Glu Leu Val Gly Gly Ala Ala Thr Ser Asp
305 310 315 320
Ala His His Met Thr Glu Pro Asp Pro Gln Gly Arg Gly Val Arg Leu
325 330 335
Cys Leu Glu Arg Ala Leu Glu Arg Ala Arg Leu Ala Pro Glu Arg Val
340 345 350
Gly Tyr Val Asn Ala His Gly Thr Ser Thr Pro Ala Gly Asp Val Ala

355 360 365
Glu Tyr Arg Ala Ile Arg Ala Val Ile Pro Gln Asp Ser Leu Arg Ile
370 375 380
Asn Ser Thr Lys Ser Met Ile Gly His Leu Leu Gly Gly Ala Gly Ala
385 390 395 400
Val Glu Ala Val Ala Ala Ile Gln Ala Leu Arg Thr Gly Trp Leu His
405 410 415
Pro Asn Leu Asn Leu Glu Asn Pro Ala Pro Gly Val Asp Pro Val Val
420 425 430
Leu Val Gly Pro Arg Lys Glu Arg Ala Glu Asp Leu Asp Val Val Leu
435 440 445
Ser Asn Ser Phe Gly Phe Gly Gly His Asn Ser Cys Val Ile Phe Arg
450 455 460
Lys Tyr Asp Glu Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp
465 470 475 480
His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys
485 490

转化构建体pSZ2014的相关序列

<210> 237

<211> 3602

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 237

gctcttcgct caccgcgtga attgctgtcc caaacgtaag catcatcgtg gctcggtcac 60
gcgatcctgg atccggggat cctagaccgc tggtggagag cgctgccgtc ggattggtgg 120
caagtaagat tgcgcaggtt ggcgaaggga gagaccaaaa ccggaggctg gaagcgggca 180
caacatcgta ttattgcgta tagtagagca gtggcagtcg catttcgagg tccgcaacgg 240
atctcgcaag ctcgctacgc tcacagtagg agaaagggga ccactgcccc tgccagaatg 300
gtcgcgaccc tctccctcgc cggccccgcc tgcaacacgc agtgcgtatc cggcaagcgg 360
gctgtcgcct tcaaccgccc ccatgttggc gtccgggctc gatcaggtgc gctgaggggg 420
gtttggtgtg cccgcgcctc tgggcccgtg tcggccgtgc ggacgtgggg ccctgggcag 480
tggatcagca gggtttgcgt gcaaatgcct ataccggcga ttgaatagcg atgaacggga 540
tacggttgcg ctcactccat gcccatgcga ccccgtttct gtccgccagc cgtggtcgcc 600
cgggctgcga agcgggaccc cacccagcgc attgtgatca ccggaatggg cgtggggtac 660
cctttcttgc gctatgacac ttccagcaaa aggtagggcg ggctgcgaga cggcttcccg 720
gcgctgcatg caacaccgat gatgcttcga ccccccgaag ctccttcggg gctgcatggg 780
cgctccgatg ccgctccagg gcgagcgctg tttaaatagc caggcccccg attgcaaaga 840

cattatagcg agctaccaaa gccatattca aacacctaga tcactaccac ttctacacag 900
gccactcgag cttgtgatcg cactccgcta agggggcgcc tcttcctctt cgtttcagtc 960
acaacccgca aacggcgcgc catgctgctg caggccttcc tgttcctgct ggccggcttc 1020
gccgccaaga tcagcgcctc catgacgaac gagacgtccg accgcccccct ggtgcacttc 1080
acccccaaca agggctggat gaacgacccc aacggcctgt ggtacgacga gaaggacgcc 1140
aagtggcacc tgtacttcca gtacaacccg aacgacaccg tctgggggac gcccttgttc 1200
tggggccacg ccacgtccga cgacctgacc aactgggagg accagcccat cgccatcgcc 1260
ccgaagcgca acgactccgg cgccttctcc ggctccatgg tggtggacta caacaacacc 1320
tccggcttct tcaacgacac catcgacccg cgccagcgct gcgtggccat ctggacctac 1380
aacaccccgg agtccgagga gcagtacatc tcctacagcc tggacggcgg ctacaccttc 1440
accgagtacc agaagaaccc cgtgctggcc gccaactcca cccagttccg cgacccgaag 1500
gtcttctggt acgagccctc ccagaagtgg atcatgaccg cggccaagtc ccaggactac 1560
aagatcgaga tctactcctc cgacgacctg aagtcctgga agctggagtc cgcgttcgcc 1620
aacgagggct tcctcggcta ccagtacgag tgccccggcc tgatcgaggt ccccaccgag 1680
caggacccca gcaagtccta ctgggtgatg ttcatctcca tcaaccccgg cgccccggcc 1740
ggcggctcct tcaaccagta cttcgtcggc agcttcaacg gcacccactt cgaggccttc 1800
gacaaccagt cccgcgtggt ggacttcggc aaggactact acgccctgca gaccttcttc 1860
aacaccgacc cgacctacgg gagcgccctg ggcatcgcgt gggcctccaa ctgggagtac 1920
tccgccttcg tgcccaccaa cccctggcgc tcctccatgt ccctcgtgcg caagttctcc 1980
ctcaacaccg agtaccaggc caacccggag acggagctga tcaacctgaa ggccgagccg 2040
atcctgaaca tcagcaacgc cggcccctgg agccggttcg ccaccaacac cacgttgacg 2100
aaggccaaca gctacaacgt cgacctgtcc aacagcaccg gcaccctgga gttcgagctg 2160
gtgtacgccg tcaacaccac ccagacgatc tccaagtccg tgttcgcgga cctctccctc 2220
tggttcaagg gcctggagga ccccgaggag tacctccgca tgggcttcga ggtgtccgcg 2280
tcctccttct tcctggaccg cgggaacagc aaggtgaagt tcgtgaagga gaacccctac 2340
ttcaccaacc gcatgagcgt gaacaaccag cccttcaaga gcgagaacga cctgtcctac 2400
tacaaggtgt acggcttgct ggaccagaac atcctggagc tgtacttcaa cgacggcgac 2460
gtcgtgtcca ccaacaccta cttcatgacc accgggaacg ccctgggctc cgtgaacatg 2520
acgacggggg tggacaacct gttctacatc gacaagttcc aggtgcgcga ggtcaagtga 2580
caattggcag cagcagctcg gatagtatcg acacactctg gacgctggtc gtgtgatgga 2640
ctgttgccgc cacacttgct gccttgacct gtgaatatcc ctgccgcttt tatcaaacag 2700
cctcagtgtg tttgatcttg tgtgtacgcg cttttgcgag ttgctagctg cttgtgctat 2760
ttgcgaatac caccccccagc atcccccttcc ctcgtttcat atcgcttgca tcccaaccgc 2820
aacttatcta cgctgtcctg ctatccctca gcgctgctcc tgctcctgct cactgcccct 2880
cgcacagcct tggtttgggc tccgcctgta ttctcctggt actgcaacct gtaaaccagc 2940
actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca caaatggagg atcgtagagc 3000
tccacctgca tccgcctggc gctcgaggac gccggcgtct cgcccgacga ggtcaactac 3060
gtcaacgcgc acgccacctc caccctggtg ggcgacaagg ccgaggtgcg cgcggtcaag 3120
tcggtctttg gcgacatgaa gggcatcaag atgaacgcca ccaagtccat gatcgggcac 3180

tgcctgggcg ccgccggcgg catggaggcc gtcgccacgc tcatggccat ccgcaccggc 3240
tgggtgcacc ccaccatcaa ccacgacaac cccatcgccg aggtcgacgg cctggacgtc 3300
gtcgccaacg ccaaggccca gcacaaaatc aacgtcgcca tctccaactc cttcggcttc 3360
ggcgggcaca actccgtcgt cgcctttgcg cccttccgcg agtaggcgga gcgagcgcgc 3420
ttggctgagg agggaggcgg ggtgcgagcc ctttggctgc gcgcgatact ctccccgcac 3480
gagcagactc cacgcgcctg aatctacttg tcaacgagca accgtgtgtt ttgtccgtgg 3540
ccattcttat tatttctccg actgtggccg tactctgttt ggctgtgcaa gcaccgaaga 3600
gc 3602

<210> 238
<211> 661
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 238

gctcttcgct caccgcgtga attgctgtcc caaacgtaag catcatcgtg gctcggtcac 60
gcgatcctgg atccggggat cctagaccgc tggtggagag cgctgccgtc ggattggtgg 120
caagtaagat tgcgcaggtt ggcgaaggga gagaccaaaa ccggaggctg gaagcgggca 180
caacatcgta ttattgcgta tagtagagca gtggcagtcg catttcgagg tccgcaacgg 240
atctcgcaag ctcgctacgc tcacagtagg agaaagggga ccactgcccc tgccagaatg 300
gtcgcgaccc tctccctcgc cggccccgcc tgcaacacgc agtgcgtatc cggcaagcgg 360
gctgtcgcct tcaaccgccc ccatgttggc gtccgggctc gatcaggtgc gctgaggggg 420
gtttggtgtg cccgcgcctc tgggcccgtg tcggccgtgc ggacgtgggg ccctgggcag 480
tggatcagca gggtttgcgt gcaaatgcct ataccggcga ttgaatagcg atgaacggga 540
tacggttgcg ctcactccat gcccatgcga ccccgtttct gtccgccagc cgtggtcgcc 600
cgggctgcga agcgggaccc cacccagcgc attgtgatca ccggaatggg cgtggggtac 660
c 661

<210> 239
<211> 605
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 239

gagctccacc tgcatccgcc tggcgctcga ggacgccggc gtctcgcccg acgaggtcaa 60
ctacgtcaac gcgcacgcca cctccaccct ggtgggcgac aaggccgagg tgcgcgcggt 120
caagtcggtc tttggcgaca tgaagggcat caagatgaac gccaccaagt ccatgatcgg 180
gcactgcctg ggcgccgccg gcggcatgga ggccgtcgcc acgctcatgg ccatccgcac 240
cggctgggtg caccccacca tcaaccacga caaccccatc gccgaggtcg acggcctgga 300

cgtcgtcgcc aacgccaagg cccagcacaa aatcaacgtc gccatctcca actccttcgg 360
cttcggcggg cacaactccg tcgtcgcctt tgcgcccttc cgcgagtagg cggagcgagc 420
gcgcttggct gaggagggag gcggggtgcg agccctttgg ctgcgcgcga tactctcccc 480
gcacgagcag actccacgcg cctgaatcta cttgtcaacg agcaaccgtg tgttttgtcc 540
gtggccattc ttattatttc tccgactgtg gccgtactct gtttggctgt gcaagcaccg 600
aagag 605

pSZ2110,KASII'5::CrbTub:NeoR:nr::KASII-'3的相关序列

<210> 240
<211> 3124
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 240

gctcttcgct ggcctgttgc cggacgatcc gtgtcgtcga gactgcattt tgttttgggt 60
gtggggctgg ggtactggat ggcttgaggg catgactttt tctgatggag aagattgcaa 120
tgagatcatt tgggtcgtct atttgtttgc tgtgcaagag ggtttactgg tatctggcac 180
cagcttttgg cccgtgcccg tttgatggac gcgtgacagg caggcgtcct ggaaagcaca 240
gacaccgtac gtacgacctt gacctccccc ccttctccac acggcaggtg cgaggctgcc 300
cacggcgtcg aggcgggcgg tgcgccgggc atggtcccgc atcgcgcgcg cggcggccgc 360
ggccgacgca aaccccgccc gccctgagcg ccgcgtggtc atcacgggcc agggcgtggt 420
gaccagcctg ggccagacga tcgagcagtt ttacagcagc ctgctggagg gcgtgagcgg 480
catctcgcag atacagaagt tcgacaccac gggctacacg acgacgatcg cgggcgagat 540
caagtcgctg cagctggacc cgtacgtgcc caagcgctgg gcgaagcgcg tggacgacgt 600
gataaagtac gtctacatcg cgggcaagca ggcgctggag agcgccggcc tgccgatcga 660
ggcggcgggg ctggcgggcg cggggctgga cccggcgctg tgcggcgtgc tcatcggcac 720
cgccatggcg ggcatgacgt ctttcgcggc gggcgtggag gcgctgacgc gcggtaccct 780
ttcttgcgct atgacacttc cagcaaaagg tagggcgggc tgcgagacgg cttcccggcg 840
ctgcatgcaa caccgatgat gcttcgaccc cccgaagctc cttcggggct gcatgggcgc 900
tccgatgccg ctccagggcg agcgctgttt aaatagccag gcccccgatt gcaaagacat 960
tatagcgagc taccaaagcc atattcaaac acctagatca ctaccacttc tacacaggcc 1020
actcgagctt gtgatcgcac tccgctaagg gggcgcctct tcctcttcgt ttcagtcaca 1080
acccgcaaac tctagaatat caatgatcga gcaggacggc ctccacgccg gctccccccgc 1140
cgcctgggtg gagcgcctgt tcggctacga ctgggcccag cagaccatcg gctgctccga 1200
cgccgccgtg ttccgcctgt ccgcccaggg ccgccccgtg ctgttcgtga agaccgacct 1260
gtccggcgcc ctgaacgagc tgcaggacga ggccgcccgc ctgtcctggc tggccaccac 1320
cggcgtgccc tgcgccgccg tgctggacgt ggtgaccgag gccggccgcg actggctgct 1380
gctgggcgag gtgcccggcc aggacctgct gtcctcccac ctggcccccg ccgagaaggt 1440
gtccatcatg gccgacgcca tgcgccgcct gcacaccctg gaccccgcca cctgcccctt 1500

cgaccaccag gccaagcacc gcatcgagcg cgcccgcacc cgcatggagg ccggcctggt 1560
ggaccaggac gacctggacg aggagcacca gggcctggcc cccgccgagc tgttcgcccg 1620
cctgaaggcc cgcatgcccg acggcgagga cctggtggtg acccacggcg acgcctgcct 1680
gcccaacatc atggtggaga acggccgctt ctccggcttc atcgactgcg gccgcctggg 1740
cgtggccgac cgctaccagg acatcgccct ggccacccgc gacatcgccg aggagctggg 1800
cggcgagtgg gccgaccgct tcctggtgct gtacggcatc gccgcccccg actcccagcg 1860
catcgccttc taccgcctgc tggacgagtt cttctgacaa ttggcagcag cagctcggat 1920
agtatcgaca cactctggac gctggtcgtg tgatggactg ttgccgccac acttgctgcc 1980
ttgacctgtg aatatccctg ccgcttttat caaacagcct cagtgtgttt gatcttgtgt 2040
gtacgcgctt ttgcgagttg ctagctgctt gtgctatttg cgaataccac ccccagcatc 2100
cccttccctc gtttcatatc gcttgcatcc caaccgcaac ttatctacgc tgtcctgcta 2160
tccctcagcg ctgctcctgc tcctgctcac tgcccctcgc acagccttgg tttgggctcc 2220
gcctgtattc tcctggtact gcaacctgta aaccagcact gcaatgctga tgcacgggaa 2280
gtagtgggat gggaacacaa atggaggatc cactagttct agagcggccg ccaccgcggt 2340
ggagctcggc ggcgtgcgca agatgaaccc cttttgcatc cccttctcca tctccaacat 2400
gggcggcgcg atgctggcga tggacatcgg cttcatgggc cccaactact ccatctccac 2460
ggcctgcgcg acgggcaact actgcatcct gggcgcggcg gaccacatcc ggcgcggcga 2520
cgcaaacgtg atgctggccg gcggcgcgga cgcggccatc atcccctcgg gcatcggcgg 2580
cttcatcgcg tgcaaggcgc tgagcaagcg caacgacgag cccgagcgcg cgagccggcc 2640
ctgggacgcc gaccgcgacg gcttcgtcat gggcgagggc gccggcgtgc tggtgctgga 2700
ggagctggag cacgccaagc gccgcggcgc gaccattttg gctgaattag ttggcggcgc 2760
ggccacctcg gacgcgcacc acatgaccga gcccgacccg cagggccgcg gcgtgcgcct 2820
ctgcctcgag cgcgcgctcg agcgcgcgcg cctcgcgccc gagcgcgtcg gctacgtcaa 2880
cgcgcacggc accagcacgc ccgcgggcga cgtggccgag taccgcgcca tccgcgccgt 2940
catcccgcag gactcactac gcatcaactc cacaaagtcc atgatcgggc acctgctcgg 3000
cggcgccggc gcggtcgagg ccgtggccgc catccaggcc ctgcgcaccg gctggctcca 3060
ccccaacttg aacctcgaga accccgcgcc tggcgtcgac cccgtcgtgc tcgtgggctc 3120
ttcc 3124

包括CtFad2-2的密码子优化的核苷酸序列的pSZ2150的部分

<210> 241
<211> 1173
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 241

actagtatgg gcgccggcgg ccgcatgtcc aacccctccg agggcgagaa gaagaccgag 60
ctggagggca tccagcgcgt gccctaccag aagcccccct tcaccgtggg cgacgggcgc 120
gccgtgaaga aggccatccc cccccactgc ttcaaccgct ccgtgatccg ctccttctcc 180

tacgtggtgt acgacctgac catcgcctcc atcctgtact acatcgccac caccttcatc 240
cccctgctgc cccacccccc ggcctacgtg gcctggccca tctactgggc cgtgcagggc 300
tgcgtgatga ccggcgtgtg ggtgatcgcc cacgagtgcg gccaccacgc cttctccgac 360
taccagtggc tggacgacac cgtgggcctg atcctgcact ccgtgctgct ggtgccctac 420
ttctcctgga agtactccca ccgccgccac cactccaaca ccggctccat cgagcacgac 480
gaggtgttcg tgcccaagct gaagtccggc gtgcgctcca ccgccaagta cctgaacaac 540
ccccccggcc gcatcctgac cctgctggtg accctgaccc tgggctggcc cctgtacctg 600
atgttcaacg tgtccggccg ctactacgac cgcttcgcct gccacttcga ccccaactcc 660
cccatctact ccaaccgcga gcgcgcccag atcttcatct ccgacgccgg catcttcgcc 720
gtgctgtacg gcctgtaccg cctggccgcc gtgaagggcc tggtgtgggt gctgaccgtg 780
tacgccggcc ccctgctggt ggtgaacggc ttcctggtgc tgatcacctt cctgcagcac 840
acccaccccc ccctgcccca ctacgactcc accgagtggg actggctgcg cggcgccctg 900
gccaccatcg accgcgacta cggcatcctg aacaaggtgt tccacaacat caccgacacc 960
cacgtgaccc accacctgtt ctccaccatg ccccactacc acgccatgga ggccaccaag 1020
gccatcatcc ccatcctggg cgactactac cagttcgacg gcacctccgt gttcaaggcc 1080
atgtaccgcg agaccaagga gtgcatctac gtggacaagg acgaggaggt gaaggacggc 1140
gtgtactggt accgcaacaa gatctagatc gat 1173

包括GlmFad2-2的密码子优化的核苷酸序列的pSZ2151的部分

<210> 242
<211> 1173
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 242

actagtatgg gcgccggcgg ccgcaccgac gtgccccccg ccaaccgcaa gtccgaggtg 60
gacccccctga agcgcgtgcc cttcgagaag ccccccttct ccctgtccca ggggcgcgcc 120
atcaagaagg tgatcccccc ccactgcttc cagcgctccg tgttccgctc cttctcctac 180
gtggtgtacg acctgaccat cgccttctgc ctgtactacg tggccaccca ctacttccac 240
ctgctgccct ccccccctgtc cttcctggcc tggcccatct actgggccgt gcagggctgc 300
atcctgaccg gcgtgtgggt gatcgcccac gagtgcggcc accacgcctt ctccgactac 360
cagctgctgg acgacatcgt gggcctggtg ctgcactccg gcctgctggt gccctacttc 420
tcctggaagt actcccaccg ccgccaccac tccaacaccg gctccctgga gcgcgacgag 480
gtgttcgtgc ccaagcagaa gtcctgcatc aagtggtact ccaagtacct gaacaacccc 540
cccggccgcg tgctgaccct ggccgtgacc ctgaccctgg gctggcccct gtacctggcc 600
ctgaacgtgt ccggccgccc ctacgaccgc ttcgcctgcc actacgaccc ctacggcccc 660
atctactccg accgcgagcg cctgcagatc tacatctccg acgccggcgt gctggccgtg 720
tgctacggcc tgttccgcct ggccatggcc aagggcctgg cctgggtggt gtgcgtgtac 780
ggcgtgcccc tgctggtggt gaacggcttc ctggtgctga tcaccttcct gcagcacacc 840

caccccgccc tgccccacta cacctcctcc gagtgggact ggctgcgcgg cgccctggcc 900
accgtggacc gcgactacgg catcctggac aaggtgttcc acaacatcac cgacacccac 960
gtggcccacg agctgttctc caccatgccc cactaccacg ccatggaggc caccaaggcc 1020
atcaagccca tcctgggcga gtactaccgc ttcgacgaga ccccccttcgt gaaggccatg 1080
tggcgcgagg cccgcgagtg catctacgtg gagcccgacc agtccaccga gtccaagggc 1140
gtgttctggt acaacaacaa gctgtagatc gat 1173

包括HaFad2的密码子优化的核苷酸序列的pSZ2152的部分

<210> 243
<211> 1170
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 243

actagtatgg gcgccggcgg ccgcatgtcc tcccccaacg gcaaggagaa ggacggcccc 60
aagcccctgg agcgcgccct gcacgagaag ccccccttca ccgtgggcga cgggcgcgcc 120
atcaagaagg tgatcccccc ccactgcttc aagcgctccg tgatccgctc cttctcctac 180
gtggtgtacg acctgaccat cgcctccatc ttctactacc tggccaacaa ctacatcccc 240
ctgctgccca actccctggc ctacgtggcc tggcccgtgt actggatctt ccagggctgc 300
gtgctgaccg gcgtgtgggt gatcgcccac gagtgcggcc accacgcctt ctccgactac 360
cagtggctgg acgacaccgt gggcctgatc ctgcactccg ccctgctggt gccctacttc 420
tcctggaagt actcccaccg ccgccaccac tccaacaccg gctccatcga gcacgacgag 480
gtgttcgtgc ccaagctgaa gtcctccgtg cgctccaccg ccaagtacct gaacaacccc 540
cccggccgca tcctgaccct gctggtgacc ctgaccatgg gctggcccct gtacctgatg 600
ttcaacgtgt ccggccgcta ctacgaccgc ttcgcctgcc acttcgaccc caactccccc 660
atctactcca accgcgagcg cgcccagatc ttcatctccg acgccggcat cctgaccgtg 720
ttctacatcc tgttccgcct ggcctccacc aagggcctgg tgtgggtgct gaccatgtac 780
ggcggccccc tgctggtggt gaacggcttc ctggtgctga tcaccttcct gcagcacacc 840
cacccctccc tgccccacta cgactccacc gagtgggact ggctgcgcgg cgccctggcc 900
accgtggacc gcgactacgg catcctgaac aaggtgttcc acaacatcac cgacacccac 960
gtgacccacc acctgttctc caccatgccc cactaccacg ccatggaggc caccaaggcc 1020
atcaagccca tcctgggcga ctactaccag ttcgacggca cctccatctt caaggccatg 1080
taccgcgaga ccaaggagtg catctacgtg gacaaggacg aggacgtgaa ggacggcgtg 1140
tactggtacc gcaacaagat ctagatcgat 1170

包括ZmFad2的密码子优化的核苷酸序列的pSZ2153的部分

<210> 244
<211> 1185
<212> DNA
<213> 人工序列

<220>

<223> 合成的构建体

<400> 244

actagtatgg gcgccggcgg ccgcatgacc gagaaggagc gcgagaagca ggagcagctg 60
gcccgcgcca ccggcggcgc cgccatgcag cgctcccccg tggagaagcc ccccttcacc 120
ctgggccagg ggcgcgccat caagaaggcc atccccccccc actgcttcga gcgctccgtg 180
ctgaagtcct tctcctacgt ggtgcacgac ctggtgatcg ccgccgccct gctgtacttc 240
gccctggcca tcatccccgc cctgccctcc cccctgggct acgccgcctg gcccctgtac 300
tggatcgccc agggctgcgt gtgcaccggc gtgtgggtga tcgcccacga gtgcggccac 360
cacgccttct ccgactactc cctgctggac gacgtggtgg gcctggtgct gcactcctcc 420
ctgatggtgc cctacttctc ctggaagtac tcccaccgcc gccaccactc caacaccggc 480
tccctggagc gcgacgaggt gttcgtgccc aagaagaagg aggccctgcc ctggtacacc 540
ccctacgtgt acaacaaccc cgtgggccgc ctggcccaca tcgtggtgca gctgaccctg 600
ggctggcccc tgtacctggc caccaacgcc tccggccgcc cctacccccg cttcgcctgc 660
cacttcgacc cctacggccc catctacaac gaccgcgagc gcgcccagat cttcgtgtcc 720
gacgccggcg tggccgccgt ggccttcggc ctgtacaagc tggccgccgc cttcggcgtg 780
tggtgggtgg tgcgcgtgta cgccgtgccc ctgctgatcg tgaacgcctg gctggtgctg 840
atcacctacc tgcagcacac ccacccctcc ctgccccact acgactcctc cgagtgggac 900
tggctgcgcg gcgccctggc caccatggac cgcgactacg gcatcctgaa ccgcgtgttc 960
cacaacatca ccgacaccca cgtggcccac cacctgttct ccaccatgcc ccactaccac 1020
gccatggagg ccaccaaggc catccgcccc atcctgggcg actactacca cttcgacccc 1080
acccccgtgg ccaaggccac ctggcgcgag gccggcgagt gcatctacgt ggagcccgag 1140
gaccgcaagg gcgtgttctg gtacaacaag aagttctaga tcgat 1185

包含在质粒pSZ2172中的转化DNA的核苷酸序列

<210> 245

<211> 6756

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 245

gctcttcgcc gccgccactc ctgctcgagc gcgcccgcgc gtgcgccgcc agcgccttgg 60
ccttttcgcc gcgctcgtgc gcgtcgctga tgtccatcac caggtccatg aggtctgcct 120
tgcgccggct gagccactgc ttcgtccggg cggccaagag gagcatgagg gaggactcct 180
ggtccagggt cctgacgtgg tcgcggctct gggagcgggc cagcatcatc tggctctgcc 240
gcaccgaggc cgcctccaac tggtcctcca gcagccgcag tcgccgccga ccctggcaga 300
ggaagacagg tgaggggggt atgaattgta cagaacaacc acgagccttg tctaggcaga 360
atccctacca gtcatggctt tacctggatg acggcctgcg aacagctgtc cagcgaccct 420
cgctgccgcc gcttctcccg cacgcttctt tccagcaccg tgatggcgcg agccagcgcc 480

gcacgctggc gctgcgcttc gccgatctga ggacagtcgg ggaactctga tcagtctaaa 540
ccccccttgcg cgttagtgtt gccatccttt gcagaccggt gagagccgac ttgttgtgcg 600
ccacccccca caccacctcc tcccagacca attctgtcac ctttttggcg aaggcatcgg 660
cctcggcctg cagagaggac agcagtgccc agccgctggg ggttggcgga tgcacgctca 720
ggtacccttt cttgcgctat gacacttcca gcaaaaggta gggcgggctg cgagacggct 780
tcccggcgct gcatgcaaca ccgatgatgc ttcgaccccc cgaagctcct tcggggctgc 840
atgggcgctc cgatgccgct ccagggcgag cgctgtttaa atagccaggc ccccgattgc 900
aaagacatta tagcgagcta ccaaagccat attcaaacac ctagatcact accacttcta 960
cacaggccac tcgagcttgt gatcgcactc cgctaagggg gcgcctcttc ctcttcgttt 1020
cagtcacaac ccgcaaactc tagaatatca atgctgctgc aggccttcct gttcctgctg 1080
gccggcttcg ccgccaagat cagcgcctcc atgacgaacg agacgtccga ccgccccctg 1140
gtgcacttca cccccaacaa gggctggatg aacgacccca acggcctgtg gtacgacgag 1200
aaggacgcca agtggcacct gtacttccag tacaacccga acgacaccgt ctgggggacg 1260
cccttgttct ggggccacgc cacgtccgac gacctgacca actgggagga ccagcccatc 1320
gccatcgccc cgaagcgcaa cgactccggc gccttctccg gctccatggt ggtggactac 1380
aacaacacct ccggcttctt caacgacacc atcgacccgc gccagcgctg cgtggccatc 1440
tggacctaca acaccccgga gtccgaggag cagtacatct cctacagcct ggacggcggc 1500
tacaccttca ccgagtacca gaagaacccc gtgctggccg ccaactccac ccagttccgc 1560
gacccgaagg tcttctggta cgagccctcc cagaagtgga tcatgaccgc ggccaagtcc 1620
caggactaca agatcgagat ctactcctcc gacgacctga agtcctggaa gctggagtcc 1680
gcgttcgcca acgagggctt cctcggctac cagtacgagt gccccggcct gatcgaggtc 1740
cccaccgagc aggaccccag caagtcctac tgggtgatgt tcatctccat caaccccggc 1800
gccccggccg gcggctcctt caaccagtac ttcgtcggca gcttcaacgg cacccacttc 1860
gaggccttcg acaaccagtc ccgcgtggtg gacttcggca aggactacta cgccctgcag 1920
accttcttca acaccgaccc gacctacggg agcgccctgg gcatcgcgtg ggcctccaac 1980
tgggagtact ccgccttcgt gcccaccaac ccctggcgct cctccatgtc cctcgtgcgc 2040
aagttctccc tcaacaccga gtaccaggcc aacccggaga cggagctgat caacctgaag 2100
gccgagccga tcctgaacat cagcaacgcc ggcccctgga gccggttcgc caccaacacc 2160
acgttgacga aggccaacag ctacaacgtc gacctgtcca acagcaccgg caccctggag 2220
ttcgagctgg tgtacgccgt caacaccacc cagacgatct ccaagtccgt gttcgcggac 2280
ctctccctct ggttcaaggg cctggaggac cccgaggagt acctccgcat gggcttcgag 2340
gtgtccgcgt cctccttctt cctggaccgc gggaacagca aggtgaagtt cgtgaaggag 2400
aacccctact tcaccaaccg catgagcgtg aacaaccagc ccttcaagag cgagaacgac 2460
ctgtcctact acaaggtgta cggcttgctg gaccagaaca tcctggagct gtacttcaac 2520
gacggcgacg tcgtgtccac caacacctac ttcatgacca ccgggaacgc cctgggctcc 2580
gtgaacatga cgacgggggt ggacaacctg ttctacatcg acaagttcca ggtgcgcgag 2640
gtcaagtgac aattggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg 2700
tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt 2760
atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc 2820

ttgtgctatt tgcgaatacc acccccagca tccccttccc tcgtttcata tcgcttgcat 2880
cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc 2940
actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg 3000
taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggagga 3060
tcccgcgtct cgaacagagc gcgcagagga acgctgaagg tctcgcctct gtcgcacctc 3120
agcgcggcat acaccacaat aaccacctga cgaatgcgct tggttcttcg tccattagcg 3180
aagcgtccgg ttcacacacg tgccacgttg gcgaggtggc aggtgacaat gatcggtgga 3240
gctgatggtc gaaacgttca cagcctaggg atatcgaatt cggccgacag gacgcgcgtc 3300
aaaggtgctg gtcgtgtatg ccctggccgg caggtcgttg ctgctgctgg ttagtgattc 3360
cgcaaccctg attttggcgt cttattttgg cgtggcaaac gctggcgccc gcgagccggg 3420
ccggcggcga tgcggtgccc cacggctgcc ggaatccaag ggaggcaaga gcgcccgggt 3480
cagttgaagg gctttacgcg caaggtacag ccgctcctgc aaggctgcgt ggtggaattg 3540
gacgtgcagg tcctgctgaa gttcctccac cgcctcacca gcggacaaag caccggtgta 3600
tcaggtccgt gtcatccact ctaaagaact cgactacgac ctactgatgg ccctagattc 3660
ttcatcaaaa acgcctgaga cacttgccca ggattgaaac tccctgaagg gaccaccagg 3720
ggccctgagt tgttccttcc ccccgtggcg agctgccagc caggctgtac ctgtgatcga 3780
ggctggcggg aaaataggct tcgtgtgctc aggtcatggg aggtgcagga cagctcatga 3840
aacgccaaca atcgcacaat tcatgtcaag ctaatcagct atttcctctt cacgagctgt 3900
aattgtccca aaattctggt ctaccggggg tgatccttcg tgtacgggcc cttccctcaa 3960
ccctaggtat gcgcgcatgc ggtcgccgcg caactcgcgc gagggccgag ggtttgggac 4020
gggccgtccc gaaatgcagt tgcacccgga tgcgtggcac ctttttgcg ataatttatg 4080
caatggactg ctctgcaaaa ttctggctct gtcgccaacc ctaggatcag cggcgtagga 4140
tttcgtaatc attcgtcctg atggggagct accgactacc ctaatatcag cccgactgcc 4200
tgacgccagc gtccactttt gtgcacacat tccattcgtg cccaagacat ttcattgtgg 4260
tgcgaagcgt ccccagttac gctcacctgt ttcccgacct ccttactgtt ctgtcgacag 4320
agcgggccca caggccggtc gcagccacta gtatgggcgc cggcggccgc atgtccgtgc 4380
ccacctcccc caagaagccc gagttcaact ccctgaagcg cgtgccctac tccaagcccc 4440
ccttcaccct gtccgagggg cgcgccatgg actacaagga ccacgacggc gactacaagg 4500
accacgacat cgactacaag gacgacgacg acaagatcaa gaaggccatc ccccccccact 4560
gcttccagcg ctccgtgctg cgctccttct cctacctgct gtacgacttc atcctggcct 4620
ccctgttcta ccacgtggcc accaactact tccccaacct gccccaggcc ctgtccaacg 4680
tggcctggcc cctgtactgg gccatgcagg gctgcatcct gaccggcgtg tgggtgatcg 4740
cccacgagtg cggccaccac gccttctccg actaccagtg gctggacgac accgtgggcc 4800
tgatcctgca ctcctccctg ctggtgccct acttctcctg gaagtactcc caccgccgcc 4860
accactccaa caccggctcc ctggagcgcg acgaggtgtt cgtgcccaag aagaagtccg 4920
gcctgcgctg gtgggccaag cacttcaaca acccccccgg ccgcttcctg tccatcacca 4980
tccagctgac cctgggctgg cccctgtacc tggccttcaa cgtggccggc cgcccctacg 5040
accgcttcgc ctgccactac gacccctacg gccccatctt ctccgaccgc gagcgcctgc 5100
agatctacat ctccgacgcc ggcgtgctgg ccgtggccta cgccctgtac cgcctggtgc 5160

tggccaaggg cgtgggctgg gtgatctccg tgtacggcgt gcccctgctg gtggtgaacg 5220
ccttcctggt gatgatcacc tacctgcagc acacccaccc ctccctgccc cactacgact 5280
cctccgagtg ggactggatg cgcggcgccc tgtccaccgt ggaccgcgac tacggcatcc 5340
tgaacaaggt gttccacaac atcaccgaca cccacgtggc ccaccacctg ttctccacca 5400
tgccccacta ccacgccatg gtggccacca aggccatcaa gcccatcctg ggcgagtact 5460
accagttcga cggcatgccc gtgtacaagg ccatctggcg cgaggccaag gagtgcctgt 5520
acgtggagcc cgacgagggc gacaaggaca agggcgtgtt ctggttccgc aacaagctgt 5580
agatcgatct taaggcagca gcagctcgga tagtatcgac acactctgga cgctggtcgt 5640
gtgatggact gttgccgcca cacttgctgc cttgacctgt gaatatccct gccgctttta 5700
tcaaacagcc tcagtgtgtt tgatcttgtg tgtacgcgct tttgcgagtt gctagctgct 5760
tgtgctattt gcgaatacca cccccagcat ccccttccct cgtttcatat cgcttgcatc 5820
ccaaccgcaa cttatctacg ctgtcctgct atccctcagc gctgctcctg ctcctgctca 5880
ctgcccctcg cacagccttg gtttgggctc cgcctgtatt ctcctggtac tgcaacctgt 5940
aaaccagcac tgcaatgctg atgcacggga agtagtggga tgggaacaca aatggaaagc 6000
ttaattaaga gctcttgttt tccagaagga gttgctcctt gagcctttca ttctcagcct 6060
cgataacctc caaagccgct ctaattgtgg agggggttcg aatttaaaag cttggaatgt 6120
tggttcgtgc gtctggaaca agcccagact tgttgctcac tgggaaaagg accatcagct 6180
ccaaaaaact tgccgctcaa accgcgtacc tctgctttcg cgcaatctgc cctgttgaaa 6240
tcgccaccac attcatattg tgacgcttga gcagtctgta attgcctcag aatgtggaat 6300
catctgcccc ctgtgcgagc ccatgccagg catgtcgcgg gcgaggacac ccgccactcg 6360
tacagcagac cattatgcta cctcacaata gttcataaca gtgaccatat ttctcgaagc 6420
tccccaacga gcacctccat gctctgagtg gccacccccc ggccctggtg cttgcggagg 6480
gcaggtcaac cggcatgggg ctaccgaaat ccccgaccgg atcccaccac ccccgcgatg 6540
ggaagaatct ctccccggga tgtgggccca ccaccagcac aacctgctgg cccaggcgag 6600
cgtcaaacca taccacacaa atatccttgg catcggccct gaattccttc tgccgctctg 6660
ctacccggtg cttctgtccg aagcaggggt tgctagggat cgctccgagt ccgcaaaccc 6720
ttgtcgcgtg gcggggcttg ttcgagcttg aagagc 6756

pSZ1499

<210> 246

<211> 9416

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 246

cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag 60
agttcttgaa gtggtggcct aactacggct acactagaag gacagtattt ggtatctgcg 120
ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa 180
ccaccgctgg tagcggtggt ttttttgttt gcaagcagca gattacgcgc agaaaaaaag 240

gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg aacgaaaact 300
cacgttaagg gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa 360
attaaaaatg aagttttaaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt 420
accaatgctt aatcagtgag gcacctatct cagcgatctg tctatttcgt tcatccatag 480
ttgcctgact ccccgtcgtg tagataacta cgatacggga gggcttacca tctggcccca 540
gtgctgcaat gataccgcga gacccacgct caccggctcc agatttatca gcaataaacc 600
agccagccgg aagggccgag cgcagaagtg gtcctgcaac tttatccgcc tccatccagt 660
ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg 720
ttgttgccat tgctacaggc atcgtggtgt cacgctcgtc gtttggtatg gcttcattca 780
gctccggttc ccaacgatca aggcgagtta catgatcccc catgttgtgc aaaaaagcgg 840
ttagctcctt cggtcctccg atcgttgtca gaagtaagtt ggccgcagtg ttatcactca 900
tggttatggc agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg 960
tgactggtga gtactcaacc aagtcattct gagaatagtg tatgcggcga ccgagttgct 1020
cttgcccggc gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgctca 1080
tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca 1140
gttcgatgta acccactcgt gcacccaact gatcttcagc atcttttact ttcaccagcg 1200
tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata agggcgacac 1260
ggaaatgttg aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt 1320
attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacaa ataggggttc 1380
cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat 1440
taacctataa aaataggcgt atcacgaggc cctttcgtct cgcgcgtttc ggtgatgacg 1500
gtgaaaacct ctgacacatg cagctcccgg agacggtcac agcttgtctg taagcggatg 1560
ccgggagcag acaagcccgt cagggcgcgt cagcgggtgt tggcgggtgt cggggctggc 1620
ttaactatgc ggcatcagag cagattgtac tgagagtgca ccatatgcgg tgtgaaatac 1680
cgcacagatg cgtaaggaga aaataccgca tcaggcgcca ttcgccattc aggctgcgca 1740
actgttggga agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg 1800
gatgtgctgc aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta 1860
aaacgacggc cagtgaattg atgcatgctc ttcgggctgg tctgaatcct tcaggcgggt 1920
gttacccgag aaagaaaggg tgccgatttc aaagcagacc catgtgccgg gccctgtggc 1980
ctgtgttggc gcctatgtag tcacccccc tcacccaatt gtcgccagtt tgcgcactcc 2040
ataaactcaa aacagcagct tctgagctgc gctgttcaag aacacctctg gggtttgctc 2100
acccgcgagg tcgacgccca gcatggctat caagacgaac aggcagcctg tggagaagcc 2160
tccgttcacg atcgggacgc tgcgcaaggc catccccgcg cactgtttcg agcgctcggc 2220
gcttcgtagc agcatgtacc tggcctttga catcgcggtc atgtccctgc tctacgtcgc 2280
gtcgacgtac atcgaccctg caccggtgcc tacgtgggtc aagtacggca tcatgtggcc 2340
gctctactgg ttcttccagg tgtgtttgag ggttttggtt gcccgtattg aggtcctggt 2400
ggcgcgcatg gaggagaagg cgcctgtccc gctgaccccc ccggctaccc tcccggcacc 2460
ttccagggcg ccttcggcac gggtgtctgg gtgtgcgcgc acgagtgcgg ccaccaggcc 2520
ttttcctcca gccaggccat caacgacggc gtgggcctgg tgttccacag cctgctgctg 2580

gtgccctact actcctggaa gcactcgcac cgggtaccct ttcttgcgct atgacacttc 2640
cagcaaaagg tagggcgggc tgcgagacgg cttcccggcg ctgcatgcaa caccgatgat 2700
gcttcgaccc cccgaagctc cttcggggct gcatgggcgc tccgatgccg ctccagggcg 2760
agcgctgttt aaatagccag gccccccgatt gcaaagacat tatagcgagc taccaaagcc 2820
atattcaaac acctagatca ctaccacttc tacacaggcc actcgagctt gtgatcgcac 2880
tccgctaagg gggcgcctct tcctcttcgt ttcagtcaca acccgcaaac tctagaatat 2940
caatgctgct gcaggccttc ctgttcctgc tggccggctt cgccgccaag atcagcgcct 3000
ccatgacgaa cgagacgtcc gaccgccccc tggtgcactt cacccccaac aagggctgga 3060
tgaacgaccc caacggcctg tggtacgacg agaaggacgc caagtggcac ctgtacttcc 3120
agtacaaccc gaacgacacc gtctggggga cgcccttgtt ctggggccac gccacgtccg 3180
acgacctgac caactgggag gaccagccca tcgccatcgc cccgaagcgc aacgactccg 3240
gcgccttctc cggctccatg gtggtggact acaacaacac ctccggcttc ttcaacgaca 3300
ccatcgaccc gcgccagcgc tgcgtggcca tctggaccta caacaccccg gagtccgagg 3360
agcagtacat ctcctacagc ctggacggcg gctacacctt caccgagtac cagaagaacc 3420
ccgtgctggc cgccaactcc acccagttcc gcgacccgaa ggtcttctgg tacgagccct 3480
cccagaagtg gatcatgacc gcggccaagt cccaggacta caagatcgag atctactcct 3540
ccgacgacct gaagtcctgg aagctggagt ccgcgttcgc caacgagggc ttcctcggct 3600
accagtacga gtgccccggc ctgatcgagg tccccaccga gcaggacccc agcaagtcct 3660
actgggtgat gttcatctcc atcaaccccg gcgccccggc cggcggctcc ttcaaccagt 3720
acttcgtcgg cagcttcaac ggcacccact tcgaggcctt cgacaaccag tcccgcgtgg 3780
tggacttcgg caaggactac tacgccctgc agaccttctt caacaccgac ccgacctacg 3840
ggagcgccct gggcatcgcg tgggcctcca actgggagta ctccgccttc gtgcccacca 3900
acccctggcg ctcctccatg tccctcgtgc gcaagttctc cctcaacacc gagtaccagg 3960
ccaacccgga gacggagctg atcaacctga aggccgagcc gatcctgaac atcagcaacg 4020
ccggcccctg gagccggttc gccaccaaca ccacgttgac gaaggccaac agctacaacg 4080
tcgacctgtc caacagcacc ggcaccctgg agttcgagct ggtgtacgcc gtcaacacca 4140
cccagacgat ctccaagtcc gtgttcgcgg acctctccct ctggttcaag ggcctggagg 4200
accccgagga gtacctccgc atgggcttcg aggtgtccgc gtcctccttc ttcctggacc 4260
gcgggaacag caaggtgaag ttcgtgaagg agaaccccta cttcaccaac cgcatgagcg 4320
tgaacaacca gcccttcaag agcgagaacg acctgtccta ctacaaggtg tacggcttgc 4380
tggaccagaa catcctggag ctgtacttca acgacggcga cgtcgtgtcc accaacacct 4440
acttcatgac caccgggaac gccctgggct ccgtgaacat gacgacgggg gtggacaacc 4500
tgttctacat cgacaagttc caggtgcgcg aggtcaagtg acaattggca gcagcagctc 4560
ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg ccacacttgc 4620
tgccttgacc tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt gtttgatctt 4680
gtgtgtacgc gcttttgcga gttgctagct gcttgtgcta tttgcgaata ccacccccag 4740
catccccttc cctcgtttca tatcgcttgc atcccaaccg caacttatct acgctgtcct 4800
gctatccctc agcgctgctc ctgctcctgc tcactgcccc tcgcacagcc ttggtttggg 4860
ctccgcctgt attctcctgg tactgcaacc tgtaaaccag cactgcaatg ctgatgcacg 4920

ggaagtagtg ggatgggaac acaaatggag gatcccgcgt ctcgaacaga gcgcgcagag 4980
gaacgctgaa ggtctcgcct ctgtcgcacc tcagcgcggc atacaccaca ataaccacct 5040
gacgaatgcg cttggttctt cgtccattag cgaagcgtcc ggttcacaca cgtgccacgt 5100
tggcgaggtg gcaggtgaca atgatcggtg gagctgatgg tcgaaacgtt cacagcctag 5160
ggatatcgaa ttcggccgac aggacgcgcg tcaaaggtgc tggtcgtgta tgccctggcc 5220
ggcaggtcgt tgctgctgct ggttagtgat tccgcaaccc tgattttggc gtcttatttt 5280
ggcgtggcaa acgctggcgc ccgcgagccg ggccggcggc gatgcggtgc cccacggctg 5340
ccggaatcca agggaggcaa gagcgcccgg gtcagttgaa gggctttacg cgcaaggtac 5400
agccgctcct gcaaggctgc gtggtggaat tggacgtgca ggtcctgctg aagttcctcc 5460
accgcctcac cagcggacaa agcaccggtg tatcaggtcc gtgtcatcca ctctaaagaa 5520
ctcgactacg acctactgat ggccctagat tcttcatcaa aaacgcctga gacacttgcc 5580
caggattgaa actccctgaa gggaccacca ggggccctga gttgttcctt ccccccgtgg 5640
cgagctgcca gccaggctgt acctgtgatc gaggctggcg ggaaaatagg cttcgtgtgc 5700
tcaggtcatg ggaggtgcag gacagctcat gaaacgccaa caatcgcaca attcatgtca 5760
agctaatcag ctatttcctc ttcacgagct gtaattgtcc caaaattctg gtctaccggg 5820
ggtgatcctt cgtgtacggg cccttccctc aaccctaggt atgcgcgcat gcggtcgccg 5880
cgcaactcgc gcgagggccg agggtttggg acgggccgtc ccgaaatgca gttgcacccg 5940
gatgcgtggc acctttttg cgataattta tgcaatggac tgctctgcaa aattctggct 6000
ctgtcgccaa ccctaggatc agcggcgtag gatttcgtaa tcattcgtcc tgatggggag 6060
ctaccgacta ccctaatatc agcccgactg cctgacgcca gcgtccactt ttgtgcacac 6120
attccattcg tgcccaagac atttcattgt ggtgcgaagc gtccccagtt acgctcacct 6180
gtttcccgac ctccttactg ttctgtcgac agagcgggcc cacaggccgg tcgcagccac 6240
tagtatggcc accgcatcca ctttctcggc gttcaatgcc cgctgcggcg acctgcgtcg 6300
ctcggcgggc tccgggcccc ggcgcccagc gaggcccctc cccgtgcgcg ggcgcgccga 6360
ggtgcacgtg caggtgaccc actccctggc ccccgagaag cgcgagatct tcaactccct 6420
gaacaactgg gcccaggaga acatcctggt gctgctgaag gacgtggaca agtgctggca 6480
gccctccgac ttcctgcccg actccgcctc cgagggcttc gacgagcagg tgatggagct 6540
gcgcaagcgc tgcaaggaga tccccgacga ctacttcatc gtgctggtgg gcgacatgat 6600
caccgaggag gccctgccca cctaccagac catgctgaac accctggacg gcgtgcgcga 6660
cgagaccggc gcctccctga ccccctgggc catctggacc cgcgcctgga ccgccgagga 6720
gaaccgccac ggcgacctgc tgaacaagta cctgtacctg tccggccgcg tggacatgaa 6780
gcagatcgag aagaccatcc agtacctgat cggctccggc atggaccccc gcaccgagaa 6840
caacccctac ctgggcttca tctacacctc cttccaggag cgcgccacct tcatctccca 6900
cggcaacacc gcccgcctgg ccaaggagca cggcgacctg aagctggccc agatctgcgg 6960
catcatcgcc gccgacgaga agcgccacga gaccgcctac accaagatcg tggagaagct 7020
gttcgagatc gaccccgacg gcaccgtgct ggccctggcc gacatgatgc gcaagaaggt 7080
gtccatgccc gcccacctga tgtacgacgg ccaggacgac aacctgttcg agaacttctc 7140
ctccgtggcc cagcgcctgg gcgtgtacac cgccaaggac tacgccgaca tcctggagtt 7200
cctggtgggc cgctgggaca tcgagaagct gaccggcctg tccggcgagg gccgcaaggc 7260

ccaggactac gtgtgcaccc tgccccccg catccgccgc ctggaggagc gcgcccagtc 7320
ccgcgtgaag aaggcctccg ccacccccctt ctcctggatc ttcggccgcg agatcaacct 7380
gatggactac aaggaccacg acggcgacta caaggaccac gacatcgact acaaggacga 7440
cgacgacaag tgaatcgata gatctcttaa ggcagcagca gctcggatag tatcgacaca 7500
ctctggacgc tggtcgtgtg atggactgtt gccgccacac ttgctgcctt gacctgtgaa 7560
tatccctgcc gcttttatca aacagcctca gtgtgtttga tcttgtgtgt acgcgctttt 7620
gcgagttgct agctgcttgt gctatttgcg aataccaccc ccagcatccc cttccctcgt 7680
ttcatatcgc ttgcatccca accgcaactt atctacgctg tcctgctatc cctcagcgct 7740
gctcctgctc ctgctcactg cccctcgcac agccttggtt tgggctccgc ctgtattctc 7800
ctggtactgc aacctgtaaa ccagcactgc aatgctgatg cacgggaagt agtgggatgg 7860
gaacacaaat ggaaagctta attaagagct cccgccacca ctccaacacg gggtgcctgg 7920
acaaggacga ggtgtttgtg ccgccgcacc gcgcagtggc gcacgagggc ctggagtggg 7980
aggagtggct gcccatccgc atgggcaagg tgctggtcac cctgaccctg ggctggccgc 8040
tgtacctcat gttcaacgtc gcctcgcggc cgtacccgcg cttcgccaac cactttgacc 8100
cgtggtcgcc catcttcagc aagcgcgagc gcatcgaggt ggtcatctcc gacctggcgc 8160
tggtggcggt gctcagcggg ctcagcgtgc tgggccgcac catgggctgg gcctggctgg 8220
tcaagaccta cgtggtgccc tacctgatcg tgaacatgtg gctcgtgctc atcacgctgc 8280
tccagcacac gcacccggcg ctgccgcact acttcgagaa ggactgggac tggctgcgcg 8340
gcgccatggc caccgtggac cgctccatgg gcccgccctt catggacaac atcctgcacc 8400
acatctccga cacccacgtg ctgcaccacc tcttcagcac catcccgcac taccacgccg 8460
aggaggcctc cgccgccatc aggcccatcc tgggcaagta ctaccagtcc gacagccgct 8520
gggtcggccg cgccctgtgg gaggactggc gcgactgccg ctacgtcgtc ccggacgcgc 8580
ccgaggacga ctccgcgctc tggttccaca agtgagtgag tgagaagagc ctctagagtc 8640
gacctgcagg catgcaagct tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg 8700
ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgta aagcctgggg 8760
tgcctaatga gtgagctaac tcacattaat tgcgttgcgc tcactgcccg ctttccagtc 8820
gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt 8880
gcgtattggg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct 8940
gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 9000
taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 9060
cgcgttgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg 9120
ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttcccccctgg 9180
aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 9240
tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggt 9300
gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 9360
cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgc 9416

<210> 247
<211> 719
<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 247

gggctggtct gaatccttca ggcgggtgtt acccgagaaa gaaagggtgc cgatttcaaa 60

gcagacccat gtgccgggcc ctgtggcctg tgttggcgcc tatgtagtca ccccccctca 120

cccaattgtc gccagtttgc gcactccata aactcaaaac agcagcttct gagctgcgct 180

gttcaagaac acctctgggg tttgctcacc cgcgaggtcg acgcccagca tggctatcaa 240

gacgaacagg cagcctgtgg agaagcctcc gttcacgatc gggacgctgc gcaaggccat 300

ccccgcgcac tgtttcgagc gctcggcgct tcgtagcagc atgtacctgg cctttgacat 360

cgcggtcatg tccctgctct acgtcgcgtc gacgtacatc gaccctgcac cggtgcctac 420

gtgggtcaag tacggcatca tgtggccgct ctactggttc ttccaggtgt gtttgagggt 480

tttggttgcc cgtattgagg tcctggtggc gcgcatggag gagaaggcgc ctgtcccgct 540

gacccccccg gctaccctcc cggcaccttc cagggcgcct tcggcacggg tgtctgggtg 600

tgcgcgcacg agtgcggcca ccaggccttt tcctccagcc aggccatcaa cgacggcgtg 660

ggcctggtgt tccacagcct gctgctggtg ccctactact cctggaagca ctcgcaccg 719

<210> 248

<211> 732

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 248

ccgccaccac tccaacacgg ggtgcctgga caaggacgag gtgtttgtgc cgccgcaccg 60

cgcagtggcg cacgagggcc tggagtggga ggagtggctg cccatccgca tgggcaaggt 120

gctggtcacc ctgaccctgg gctggccgct gtacctcatg ttcaacgtcg cctcgcggcc 180

gtacccgcgc ttcgccaacc actttgaccc gtggtcgccc atcttcagca agcgcgagcg 240

catcgaggtg gtcatctccg acctggcgct ggtggcggtg ctcagcgggc tcagcgtgct 300

gggccgcacc atgggctggg cctggctggt caagacctac gtggtgccct acctgatcgt 360

gaacatgtgg ctcgtgctca tcacgctgct ccagcacacg cacccggcgc tgccgcacta 420

cttcgagaag gactgggact ggctgcgcgg cgccatggcc accgtggacc gctccatggg 480

cccgcccttc atggacaaca tcctgcacca catctccgac acccacgtgc tgcaccacct 540

cttcagcacc atcccgcact accacgccga ggaggcctcc gccgccatca ggcccatcct 600

gggcaagtac taccagtccg acagccgctg ggtcggccgc gccctgtggg aggactggcg 660

cgactgccgc tacgtcgtcc cggacgcgcc cgaggacgac tccgcgctct ggttccacaa 720

gtgagtgagt ga 732

pSZ2004

<210> 249

<211> 12342

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 249

tagaaaaact catcgagcat caaatgaaac tgcaatttat tcatatcagg attatcaata 60
ccatattttt gaaaaagccg tttctgtaat gaaggagaaa actcaccgag gcagttccat 120
aggatggcaa gatcctggta tcggtctgcg attccgactc gtccaacatc aatacaacct 180
attaatttcc cctcgtcaaa aataaggtta tcaagtgaga aatcaccatg agtgacgact 240
gaatccggtg agaatggcaa aagtttatgc atttctttcc agacttgttc aacaggccag 300
ccattacgct cgtcatcaaa atcactcgca tcaaccaaac cgttattcat tcgtgattgc 360
gcctgagcga ggcgaaatac gcgatcgctg ttaaaaggac aattacaaac aggaatcgag 420
tgcaaccggc gcaggaacac tgccagcgca tcaacaatat tttcacctga atcaggatat 480
tcttctaata cctggaacgc tgtttttccg gggatcgcag tggtgagtaa ccatgcatca 540
tcaggagtac ggataaaatg cttgatggtc ggaagtggca taaattccgt cagccagttt 600
agtctgacca tctcatctgt aacatcattg gcaacgctac ctttgccatg tttcagaaac 660
aactctggcg catcgggctt cccatacaag cgatagattg tcgcacctga ttgcccgaca 720
ttatcgcgag cccatttata cccatataaa tcagcatcca tgttggaatt taatcgcggc 780
ctcgacgttt cccgttgaat atggctcata ttcttccttt ttcaatatta ttgaagcatt 840
tatcagggtt attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacaa 900
ataggggtca gtgttacaac caattaacca attctgaaca ttatcgcgag cccatttata 960
cctgaatatg gctcataaca ccccttgttt gcctggcggc agtagcgcgg tggtcccacc 1020
tgaccccatg ccgaactcag aagtgaaacg ccgtagcgcc gatggtagtg tggggactcc 1080
ccatgcgaga gtagggaact gccaggcatc aaataaaacg aaaggctcag tcgaaagact 1140
gggcctttcg cccgggctaa ttagggggtg tcgcccttat tcgactctat agtgaagttc 1200
ctattctcta gaaagtatag gaacttctga agtggggaag aatgcatggc tcttccggcc 1260
tgttgccgga cgatccgtgt cgtcgagact gcattttgtt ttgggtgtgg ggctggggta 1320
ctggatggct tgagggcatg actttttctg atggagaaga ttgcaatgag atcatttggg 1380
tcgtctattt gtttgctgtg caagagggtt tactggtatc tggcaccagc ttttggcccg 1440
tgcccgtttg atggacgcgt gacaggcagg cgtcctggaa agcacagaca ccgtacgtac 1500
gaccttgacc tccccccctt ctccacacgg caggtgcgag gctgcccacg gcgtcgaggc 1560
gggcggtgcg ccgggcatgg tcccgcatcg cgcgcgcggc ggccgcggcc gacgcaaacc 1620
ccgcccgccc tgagcgccgc gtggtcatca cgggccaggg cgtggtgacc agcctgggcc 1680
agacgatcga gcagttttac agcagcctgc tggagggcgt gagcggcatc tcgcagatac 1740
agaagttcga caccacgggc tacacgacga cgatcgcggg cgagatcaag tcgctgcagc 1800
tggacccgta cgtgcccaag cgctgggcga agcgcgtgga cgacgtgata aagtacgtct 1860
acatcgcggg caagcaggcg ctggagagcg ccggcctgcc gatcgaggcg gcggggctgg 1920
cgggcgcggg gctggacccg gcgctgtgcg gcgtgctcat cggcaccgcc atggcgggca 1980
tgacgtcttt cgcggcgggc gtggaggcgc tgacgcgcgg taccctttct tgcgctatga 2040

cacttccagc aaaaggtagg gcgggctgcg agacggcttc ccggcgctgc atgcaacacc 2100
gatgatgctt cgacccccg aagctccttc ggggctgcat gggcgctccg atgccgctcc 2160
agggcgagcg ctgtttaaat agccaggccc ccgattgcaa agacattata gcgagctacc 2220
aaagccatat tcaaacacct agatcactac cacttctaca caggccactc gagcttgtga 2280
tcgcactccg ctaaggggc gcctcttcct cttcgtttca gtcacaaccc gcaaacggcg 2340
cgccatgctg ctgcaggcct tcctgttcct gctggccggc ttcgccgcca agatcagcgc 2400
ctccatgacg aacgagacgt ccgaccgccc cctggtgcac ttcaccccca acaagggctg 2460
gatgaacgac cccaacggcc tgtggtacga cgagaaggac gccaagtggc acctgtactt 2520
ccagtacaac ccgaacgaca ccgtctgggg gacgcccttg ttctggggcc acgccacgtc 2580
cgacgacctg accaactggg aggaccagcc catcgccatc gccccgaagc gcaacgactc 2640
cggcgccttc tccggctcca tggtggtgga ctacaacaac acctccggct tcttcaacga 2700
caccatcgac ccgcgccagc gctgcgtggc catctggacc tacaacaccc cggagtccga 2760
ggagcagtac atctcctaca gcctggacgg cggctacacc ttcaccgagt accagaagaa 2820
ccccgtgctg gccgccaact ccacccagtt ccgcgacccg aaggtcttct ggtacgagcc 2880
ctcccagaag tggatcatga ccgcggccaa gtcccaggac tacaagatcg agatctactc 2940
ctccgacgac ctgaagtcct ggaagctgga gtccgcgttc gccaacgagg gcttcctcgg 3000
ctaccagtac gagtgccccg gcctgatcga ggtccccacc gagcaggacc ccagcaagtc 3060
ctactgggtg atgttcatct ccatcaaccc cggcgccccg gccggcggct ccttcaacca 3120
gtacttcgtc ggcagcttca acggcaccca cttcgaggcc ttcgacaacc agtcccgcgt 3180
ggtggacttc ggcaaggact actacgccct gcagaccttc ttcaacaccg acccgaccta 3240
cgggagcgcc ctgggcatcg cgtgggcctc caactgggag tactccgcct tcgtgcccac 3300
caacccctgg cgctcctcca tgtccctcgt gcgcaagttc tccctcaaca ccgagtacca 3360
ggccaacccg gagacggagc tgatcaacct gaaggccgag ccgatcctga acatcagcaa 3420
cgccggcccc tggagccggt tcgccaccaa caccacgttg acgaaggcca acagctacaa 3480
cgtcgacctg tccaacagca ccggcaccct ggagttcgag ctggtgtacg ccgtcaacac 3540
cacccagacg atctccaagt ccgtgttcgc ggacctctcc ctctggttca agggcctgga 3600
ggaccccgag gagtacctcc gcatgggctt cgaggtgtcc gcgtcctcct tcttcctgga 3660
ccgcgggaac agcaaggtga agttcgtgaa ggagaacccc tacttcacca accgcatgag 3720
cgtgaacaac cagcccttca agagcgagaa cgacctgtcc tactacaagg tgtacggctt 3780
gctggaccag aacatcctgg agctgtactt caacgacggc gacgtcgtgt ccaccaacac 3840
ctacttcatg accaccggga acgccctggg ctccgtgaac atgacgacgg gggtggacaa 3900
cctgttctac atcgacaagt tccaggtgcg cgaggtcaag tgacaattgg cagcagcagc 3960
tcggatagta tcgacacact ctggacgctg gtcgtgtgat ggactgttgc cgccacactt 4020
gctgccttga cctgtgaata tccctgccgc ttttatcaaa cagcctcagt gtgtttgatc 4080
ttgtgtgtac gcgcttttgc gagttgctag ctgcttgtgc tatttgcgaa taccaccccc 4140
agcatcccct tccctcgttt catatcgctt gcatcccaac cgcaacttat ctacgctgtc 4200
ctgctatccc tcagcgctgc tcctgctcct gctcactgcc cctcgcacag ccttggtttg 4260
ggctccgcct gtattctcct ggtactgcaa cctgtaaacc agcactgcaa tgctgatgca 4320
cgggaagtag tgggatggga acacaaatgg aggatcgtag agctggatcc tctaaagctt 4380

ggccgacagg acgcgcgtca aaggtgctgg tcgtgtatgc cctggccggc aggtcgttgc 4440
tgctgctggt tagtgattcc gcaaccctga ttttggcgtc ttattttggc gtggcaaacg 4500
ctggcgcccg cgagccgggc cggcggcgat gcggtgcccc acggctgccg gaatccaagg 4560
gaggcaagag cgcccgggtc agttgaaggg ctttacgcgc aaggtacagc cgctcctgca 4620
aggctgcgtg gtggaattgg acgtgcaggt cctgctgaag ttcctccacc gcctcaccag 4680
cggacaaagc accggtgtat caggtccgtg tcatccactc taaagaactc gactacgacc 4740
tactgatggc cctagattct tcatcaaaaa cgcctgagac acttgcccag gattgaaact 4800
ccctgaaggg accaccaggg gccctgagtt gttccttccc cccgtggcga gctgccagcc 4860
aggctgtacc tgtgatcgag gctggcggga aaataggctt cgtgtgctca ggtcatggga 4920
ggtgcaggac agctcatgaa acgccaacaa tcgcacaatt catgtcaagc taatcagcta 4980
tttcctcttc acgagctgta attgtcccaa aattctggtc taccgggggt gatccttcgt 5040
gtacgggccc ttccctcaac cctaggtatg cgcgcatgcg gtcgccgcgc aactcgcgcg 5100
agggccgagg gtttgggacg ggccgtcccg aaatgcagtt gcacccggat gcgtggcacc 5160
ttttttgcga taatttatgc aatggactgc tctgcaaaat tctggctctg tcgccaaccc 5220
taggatcagc ggcgtaggat ttcgtaatca ttcgtcctga tggggagcta ccgactaccc 5280
taatatcagc ccgactgcct gacgccagcg tccacttttg tgcacacatt ccattcgtgc 5340
ccaagacatt tcattgtggt gcgaagcgtc cccagttacg ctcacctgtt tcccgacctc 5400
cttactgttc tgtcgacaga gcgggcccac aggccggtcg cagccactag tatggtggcc 5460
accgccgcct cctccgcctt cttccccctg ccctccgccg acacctcctc ccgccccggc 5520
aagctgggca acaagccctc ctccctgtcc cccctgaagc ccaagtccac ccccaacggc 5580
ggcctgcagg tgaaggccaa cgcctccgcc ccccccaaga tcaacggctc ccccgtgggc 5640
ctgaagtccg gcggcctgaa ggggcgcgcc acccaggagg acgcccactc cgccccccccc 5700
ccccgcacct tcatcaacca gctgcccgac tggtccatgc tgctggccgc catcaccacc 5760
gtgttcctgg ccgccgagaa gcagtggatg atgctggact ggaagcccaa gcgccccgac 5820
atgctggtgg acccccttcgg cctgggctcc atcgtgcagg acggcctggt gttccgccag 5880
aacttctcca tccgctccta cgagatcggc gccgaccgca ccgcctccat cgagaccgtg 5940
atgaaccacc tgcaggagac cgccctgaac cacgtgaaga tcgccggcct gtccaacgac 6000
ggcttcggcc gcacccccga gatgtacaag cgcgacctga tctgggtggt ggccaagatg 6060
caggtgatgg tgaaccgcta ccccacctgg ggcgacaccg tggaggtgaa cacctgggtg 6120
gccaagtccg gcaagaacgg catgcgccgc gactggctga tctccgactg caacaccggc 6180
gagatcctga cccgcgcctc ctccgtgtgg gtgatgatga accagaagac ccgccgcctg 6240
tccaagatcc ccgacgaggt gcgcaacgag atcgagcccc acttcgtgga ctccccccccc 6300
gtgatcgagg acgacgaccg caagctgccc aagctggacg agaagaccgc cgactccatc 6360
cgcaagggcc tgacccccccg ctggaacgac ctggacgtga accagcacgt gaacaacgtg 6420
aagtacatcg gctggatcct ggagtccacc ccccccgagg tgctggagac ccaggagctg 6480
tgctccctga ccctggagta ccgccgcgag tgcggccgcg agtccgtgct ggagtccctg 6540
accgccatgg acccctccgg cggcggctac ggctcccagt tccagcacct gctgcgcctg 6600
gaggacggcg gcgagatcgt gaagggccgc accgagtggc gccccaagaa cggcgtgatc 6660
aacggcgtgg tgcccaccgg cgagtcctcc cccggcgact actccatgga ctacaaggac 6720

cacgacggcg actacaagga ccacgacatc gactacaagg acgacgacga caagtgaatc 6780
gatagatctc ttaaggcagc agcagctcgg atagtatcga cacactctgg acgctggtcg 6840
tgtgatggac tgttgccgcc acacttgctg ccttgacctg tgaatatccc tgccgctttt 6900
atcaaacagc ctcagtgtgt ttgatcttgt gtgtacgcgc ttttgcgagt tgctagctgc 6960
ttgtgctatt tgcgaatacc acccccagca tccccttccc tcgtttcata tcgcttgcat 7020
cccaaccgca acttatctac gctgtcctgc tatccctcag cgctgctcct gctcctgctc 7080
actgcccctc gcacagcctt ggtttgggct ccgcctgtat tctcctggta ctgcaacctg 7140
taaaccagca ctgcaatgct gatgcacggg aagtagtggg atgggaacac aaatggaaag 7200
ctgtagaatt cggccgacag gacgcgcgtc aaaggtgctg gtcgtgtatg ccctggccgg 7260
caggtcgttg ctgctgctgg ttagtgattc cgcaaccctg attttggcgt cttattttgg 7320
cgtggcaaac gctggcgccc gcgagccggg ccggcggcga tgcggtgccc cacggctgcc 7380
ggaatccaag ggaggcaaga gcgcccgggt cagttgaagg gctttacgcg caaggtacag 7440
ccgctcctgc aaggctgcgt ggtggaattg gacgtgcagg tcctgctgaa gttcctccac 7500
cgcctcacca gcggacaaag caccggtgta tcaggtccgt gtcatccact ctaaagaact 7560
cgactacgac ctactgatgg ccctagattc ttcatcaaaa acgcctgaga cacttgccca 7620
ggattgaaac tccctgaagg gaccaccagg ggccctgagt tgttccttcc ccccgtggcg 7680
agctgccagc caggctgtac ctgtgatcga ggctggcggg aaaataggct tcgtgtgctc 7740
aggtcatggg aggtgcagga cagctcatga aacgccaaca atcgcacaat tcatgtcaag 7800
ctaatcagct atttcctctt cacgagctgt aattgtccca aaattctggt ctaccggggg 7860
tgatccttcg tgtacgggcc cttccctcaa ccctaggtat gcgcgcatgc ggtcgccgcg 7920
caactcgcgc gagggccgag ggtttgggac gggccgtccc gaaatgcagt tgcacccgga 7980
tgcgtggcac ctttttgcg ataatttatg caatggactg ctctgcaaaa ttctggctct 8040
gtcgccaacc ctaggatcag cggcgtagga tttcgtaatc attcgtcctg atggggagct 8100
accgactacc ctaatatcag cccgactgcc tgacgccagc gtccactttt gtgcacacat 8160
tccattcgtg cccaagacat ttcattgtgg tgcgaagcgt ccccagttac gctcacctgt 8220
ttcccgacct ccttactgtt ctgtcgacag agcgggccca caggccggtc gcagccacta 8280
gtatggtggc caccgccgcc tcctccgcct tcttcccccc gccctccgcc gacacctcct 8340
cccgccccgg caagctgggc aacaagccct cctccctgtc ccccctgaag cccaagtcca 8400
cccccaacgg cggcctgcag gtgaaggcca acgcctccgc ccccccccaag atcaacggct 8460
cccccgtggg cctgaagtcc ggcggcctga aggggcgcgc cacccaggag gacgcccact 8520
ccgccccccc cccccgcacc ttcatcaacc agctgcccga ctggtccatg ctgctggccg 8580
ccatcaccac cgtgttcctg gccgccgaga agcagtggat gatgctggac tggaagccca 8640
agcgccccga catgctggtg gacccccttcg gcctgggctc catcgtgcag gacggcctgg 8700
tgttccgcca gaacttctcc atccgctcct acgagatcgg cgccgaccgc accgcctcca 8760
tcgagaccgt gatgaaccac ctgcaggaga ccgccctgaa ccacgtgaag atcgccggcc 8820
tgtccaacga cggcttcggc cgcacccccg agatgtacaa gcgcgacctg atctgggtgg 8880
tggccaagat gcaggtgatg gtgaaccgct accccacctg gggcgacacc gtggaggtga 8940
acacctgggt ggccaagtcc ggcaagaacg gcatgcgccg cgactggctg atctccgact 9000
gcaacaccgg cgagatcctg acccgcgcct cctccgtgtg ggtgatgatg aaccagaaga 9060

cccgccgcct gtccaagatc cccgacgagg tgcgcaacga gatcgagccc cacttcgtgg 9120
actccccccc cgtgatcgag gacgacgacc gcaagctgcc caagctggac gagaagaccg 9180
ccgactccat ccgcaagggc ctgacccccc gctggaacga cctggacgtg aaccagcacg 9240
tgaacaacgt gaagtacatc ggctggatcc tggagtccac ccccccccgag gtgctggaga 9300
cccaggagct gtgctccctg accctggagt accgccgcga gtgcggccgc gagtccgtgc 9360
tggagtccct gaccgccatg gacccctccg gcggcggcta cggctcccag ttccagcacc 9420
tgctgcgcct ggaggacggc ggcgagatcg tgaagggccg caccgagtgg cgccccaaga 9480
acggcgtgat caacggcgtg gtgcccaccg gcgagtcctc ccccggcgac tactccatgg 9540
actacaagga ccacgacggc gactacaagg accacgacat cgactacaag gacgacgacg 9600
acaagtgaat cgatagatct cttaaggcag cagcagctcg gatagtatcg acacactctg 9660
gacgctggtc gtgtgatgga ctgttgccgc cacacttgct gccttgacct gtgaatatcc 9720
ctgccgcttt tatcaaacag cctcagtgtg tttgatcttg tgtgtacgcg cttttgcgag 9780
ttgctagctg cttgtgctat ttgcgaatac cacccccagc atcccccttcc ctcgtttcat 9840
atcgcttgca tcccaaccgc aacttatcta cgctgtcctg ctatccctca gcgctgctcc 9900
tgctcctgct cactgcccct cgcacagcct tggtttgggc tccgcctgta ttctcctggt 9960
actgcaacct gtaaaccagc actgcaatgc tgatgcacgg gaagtagtgg gatgggaaca 10020
caaatggaaa gctgtagagc tcggcggcgt gcgcaagatg aacccctttt gcatcccctt 10080
ctccatctcc aacatgggcg gcgcgatgct ggcgatggac atcggcttca tgggccccaa 10140
ctactccatc tccacggcct gcgcgacggg caactactgc atcctgggcg cggcggacca 10200
catccggcgc ggcgacgcaa acgtgatgct ggccggcggc gcggacgcgg ccatcatccc 10260
ctcgggcatc ggcggcttca tcgcgtgcaa ggcgctgagc aagcgcaacg acgagcccga 10320
gcgcgcgagc cggccctggg acgccgaccg cgacggcttc gtcatgggcg agggcgccgg 10380
cgtgctggtg ctggaggagc tggagcacgc caagcgccgc ggcgcgacca ttttggctga 10440
attagttggc ggcgcggcca cctcggacgc gcaccacatg accgagcccg acccgcaggg 10500
ccgcggcgtg cgcctctgcc tcgagcgcgc gctcgagcgc gcgcgcctcg cgcccgagcg 10560
cgtcggctac gtcaacgcgc acggcaccag cacgcccgcg ggcgacgtgg ccgagtaccg 10620
cgccatccgc gccgtcatcc cgcaggactc actacgcatc aactccacaa agtccatgat 10680
cgggcacctg ctcggcggcg ccggcgcggt cgaggccgtg gccgccatcc aggccctgcg 10740
caccggctgg ctccacccca acttgaacct cgagaacccc gcgcctggcg tcgaccccgt 10800
cgtgctcgtg gggaagagcc tctagatctt aaaatgaagt gaagttccta tactttctag 10860
agaataggaa cttctatagt gagtcgaata agggcgacac aaaatttatt ctaaatgcat 10920
aataaatact gataacatct tatagtttgt attatatttt gtattatcgt tgacatgtat 10980
aattttgata tcaaaaactg attttccctt tattattttc gagatttatt ttcttaattc 11040
tctttaacaa actagaaata ttgtatatac aaaaaatcat aaataataga tgaatagttt 11100
aattataggt gttcatcaat cgaaaaagca acgtatctta tttaaagtgc gttgcttttt 11160
tctcatttat aaggttaaat aattctcata tatcaagcaa agtgacaggc gcccttaaat 11220
attctgacaa atgctctttc cctaaactcc ccccataaaa aaacccgccg aagcgggttt 11280
ttacgttatt tgcggattaa cgattactcg ttatcagaac cgcccagggg gcccgagctt 11340
aagactggcc gtcgttttac aacacagaaa gagtttgtag aaacgcaaaa aggccatccg 11400

tcaggggcct tctgcttagt ttgatgcctg gcagttccct actctcgcct tccgcttcct 11460
cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa 11520
aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa 11580
aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt ttccataggc 11640
tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg cgaaacccga 11700
caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc tctcctgttc 11760
cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc gtggcgcttt 11820
ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc aagctgggct 11880
gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac tatcgtcttg 11940
agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta 12000
gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtgggct aactacggct 12060
acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc ttcggaaaaa 12120
gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt ttttttgttt 12180
gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta 12240
cggggtctga cgctcagtgg aacgacgcgc gcgtaactca cgttaaggga ttttggtcat 12300
gagcttgcgc cgtcccgtca agtcagcgta atgctctgct tt 12342

pSZ1500

<210> 250

<211> 9362

<212> DNA

<213> 人工序列

<220>

<223> 合成的构建体

<400> 250

cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag 60
agttcttgaa gtggtggcct aactacggct acactagaag gacagtattt ggtatctgcg 120
ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa 180
ccaccgctgg tagcggtggt ttttttgttt gcaagcagca gattacgcgc agaaaaaaag 240
gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg aacgaaaact 300
cacgttaagg gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa 360
attaaaaatg aagttttaaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt 420
accaatgctt aatcagtgag gcacctatct cagcgatctg tctatttcgt tcatccatag 480
ttgcctgact ccccgtcgtg tagataacta cgatacggga gggcttacca tctggcccca 540
gtgctgcaat gataccgcga gacccacgct caccggctcc agatttatca gcaataaacc 600
agccagccgg aagggccgag cgcagaagtg gtcctgcaac tttatccgcc tccatccagt 660
ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg 720
ttgttgccat tgctacaggc atcgtggtgt cacgctcgtc gtttggtatg gcttcattca 780
gctccggttc ccaacgatca aggcgagtta catgatcccc catgttgtgc aaaaaagcgg 840
ttagctcctt cggtcctccg atcgttgtca gaagtaagtt ggccgcagtg ttatcactca 900

tggttatggc agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg 960
tgactggtga gtactcaacc aagtcattct gagaatagtg tatgcggcga ccgagttgct 1020
cttgcccggc gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgctca 1080
tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca 1140
gttcgatgta acccactcgt gcacccaact gatcttcagc atcttttact ttcaccagcg 1200
tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata agggcgacac 1260
ggaaatgttg aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt 1320
attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacaa ataggggttc 1380
cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat 1440
taacctataa aaataggcgt atcacgaggc cctttcgtct cgcgcgtttc ggtgatgacg 1500
gtgaaaacct ctgacacatg cagctcccgg agacggtcac agcttgtctg taagcggatg 1560
ccgggagcag acaagcccgt cagggcgcgt cagcgggtgt tggcgggtgt cggggctggc 1620
ttaactatgc ggcatcagag cagattgtac tgagagtgca ccatatgcgg tgtgaaatac 1680
cgcacagatg cgtaaggaga aaataccgca tcaggcgcca ttcgccattc aggctgcgca 1740
actgttggga agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg 1800
gatgtgctgc aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta 1860
aaacgacggc cagtgaattg atgcatgctc ttcgggctgg tctgaatcct tcaggcgggt 1920
gttacccgag aaagaaaggg tgccgatttc aaagcagacc catgtgccgg gccctgtggc 1980
ctgtgttggc gcctatgtag tcaccccccc tcacccaatt gtcgccagtt tgcgcactcc 2040
ataaactcaa aacagcagct tctgagctgc gctgttcaag aacacctctg gggtttgctc 2100
acccgcgagg tcgacgccca gcatggctat caagacgaac aggcagcctg tggagaagcc 2160
tccgttcacg atcgggacgc tgcgcaaggc catccccgcg cactgtttcg agcgctcggc 2220
gcttcgtagc agcatgtacc tggcctttga catcgcggtc atgtccctgc tctacgtcgc 2280
gtcgacgtac atcgaccctg caccggtgcc tacgtgggtc aagtacggca tcatgtggcc 2340
gctctactgg ttcttccagg tgtgtttgag ggttttggtt gcccgtattg aggtcctggt 2400
ggcgcgcatg gaggagaagg cgcctgtccc gctgaccccc ccggctaccc tcccggcacc 2460
ttccagggcg ccttcggcac gggtgtctgg gtgtgcgcgc acgagtgcgg ccaccaggcc 2520
ttttcctcca gccaggccat caacgacggc gtgggcctgg tgttccacag cctgctgctg 2580
gtgccctact actcctggaa gcactcgcac cgggtaccct ttcttgcgct atgacacttc 2640
cagcaaaagg tagggcgggc tgcgagacgg cttcccggcg ctgcatgcaa caccgatgat 2700
gcttcgaccc cccgaagctc cttcggggct gcatgggcgc tccgatgccg ctccagggcg 2760
agcgctgttt aaatagccag gcccccgatt gcaaagacat tatagcgagc taccaaagcc 2820
atattcaaac acctagatca ctaccacttc tacacaggcc actcgagctt gtgatcgcac 2880
tccgctaagg gggcgcctct tcctcttcgt ttcagtcaca acccgcaaac tctagaatat 2940
caatgctgct gcaggccttc ctgttcctgc tggccggctt cgccgccaag atcagcgcct 3000
ccatgacgaa cgagacgtcc gaccgccccc tggtgcactt cacccccaac aagggctgga 3060
tgaacgaccc caacggcctg tggtacgacg agaaggacgc caagtggcac ctgtacttcc 3120
agtacaaccc gaacgacacc gtctggggga cgcccttgtt ctggggccac gccacgtccg 3180
acgacctgac caactgggag gaccagccca tcgccatcgc cccgaagcgc aacgactccg 3240

gcgccttctc cggctccatg gtggtggact acaacaacac ctccggcttc ttcaacgaca 3300
ccatcgaccc gcgccagcgc tgcgtggcca tctggaccta caacaccccg gagtccgagg 3360
agcagtacat ctcctacagc ctggacggcg gctacacctt caccgagtac cagaagaacc 3420
ccgtgctggc cgccaactcc acccagttcc gcgacccgaa ggtcttctgg tacgagccct 3480
cccagaagtg gatcatgacc gcggccaagt cccaggacta caagatcgag atctactcct 3540
ccgacgacct gaagtcctgg aagctggagt ccgcgttcgc caacgagggc ttcctcggct 3600
accagtacga gtgccccggc ctgatcgagg tccccaccga gcaggacccc agcaagtcct 3660
actgggtgat gttcatctcc atcaaccccg gcgccccggc cggcggctcc ttcaaccagt 3720
acttcgtcgg cagcttcaac ggcacccact tcgaggcctt cgacaaccag tcccgcgtgg 3780
tggacttcgg caaggactac tacgccctgc agaccttctt caacaccgac ccgacctacg 3840
ggagcgccct gggcatcgcg tgggcctcca actgggagta ctccgccttc gtgcccacca 3900
acccctggcg ctcctccatg tccctcgtgc gcaagttctc cctcaacacc gagtaccagg 3960
ccaacccgga gacggagctg atcaacctga aggccgagcc gatcctgaac atcagcaacg 4020
ccggcccctg gagccggttc gccaccaaca ccacgttgac gaaggccaac agctacaacg 4080
tcgacctgtc caacagcacc ggcaccctgg agttcgagct ggtgtacgcc gtcaacacca 4140
cccagacgat ctccaagtcc gtgttcgcgg acctctccct ctggttcaag ggcctggagg 4200
accccgagga gtacctccgc atgggcttcg aggtgtccgc gtcctccttc ttcctggacc 4260
gcgggaacag caaggtgaag ttcgtgaagg agaaccccta cttcaccaac cgcatgagcg 4320
tgaacaacca gcccttcaag agcgagaacg acctgtccta ctacaaggtg tacggcttgc 4380
tggaccagaa catcctggag ctgtacttca acgacggcga cgtcgtgtcc accaacacct 4440
acttcatgac caccgggaac gccctgggct ccgtgaacat gacgacgggg gtggacaacc 4500
tgttctacat cgacaagttc caggtgcgcg aggtcaagtg acaattggca gcagcagctc 4560
ggatagtatc gacacactct ggacgctggt cgtgtgatgg actgttgccg ccacacttgc 4620
tgccttgacc tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt gtttgatctt 4680
gtgtgtacgc gcttttgcga gttgctagct gcttgtgcta tttgcgaata ccacccccag 4740
catccccttc cctcgtttca tatcgcttgc atcccaaccg caacttatct acgctgtcct 4800
gctatccctc agcgctgctc ctgctcctgc tcactgcccc tcgcacagcc ttggtttggg 4860
ctccgcctgt attctcctgg tactgcaacc tgtaaaccag cactgcaatg ctgatgcacg 4920
ggaagtagtg ggatgggaac acaaatggag gatcccgcgt ctcgaacaga gcgcgcagag 4980
gaacgctgaa ggtctcgcct ctgtcgcacc tcagcgcggc atacaccaca ataaccacct 5040
gacgaatgcg cttggttctt cgtccattag cgaagcgtcc ggttcacaca cgtgccacgt 5100
tggcgaggtg gcaggtgaca atgatcggtg gagctgatgg tcgaaacgtt cacagcctag 5160
ggatatcgaa ttcggccgac aggacgcgcg tcaaaggtgc tggtcgtgta tgccctggcc 5220
ggcaggtcgt tgctgctgct ggttagtgat tccgcaaccc tgattttggc gtcttatttt 5280
ggcgtggcaa acgctggcgc ccgcgagccg ggccggcggc gatgcggtgc cccacggctg 5340
ccggaatcca agggaggcaa gagcgcccgg gtcagttgaa gggctttacg cgcaaggtac 5400
agccgctcct gcaaggctgc gtggtggaat tggacgtgca ggtcctgctg aagttcctcc 5460
accgcctcac cagcggacaa agcaccggtg tatcaggtcc gtgtcatcca ctctaaagaa 5520
ctcgactacg acctactgat ggccctagat tcttcatcaa aaacgcctga gacacttgcc 5580

caggattgaa actccctgaa gggaccacca ggggccctga gttgttcctt ccccccgtgg 5640
cgagctgcca gccaggctgt acctgtgatc gaggctggcg ggaaaatagg cttcgtgtgc 5700
tcaggtcatg ggaggtgcag gacagctcat gaaacgccaa caatcgcaca attcatgtca 5760
agctaatcag ctatttcctc ttcacgagct gtaattgtcc caaaattctg gtctaccggg 5820
ggtgatcctt cgtgtacggg cccttccctc aaccctaggt atgcgcgcat gcggtcgccg 5880
cgcaactcgc gcgagggccg agggtttggg acgggccgtc ccgaaatgca gttgcacccg 5940
gatgcgtggc acctttttg cgataattta tgcaatggac tgctctgcaa aattctggct 6000
ctgtcgccaa ccctaggatc agcggcgtag gatttcgtaa tcattcgtcc tgatggggag 6060
ctaccgacta ccctaatatc agcccgactg cctgacgcca gcgtccactt ttgtgcacac 6120
attccattcg tgcccaagac atttcattgt ggtgcgaagc gtccccagtt acgctcacct 6180
gtttcccgac ctccttactg ttctgtcgac agagcgggcc cacaggccgg tcgcagccac 6240
tagtatggcc accgcatcca ctttctcggc gttcaatgcc cgctgcggcg acctgcgtcg 6300
ctcggcgggc tccgggcccc ggcgcccagc gaggcccctc cccgtgcgcg ggcgcgccgc 6360
caccggcgag cagccctccg gcgtggcctc cctgcgcgag gccgacaagg agaagtccct 6420
gggcaaccgc ctgcgcctgg gctccctgac cgaggacggc ctgtcctaca aggagaagtt 6480
cgtgatccgc tgctacgagg tgggcatcaa caagaccgcc accatcgaga ccatcgccaa 6540
cctgctgcag gaggtgggcg gcaaccacgc ccagggcgtg ggcttctcca ccgacggctt 6600
cgccaccacc accaccatgc gcaagctgca cctgatctgg gtgaccgccc gcatgcacat 6660
cgagatctac cgctaccccg cctggtccga cgtgatcgag atcgagacct gggtgcaggg 6720
cgagggcaag gtgggcaccc gccgcgactg gatcctgaag gactacgcca acggcgaggt 6780
gatcggccgc gccacctcca agtgggtgat gatgaacgag gacacccgcc gcctgcagaa 6840
ggtgtccgac gacgtgcgcg aggagtacct ggtgttctgc ccccgcaccc tgcgcctggc 6900
cttccccgag gagaacaaca actccatgaa gaagatcccc aagctggagg accccgccga 6960
gtactcccgc ctgggcctgg tgccccgccg ctccgacctg gacatgaaca agcacgtgaa 7020
caacgtgacc tacatcggct gggccctgga gtccatcccc cccgagatca tcgacaccca 7080
cgagctgcag gccatcaccc tggactaccg ccgcgagtgc cagcgcgacg acatcgtgga 7140
ctccctgacc tcccgcgagc ccctgggcaa cgccgccggc gtgaagttca aggagatcaa 7200
cggctccgtg tcccccaaga aggacgagca ggacctgtcc cgcttcatgc acctgctgcg 7260
ctccgccggc tccggcctgg agatcaaccg ctgccgcacc gagtggcgca agaagcccgc 7320
caagcgcatg gactacaagg accacgacgg cgactacaag gaccacgaca tcgactacaa 7380
ggacgacgac gacaagtgaa tcgatagatc tcttaaggca gcagcagctc ggatagtatc 7440
gacacactct ggacgctggt cgtgtgatgg actgttgccg ccacacttgc tgccttgacc 7500
tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt gtttgatctt gtgtgtacgc 7560
gcttttgcga gttgctagct gcttgtgcta tttgcgaata ccacccccag catcccccttc 7620
cctcgtttca tatcgcttgc atcccaaccg caacttatct acgctgtcct gctatccctc 7680
agcgctgctc ctgctcctgc tcactgcccc tcgcacagcc ttggtttggg ctccgcctgt 7740
attctcctgg tactgcaacc tgtaaaccag cactgcaatg ctgatgcacg ggaagtagtg 7800
ggatgggaac acaaatggaa agcttaatta agagctcccg ccaccactcc aacacggggt 7860
gcctggacaa ggacgaggtg tttgtgccgc cgcaccgcgc agtggcgcac gagggcctgg 7920

agtgggagga gtggctgccc atccgcatgg gcaaggtgct ggtcaccctg accctgggct 7980
ggccgctgta cctcatgttc aacgtcgcct cgcggccgta cccgcgcttc gccaaccact 8040
ttgacccgtg gtcgcccatc ttcagcaagc gcgagcgcat cgaggtggtc atctccgacc 8100
tggcgctggt ggcggtgctc agcgggctca gcgtgctggg ccgcaccatg ggctgggcct 8160
ggctggtcaa gacctacgtg gtgccctacc tgatcgtgaa catgtggctc gtgctcatca 8220
cgctgctcca gcacacgcac ccggcgctgc cgcactactt cgagaaggac tgggactggc 8280
tgcgcggcgc catggccacc gtggaccgct ccatgggccc gcccttcatg gacaacatcc 8340
tgcaccacat ctccgacacc cacgtgctgc accacctctt cagcaccatc ccgcactacc 8400
acgccgagga ggcctccgcc gccatcaggc ccatcctggg caagtactac cagtccgaca 8460
gccgctgggt cggccgcgcc ctgtgggagg actggcgcga ctgccgctac gtcgtcccgg 8520
acgcgcccga ggacgactcc gcgctctggt tccacaagtg agtgagtgag aagagcctct 8580
agagtcgacc tgcaggcatg caagcttggc gtaatcatgg tcatagctgt ttcctgtgtg 8640
aaattgttat ccgctcacaa ttccacacaa catacgagcc ggaagcataa agtgtaaagc 8700
ctggggtgcc taatgagtga gctaactcac attaattgcg ttgcgctcac tgcccgcttt 8760
ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg 8820
cggtttgcgt attgggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt 8880
tcggctgcgg cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc 8940
aggggataac gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa 9000
aaaggccgcg ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa 9060
tcgacgctca agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc 9120
ccctggaagc tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc 9180
cgcctttctc ccttcgggaa gcgtggcgct ttctcatagc tcacgctgta ggtatctcag 9240
ttcggtgtag gtcgttcgct ccaagctggg ctgtgtgcac gaacccccg ttcagcccga 9300
ccgctgcgcc ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc 9360
gc 9362

Prototheca moriformis FATA1敲除同源重组靶向构建体的5'供体DNA序列

<210> 251
<211> 612
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 251

gctcttcgga gtcactgtgc cactgagttc gactggtagc tgaatggagt cgctgctcca 60
ctaaacgaat tgtcagcacc gccagccggc cgaggacccg agtcatagcg agggtagtag 120
cgcgccatgg caccgaccag cctgcttgcc agtactggcg tctcttccgc ttctctgtgg 180
tcctctgcgc gctccagcgc gtgcgctttt ccggtggatc atgcggtccg tggcgcaccg 240
cagcggccgc tgcccatgca gcgccgctgc ttccgaacag tggcggtcag ggccgcaccc 300
gcggtagccg tccgtccgga acccgcccaa gagttttggg agcagcttga gccctgcaag 360

atggcggagg acaagcgcat cttcctggag gagcaccggt gcgtggaggt ccggggctga 420
ccggccgtcg cattcaacgt aatcaatcgc atgatgatca gaggacacga agtcttggtg 480
gcggtggcca gaaacactgt ccattgcaag ggcataggga tgcgttcctt cacctctcat 540
ttctcatttc tgaatccctc cctgctcact ctttctcctc ctccttcccg ttcacgcagc 600
attcggggta cc 612

Prototheca moriformis FATA1敲除同源重组靶向构建体的3'供体DNA序列

<210> 252
<211> 528
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 252

gacagggtgg ttggctggat ggggaaacgc tggtcgcggg attcgatcct gctgcttata 60
tcctccctgg aagcacaccc acgactctga agaagaaaac gtgcacacac acaacccaac 120
cggccgaata tttgcttcct tatcccgggt ccaagagaga ctgcgatgcc cccctcaatc 180
agcatcctcc tccctgccgc ttcaatcttc cctgcttgcc tgcgcccgcg gtgcgccgtc 240
tgcccgccca gtcagtcact cctgcacagg ccccttgtgc gcagtgctcc tgtacccttt 300
accgctcctt ccattctgcg aggcccccta ttgaatgtat tcgttgcctg tgtggccaag 360
cgggctgctg ggcgcgccgc cgtcgggcag tgctcggcga ctttggcgga agccgattgt 420
tcttctgtaa gccacgcgct tgctgctttg ggaagagaag ggggggggta ctgaatggat 480
gaggaggaga aggaggggta ttggtattat ctgagttggg tgaagagc 528

<210> 253
<211> 3499
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 253

gctcttcgga gtcactgtgc cactgagttc gactggtagc tgaatggagt cgctgctcca 60
ctaaacgaat tgtcagcacc gccagccggc cgaggacccg agtcatagcg agggtagtag 120
cgcgccatgg caccgaccag cctgcttgcc agtactggcg tctcttccgc ttctctgtgg 180
tcctctgcgc gctccagcgc gtgcgctttt ccggtggatc atgcggtccg tggcgcaccg 240
cagcggccgc tgcccatgca gcgccgctgc ttccgaacag tggcggtcag ggccgcaccc 300
gcggtagccg tccgtccgga acccgcccaa gagttttggg agcagcttga gccctgcaag 360
atggcggagg acaagcgcat cttcctggag gagcaccggt gcgtggaggt ccggggctga 420
ccggccgtcg cattcaacgt aatcaatcgc atgatgatca gaggacacga agtcttggtg 480
gcggtggcca gaaacactgt ccattgcaag ggcataggga tgcgttcctt cacctctcat 540
ttctcatttc tgaatccctc cctgctcact ctttctcctc ctccttcccg ttcacgcagc 600

attcggggta ccctttcttg cgctatgaca cttccagcaa aaggtagggc gggctgcgag 660
acggcttccc ggcgctgcat gcaacaccga tgatgcttcg acccccccgaa gctccttcgg 720
ggctgcatgg gcgctccgat gccgctccag ggcgagcgct gtttaaatag ccaggccccc 780
gattgcaaag acattatagc gagctaccaa agccatattc aaacacctag atcactacca 840
cttctacaca ggccactcga gcttgtgatc gcactccgct aaggggcgc ctcttcctct 900
tcgtttcagt cacaacccgc aaacggcgcg ccatgctgct gcaggccttc ctgttcctgc 960
tggccggctt cgccgccaag atcagcgcct ccatgacgaa cgagacgtcc gaccgccccc 1020
tggtgcactt cacccccaac aagggctgga tgaacgaccc caacggcctg tggtacgacg 1080
agaaggacgc caagtggcac ctgtacttcc agtacaaccc gaacgacacc gtctggggga 1140
cgcccttgtt ctggggccac gccacgtccg acgacctgac caactgggag gaccagccca 1200
tcgccatcgc cccgaagcgc aacgactccg gcgccttctc cggctccatg gtggtggact 1260
acaacaacac ctccggcttc ttcaacgaca ccatcgaccc gcgccagcgc tgcgtggcca 1320
tctggaccta caacaccccg gagtccgagg agcagtacat ctcctacagc ctggacggcg 1380
gctacacctt caccgagtac cagaagaacc ccgtgctggc cgccaactcc acccagttcc 1440
gcgacccgaa ggtcttctgg tacgagccct cccagaagtg gatcatgacc gcggccaagt 1500
cccaggacta caagatcgag atctactcct ccgacgacct gaagtcctgg aagctggagt 1560
ccgcgttcgc caacgagggc ttcctcggct accagtacga gtgccccggc ctgatcgagg 1620
tccccaccga gcaggacccc agcaagtcct actgggtgat gttcatctcc atcaaccccg 1680
gcgccccggc cggcggctcc ttcaaccagt acttcgtcgg cagcttcaac ggcacccact 1740
tcgaggcctt cgacaaccag tcccgcgtgg tggacttcgg caaggactac tacgccctgc 1800
agaccttctt caacaccgac ccgacctacg ggagcgccct gggcatcgcg tgggcctcca 1860
actgggagta ctccgccttc gtgcccacca acccctggcg ctcctccatg tccctcgtgc 1920
gcaagttctc cctcaacacc gagtaccagg ccaacccgga gacggagctg atcaacctga 1980
aggccgagcc gatcctgaac atcagcaacg ccggcccctg gagccggttc gccaccaaca 2040
ccacgttgac gaaggccaac agctacaacg tcgacctgtc caacagcacc ggcaccctgg 2100
agttcgagct ggtgtacgcc gtcaacacca cccagacgat ctccaagtcc gtgttcgcgg 2160
acctctccct ctggttcaag ggcctggagg accccgagga gtacctccgc atgggcttcg 2220
aggtgtccgc gtcctccttc ttcctggacc gcgggaacag caaggtgaag ttcgtgaagg 2280
agaaccccta cttcaccaac cgcatgagcg tgaacaacca gcccttcaag agcgagaacg 2340
acctgtccta ctacaaggtg tacggcttgc tggaccagaa catcctggag ctgtacttca 2400
acgacggcga cgtcgtgtcc accaacacct acttcatgac caccgggaac gccctgggct 2460
ccgtgaacat gacgacgggg gtggacaacc tgttctacat cgacaagttc caggtgcgcg 2520
aggtcaagtg acaattggca gcagcagctc ggatagtatc gacacactct ggacgctggt 2580
cgtgtgatgg actgttgccg ccacacttgc tgccttgacc tgtgaatatc cctgccgctt 2640
ttatcaaaca gcctcagtgt gtttgatctt gtgtgtacgc gcttttgcga gttgctagct 2700
gcttgtgcta tttgcgaata ccacccccag catcccccttc cctcgtttca tatcgcttgc 2760
atcccaaccg caacttatct acgctgtcct gctatccctc agcgctgctc ctgctcctgc 2820
tcactgcccc tcgcacagcc ttggtttggg ctccgcctgt attctcctgg tactgcaacc 2880
tgtaaaccag cactgcaatg ctgatgcacg ggaagtagtg ggatgggaac acaaatggag 2940

gatcgtagag ctcactagta tcgatttcga agacagggtg gttggctgga tggggaaacg 3000
ctggtcgcgg gattcgatcc tgctgcttat atcctccctg gaagcacacc cacgactctg 3060
aagaagaaaa cgtgcacaca cacaacccaa ccggccgaat atttgcttcc ttatcccggg 3120
tccaagagag actgcgatgc ccccctcaat cagcatcctc ctccctgccg cttcaatctt 3180
ccctgcttgc ctgcgcccgc ggtgcgccgt ctgcccgccc agtcagtcac tcctgcacag 3240
gccccttgtg cgcagtgctc ctgtaccctt taccgctcct tccattctgc gaggcccccc 3300
attgaatgta ttcgttgcct gtgtggccaa gcgggctgct gggcgcgccg ccgtcgggca 3360
gtgctcggcg actttggcgg aagccgattg ttcttctgta agccacgcgc ttgctgcttt 3420
gggaagagaa gggggggggt actgaatgga tgaggaggag aaggaggggt attggtatta 3480
tctgagttgg gtgaagagc 3499

<210> 254
<211> 6514
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 254

gctcttcgga gtcactgtgc cactgagttc gactggtagc tgaatggagt cgctgctcca 60
ctaaacgaat tgtcagcacc gccagccggc cgaggacccg agtcatagcg agggtagtag 120
cgcgccatgg caccgaccag cctgcttgcc agtactggcg tctcttccgc ttctctgtgg 180
tcctctgcgc gctccagcgc gtgcgctttt ccggtggatc atgcggtccg tggcgcaccg 240
cagcggccgc tgcccatgca gcgccgctgc ttccgaacag tggcggtcag ggccgcaccc 300
gcggtagccg tccgtccgga acccgcccaa gagttttggg agcagcttga gccctgcaag 360
atggcggagg acaagcgcat cttcctggag gagcaccggt gcgtggaggt ccggggctga 420
ccggccgtcg cattcaacgt aatcaatcgc atgatgatca gaggacacga agtcttggtg 480
gcggtggcca gaaacactgt ccattgcaag ggcataggga tgcgttcctt cacctctcat 540
ttctcatttc tgaatccctc cctgctcact ctttctcctc ctccttcccg ttcacgcagc 600
attcggggta ccctttcttg cgctatgaca cttccagcaa aaggtagggc gggctgcgag 660
acggcttccc ggcgctgcat gcaacaccga tgatgcttcg acccccccgaa gctccttcgg 720
ggctgcatgg gcgctccgat gccgctccag ggcgagcgct gtttaaatag ccaggccccc 780
gattgcaaag acattatagc gagctaccaa agccatattc aaacacctag atcactacca 840
cttctacaca ggccactcga gcttgtgatc gcactccgct aaggggcgc ctcttcctct 900
tcgtttcagt cacaacccgc aaacggcgcg ccatgctgct gcaggccttc ctgttcctgc 960
tggccggctt cgccgccaag atcagcgcct ccatgacgaa cgagacgtcc gaccgccccc 1020
tggtgcactt cacccccaac aagggctgga tgaacgaccc caacggcctg tggtacgacg 1080
agaaggacgc caagtggcac ctgtacttcc agtacaaccc gaacgacacc gtctggggga 1140
cgcccttgtt ctggggccac gccacgtccg acgacctgac caactgggag gaccagccca 1200
tcgccatcgc cccgaagcgc aacgactccg gcgccttctc cggctccatg gtggtggact 1260
acaacaacac ctccggcttc ttcaacgaca ccatcgaccc gcgccagcgc tgcgtggcca 1320

tctggaccta caacaccccg gagtccgagg agcagtacat ctcctacagc ctggacggcg 1380
gctacacctt caccgagtac cagaagaacc ccgtgctggc cgccaactcc acccagttcc 1440
gcgacccgaa ggtcttctgg tacgagccct cccagaagtg gatcatgacc gcggccaagt 1500
cccaggacta caagatcgag atctactcct ccgacgacct gaagtcctgg aagctggagt 1560
ccgcgttcgc caacgagggc ttcctcggct accagtacga gtgccccggc ctgatcgagg 1620
tccccaccga gcaggacccc agcaagtcct actgggtgat gttcatctcc atcaaccccg 1680
gcgccccggc cggcggctcc ttcaaccagt acttcgtcgg cagcttcaac ggcacccact 1740
tcgaggcctt cgacaaccag tcccgcgtgg tggacttcgg caaggactac tacgccctgc 1800
agaccttctt caacaccgac ccgacctacg ggagcgccct gggcatcgcg tgggcctcca 1860
actgggagta ctccgccttc gtgcccacca acccctggcg ctcctccatg tccctcgtgc 1920
gcaagttctc cctcaacacc gagtaccagg ccaacccgga gacggagctg atcaacctga 1980
aggccgagcc gatcctgaac atcagcaacg ccggcccctg gagccggttc gccaccaaca 2040
ccacgttgac gaaggccaac agctacaacg tcgacctgtc caacagcacc ggcaccctgg 2100
agttcgagct ggtgtacgcc gtcaacacca cccagacgat ctccaagtcc gtgttcgcgg 2160
acctctccct ctggttcaag ggcctggagg accccgagga gtacctccgc atgggcttcg 2220
aggtgtccgc gtcctccttc ttcctggacc gcgggaacag caaggtgaag ttcgtgaagg 2280
agaaccccta cttcaccaac cgcatgagcg tgaacaacca gcccttcaag agcgagaacg 2340
acctgtccta ctacaaggtg tacggcttgc tggaccagaa catcctggag ctgtacttca 2400
acgacggcga cgtcgtgtcc accaacacct acttcatgac caccgggaac gccctgggct 2460
ccgtgaacat gacgacgggg gtggacaacc tgttctacat cgacaagttc caggtgcgcg 2520
aggtcaagtg acaattggca gcagcagctc ggatagtatc gacacactct ggacgctggt 2580
cgtgtgatgg actgttgccg ccacacttgc tgccttgacc tgtgaatatc cctgccgctt 2640
ttatcaaaca gcctcagtgt gtttgatctt gtgtgtacgc gcttttgcga gttgctagct 2700
gcttgtgcta tttgcgaata ccacccccag catccccttc cctcgtttca tatcgcttgc 2760
atcccaaccg caacttatct acgctgtcct gctatccctc agcgctgctc ctgctcctgc 2820
tcactgcccc tcgcacagcc ttggtttggg ctccgcctgt attctcctgg tactgcaacc 2880
tgtaaaccag cactgcaatg ctgatgcacg ggaagtagtg ggatgggaac acaaatggag 2940
gatcccgcgt ctcgaacaga gcgcgcagag gaacgctgaa ggtctcgcct ctgtcgcacc 3000
tcagcgcggc atacaccaca ataaccacct gacgaatgcg cttggttctt cgtccattag 3060
cgaagcgtcc ggttcacaca cgtgccacgt tggcgaggtg gcaggtgaca atgatcggtg 3120
gagctgatgg tcgaaacgtt cacagcctag ggatatcgaa ttcggccgac aggacgcgcg 3180
tcaaaggtgc tggtcgtgta tgccctggcc ggcaggtcgt tgctgctgct ggttagtgat 3240
tccgcaaccc tgattttggc gtcttatttt ggcgtggcaa acgctggcgc ccgcgagccg 3300
ggccggcggc gatgcggtgc cccacggctg ccggaatcca agggaggcaa gagcgcccgg 3360
gtcagttgaa gggctttacg cgcaaggtac agccgctcct gcaaggctgc gtggtggaat 3420
tggacgtgca ggtcctgctg aagttcctcc accgcctcac cagcggacaa agcaccggtg 3480
tatcaggtcc gtgtcatcca ctctaaagag ctcgactacg acctactgat ggccctagat 3540
tcttcatcaa aaacgcctga gacacttgcc caggattgaa actccctgaa gggaccacca 3600
ggggccctga gttgttcctt ccccccgtgg cgagctgcca gccaggctgt acctgtgatc 3660

gaggctggcg ggaaaatagg cttcgtgtgc tcaggtcatg ggaggtgcag gacagctcat 3720
gaaacgccaa caatcgcaca attcatgtca agctaatcag ctatttcctc ttcacgagct 3780
gtaattgtcc caaaattctg gtctaccggg ggtgatcctt cgtgtacggg cccttccctc 3840
aaccctaggt atgcgcgcat gcggtcgccg cgcaactcgc gcgagggccg agggtttggg 3900
acgggccgtc ccgaaatgca gttgcacccg gatgcgtggc acctttttg cgataattta 3960
tgcaatggac tgctctgcaa aattctggct ctgtcgccaa ccctaggatc agcggcgtag 4020
gatttcgtaa tcattcgtcc tgatggggag ctaccgacta ccctaatatc agcccgactg 4080
cctgacgcca gcgtccactt ttgtgcacac attccattcg tgcccaagac atttcattgt 4140
ggtgcgaagc gtccccagtt acgctcacct gtttcccgac ctccttactg ttctgtcgac 4200
agagcgggcc cacaggccgg tcgcagccac tagtatggtg gtggccgccg ccgccagcag 4260
cgccttcttc cccgtgcccg ccccccgccc caccccccaag cccggcaagt tcggcaactg 4320
gcccagcagc ctgagccagc ccttcaagcc caagagcaac cccaacggcc gcttccaggt 4380
gaaggccaac gtgagccccc acgggcgcgc ccccaaggcc aacggcagcg ccgtgagcct 4440
gaagtccggc agcctgaaca ccctggagga ccccccccagc agccccccc cccgcacctt 4500
cctgaaccag ctgcccgact ggagccgcct gcgcaccgcc atcaccaccg tgttcgtggc 4560
cgccgagaag cagttcaccc gcctggaccg caagagcaag cgccccgaca tgctggtgga 4620
ctggttcggc agcgagacca tcgtgcagga cggcctggtg ttccgcgagc gcttcagcat 4680
ccgcagctac gagatcggcg ccgaccgcac cgccagcatc gagaccctga tgaaccacct 4740
gcaggacacc agcctgaacc actgcaagag cgtgggcctg ctgaacgacg gcttcggccg 4800
caccccgag atgtgcaccc gcgacctgat ctgggtgctg accaagatgc agatcgtggt 4860
gaaccgctac cccacctggg gcgacaccgt ggagatcaac agctggttca gccagagcgg 4920
caagatcggc atgggccgcg agtggctgat cagcgactgc aacaccggcg agatcctggt 4980
gcgcgccacc agcgcctggg ccatgatgaa ccagaagacc cgccgcttca gcaagctgcc 5040
ctgcgaggtg cgccaggaga tcgcccccca cttcgtggac gccccccccg tgatcgagga 5100
caacgaccgc aagctgcaca agttcgacgt gaagaccggc gacagcatct gcaagggcct 5160
gacccccggc tggaacgact tcgacgtgaa ccagcacgtg agcaacgtga agtacatcgg 5220
ctggattctg gagagcatgc ccaccgaggt gctggagacc caggagctgt gcagcctgac 5280
cctggagtac cgccgcgagt gcggccgcga gagcgtggtg gagagcgtga ccagcatgaa 5340
ccccagcaag gtgggcgacc gcagccagta ccagcacctg ctgcgcctgg aggacggcgc 5400
cgacatcatg aagggccgca ccgagtggcg ccccaagaac gccggcacca accgcgccat 5460
cagcacctga ttaattaact cgaggcagca gcagctcgga tagtatcgac acactctgga 5520
cgctggtcgt gtgatggact gttgccgcca cacttgctgc cttgacctgt gaatatccct 5580
gccgctttta tcaaacagcc tcagtgtgtt tgatcttgtg tgtacgcgct tttgcgagtt 5640
gctagctgct tgtgctattt gcgaatacca cccccagcat ccccttccct cgtttcatat 5700
cgcttgcatc ccaaccgcaa cttatctacg ctgtcctgct atccctcagc gctgctcctg 5760
ctcctgctca ctgcccctcg cacagccttg gtttgggctc cgcctgtatt ctcctggtac 5820
tgcaacctgt aaaccagcac tgcaatgctg atgcacggga agtagtggga tgggaacaca 5880
aatggaaagc ttgagctctt gttttccaga aggagttgct ccttgagcct ttcattctca 5940
gcctcgataa cctccaaagc cgctctaatt gtggaggggg ttcgaagaca gggtggttgg 6000

ctggatgggg aaacgctggt cgcgggattc gatcctgctg cttatatcct ccctggaagc 6060
acacccacga ctctgaagaa gaaaacgtgc acacacacaa cccaaccggc cgaatatttg 6120
cttccttatc ccgggtccaa gagagactgc gatgcccccc tcaatcagca tcctcctccc 6180
tgccgcttca atcttccctg cttgcctgcg cccgcggtgc gccgtctgcc cgcccagtca 6240
gtcactcctg cacaggcccc ttgtgcgcag tgctcctgta ccctttaccg ctccttccat 6300
tctgcgaggc cccctattga atgtattcgt tgcctgtgtg gccaagcggg ctgctgggcg 6360
cgccgccgtc gggcagtgct cggcgacttt ggcggaagcc gattgttctt ctgtaagcca 6420
cgcgcttgct gctttgggaa gagaaggggg ggggtactga atggatgagg aggagaagga 6480
ggggtattgg tattatctga gttgggtgaa gagc 6514

RNAi表达盒的FATA部分

<210> 255
<211> 767
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 255

atggcaccga ccagcctgct tgccagtact ggcgtctctt ccgcttctct gtggtcctct 60
gcgcgctcca gcgcgtgcgc ttttccggtg gatcatgcgg tccgtggcgc accgcagcgg 120
ccgctgccca tgcagcgccg ctgcttccga acagtggcgg tcagggccgc acccgcggta 180
gccgtccgtc cggaacccgc ccaagagttt tgggagcagc ttgagccctg caagatggcg 240
gaggacaagc gcatcttcct ggaggagcac cggtgcgtgg aggtccgggg ctgaccggcc 300
gtcgcattca acgtaatcaa tcgcatgatg atcagaggac acgaagtctt ggtggcggtg 360
gccagaaaca ctgtccattg caagggcata gggatgcgtt ccttcacctc tcatttctca 420
tttctgaatc cctccctgct cactctttct cctcctcctt cccgttcacg cagcattcgg 480
ggcaacgagg tgggcccgtg ctcctccagg aagatgcgct tgtcctccgc catcttgcag 540
ggctcaagct gctcccaaaa ctcttgggcg ggttccggac ggacggctac cgcgggtgcg 600
gccctgaccg ccactgttcg gaagcagcgg cgctgcatgg gcagcggccg ctgcggtgcg 660
ccacggaccg catgatccac cggaaaagcg cacgcgctgg agcgcgcaga ggaccacaga 720
gaagcggaag agacgccagt actggcaagc aggctggtcg gtgccat 767

pSZ1283

<210> 256
<211> 9426
<212> DNA
<213> 人工序列
<220>
<223> 合成的构建体
<400> 256

cactggcagc agccactggt aacaggatta gcagagcgag gtatgtaggc ggtgctacag 60

agttcttgaa gtggtggcct aactacggct acactagaag gacagtattt ggtatctgcg 120
ctctgctgaa gccagttacc ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa 180
ccaccgctgg tagcggtggt ttttttgttt gcaagcagca gattacgcgc agaaaaaaag 240
gatctcaaga agatcctttg atcttttcta cggggtctga cgctcagtgg aacgaaaact 300
cacgttaagg gattttggtc atgagattat caaaaaggat cttcacctag atccttttaa 360
attaaaaatg aagttttaaa tcaatctaaa gtatatatga gtaaacttgg tctgacagtt 420
accaatgctt aatcagtgag gcacctatct cagcgatctg tctatttcgt tcatccatag 480
ttgcctgact ccccgtcgtg tagataacta cgatacggga gggcttacca tctggcccca 540
gtgctgcaat gataccgcga gacccacgct caccggctcc agatttatca gcaataaacc 600
agccagccgg aagggccgag cgcagaagtg gtcctgcaac tttatccgcc tccatccagt 660
ctattaattg ttgccgggaa gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg 720
ttgttgccat tgctacaggc atcgtggtgt cacgctcgtc gtttggtatg gcttcattca 780
gctccggttc ccaacgatca aggcgagtta catgatcccc catgttgtgc aaaaaagcgg 840
ttagctcctt cggtcctccg atcgttgtca gaagtaagtt ggccgcagtg ttatcactca 900
tggttatggc agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg 960
tgactggtga gtactcaacc aagtcattct gagaatagtg tatgcggcga ccgagttgct 1020
cttgcccggc gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgctca 1080
tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca 1140
gttcgatgta acccactcgt gcacccaact gatcttcagc atcttttact ttcaccagcg 1200
tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata agggcgacac 1260
ggaaatgttg aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt 1320
attgtctcat gagcggatac atatttgaat gtatttagaa aaataaacaa ataggggttc 1380
cgcgcacatt tccccgaaaa gtgccacctg acgtctaaga aaccattatt atcatgacat 1440
taacctataa aaataggcgt atcacgaggc cctttcgtct cgcgcgtttc ggtgatgacg 1500
gtgaaaacct ctgacacatg cagctcccgg agacggtcac agcttgtctg taagcggatg 1560
ccgggagcag acaagcccgt cagggcgcgt cagcgggtgt tggcgggtgt cggggctggc 1620
ttaactatgc ggcatcagag cagattgtac tgagagtgca ccatatgcgg tgtgaaatac 1680
cgcacagatg cgtaaggaga aaataccgca tcaggcgcca ttcgccattc aggctgcgca 1740
actgttggga agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg 1800
gatgtgctgc aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta 1860
aaacgacggc cagtgaattg atgcatgctc ttcgccgccg ccactcctgc tcgagcgcgc 1920
ccgcgcgtgc gccgccagcg ccttggcctt ttcgccgcgc tcgtgcgcgt cgctgatgtc 1980
catcaccagg tccatgaggt ctgccttgcg ccggctgagc cactgcttcg tccgggcggc 2040
caagaggagc atgagggagg actcctggtc cagggtcctg acgtggtcgc ggctctggga 2100
gcgggccagc atcatctggc tctgccgcac cgaggccgcc tccaactggt cctccagcag 2160
ccgcagtcgc cgccgaccct ggcagaggaa gacaggtgag gggggtatga attgtacaga 2220
acaaccacga gccttgtcta ggcagaatcc ctaccagtca tggctttacc tggatgacgg 2280
cctgcgaaca gctgtccagc gaccctcgct gccgccgctt ctcccgcacg cttctttcca 2340
gcaccgtgat ggcgcgagcc agcgccgcac gctggcgctg cgcttcgccg atctgaggac 2400

agtcggggaa ctctgatcag tctaaacccc cttgcgcgtt agtgttgcca tcctttgcag 2460
accggtgaga gccgacttgt tgtgcgccac cccccacacc acctcctccc agaccaattc 2520
tgtcaccttt ttggcgaagg catcggcctc ggcctgcaga gaggacagca gtgcccagcc 2580
gctgggggtt ggcggatgca cgctcaggta ccctttcttg cgctatgaca cttccagcaa 2640
aaggtagggc gggctgcgag acggcttccc ggcgctgcat gcaacaccga tgatgcttcg 2700
acccccccgaa gctccttcgg ggctgcatgg gcgctccgat gccgctccag ggcgagcgct 2760
gtttaaatag ccaggccccc gattgcaaag acattatagc gagctaccaa agccatattc 2820
aaacacctag atcactacca cttctacaca ggccactcga gcttgtgatc gcactccgct 2880
aaggggcgcc ctcttcctct tcgtttcagt cacaacccgc aaacggcgcg ccatgctgct 2940
gcaggccttc ctgttcctgc tggccggctt cgccgccaag atcagcgcct ccatgacgaa 3000
cgagacgtcc gaccgccccc tggtgcactt caccccccaac aagggctgga tgaacgaccc 3060
caacggcctg tggtacgacg agaaggacgc caagtggcac ctgtacttcc agtacaaccc 3120
gaacgacacc gtctgggggga cgcccttgtt ctggggccac gccacgtccg acgacctgac 3180
caactgggag gaccagccca tcgccatcgc cccgaagcgc aacgactccg gcgccttctc 3240
cggctccatg gtggtggact acaacaacac ctccggcttc ttcaacgaca ccatcgaccc 3300
gcgccagcgc tgcgtggcca tctggaccta caacaccccg gagtccgagg agcagtacat 3360
ctcctacagc ctggacggcg gctacacctt caccgagtac cagaagaacc ccgtgctggc 3420
cgccaactcc acccagttcc gcgacccgaa ggtcttctgg tacgagccct cccagaagtg 3480
gatcatgacc gcggccaagt cccaggacta caagatcgag atctactcct ccgacgacct 3540
gaagtcctgg aagctggagt ccgcgttcgc caacgagggc ttcctcggct accagtacga 3600
gtgccccggc ctgatcgagg tccccaccga gcaggacccc agcaagtcct actgggtgat 3660
gttcatctcc atcaaccccg gcgccccggc cggcggctcc ttcaaccagt acttcgtcgg 3720
cagcttcaac ggcacccact tcgaggcctt cgacaaccag tcccgcgtgg tggacttcgg 3780
caaggactac tacgccctgc agaccttctt caacaccgac ccgacctacg ggagcgccct 3840
gggcatcgcg tgggcctcca actgggagta ctccgccttc gtgcccacca acccctggcg 3900
ctcctccatg tccctcgtgc gcaagttctc cctcaacacc gagtaccagg ccaacccgga 3960
gacggagctg atcaacctga aggccgagcc gatcctgaac atcagcaacg ccggcccctg 4020
gagccggttc gccaccaaca ccacgttgac gaaggccaac agctacaacg tcgacctgtc 4080
caacagcacc ggcaccctgg agttcgagct ggtgtacgcc gtcaacacca cccagacgat 4140
ctccaagtcc gtgttcgcgg acctctccct ctggttcaag ggcctggagg accccgagga 4200
gtacctccgc atgggcttcg aggtgtccgc gtcctccttc ttcctggacc gcgggaacag 4260
caaggtgaag ttcgtgaagg agaaccccta cttcaccaac cgcatgagcg tgaacaacca 4320
gcccttcaag agcgagaacg acctgtccta ctacaaggtg tacggcttgc tggaccagaa 4380
catcctggag ctgtacttca acgacggcga cgtcgtgtcc accaacacct acttcatgac 4440
caccgggaac gccctgggct ccgtgaacat gacgacgggg gtggacaacc tgttctacat 4500
cgacaagttc caggtgcgcg aggtcaagtg acaattggca gcagcagctc ggatagtatc 4560
gacacactct ggacgctggt cgtgtgatgg actgttgccg ccacacttgc tgccttgacc 4620
tgtgaatatc cctgccgctt ttatcaaaca gcctcagtgt gtttgatctt gtgtgtacgc 4680
gcttttgcga gttgctagct gcttgtgcta tttgcgaata ccacccccag catcccccttc 4740

cctcgtttca tatcgcttgc atcccaaccg caacttatct acgctgtcct gctatccctc 4800
agcgctgctc ctgctcctgc tcactgcccc tcgcacagcc ttggtttggg ctccgcctgt 4860
attctcctgg tactgcaacc tgtaaaccag cactgcaatg ctgatgcacg ggaagtagtg 4920
ggatgggaac acaaatggag gatcccgcgt ctcgaacaga gcgcgcagag gaacgctgaa 4980
ggtctcgcct ctgtcgcacc tcagcgcggc atacaccaca ataaccacct gacgaatgcg 5040
cttggttctt cgtccattag cgaagcgtcc ggttcacaca cgtgccacgt tggcgaggtg 5100
gcaggtgaca atgatcggtg gagctgatgg tcgaaacgtt cacagcctag ggatatcgaa 5160
ttcggccgac aggacgcgcg tcaaaggtgc tggtcgtgta tgccctggcc ggcaggtcgt 5220
tgctgctgct ggttagtgat tccgcaaccc tgattttggc gtcttatttt ggcgtggcaa 5280
acgctggcgc ccgcgagccg ggccggcggc gatgcggtgc cccacggctg ccggaatcca 5340
agggaggcaa gagcgcccgg gtcagttgaa gggctttacg cgcaaggtac agccgctcct 5400
gcaaggctgc gtggtggaat tggacgtgca ggtcctgctg aagttcctcc accgcctcac 5460
cagcggacaa agcaccggtg tatcaggtcc gtgtcatcca ctctaaagag ctcgactacg 5520
acctactgat ggccctagat tcttcatcaa aaacgcctga gacacttgcc caggattgaa 5580
actccctgaa gggaccacca ggggccctga gttgttcctt ccccccgtgg cgagctgcca 5640
gccaggctgt acctgtgatc gaggctggcg ggaaaatagg cttcgtgtgc tcaggtcatg 5700
ggaggtgcag gacagctcat gaaacgccaa caatcgcaca attcatgtca agctaatcag 5760
ctatttcctc ttcacgagct gtaattgtcc caaaattctg gtctaccggg ggtgatcctt 5820
cgtgtacggg cccttccctc aaccctaggt atgcgcgcat gcggtcgccg cgcaactcgc 5880
gcgagggccg agggtttggg acgggccgtc ccgaaatgca gttgcacccg gatgcgtggc 5940
acctttttg cgataattta tgcaatggac tgctctgcaa aattctggct ctgtcgccaa 6000
ccctaggatc agcggcgtag gatttcgtaa tcattcgtcc tgatggggag ctaccgacta 6060
ccctaatatc agcccgactg cctgacgcca gcgtccactt ttgtgcacac attccattcg 6120
tgcccaagac atttcattgt ggtgcgaagc gtccccagtt acgctcacct gtttcccgac 6180
ctccttactg ttctgtcgac agagcgggcc cacaggccgg tcgcagccac tagtatggtg 6240
gtggccgccg ccgccagcag cgccttcttc cccgtgcccg ccccccgccc cacccccaag 6300
cccggcaagt tcggcaactg gcccagcagc ctgagccagc ccttcaagcc caagagcaac 6360
cccaacggcc gcttccaggt gaaggccaac gtgagccccc acgggcgcgc ccccaaggcc 6420
aacggcagcg ccgtgagcct gaagtccggc agcctgaaca ccctggagga ccccccccagc 6480
agccccccccc cccgcacctt cctgaaccag ctgcccgact ggagccgcct gcgcaccgcc 6540
atcaccaccg tgttcgtggc cgccgagaag cagttcaccc gcctggaccg caagagcaag 6600
cgccccgaca tgctggtgga ctggttcggc agcgagacca tcgtgcagga cggcctggtg 6660
ttccgcgagc gcttcagcat ccgcagctac gagatcggcg ccgaccgcac cgccagcatc 6720
gagaccctga tgaaccacct gcaggacacc agcctgaacc actgcaagag cgtgggcctg 6780
ctgaacgacg gcttcggccg caccccccgag atgtgcaccc gcgacctgat ctgggtgctg 6840
accaagatgc agatcgtggt gaaccgctac cccacctggg gcgacaccgt ggagatcaac 6900
agctggttca gccagagcgg caagatcggc atgggccgcg agtggctgat cagcgactgc 6960
aacaccggcg agatcctggt gcgcgccacc agcgcctggg ccatgatgaa ccagaagacc 7020
cgccgcttca gcaagctgcc ctgcgaggtg cgccaggaga tcgcccccca cttcgtggac 7080

gccccccccg tgatcgagga caacgaccgc aagctgcaca agttcgacgt gaagaccggc 7140
gacagcatct gcaagggcct gacccccggc tggaacgact tcgacgtgaa ccagcacgtg 7200
agcaacgtga agtacatcgg ctggattctg gagagcatgc ccaccgaggt gctggagacc 7260
caggagctgt gcagcctgac cctggagtac cgccgcgagt gcggccgcga gagcgtggtg 7320
gagagcgtga ccagcatgaa ccccagcaag gtgggcgacc gcagccagta ccagcacctg 7380
ctgcgcctgg aggacggcgc cgacatcatg aagggccgca ccgagtggcg ccccaagaac 7440
gccggcacca accgcgccat cagcacctga ttaattaact cgaggcagca gcagctcgga 7500
tagtatcgac acactctgga cgctggtcgt gtgatggact gttgccgcca cacttgctgc 7560
cttgacctgt gaatatccct gccgctttta tcaaacagcc tcagtgtgtt tgatcttgtg 7620
tgtacgcgct tttgcgagtt gctagctgct tgtgctattt gcgaatacca cccccagcat 7680
ccccttccct cgtttcatat cgcttgcatc ccaaccgcaa cttatctacg ctgtcctgct 7740
atccctcagc gctgctcctg ctcctgctca ctgcccctcg cacagccttg gtttgggctc 7800
cgcctgtatt ctcctggtac tgcaacctgt aaaccagcac tgcaatgctg atgcacggga 7860
agtagtggga tgggaacaca aatggaaagc ttgagctctt gttttccaga aggagttgct 7920
ccttgagcct ttcattctca gcctcgataa cctccaaagc cgctctaatt gtggaggggg 7980
ttcgaattta aaagcttgga atgttggttc gtgcgtctgg aacaagccca gacttgttgc 8040
tcactgggaa aaggaccatc agctccaaaa aacttgccgc tcaaaccgcg tacctctgct 8100
ttcgcgcaat ctgccctgtt gaaatcgcca ccacattcat attgtgacgc ttgagcagtc 8160
tgtaattgcc tcagaatgtg gaatcatctg ccccctgtgc gagcccatgc caggcatgtc 8220
gcgggcgagg acacccgcca ctcgtacagc agaccattat gctacctcac aatagttcat 8280
aacagtgacc atatttctcg aagctcccca acgagcacct ccatgctctg agtggccacc 8340
ccccggccct ggtgcttgcg gagggcaggt caaccggcat ggggctaccg aaatccccga 8400
ccggatccca ccacccccgc gatgggaaga atctctcccc gggatgtggg cccaccacca 8460
gcacaacctg ctggcccagg cgagcgtcaa accataccac acaaatatcc ttggcatcgg 8520
ccctgaattc cttctgccgc tctgctaccc ggtgcttctg tccgaagcag gggttgctag 8580
ggatcgctcc gagtccgcaa acccttgtcg cgtggcgggg cttgttcgag cttgaagagc 8640
ctctagagtc gacctgcagg catgcaagct tggcgtaatc atggtcatag ctgtttcctg 8700
tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgta 8760
aagcctgggg tgcctaatga gtgagctaac tcacattaat tgcgttgcgc tcactgcccg 8820
ctttccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga 8880
gaggcggttt gcgtattggg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg 8940
tcgttcggct gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag 9000
aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc 9060
gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca 9120
aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt 9180
ttcccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 9240
tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc 9300
tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc 9360
ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact 9420

tatcgc 9426

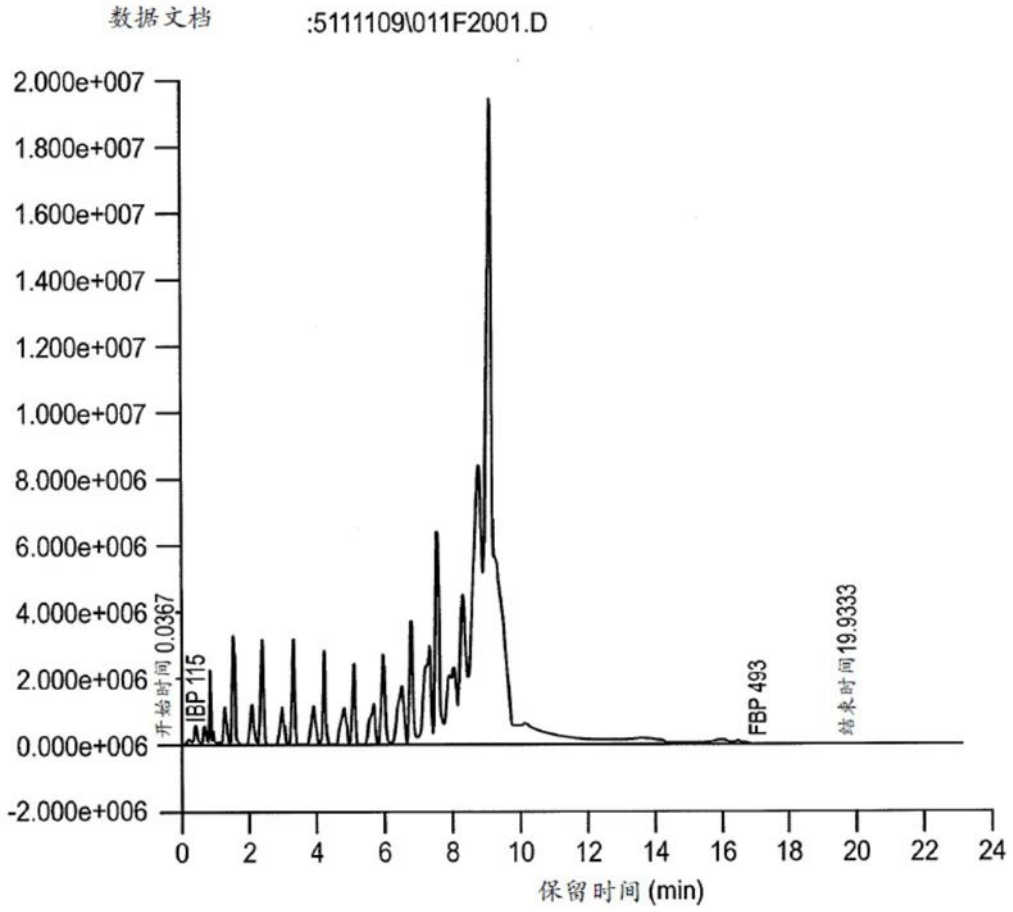

数据文档
:5111109\011F2001.D
2.000e+007
1.800e+007
1.600e+007
1.400e+007
1.200e+007
1.000e+007
8.000e+006
6.000e+006
4.000e+006
2.000e+006
0.000e+006
-2.000e+006
开始时间 0.0367
IBP 115
FBP 493
结束时间19.9333
0
2
4
6
8
10
12
14
16
18
20
22
24
保留时间 (min)

图1

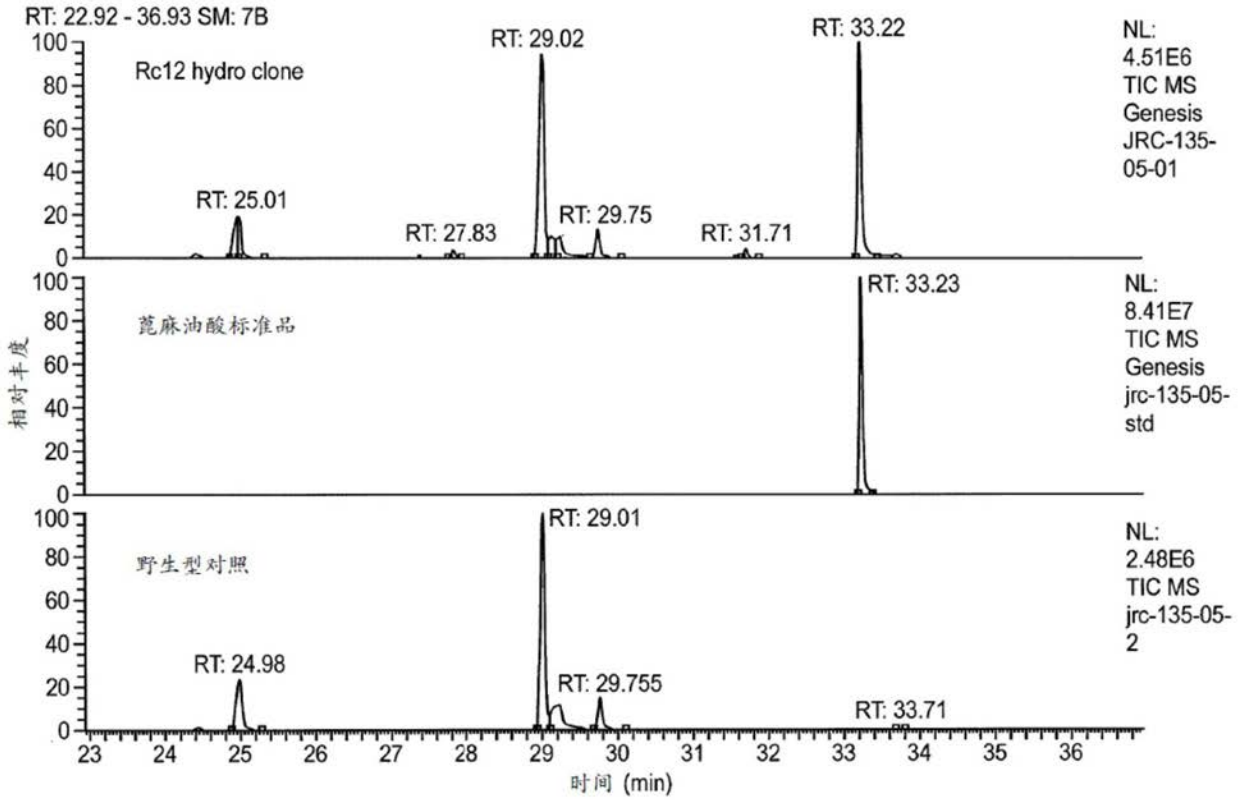

RT: 22.92 - 36.93 SM: 7B
Rc12 hydro clone
RT: 29.02
RT: 33.22
NL:
4.51E6
TIC MS
Genesis
JRC-135-
05-01
RT: 25.01
RT: 27.83
RT: 29.75
RT: 31.71
蓖麻油酸标准品
RT: 33.23
NL:
8.41E7
TIC MS
Genesis
jrc-135-05-
std
相对丰度
野生型对照
RT: 29.01
NL:
2.48E6
TIC MS
jrc-135-05-
2
RT: 24.98
RT: 29.755
RT: 33.71
时间 (min)

图2

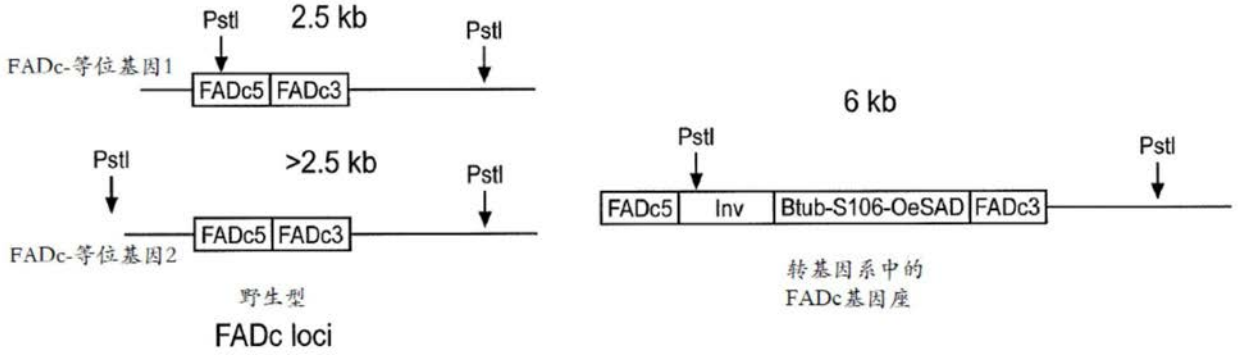
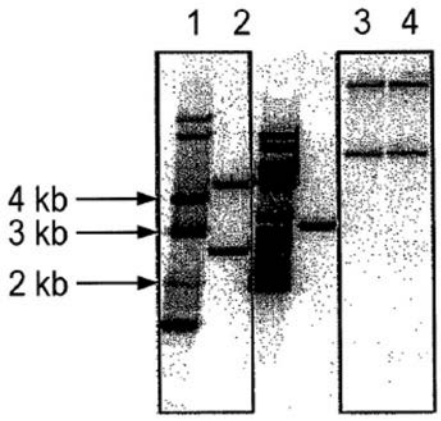

PstI
2.5 kb
PstI
FADc-等位基因1
FADc5
FADc3
PstI
>2.5 kb
PstI
FADc-等位基因2
FADc5
FADc3
野生型
FADc loci
6 kb
PstI
PstI
FADc5
Inv
Btub-S106-OeSAD
FADc3
转基因系中的
FADc基因座

图3



图例

1. 标记

2. 野生型（未转化的），显示出两个FADc等位基因

3. 经转化

4. 转化的

图4