



[12] 发明专利说明书

专利号 ZL 200480019580.3

[45] 授权公告日 2010年2月24日

[11] 授权公告号 CN 100591773C

[22] 申请日 2004.5.7

[21] 申请号 200480019580.3

[30] 优先权

[32] 2003.5.7 [33] US [31] 60/468,677

[86] 国际申请 PCT/US2004/014541 2004.5.7

[87] 国际公布 WO2004/101757 英 2004.11.25

[85] 进入国家阶段日期 2006.1.9

[73] 专利权人 纳慕尔杜邦公司

地址 美国特拉华州

[72] 发明人 S·K·皮卡塔吉奥

N·S·亚达夫 Q·Q·朱

[56] 参考文献

US6432684B 2002.8.13

US5968809A 1999.10.19

审查员 高雁

[74] 专利代理机构 中国专利代理(香港)有限公司
代理人 李波 王景朝

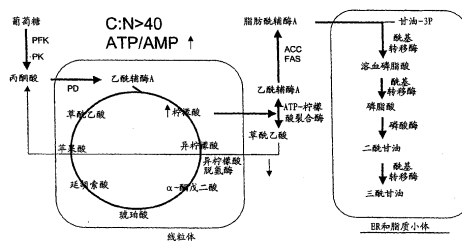
权利要求书1页 说明书144页 附图17页

[54] 发明名称

在含油酵母中生产多不饱和脂肪酸

[57] 摘要

本发明涉及在含油酵母中生产 ω -3 和/或 ω -6 脂肪酸的方法。因此, 将能够催化亚油酸(LA) 转化成 γ -亚麻酸(GLA); α -亚油酸(ALA) 转化成十八碳四烯酸(STA); GLA 转化成二高- γ -亚油酸(DGLA); STA 转化成二十碳四烯酸(ETA); DGLA 转化成花生四烯酸(ARA); ETA 转化成二十碳五烯酸(EPA); DGLA 转化成 ETA; EPA 转化成二十二碳五烯酸(DPA); 以及将 ARA 转化成 EPA 的去饱和酶和延伸酶导入了耶氏酵母属中, 用于合成 ARA 和 EPA。



1. 生产 ω -3 或 ω -6 脂肪酸的方法，包括：
 - a)提供包括编码功能性 ω -3/ ω -6 脂肪酸生物合成途径的基因的含油酵母，所述含油酵母产生占干细胞重量至少 25%的油；
 - b)在可发酵碳源的存在下生长步骤(a)的酵母，从而生产一种或多种 ω -3 或 ω -6 脂肪酸；并且
 - c)任选回收 ω -3 或 ω -6 脂肪酸。
2. 权利要求 1 的方法，其中所述编码功能性 ω -3/ ω -6 脂肪酸生物合成途径的基因选自： Δ 12 去饱和酶， Δ 6 去饱和酶，延伸酶， Δ 5 去饱和酶， Δ 17 去饱和酶， Δ 15 去饱和酶， Δ 9 去饱和酶和 Δ 4 去饱和酶。
3. 权利要求 1 的方法，其中所述可发酵碳源选自单糖，寡糖，多糖，甘油单酯，甘油二酯，甘油三酯，甲醇以及含碳的胺。
4. 权利要求 1 的方法，其中所述含油酵母选自下组：耶氏酵母属，假丝酵母属，红酵母属，红冬孢属，隐球酵母属，丝孢酵母属和油脂酵母属。
5. 权利要求 4 的方法，其中所述含油酵母是解脂耶氏酵母。

在含油酵母中生产多不饱和脂肪酸

本申请要求申请日为2003年5月7日的美国临时申请号60/468677的权益。

发明领域

本发明属于生物技术领域。更具体地讲，本发明涉及编码可用于在含油酵母中生产长链多不饱和脂肪酸（PUFAs）的酶的核酸片段的合成。

发明背景

很久以来就已经认识到，某些多不饱和脂肪酸，或 PUFAs，是健康细胞的重要的生物学成分。例如，这样的 PUFAs 被认为是：

- "必需"脂肪酸，它不能够在哺乳动物体内从头合成，相反，必须从饮食中获得或者通过亚油酸（LA）或 α -亚麻酸（ALA）的进一步去饱和及延伸产生；

- 细胞质膜的成分，其中，它们能够以诸如磷脂或甘油三酯的形式存在；

- 为正常发育所必需，特别是为婴儿脑发育和组织形成和修复所必需；和，

- 在哺乳动物体内起着重要作用的若干种生物学活性类二十烷酸的前体，包括前列环素，类二十烷酸，白三烯和前列腺素。

在二十世纪70年代，发现格陵兰岛爱斯基摩人的心脏病的低发病率和大量摄入长链 ω -3 PUFAs 相关(Dyerberg, J.等, Amer. J. Clin Nutr. 28: 958-966 (1975) ; Dyerberg, J.等, Lancet 2 (8081) : 117-119 (July 15, 1978))。更近一些的研究业已证实了 ω -3 PUFAs 的心血管保护作用(Shimokawa, H., World Rev Nutr Diet, 88: 100-108(2001) ; von Schacky, C., 和 Dyerberg, J., World Rev Nutr Diet, 88: 90-99 (2001))。另外，业已发现，某些病症对用 ω -3 脂肪酸治疗有反应，如在血管成形术之后的再狭窄率，炎症症状和类风湿性关节炎，哮喘，牛皮癣和湿疹。业已证实 γ -亚麻酸（GLA，一种 ω -6 PUFA）能降低与

应激相关的血压升高,并且改善算术测验的表现。业已证实 GLA 和二高- γ -亚麻酸 (DGLA, 另一种 ω -6 PUFA) 能抑制血小板聚集,导致血管舒张,降低胆固醇含量,并且抑制血管壁平滑肌和纤维组织的增殖 (Brenner 等, *Adv. Exp. Med. Biol.* 83: 85-101 (1976))。施用单独的或者与二十碳五烯酸 (EPA, ω -3 PUFA) 组合的 GLA 或 DGLA, 业已表现出能减轻或防止胃肠出血,和由非甾体类抗炎药物导致的其他副作用 (U. S. 4, 666, 701)。另外,业已证实 GLA 和 DGLA 能预防或治疗子宫内膜异位和经前综合征 (U. S. 4, 758, 592), 并且治疗肌痛性脑脊髓炎和在病毒感染之后的慢性疲劳 (U. S. 5, 116, 871)。其他证据表明, PUFAs 可能与钙代谢的调控有关,表明了它们可用于治疗或预防骨质疏松症和肾或尿道结石。最后, PUFAs 可用于治疗癌症和糖尿病 (U. S. 4, 826, 877; Horrobin 等, *Am. J. Clin. Nutr.* 57 (Suppl.): 732S-737S (1993))。

PUFAs 通常被划分成两种主要类型 (由 ω -6 和 ω -3 脂肪酸组成), 它们是分别通过必需脂肪酸, LA 和 ALA 的去饱和和延伸产生的。尽管具有从“必需”脂肪酸的这种常见的衍生,越来越明确饮食中 ω -6 与 ω -3 的比例对维持良好的健康状态很重要。由于人类饮食习惯的改变,目前 ω -6 与 ω -3 的比例是大约 10:1, 而优选的比例是 2:1 (Kris-Etherton, P.M. et al., *Am. J. Clin. Nutr.* 71(1 Suppl.):179S - 88S (2000); Simopoulos, A. P. et al., *Ann. Nutr. Metab.* 43:127-130 (1999); Krauss, R. M. et al. *AHA Circulation* 102:2284-2299 (2000))。

ω -6 脂肪酸的主要来源是含有大量 LA 的植物油 (如玉米油、大豆油)。GLA 存在于许多植物的种子, 包括月见草 (*Oenothera biennis*), 琉璃苣 (*Borago officinalis*) 和黑醋栗 (*Ribes nigrum*)。被孢霉属 (丝状真菌) 的微生物、*Entomophthora*、腐酶属和 *Porphyridium* (红藻) 可以用于 ω -6 脂肪酸花生四烯酸 (ARA) 的商业生产。例如, 用一种真菌高山被孢霉生产含 ARA 的油, 而 U.S. 5,658,767 (Martek Corporation) 教导了生产含 ARA 的油的方法, 包括在含有碳和氮源的培养基中培养 *Pythium insidiosum*。

重要的 ω -3 PUFAs 包括 EPA 和二十二碳六烯酸 (DHA), 两者都存在于不同类型的鱼油和海洋浮游生物中。U.S. 5,244,921 (Martek Corporation) 描述了用于生产含 EPA 的可食用油的方法, 包括在发酵

罐中培养异养硅藻，具体是 *Cyclotella* sp. 和 *Nitzschia* sp. DHA 可以获自冷水海洋鱼类、蛋黄部分，以及通过培养 Dinophyceae 纲的某些异养微藻类，具体是 *Cryptocodinium* sp.，如 *C. cohnii* (U.S. 5,492,938 和 U.S. 5,407,957)。十八碳四烯酸 (STA)，即 EPA 和 DHA 的一种前体，可以存在于海洋油类和植物种子中；其商业来源包括在 *Trichodesma* 和 *Echium* 属中的生产。 ω -3 酸的其它来源存在于亚麻籽油和胡桃油，它们均含有占绝对优势的 ALA。

尽管 PUFAs 的多种商业来源来自于天然来源，但是存在与所述生产方法相关的若干缺陷。首先，诸如鱼和植物的天然来源倾向于具有高度的不一致的油类组成。因此，从这些来源获得的油，可能需要高度纯化，以便分离或者富集一种或多种需要的 PUFAs。鱼油通常具有不佳的味道和气味，这可能不能与需要的产物经济地分开，并且可导致所述产物不能被接受为食物补充物。不佳的味道和气味可以导致基于摄入高剂量的医学方案不被接受，并且可能降低患者的依从性。此外，鱼类可能积累环境污染物，并且摄入鱼油胶囊作为饮食补充物可能导致摄入不良的污染物。天然来源同样会出现可利用性的无法控制的波动（例如，由于天气，疾病，或者对于鱼类资源来说的过度捕捞）；和产生 PUFAs 的作物通常在经济上无法与为了食物生产而培育的杂交作物竞争。能天然产生 PUFAs 的某些生物（例如，*Porphyridium*，被孢霉）的大规模发酵同样可能是昂贵的，和/或难以以商业化规模培养。

上述限制的结果是，为了达到以下目的必须进行大量的研究：1.) 开发便于商业化生产的 PUFAs 的重组资源；和 2.) 修饰脂肪酸生物合成途径，以便能够生产需要的 PUFAs。

在过去若干年中在从各种生物中分离，克隆并且操作脂肪酸去饱和酶和延伸酶基因方面取得了进展。对这些基因序列的了解，提供了在不能天然产生 PUFAs 的新的宿主生物中生产需要的脂肪酸和/或脂肪酸组合物的前景。以下文献披露了在酿酒酵母中的多种例子，如：

1. Domergue, F., et al. (*Eur. J. Biochem.* 269:4105-4113 (2002)), 其中海洋硅藻三角褐指藻来源的两种去饱和酶被克隆进酿酒酵母内，导致生成了 EPA;
2. Beaudoin F., et al. (*Proc. Natl. Acad. Sci. U.S.A.* 97 (12):6421-6426 (2000)), 其中利用了秀丽新杆线虫来源的基因在酿酒酵母内重构了

ω -3 和 ω -6 PUFA 生物合成途径;

3. Dyer, J.M. et al. (Appl. Environ. Microbiol., 59:224-230 (2002)), 其中在酿酒酵母内表达了植物脂肪酸去饱和酶 (FAD2 和 FAD3), 导致生成了 ALA; 以及

4. U.S. 6,136,574 (Knutzon et al., Abbott Laboratories), 其中将甘蓝型油菜来源的一种去饱和酶和真菌高山被孢霉来源的两种去饱和酶克隆进酿酒酵母内, 导致生成了 LA、GLA、ALA 和 STA.

不过, 仍然需要合适的微生物系统, 其中, 上述类型的基因可以表达, 以便提供商业化数量的一种或多种 PUFAs 的经济生产。另外, 需要富集了特殊 PUFAs, 特别是 EPA 和 DHA 的油。

许多微生物 (包括藻类、细菌、霉菌和真菌) 能够在细胞代谢的普通过程中合成油类。因此, 油生产包括在合适的培养基中培养微生物以合成油, 然后从发酵培养基分离微生物, 并且进行处理以回收细胞内的油。已经进行了多种尝试以便通过发酵手段优化脂肪酸的生产, 所述手段包括改变如使用的微生物、允许油生产的培养基和条件等参数。但是, 已经证明这些努力在改进油产率或控制产生的油组合物的特征方面非常不成功。

以前尚未作为 PUFAs 的生产平台检验的一种类型的微生物是含油酵母。这种微生物能够积累最多占它们的干细胞重量 80% 的油类。业已完善了用于生长具有高油含量的含油酵母的技术 (例如, 参见 EP 0 005 277 B1; Ratledge, C., Prog. Ind. Microbiol. 16: 119-206 (1982)), 并且可以提供与生产 ω -3 或 ω -6 PUFAs 的商业化微藻发酵相比的成本优势。完整的酵母细胞可能还体现了对 ω -3 或 ω -6 PUFA-富集的油进行包囊以便用于功能性食品和动物饲料添加剂的方便途径。

尽管具有上述优点, 含油酵母天然上是 ω -6 和 ω -3 PUFAs 缺陷型的, 因为在这种微生物中天然产生的 PUFAs 局限于 18: 2 脂肪酸 (以及更不常见的是 18: 3 脂肪酸)。因此, 要解决的问题是, 开发能积累富含 ω -3 和/或 ω -6 脂肪酸的油类的含油酵母。为此, 必须导入去饱和酶和延伸酶, 这两种酶能够在含油酵母中合成和积累 ω -3 和/或 ω -6 脂肪酸。尽管在基因工程领域已经有了进展, 但是这类技术没有被开发用于含油酵母。因此, 必须克服与将这些特定宿主生物用于生产 PUFAs 相关的问题。

申请人通过导入异源 ω -3 和/或 ω -6 生物合成途径,证明了宿主解脂耶氏酵母 (*Yarrowia Lipolytica*) 中 PUFAs 的生产,从而解决了上述问题。具体地,在此生产了 ARA (代表 ω -6 脂肪酸) 和 EPA (代表 ω -3 脂肪酸), 以例证本发明的技术。

发明概述

本发明提供了在含油酵母宿主中表达包括 ω -3/ ω -6 脂肪酸生物合成途径的酶,以便生产 ω -3 和/或 ω -6 脂肪酸的方法。因此,本发明提供了生产 ω -3 和/或 ω -6 脂肪酸的方法,包括:

- a) 提供包括功能性 ω -3/ ω -6 脂肪酸生物合成途径的含油酵母;
- b) 在可发酵碳源的存在下生长步骤(a)的酵母,从而生产 ω -3 或 ω -6 脂肪酸; 并且
- c) 任选回收 ω -3 或 ω -6 脂肪酸。

在一种具体实施方案中,本发明提供生产亚油酸的方法,包括:

- a) 提供包括以下成分的含油酵母:
 - (i) 编码 Δ 12 去饱和酶多肽的基因; 和
 - (ii) 油酸的内源来源;
- b) 在合适的可发酵碳源的存在下生长步骤(a)的酵母,其中表达编码 Δ 12 去饱和酶多肽的基因并且将油酸转化为亚油酸; 并且
- c) 任选回收步骤(b)的亚油酸。

在具体实施方案中,本发明提供了通过从头生物合成或从合适的前体进行单步骤酶促反应而生产特殊 ω -6 脂肪酸,如亚油酸(LA), γ -亚麻酸(GLA), 二高- γ -亚油酸(DGLA), 和花生四烯酸(ARA)的方法。类似地,本发明提供了通过单步骤酶促反应从合适的前体生产特殊 ω -3 脂肪酸,如 α -亚油酸(ALA), 十八碳四烯酸(STA), 二十碳四烯酸(ETA), 二十碳五烯酸(EPA), 二十二碳五烯酸(DPA) 和二十二碳六烯酸的方法。

附图和序列描述的简要说明

图 1 表示含油酵母内脂质积累生化机制的示意图。

图 2 表示 ω -3/ ω -6 脂肪酸生物合成途径。

图 3 表示用于在解脂耶氏酵母中进行基因表达的质粒载体 pY5 的

构建。

图 4 表示用于在解脂耶氏酵母中进行基因表达的质粒载体 pY5-4 和 pY5-13 的构建。

图 5 表示中间载体 pYZM5CHPPA 的构建示意图。

图 6 表示异丝水霉 $\Delta 17$ 去饱和酶基因和密码子优化以便在解脂耶氏酵母中表达的合成基因的 DNA 序列的比较。

图 7 表示在解脂耶氏酵母中位于翻译起始密码子 'ATG' 周围的优选的共有序列。

图 8 表示体外合成密码子优化的 $\Delta 17$ 去饱和酶基因的方案。

图 9 表示用于在解脂耶氏酵母内表达合成的密码子优化的 $\Delta 17$ 去饱和酶基因和野生型 $\Delta 17$ 去饱和酶基因的质粒。

图 10A 和 10B 所示为对在分别经过野生型 $\Delta 17$ 去饱和酶基因和合成的密码子优化的 $\Delta 17$ 去饱和酶基因转化的解脂耶氏酵母内生成的脂肪酸进行气相色谱分析的结果。

图 11 所示为中间载体 pY24-4 的构建示意图。

图 12 所示为中间载体 pYZV16 的构建示意图。

图 13 所示为整合载体 pYZM5EL6 的构建示意图。

图 14 所示为整合载体 pYZV5EL6 和 pYZV5EL6/17 的构建示意图。

图 15 是说明由工程改造的解脂耶氏酵母生产 ARA 的色谱图。

图 16 是说明由工程改造的解脂耶氏酵母生产 EPA 的色谱图。

通过以下的详细说明和所附的序列说明能够更全面地理解本发明，这些内容构成了本发明的一部分。

下列序列遵循 37 C.F.R. §1.821-1.825 (“对含有核苷酸序列和/或氨基酸序列公开内容的专利申请的要求—序列规则”)，并符合世界知识产权组织 (WIPO) 标准 ST.25 (1998) 以及 EPO 和 PCT 的序列列表要求 (管理细则的规则 5.2 和 49.5 (a-bis)，以及第 208 节和附录 C)。应用于核苷酸和氨基酸序列信息的符号和格式遵循 37 C.F.R. §1.822 所述规则。

SEQ ID NO: 1 表示高山被孢霉 $\Delta 6$ 去饱和酶基因的 DNA 序列，而 SEQ ID NO: 2 表示高山被孢霉 $\Delta 6$ 去饱和酶的氨基酸序列。

SEQ ID NO: 3 表示高山被孢霉 $\Delta 5$ 去饱和酶基因的 DNA 序列，而

SEQ ID NO: 4 表示高山被孢霉 $\Delta 5$ 去饱和酶的相应的氨基酸序列。

SEQ ID NO: 5 表示异丝水霉 $\Delta 17$ 去饱和酶基因的 DNA 序列，而 **SEQ ID NO: 6** 表示异丝水霉 $\Delta 17$ 去饱和酶的相应的氨基酸序列。

SEQ ID NO: 7 表示高山被孢霉高亲和力延伸酶基因的 DNA 序列，而 **SEQ ID NO: 8** 表示高山被孢霉高亲和力延伸酶的氨基酸序列。

SEQ ID NO: 9 表示密码子优化以便在解脂耶氏酵母中表达的合成的 $\Delta 17$ 去饱和酶基因的 DNA 序列。

SEQ ID NOs: 10-31 相当于 11 对寡核苷酸，它们共同组成了异丝水霉 $\Delta 17$ 去饱和酶基因的完整的密码子优化的编码区（例如，分别为 **D17-1A, D17-1B, D17-2A, D17-2B, D17-3A, D17-3B, D17-4A, D17-4B, D17-5A, D17-5B, D17-6A, D17-6B, D17-7A, D17-7B, D17-8A, D17-8B, D17-9A, D17-9B, D17-10A, D17-10B, D17-11A** 和 **D17-11B**）。

SEQ ID NOs: 32 - 37 分别相当于引物 **D17-1, D17-4R, D17-5, D17-8D, D17-8U** 和 **D17-11**，用于在合成密码子优化的 $\Delta 17$ 去饱和酶基因期间进行 PCR 扩增。

SEQ ID NOs: 38 和 **39** 分别相当于用于分离 TEF 启动子的引物 **TEF5'** 和 **TEF3'**。

SEQ ID NOs: 40 和 **41** 分别相当于用于分离 XPR2 转录终止子的引物 **XPR5'** 和 **XPR3'**。

SEQ ID NOs: 42 和 **43** 相当于引物 **YL21A** 和 **YL22**，用于从质粒 pRSP19 扩增异丝水霉的野生型 $\Delta 17$ 去饱和酶基因。

SEQ ID NOs: 44 和 **45** 分别相当于引物 **YL53** 和 **YL54**，用于定点诱变，以便产生 pYSD17M。

SEQ ID NO:46 和 **47** 分别对应于引物 **KU5** 和 **KU3**，被用于扩增含有耶氏酵母 URA3 基因的 1.7kB DNA 片段（**SEQ ID NO:48**；氨基酸序列如 **SEQ ID NO:49** 所示）。

SEQ ID NO:50 和 **51** 分别对应于引物 **KI5** 和 **KI3**，被用于扩增含有 *Impatiens balsama* 的接合酶基因的 1.1kB DNA 片段（**SEQ ID NO:52**；氨基酸序列如 **SEQ ID NO:53** 所示）。

SEQ ID NO:54 和 **55** 分别对应于引物 **KTI5** 和 **KTI3**，被用于扩增含有 TEF::接合酶::XPR 嵌合基因的 1.7kB DNA 片段（**SEQ ID NO:56**；氨基酸序列如 **SEQ ID NO:57** 所示）。

SEQ ID NO:58 和 59 分别对应于引物 KH5 和 KH3, 被用于扩增含有大肠杆菌潮霉素抗性基因的 1kB DNA 片段 (SEQ ID NO:60; 氨基酸序列如 SEQ ID NO:61 所示)。

SEQ ID NO:62 和 63 分别对应于引物 KTH5 和 KTH3, 被用于扩增含有 TEF::HPT::XPR 融合基因的 1.6kB DNA 片段 (SEQ ID NO:64; 氨基酸序列如 SEQ ID NO:65 所示)。

SEQ ID NO:66 和 67 分别对应于解脂耶氏酵母 URA3 基因中 401bp 的 5'-序列和 568bp 的 3'-序列, 被用于将表达盒直接整合进耶氏酵母基因组的 Ura 位点。

SEQ ID NO:68-71 分别对应于引物 YL63、YL64、YL65 和 YL66, 被用于定点诱变, 以生成 pY24-4。

SEQ ID NO:72 和 73 分别对应于引物 YL11 和 YL12, 被用于扩增高山被孢霉 $\Delta 5$ 去饱和酶。

SEQ ID NO:74-77 分别对应于引物 YL81、YL82、YL83 和 YL84, 被用于定点诱变, 以生成 pYZM5CH。

SEQ ID NO:78 和 79 分别对应于引物 YL105 和 YL106, 被用于定点诱变, 以生成 pYZM5CHPP。

SEQ ID NO:80 和 81 分别对应于引物 YL119 和 YL120, 被用于定点诱变, 以生成 pYZM5CHPPA。

SEQ ID NO:82 和 83 分别对应于引物 YL121 和 YL122, 被用于扩增解脂耶氏酵母 URA3 基因上游 440bp 的 5'-非编码 DNA 序列 (SEQ ID NO:84)。

SEQ ID NO:85 和 86 分别对应于引物 YL114 和 YL115, 被用于定点诱变, 以生成 pYZV5 和 pYZV5P。

SEQ ID NO:87 对应于适合在解脂耶氏酵母基因组内整合和表达 $\Delta 5$ 去饱和酶基因的 5.2kB DNA 片段。

SEQ ID NO:88 - 91 分别对应于引物 YL61、YL62、YL69 和 YL70, 被用于定点诱变, 以生成 pY58BH。

SEQ ID NO:92-95 分别对应于引物 YL77、YL78、YL79A 和 YL80A, 被用于定点诱变, 以生成 pY54PC。

SEQ ID NO:96 对应于适合在解脂耶氏酵母基因组内整合和同步表达 $\Delta 6$ 去饱和酶、高山被孢霉延伸酶和高山被孢霉 $\Delta 5$ 去饱和酶基因的

8.9kB DNA 片段。

SEQ ID NO:97-100 分别对应于引物 YL101、YL102、YL103 和 YL104，被用于定点诱变，以生成 pYSD17SPC。

SEQ ID NO:101 对应于适合在解脂耶氏酵母基因组内整合和同步表达 $\Delta 6$ 去饱和酶、高山被孢霉延伸酶、高山被孢霉 $\Delta 5$ 去饱和酶和密码子优化的 $\Delta 17$ 去饱和酶基因的 10.3kB DNA 片段。

SEQ ID NO:102-113 分别对应于引物 YL1、YL2、YL3、YL4、YL5、YL6、YL7、YL8、YL9、YL10、YL23 和 YL24，被用于构建质粒。

SEQ ID NO: 114 表示异丝水霉 $\Delta 5$ 去饱和酶基因的 DNA 序列，而 SEQ ID NO: 115 表示异丝水霉 $\Delta 5$ 去饱和酶的相应的氨基酸序列。

SEQ ID NO: 116、117、120、121、124 和 125 分别对应于引物 YL13A、YL14A、YL19A、YL20、YL15 和 YL16B，用于克隆多种 $\Delta 5$ 去饱和酶基因。

SEQ ID NO: 118 表示 *Isochrysis galbana* $\Delta 5$ 去饱和酶基因的 DNA 序列，而 SEQ ID NO: 119 表示 *Isochrysis galbana* $\Delta 5$ 去饱和酶的相应的氨基酸序列。

SEQ ID NO: 122 表示 *Thraustochytrium aureum* $\Delta 5$ 去饱和酶基因的 DNA 序列，而 SEQ ID NO: 119 表示 *Thraustochytrium aureum* $\Delta 5$ 去饱和酶的相应的氨基酸序列。

SEQ ID NO:126 对应于最适合在耶氏酵母种内表达的基因的密码子优化翻译起始位点。

发明的详细说明

根据本发明，申请人提供了在含油酵母中生产 ω -3 和/或 ω -6 脂肪酸的方法。具体地，申请人提供了生产亚油酸、 γ -亚麻酸、二高- γ -亚油酸、花生四烯酸、 α -亚麻酸、十八碳四烯酸、二十碳四烯酸、二十碳五烯酸、二十二碳五烯酸和二十二碳六烯酸的方法。这是通过将由赋予 $\Delta 17$ 去饱和酶、 $\Delta 6$ 去饱和酶、 $\Delta 5$ 去饱和酶、 $\Delta 9$ 去饱和酶、 $\Delta 12$ 去饱和酶、 $\Delta 15$ 去饱和酶、 $\Delta 4$ 去饱和酶和延伸酶活性的基因编码的功能性 ω -3/ ω -6 脂肪酸生物合成途径导入含油宿主酵母中以便进行重组表达而实现的。因此，本公开内容证明了可以对含油酵母进行工程改造，以便能够生产任何所需的 PUFA 组合物。

本发明具有多种用途。通过本文所披露的方法制备的 PUFAs, 或它的衍生物可被用作饮食代用品, 或补充物, 特别是婴儿配方, 用于静脉内喂饲的患者或用于预防或治疗营养不良。另外, 可以将纯化的 PUFAs (或它的衍生物) 掺入烹饪油, 脂肪, 或配制的人造奶油中, 以便在正常使用时, 受体能接受到所需数量的饮食补充。还可将 PUFAs 掺入婴儿配方, 营养补充物或其他食品中, 并且可以用作抗炎剂或降胆固醇剂。可任选将所述组合物用于药物学用途 (人或兽医)。在这种场合下, PUFAs 通常是口服的, 不过, 可以通过能够成功地吸收它们的任何途径施用, 例如, 肠胃外 (例如皮下, 肌内或静脉内), 直肠内, 阴道内或局部 (例如, 作为皮肤油膏或洗液)。

给人或动物补充通过重组方法生产的 PUFAs, 可以导致水平增加的添加的 PUFAs, 以及它们的代谢产物。例如, 用花生四烯酸 (ARA) 治疗, 不仅能够产生水平增加的 ARA, 而且还能产生 ARA 的下游产物, 如前列腺素。复杂的调控机制, 可能需要组合各种 PUFAs, 或者添加 PUFAs 的不同的缓合物, 以便预防, 控制或克服这样的机制, 以便在个体内获得理想水平的特定的 PUFAs。

定义

在本说明书中, 使用了多种术语和缩写。提供了以下定义。

"开放读框"被缩写为 ORF。

"聚合酶链反应"被缩写为 PCR。

"美国典型培养物保藏中心"被缩写为 ATCC。

"多不饱和脂肪酸"被缩写为 PUFA (s)。

术语"脂肪酸"表示具有各种链长的长链脂族酸 (链烷酸), 从大约 C₁₂-C₂₂ (不过已知可以采用更长和更短链长的酸)。主要的链长为 C₁₆ 至 C₂₂。脂肪酸的结构是通过简单的符号系统 "X: Y" 表示的, 其中, X 是在特定脂肪酸中的碳 (C) 原子的总数, 而 Y 是双键的数量。

一般, 脂肪酸被划分成饱和的或不饱和的。术语"饱和脂肪酸"表示在它们的碳主链之间没有"双键"的脂肪酸。相反, "不饱和脂肪酸"是沿它们的碳主链具有"双键"的顺式异构体。"单不饱和脂肪酸"沿碳主链只有一个"双键" (例如, 对于棕榈油酸 (16: 1) 和油酸 (18: 1) 来说通常位于第 9 和第 10 个碳原子之间), 而"多不饱和脂肪酸" (或

"PUFAs") 沿碳主链具有至少两个双键(例如, 对于亚油酸(18: 2)来说位于第9和第10和第12和第13个碳原子之间;对于 α -亚麻酸(18: 3)来说位于第9和第10, 第12和第13, 和第15和第16个碳原子之间)。

"PUFAs"可以划分成两种主要类型(根据最靠近脂肪酸碳链的甲基末端的第一个双键的位置(n))。因此, " ω -6 脂肪酸"(ω -6 或 n -6)的第一不饱和的双键距离该分子的 ω (甲基)末端六个碳原子, 并且一共还具有两个或两个以上双键, 每一个随后的不饱和出现在朝向该分子羧基末端的三个额外的碳原子。相反, " ω -3 脂肪酸"(ω -3 或 n -3)的第一不饱和的双键距离该分子的 ω 末端三个碳原子, 并且一共还具有三个或三个以上双键, 每一个随后的不饱和出现在朝向该分子羧基末端的三个额外的碳原子。

在本说明书中, 将利用 ω -参考系统表示碳原子的编号, 双键的编号, 和最接近 ω 碳原子的双键的位置, 从 ω 碳原子开始计数(它是用于这种目的的第1号)。在下面的表1中示出了这种命名方法, 表示在标题为"简化符号"的栏目中。该表的其余部分归纳了 ω -3和 ω -6脂肪酸的常用名, 在说明书中将要使用的缩写, 以及每一种化合物的化学名称。

表 1
多不饱和脂肪酸的命名

常用名	缩写	化学名称	简化符号
亚油酸	LA	顺-9,12-十八碳三烯酸	18:2 ω-6
γ-亚油酸	GLA	顺-6, 9, 12-十八碳三烯酸	18:3 ω-6
二高-γ-亚油酸	DGLA	顺-8, 11, 14-二十碳三烯酸	20:3 ω-6
花生四烯酸	ARA	顺-5, 8, 11, 14-二十碳三烯酸	20:4 ω-6
α-亚麻酸	ALA	顺-9, 12, 15-十八碳三烯酸	18:3 ω-3
十八碳四烯酸	STA	顺-6, 9, 12, 15-十八碳三烯酸	18:4 ω-3
二十碳三烯酸	ETA	顺-8, 11, 14, 17-二十碳三烯酸	20:4 ω-3
二十碳五烯酸	EPA	顺-5, 8, 11, 14, 17-二十碳三烯酸	20:5 ω-3
二十二碳五烯酸	DPA	顺-7, 10, 13, 16, 19-二十二碳五烯酸	22:5 ω-3
二十二碳六烯酸	DHA	顺-4, 7, 10, 13, 16, 19-二十二碳六烯酸	22:6 ω-3

术语"必需脂肪酸"表示个体为了生存必须摄入的 PUFA，因为所述个体不能从头合成所述特定必需脂肪酸。例如，哺乳动物不能合成必需脂肪酸亚油酸 (18: 2, ω-6)。其他必需脂肪酸包括 GLA (ω-6)，DGLA (ω-6)，ARA (ω-6)，EPA (ω-3) 和 DHA (ω-3)。

术语"脂肪"表示脂类物质，它在 25℃ 下是固体，并且通常是饱和的。

术语"油"表示在 25℃ 下是液体并且通常是多不饱和的脂类。PUFAs 存在于某些藻类，含油酵母和丝状真菌的油中。"微生物油"或"单细胞油"是由微生物在它们的生命过程中天然产生的油类。这种油类可能包括长链 PUFAs。

术语"PUFA 生物合成途径酶"表示与 PUFA 的生物合成相关的以下任何酶（以及编码所述酶的基因），包括：Δ4 去饱和酶，Δ5 去饱和酶，Δ6 去饱和酶，Δ12 去饱和酶，Δ15 去饱和酶，Δ17 去饱和酶，Δ9 去饱和酶和/或延伸酶。

术语" ω -3/ ω -6 脂肪酸生物合成途径"表示一组基因,当它们在合适的条件下表达时,编码能催化 ω -3 和/或 ω -6 脂肪酸生产的酶。通常,与 ω -3/ ω -6 脂肪酸生物合成途径相关的基因编码某些或所有以下的酶: Δ 12 去饱和酶, Δ 6 去饱和酶,延伸酶, Δ 5 去饱和酶, Δ 17 去饱和酶, Δ 15 去饱和酶, Δ 9 去饱和酶和 Δ 4 去饱和酶。在图 2 中示出了典型的途径,证实了如何由普通来源生产 ω -3 和 ω -6 脂肪酸。本文所用与 ω -3/ ω -6 脂肪酸生物合成途径相关的术语"功能性"指该途径中的所有基因或其中一部分可表达活性酶。应当理解的是," ω -3/ ω -6 脂肪酸生物合成途径"或"功能性 ω -3/ ω -6 脂肪酸生物合成途径"并不意味着该节所列的所有基因均是必需的,因为许多脂肪酸产物仅需该途径中部分基因的表达。

术语"去饱和酶"表示能够对一种或多种脂肪酸去饱和以便产生单-或多不饱和脂肪酸或感兴趣的前体的多种酶复合物的多肽成分。尽管在本说明书中 ω -参考系统被用于表示特殊的脂肪酸,更常见的是,利用 Δ -系统从底物的羧基末端开始计数以表示去饱和酶的活性。本发明特别感兴趣的是:1.) Δ 17 去饱和酶,它能使脂肪酸的位于从该分子的羧基末端开始计数的第 17 和第 18 个碳原子之间去饱和,并且,例如,催化 ARA 转化成 EPA 和/或 DGLA 转化成 ETA; 2.) Δ 6 去饱和酶,它能催化 LA 转化成 GLA 和/或 ALA 转化成 STA; 3.) Δ 5 去饱和酶,它能催化 DGLA 转化成 ARA 和/或 ETA 转化成 EPA; 4.) Δ 4 去饱和酶,它能催化 DPA 转化成 DHA; 5.) Δ 12 去饱和酶,它能催化油酸转化成 LA; 6.) Δ 15 去饱和酶,它能催化 LA 转化成 ALA; 和 7.) Δ 9 去饱和酶,它能催化棕榈酸酯或盐转化成棕榈油酸(16:1)和/或硬脂酸酯或盐转化成油酸(18:1)。

术语"延伸酶"表示能够延伸脂肪酸碳链,以便产生比延伸酶作用的脂肪酸底物长 2 个碳原子的单或多不饱和脂肪酸的多酶复合物的多肽成分。这种延伸过程是通过与脂肪酸合成酶相关的多步骤机制发生的,其中,CoA 是酰基载体(Lassner 等, *The Plant Cell* 8: 281-292 (1996))。简单地讲,丙二酰-CoA 与长链酰基-CoA 缩合,以便产生 CO_2 和 β -酮酰基-CoA(其中,酰基部分业已延长了两个碳原子)。随后的反应包括还原成 β -羟酰基-CoA,脱水形成烯酰基-CoA,并且第二次还原,以便产生延长的酰基-CoA。由延伸酶催化的反应的例子是将

GLA 转化成 DGLA, STA 转化成 ETA 和将 EPA 转化成 DPA。因此, 延伸酶可以具有不同的特异性(例如, C_{16/18} 延伸酶优选 C₁₆ 底物, C_{8/20} 延伸酶优选 C₁₈ 底物, 而 C_{20/22} 延伸酶优选 C₂₀ 底物)。

术语"高亲和力延伸酶"表示它的底物特异性优选 GLA 的延伸酶(DGLA 是延伸酶反应的产物)。在 WO 00/12720 中披露了一种这样的延伸酶。

术语"转化效率"和"底物转化百分比"表示特定的酶(例如, 去饱和酶或延伸酶)可以将底物转化成产物的效率。转化效率是按照以下公式计算的: $([\text{产物}]/[\text{底物}+\text{产物}]) * 100$, 其中'产物'包括产生它的途径上的中间产物和所有产物。

术语"含油的"表示倾向于以脂类形式储存它们的能源的生物(Weete, In: Fungal Lipid Biochemistry, 2nd ed., Plenum, 1980)。一般, 所述微生物的细胞 PUFA 含量遵循 S 形曲线, 其中, 脂类的浓度增加, 直到在后对数期或早期静止生长期的最大值, 然后在晚期静止期和死亡期逐渐减少(Yongmanitchai and Ward, Appl. Environ. Microbiol. 57: 419-25 (1991))。

术语"含油酵母"表示被分类为酵母的微生物, 它们能够积累占干细胞重量至少 25% 的油。含油酵母的例子包括, 但不局限于以下属: 耶氏酵母属, 假丝酵母属, 红酵母属, 红冬孢属, 隐球酵母属, 丝孢酵母属和油脂酵母属。

术语"可发酵的碳源"表示微生物能够代谢以便产生能量的碳源。本发明的典型的碳源包括, 但不局限于: 单糖, 寡糖, 多糖, 链烷, 脂肪酸, 脂肪酸酯, 甘油单酯, 甘油二酯, 甘油三酯, 二氧化碳, 甲醇, 甲醛, 甲酸盐或酯以及含碳的胺。

在本文中, "分离的核酸片段"是单链或双链 RNA 或 DNA, 它任选包括合成的, 非天然的或改变了的核酸碱基。DNA 聚合物形式的分离的多核酸片段可能由一个或多个 cDNA, 基因组 DNA 或合成 DNA 的片段组成。

氨基酸或核苷酸序列的"主要部分"是包括多肽的一定数目的氨基酸序列或基因的核苷酸序列的部分, 所述数目足够通过本领域技术人员对所述序列进行的人工评估, 或通过计算机自动化的序列比较, 以及使用诸如 BLAST 的算法的鉴定(Basic Local Alignment Search

Tool; Altschul, S. F., 等, J. Mol. Biol. 215: 403-410 (1993) 而推测性地鉴定该多肽或基因。一般, 核苷酸序列的"主要部分"包括足够的序列(例如, 20-30个连续的核苷酸), 以便特异性地鉴定和/或分离包括所述序列的核酸片段。本说明书披露了编码一种或多种特定蛋白的部分或全部核苷酸序列。技术人员利用本文所提供的序列, 可以将所披露序列的全部或主要部分用于本领域技术人员所公知的目的。因此, 本发明包括在所附序列表中所提供的完整的序列, 以及如上文所定义的这些序列的主要部分。

术语"互补的"被用于表示能够彼此杂交的核苷酸碱基之间的相互关系。例如, 对于DNA来说, 腺嘌呤互补于胸腺嘧啶, 而胞嘧啶互补于鸟嘌呤。因此, 本发明还包括分离的核酸片段, 它互补于在所附序列表中提供的完整序列, 以及基本上类似的核酸序列。

"密码子简并性"表示遗传密码的性质, 它允许改变核苷酸序列, 而又不影响所编码的多肽的氨基酸序列。技术人员在用核苷酸密码子表示特定氨基酸时充分理解由特定宿主细胞所表现出来的"密码子偏倚"。因此, 在合成用于改善在宿主细胞中的表达的基因时, 需要设计所述基因, 以便它的密码子选择的频率接近宿主细胞的优选密码子选择的频率。

表示DNA序列的"化学合成的"表示组成核苷酸是在体外组装的。DNA的人工化学合成, 可以通过使用业已建立的方法完成; 或可以利用多种商业化仪器中的一种进行自动化的化学合成。"合成基因"可以由寡核苷酸基本单位组装, 它们是利用本领域技术人员所公知的方法化学合成的。将这些基本单位连接在一起并且退火, 以便形成基因片段, 然后酶促组装, 以便构建完整的基因。因此, 可以对基因进行修饰, 以便根据核苷酸序列的优化优化基因表达, 以便体现宿主细胞的密码子偏倚。如果密码子选择偏向宿主优选的密码子, 技术人员可以理解成功的基因表达的可能性。优选密码子的确定可以根据对来自宿主细胞的基因的检查, 其中, 可以获得序列信息。

"基因"表示能表达特定蛋白的核酸片段, 包括位于编码序列前面(5'非编码序列)和后面(3'非编码序列)的调控序列。"天然基因"表示天然与它自身的调控序列一起出现的基因。"嵌合基因"表示不是天然基因的任何基因, 包括不是天然一起存在的调控和编码序列。因此,

嵌合基因可以包括来自不同来源的调控序列和编码序列，或来自相同的来源，但是以不同于天然存在的方式排列的调控序列和编码序列。"内源基因"表示存在于生物基因组上的天然位置上的天然基因。"外源"基因表示正常情况下不存在于宿主生物中，而是通过基因转移导入所述宿主生物基因。外源基因可以包括插入非天然生物的天然基因，或嵌合基因。"转基因"是业已通过转化方法导入所述基因组的基因。"密码子优化的基因"是具有经过设计以便模拟宿主细胞的优选的密码子选择频率的密码子选择频率的基因。

"编码序列"表示编码特定氨基酸序列的 DNA 序列。"合适的调控序列"表示位于编码序列上游（5'非编码序列），序列内，或下游（3'非编码序列）的核苷酸序列，并且，它能影响相关编码序列的转录，RNA 加工或稳定性，或翻译。调控序列可以包括启动子，翻译前导序列，内含子，聚腺苷酸化识别序列，RNA 加工位点，效应物结合位点和茎-环结构。

"启动子"表示能够控制编码序列或功能性 RNA 表达的 DNA 序列。一般，编码序列位于启动子序列的 3'侧。启动子可能完全来自天然基因，或者由来自天然存在的不同启动子的不同元件组成，或者甚至包括合成的 DNA 片段。本领域技术人员可以理解的是，不同的启动子能够指导基因在不同的组织或细胞类型中的表达，或者在发育的不同阶段的表达，或者对不同的环境或生理学条件作出反应的表达。能导致基因在大多数细胞类型中在大多数时间表达的启动子通常被称作"组成型启动子"。还应当理解的是，由于在大多数场合下，调控序列的确切边界未能完全确定，不同长度的 DNA 片段可能具有相同的启动子活性。

术语"3'非编码序列"或"转录终止子"表示位于编码序列下游的 DNA 序列。它包括聚腺苷酸化识别序列以及编码能够影响 mRNA 加工或基因表达的调控信号的其他序列。聚腺苷酸化信号通常以影响聚腺苷酸添加在 mRNA 前体的 3'末端为特征。所述 3'区能够影响相关编码序列的转录，RNA 加工或稳定性，或翻译。

"RNA 转录物"表示由 RNA 聚合酶催化的 DNA 序列转录而得到的产物。当 RNA 转录物是所述 DNA 序列的完整的互补拷贝时，它被称作初级转录物，或者它可以是来自初级转录物的转录后加工的 RNA 序

列, 并且被称作成熟的 RNA。"信使 RNA"或"mRNA"表示没有内含子, 并且能够由细胞翻译成蛋白的 RNA。"cDNA"表示双链 DNA, 它是互补于 mRNA 并且由 mRNA 衍生的。"有义"RNA 表示包括 mRNA, 并能够由细胞翻译成蛋白的 RNA 转录物。"反义"RNA 表示互补于靶初级转录物或 mRNA 的全部或部分的 RNA, 并且它能抑制靶基因的表达 (U. S. 5, 107, 065; WO 99/28508)。反义 RNA 的互补性可以是与特定基因转录物的任何部分, 即, 位于 5'非编码序列, 3'非编码序列, 或编码序列互补的。"功能性 RNA"表示反义 RNA, 核酶 RNA, 或不能翻译, 但仍然能影响细胞加工的其他 RNA。

术语"可操作地连接的"表示核酸序列缔合在一个核酸片段上, 以便其中一个的功能受到另一个的影响。例如, 当启动子能够影响编码序列的表达时, 它就是与该编码序列可操作地连接的 (即, 所述编码序列受所述启动子的转录控制)。编码序列能够沿有义或反义方向可操作地与调控序列连接。

在本文中, 术语"表达"表示来自本发明的核酸片段的有义 (mRNA) 或反义 RNA 的转录和稳定积累。表达还可以表示 mRNA 翻译成多肽。

"转化"表示将核酸分子转入宿主生物, 导致了遗传学稳定的遗传性。例如, 核酸分子可以是能自主复制的质粒; 或者它可以整合到宿主生物的基因组中。包括转化的核酸片段的宿主生物被称作"转基因"或"重组"或"转化过的"生物。

术语"质粒", "载体"和"盒"表示染色体外元件, 它通常携带有不是细胞的中央代谢部分的基因, 并且通常是环状双链 DNA 片段形式的。所述基因可以是自主复制的序列, 基因组整合序列, 噬菌体或核苷酸序列, 线性或环状的单或双链 DNA 或 RNA, 可以来自任何来源, 其中, 多个核苷酸序列业已连接或重组成独特的结构, 它能够将启动子片段和选定的基因产物的 DNA 序列, 以及合适的 3'非翻译序列导入细胞。"转化盒"表示包括外源基因, 并且除了外源基因之外还具有能促进在特定宿主细胞中转化的元件的特定载体。"表达盒"表示包括外源基因, 并且除了外源基因之外, 还具有能够增强基因在外源宿主中表达的元件的特定载体。

术语"序列分析软件"表示可用于分析核苷酸或氨基酸序列的任何

计算机算法或软件程序。"序列分析软件"可以通过商业渠道获得的或者独立开发的。典型的序列分析软件包括，但不局限于：1.) GCG 程序组 (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI); 2.) BLASTP, BLASTN, BLASTX (Altschul 等, J. Mol. Biol. 215: 403-410 (1990)); 3.) DNASTAR (DNASTAR, Inc., Madison, WI); 和 4.) 采用了 Smith-Waterman 算法的 FASTA 程序 (W. R. Pearson, Comput. Methods Genome Res., [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor (s): Suhai, Sandor. Plenum: New York, NY)。在本申请的范围内应当理解的是，在将序列分析软件用于分析时，分析的结果是基于有关程序的"预设值"的，除非另有说明。在本文中，"预设值"表示任何组的值或参数，它是在初次初始化时最初随所述软件加载的。

本发明所使用的标准重组 DNA 和分子克隆技术是本领域众所周知的，并且披露于以下文献中：Sambrook, J., Fritsch, E. F. 和 Maniatis, T., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989) (此后称作 "Maniatis"); Silhavy, T. J., Bannan, M. L. 和 Enquist, L. W., Experiments with Gene Fusions, Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1984); 以及 Ausubel, F. M. 等, Current Protocols in Molecular Biology, 由 Greene Publishing Assoc. and Wiley-Interscience (1987) 出版。

脂肪酸的微生物生物合成

一般，在含油微生物中的脂类的积累是反应于存在于生长培养基中的总的碳与氮的比例引发的 (图 1)。当细胞耗尽了可利用的氮源时 (例如，当碳与氮的比例超过大约 40 时)，细胞一磷酸腺苷 (AMP) 的消耗，导致了线粒体中 AMP-依赖型异柠檬酸脱氢酶活性的终止和柠檬酸的积累，并且将柠檬酸转移到细胞质，并且随后通过 ATP-柠檬酸裂解酶裂解，产生乙酰-CoA。乙酰-CoA 是脂肪酸从头合成的主要基本单位。尽管任何能够有效代谢产生乙酰-CoA 的化合物都可用作脂肪酸的前体，葡萄糖是这种类型反应中的主要碳源 (图 1)。通过糖酵解将葡萄糖转化成丙酮酸，并且然后将丙酮酸转入线粒体，在这里，它能够通过丙酮酸脱氢酶转化成乙酰-CoA ("PD")。由于乙酰-CoA 不能

直接通过线粒体膜转入细胞质，乙酰-CoA 上的两个碳与草酰乙酸缩合，以便产生柠檬酸（通过柠檬酸合酶催化）。柠檬酸被直接转入细胞质，在这里，它被 ATP-柠檬酸酶裂解，以便再生乙酰-CoA 和草酰乙酸。草酰乙酸通过转化成苹果酸重新进入三羧酸循环。

丙二酰-CoA 的合成是脂肪酸生物合成的第一个关键步骤，它是在细胞质中进行的。丙二酰-CoA 是通过乙酰-CoA 羧化酶催化由乙酰-CoA 的羧化产生的。脂肪酸合成是通过多种酶脂肪酸合酶复合物 ("FAS") 催化的，并且通过八种二碳片段（来自乙酰-CoA 的乙酰基团）的缩合进行，以便形成 16-碳饱和脂肪酸棕榈酸。更具体地讲，FAS 催化一系列的 7 个反应，该系列包括以下反应 (Smith, S. FASEB J, 8 (15): 1248-59 (1994)):

1. 乙酰-CoA 和丙二酰-CoA 被转移到 FAS 的酰基载体肽 (ACP) 上。所述乙酰基团然后被转移到丙二酰基团上，形成 β -酮丁酰基-ACP，并且释放 CO_2 。

2. 所述 β -酮丁酰基-ACP 发生还原（通过 β -酮酰基还原酶）和脱水（通过 β -羟酰基脱水酶），以便形成反式-单不饱和的脂肪酰基。

3. 通过 NADPH 还原双键，产生比最初的基团长两个碳原子的饱和脂肪酰基。然后再生了丁酰基-基团与新的丙二酰基团缩合，并且重复所述延伸过程的能力。

4. 当脂肪酰基变成 16 个碳原子长时，硫酯酶活性水解它，释放游离的棕榈酸。

棕榈酸 (16: 0) 是长链饱和和不饱和脂肪酸（例如，硬脂酸 (18: 0)，棕榈油酸 (16: 1) 和油酸 (18: 1)）的前体，通过存在于内质网膜中的延伸酶和去饱和酶的作用产生棕榈酸。通过 Δ^9 去饱和酶的作用分别将棕榈酸和硬脂酸转化成它们的不饱和衍生物，即棕榈油酸 (16: 1) 和油酸 (18: 1)。

三酰基甘油（脂肪酸的主要储存单位）是通过将两个分子的酰基-CoA 酯化成甘油-3-磷酸产生 1, 2-磷酸二酰基甘油的（通常被称为磷脂酸），然后通过磷脂酸磷酸酶除去磷酸，以便得到 1, 2-二酰基甘油。三酰基甘油是通过二酰基甘油-酰基转移酶的作用添加第三个脂肪酸形成的。

$\Omega 3$ 和 $\Omega 6$ 脂肪酸的生物合成

简单地讲, 将 LA 转化成 GLA, DGLA 和 ARA (ω -6 途径) 和 ALA 转化成 STA, ETA, EPA, DPA 和 DHA (ω -3 途径) 的代谢过程涉及通过添加碳原子延长碳链, 和通过添加双键使分子去饱和 (图 2)。这需要存在于内质网膜中的一系列特殊去饱和和延伸酶。

ω -6 脂肪酸

通过 Δ 12 去饱和酶的作用将油酸转化成 LA (18: 2), 它是 ω -6 脂肪酸的第一个。按以下方法产生后续的 ω -6 脂肪酸: 1.) 通过 Δ 6 去饱和酶的活性将 LA 转化成 GLA; 2.) 通过延伸酶的作用将 GLA 转化成 DGLA; 和 3.) 通过 Δ 5 去饱和酶的作用将 DGLA 转化成 ARA。

ω -3 脂肪酸

通过 Δ 15 去饱和酶的作用将亚油酸 (LA) 转化成 ALA, 即 ω -3 脂肪酸的第一个。通过一系列类似于 ω -6 脂肪酸的步骤产生后续的 ω -3 脂肪酸。具体地讲: 1.) 通过 Δ 6 去饱和酶的活性将 ALA 转化成 STA; 2.) 通过延伸酶的活性将 STA 转化成 ETA; 和 3.) 通过 Δ 5 去饱和酶的活性将 ETA 转化成 EPA。另外, ETA 和 EPA 可以通过 Δ 17 去饱和酶的活性分别由 DGLA 和 ARA 产生。还可以通过延伸酶和 Δ 4 去饱和酶的活性进一步将 EPA 转化成 DHA。

与 Ω 脂肪酸生产相关的基因

很多微生物, 包括藻类, 细菌, 霉菌和酵母都可以通过正常细胞代谢过程合成 PUFAs 和 ω 脂肪酸。进行过充分研究的是包括 *Schizochytrium aggregatum*, 破囊壶菌属的种和高山被孢菌在内的真菌。很多沟鞭藻类 (*Dinophyceae*) 都能天然产生高浓度的 PUFAs。因此, 业已通过遗传学方法鉴定了与油类产生相关的多个基因, 并且某些基因的 DNA 序列可以公开获得 (非限定性例子如下面的表 2 所示):

表 2

某些可公开获得的与 PUFA 生产相关的基因

GenBank 登录号	描述
AY131238	<i>Argania spinosa</i> $\Delta 6$ 去饱和酶
Y055118	<i>Echium pitardii</i> var. <i>pitardii</i> $\Delta 6$ 去饱和酶
AY055117	<i>Echium gentianoides</i> $\Delta 6$ 去饱和酶
AF296076	鲁克氏毛霉 $\Delta 6$ 去饱和酶
AF007561	<i>Borago officinalis</i> $\Delta 6$ 去饱和酶
L11421	<i>Synechocystis</i> sp. $\Delta 6$ 去饱和酶
NM_031344	褐家鼠 $\Delta 6$ 脂肪酸去饱和酶
AF465283, AF465281, AF110510	高山被孢霉 $\Delta 6$ 脂肪酸去饱和酶
AF465282	深黄被孢霉 $\Delta 6$ 脂肪酸去饱和酶
AF419296	畸雌腐霉 $\Delta 6$ 脂肪酸去饱和酶
AB052086	卷枝毛霉 $\Delta 6$ 脂肪酸去饱和酶的D6d mRNA
AJ250735	<i>Ceratodon purpureus</i> $\Delta 6$ 脂肪酸去饱和酶的mRNA
AF126799	人 $\Delta 6$ 脂肪酸去饱和酶
AF126798	小鼠 $\Delta 6$ 脂肪酸去饱和酶
AF199596, AF226273	人 $\Delta 5$ 去饱和酶
AF320509	褐家鼠肝 $\Delta 5$ 去饱和酶
AB072976	小鼠 $\Delta 5$ 去饱和酶的D5D mRNA
AF489588	破囊壶菌种ATCC21685 $\Delta 5$ 脂肪酸去饱和酶
AJ510244	大雄疫霉 $\Delta 5$ 脂肪酸去饱和酶的mRNA
AF419297	畸雌腐霉 $\Delta 5$ 脂肪酸去饱和酶
AF07879	秀丽新杆线虫 $\Delta 5$ 脂肪酸去饱和酶
AF067654	高山被孢霉 $\Delta 5$ 脂肪酸去饱和酶
AB022097	<i>Dictyostellum discoideum</i> $\Delta 5$ 脂肪酸去饱和酶的mRNA
AF489589.1	破囊壶菌种ATCC21685 $\Delta 4$ 脂肪酸去饱和酶
AY332747	<i>Pavlova lutheri</i> $\Delta 4$ 脂肪酸去饱和酶
AAG36933	<i>Emericella nidulans</i> 油酸 $\Delta 12$ 去饱和酶
AF110509, AB020033	高山被孢霉 $\Delta 12$ 脂肪酸去饱和酶mRNA
AAL13300	高山被孢霉 $\Delta 12$ 脂肪酸去饱和酶

GenBank 登录号	描述
AF417244	高山被孢霉ATCC16266 Δ 12脂肪酸去饱和酶基因
AF161219	鲁克氏毛霉 Δ 12去饱和酶mRNA
X86736	<i>Spirulina platensis</i> Δ 12 去饱和酶
AF240777	秀丽新杆线虫 Δ 12去饱和酶
AB007640	<i>Chlamydomonas reinhardtii</i> Δ 12 去饱和酶
AB075526	普通小球藻 Δ 12去饱和酶
AP002063	拟南芥微粒体 Δ 12去饱和酶
NP_441622, BAA18302, BAA02924	<i>Synechocystis</i> sp. PCC 6803 Δ 15 去饱和酶
AAL36934	<i>Perilla frutescens</i> Δ 15 去饱和酶
AF338466	<i>Acheta domesticus</i> Δ 9 去饱和酶 3 mRNA
AF438199	<i>Picea glauca</i> 去饱和酶 Δ 9 (Des9) mRNA
E11368	<i>Anabaena</i> Δ 9 去饱和酶
E11367	<i>Synechocystis</i> Δ 9 去饱和酶
D83185	<i>Pichia angusta</i> Δ 9 脂肪酸去饱和酶的DNA
U90417	<i>Synechococcus vulcanus</i> Δ 9 酰基-脂类脂肪酸-去饱和酶 (desC) 基因
AF085500	高山被孢霉 Δ 9去饱和酶mRNA
AY504633	<i>Emericella nidulans</i> Δ 9硬脂酸去饱和酶 (sdeB) 基因
NM_069854	秀丽新杆线虫必需脂肪酸去饱和酶, 硬脂酰 -CoA去饱和酶 (39.1KD) (脂肪-6) 完整mRNA
AF230693	<i>Brassica oleracea</i> cultivar Rapid Cycling 硬脂酰-ACP去饱和酶 (Δ 9-B0-1) 基因, 外显子序列
AX464731	高山被孢霉延伸酶基因 (也见WO 02/08401)
NM_119617	拟南芥脂肪酸延伸酶1 (FAB1) (At4g34520) mRNA
NM_134255	小鼠ELOVL家族成员5, 长链脂肪酸的延伸 (酵母) (Elov15), mRNA
NM_134383	褐家鼠脂肪酸延伸酶2 (rELO2), mRNA
NM_134382	褐家鼠脂肪酸延伸酶1 (rELO1), mRNA
NM_068396, NM_068392, NM_070713, NM_068746, NM_064685	秀丽新杆线虫脂肪酸延伸 (elo-6), (elo-5), (elo-2), (elo-3) 和 (elo-9) mRNA

另外, 专利文献提供了与 PUFA 生产相关的很多其他基因的 DNA 序列 (和/或与上述若干基因相关的细节以及它们的分离方法)。例如, 参见: U. S. 5, 968, 809 (Δ 6 去饱和酶); U. S. 2003/0196217 A1 (Δ 17 去饱和酶); U. S. 5, 972, 664 和 U. S. 6, 075, 183 (Δ 5 去饱和酶); WO 91/13972 和 U. S. 5, 057, 419 (Δ 9 去饱和酶); WO 93/11245 (Δ 15

去饱和酶)；WO 94/11516, U.S. 5, 443, 974 和 WO 03/099216 ($\Delta 12$ 去饱和酶)；WO 02/090493 ($\Delta 4$ 去饱和酶)；和 WO 00/12720 和 U.S.2002/0139974A1 (延伸酶)。上述每一份专利和专利申请都以它们的全文形式收作本文参考。

正如本领域技术人员所了解的，导入用于生产特定的 PUFA 最终产物的宿主生物所需要的特殊功能，取决于宿主细胞（以及它的天然 PUFA 谱和/或去饱和酶谱），底物的可利用性和需要的最终产物。如图 2 中所示出的，LA, GLA, DGLA, ARA, ALA, STA, ETA, EPA, DPA 和 DHA 都可以在含油酵母中生产，这是通过导入以下 PUFA 酶功能的各种组合： $\Delta 4$ 去饱和酶， $\Delta 5$ 去饱和酶， $\Delta 6$ 去饱和酶， $\Delta 12$ 去饱和酶， $\Delta 12$ 去饱和酶， $\Delta 17$ 去饱和酶， $\Delta 9$ 去饱和酶和/或延伸酶。根据可公开获得的文献（例如，GenBank），专利文献，以及对具有生产 PUFAs 的能力的微生物进行的实验分析，本领域技术人员能够鉴定编码上述每一种酶的各种候选基因。所述序列可来自任何来源，例如，从天然来源中分离（从细菌，藻类，真菌，植物，动物等中分离），通过半合成途径生产，或从头合成。在一些实施方案中，宿主内源基因的操作是优选的；对于其它目的，必须导入异源基因。

尽管导入宿主的去饱和酶和延伸酶基因的特特定来源并不重要，选择具有去饱和酶或延伸酶活性的特定多肽的考虑因素包括：1.) 所述多肽的底物特异性；2.) 所述多肽或它的成分是否是限速酶；3.) 所述去饱和酶或延伸酶是否是合成需要的 PUFA 所必需的；和/或 4.) 所述多肽所需要的辅助因子。所表达的多肽优选具有与它在宿主细胞中的部位的生物化学环境相容的参数。例如，所述多肽可能必须与宿主细胞中的其他酶竞争底物。因此，在确定特定多肽修饰 PUFA 在特定宿主细胞中的生产的合适性时，需要考虑 K_M 和所述多肽比活性的分析。用于特定宿主细胞中的多肽是这样的多肽，它能在存在于预期宿主细胞中的生物化学条件下起作用，但原本可以是具有能够修饰需要的 PUFA 的具有去饱和酶或延伸酶活性的任何多肽。

内源 PUFA 基因

在某些情况下，需要用来生产 PUFAs 的宿主生物将具有编码某些 PUFA 生物合成途径酶的内源基因。例如，含油酵母一般可以生产 18:2 脂肪酸（某些具有合成 18:3 脂肪酸的额外能力）；因此，含油酵母

一般具有天然的 $\Delta 12$ 去饱和酶活性, 并且可能也具有 $\Delta 15$ 去饱和酶活性。因此, 在某些实施方案中, 天然去饱和酶的表达相对于异源(或“外源”)酶是优选的, 因为: 1.) 天然酶最适合与细胞内其它酶和蛋白相互作用; 且 2.) 异源基因不太可能与宿主生物共有相同的密码子优先。此外, 已知天然基因序列时的优势在于, 易于通过定向破坏将内源基因破坏。

异源 PUFA 基因

在很多情况下, 合适的去饱和酶和延伸酶不存在于选择用来生产需要的 PUFA 产物的宿主生物中。因此, 必须导入异源基因。

为了此处描述的本发明的目的, 需要证明将 ω -3 和/或 ω -6 生物合成途径导入含油宿主生物的例子; 因此, 高山被孢霉 $\Delta 5$ 去饱和酶、高山被孢霉 $\Delta 6$ 去饱和酶、异丝水霉 $\Delta 17$ 去饱和酶、和高山被孢霉延伸酶被导入解脂耶氏酵母。但是, 将特定的酶(和编码这些酶的基因)导入宿主生物和产生的特定 PUFAs 并不是对本发明进行限制。

如此处所证明的, 如果需要产生 EPA, 本领域技术人员将明白, 来源于不同来源的许多其它基因将适于将 $\Delta 5$ 去饱和酶、 $\Delta 6$ 去饱和酶、 $\Delta 17$ 去饱和酶、和延伸酶活性导入优选的微生物宿主。因此, 在本发明的一个实施方案中, 也可以将与高山被孢霉 $\Delta 6$ 去饱和酶、 $\Delta 5$ 去饱和酶和高亲和力 PUFA 延伸酶以及异丝水霉 $\Delta 17$ 去饱和酶基本等同的其它 DNA 用于在含油酵母中生产 ω -6 和/或 ω -3 脂肪酸(如 EPA)。“基本等同的”是指与选定的多肽或编码氨基酸序列的核酸序列具有至少 80%、90% 或 95% 同源性的序列。对于多肽, 比较序列的长度通常为至少 16 个氨基酸, 优选至少 20 个氨基酸或最优选 35 个氨基酸。对于核酸, 比较序列通常为至少 50 个核苷酸, 优选至少 60 个核苷酸, 更优选至少 75 个核苷酸, 最优选 110 个核苷酸。

通常采用序列分析软件测量同源性, 其中“序列分析软件”表示可用于分析核苷酸或氨基酸序列的任何计算机算法或软件程序。“序列分析软件”可以通过商业渠道获得的或者独立开发的。典型的序列分析软件包括, 但不局限于: 1.) GCG 程序组 (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI); 2.) BLASTP, BLASTN, BLASTX (Altschul 等, J. Mol. Biol. 215: 403-410 (1990)); 3.) DNASTAR (DNASTAR, Inc., Madison, WI); 和 4.) 采用了

Smith-Waterman 算法的 FASTA 程序 (W. R. Pearson, *Comput. Methods Genome Res.*, [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor (s): Suhai, Sandor. Plenum: New York, NY). 在本申请的范围内应当理解的是, 在将序列分析软件用于分析时, 分析的结果是基于有关程序的"预设值"的, 除非另有说明。在本文中, "预设值"表示任何组的值或参数, 它是在初次初始化时最初随所述软件加载的。概言之, 所述计算机软件通过给多种取代、缺失和其它修饰分配同源程度而匹配相似的序列。

此外, 本领域技术人员应当理解, 多肽可以具有氨基酸保守取代, 其取代方式使得多肽的功能没有改变或受损。具有此处描述的去饱和酶和延伸酶活性并且具有所述保守取代的多肽包括在本发明的范围内。保守取代一般包括下列各组内的取代: 1.) 甘氨酸和丙氨酸; 2.) 缬氨酸、异亮氨酸和赖氨酸; 3.) 天冬氨酸、谷氨酸、天冬酰胺和谷氨酰胺 4.) 丝氨酸和苏氨酸; 5.) 赖氨酸和精氨酸; 和 6.) 苯丙氨酸和酪氨酸。也可以在保守的疏水性或亲水性基础上进行取代 (Kyte and Doolittle, *J. Mol. Biol.* 157: 105-132 (1982)), 或在假定相似多肽二级结构的能力的基础上进行取代 (Chou and Fasman, *Adv. Enzymol.* 47: 45-148 (1978))。

在本发明的备选实施方案中, 其它尽管不基本等同于高山被孢霉 $\Delta 6$ 去饱和酶、 $\Delta 5$ 去饱和酶和高亲和力 PUFA 延伸酶以及异丝水霉 $\Delta 17$ 去饱和酶的 DNA 也可以用于此处的目的 (如, 用于生产 ω -3 和/或 ω -6 PUFAs, 如 ARA 和 EPA)。例如, 用于根据本发明的教导导入含油酵母的编码 $\Delta 6$ 去饱和酶多肽的 DNA 序列可以获自具有生产 GLA 或 STA 的能力的微生物。所述微生物包括, 例如, 属于被孢霉属、耳霉属、腐霉属、疫霉属、青霉属、*Porphyridium*、*Coidosporium*、毛霉属、镰孢属、曲霉属、红酵母属和虫霉属的那些。在 *Porphyridium* 属内, 特别感兴趣的是 *P. Cruentum*。在被孢霉属内, 特别感兴趣的是长形被孢霉、微小被孢霉、喜湿被孢霉、*M. ramanniana* var. *angulispora* 和高山被孢霉。在毛霉属中, 特别感兴趣的是卷枝毛霉和爪哇毛霉。

或者, 不与高山被孢霉 $\Delta 6$ 去饱和酶基本等同, 但可以在脂肪酸分子的羧基末端开始的第 6 个碳对脂肪酸分子去饱和的相关去饱和酶也可以在本发明中用作 $\Delta 6$ 去饱和酶, 假定该去饱和酶仍然可以有效将

LA 转化为 GLA 和/或将 ALA 转化为 STA。因此，可以通过基本与此处公开的去饱和酶和延伸酶起相同作用的能力而鉴定相关的去饱和酶和延伸酶。

总之，可以从多种来源分离适于此处目的的编码 PUFA 生物合成途径酶的基因。用于此处目的的去饱和酶的特征在于以下能力：1.) 使脂肪酸的位于从该分子的羧基末端开始计数的第 17 和第 18 个碳原子之间去饱和，并且，催化 ARA 转化成 EPA 和 DGLA 转化成 ETA ($\Delta 17$ 去饱和酶)；2.) 催化 LA 转化成 GLA 和/或 ALA 转化成 STA ($\Delta 6$ 去饱和酶)；3.) 催化 DGLA 转化成 ARA 和/或 ETA 转化成 EPA ($\Delta 5$ 去饱和酶)；4.) 催化 DPA 转化成 DHA ($\Delta 4$ 去饱和酶)；5.) 催化油酸转化成 LA ($\Delta 12$ 去饱和酶)；6.) 催化 LA 转化成 ALA ($\Delta 15$ 去饱和酶)；和 7.) 催化棕榈酸酯或盐转化成棕榈油酸和/或硬脂酸酯或盐转化成油酸。以相似的方式，用于此处目的的合适的延伸酶不限于来源于特定来源的那些；而是，用于此处目的的酶的特征在于它们具有使得脂肪酸碳链相对于延伸酶作用的底物能够延长 2 个碳，从而产生单或多不饱和脂肪酸的能力。

对 Ω 脂肪酸基因进行优化以便在特定生物中表达

尽管 PUFA 去饱和酶或延伸酶的特来源在本发明中不是关键的，本领域技术人员将理解，异源基因将在备选的宿主以多种效率进行表达。因此， ω -3 和/或 ω -6 PUFA 生产可以通过选择特定去饱和酶或延伸酶而进行优化，所述去饱和酶或延伸酶在异源宿主中的表达水平相对于备选去饱和酶或延伸酶在感兴趣的宿主生物中的表达而言是优选的。此外，可能需要根据感兴趣的特定 PUFA 产物组成，修饰特定 PUFA 生物合成途径酶的表达，以达到每一种的最佳转化效力。多种基因工程技术可以用于优化特定酶的表达。两种这样的技术包括密码子优化和基因突变，如下文所描述。例如，通过这两种方法中的任意一种产生的具有去饱和酶和/或延伸酶活性的基因可以用于本发明中合成 ω -3 和/或 ω -6 脂肪酸。

密码子优化

正如本领域技术人员所理解的，通常修饰在外源宿主中表达的编码特定多肽的密码子的一部分，使得修饰的多肽使用替代的宿主优选的密码子。使用宿主优选的密码子能够显著增强编码多肽的外源基因

的表达。

一般，可以在特定的感兴趣的宿主物种中确定宿主优选的密码子，包括检查蛋白中的密码子选择（优选大量表达的蛋白），并且确定哪些密码子是以最高频率使用的。然后，可以使用在所述宿主物种中优选的密码子完全或部分合成具有去饱和酶或延伸酶活性的感兴趣的多肽的编码序列。也可以合成DNA的全部（或部分），以便除去可能出现在转录的mRNA上的任何去稳定化序列或二级结构区。还可以合成所述DNA的全部（或部分），以便改变将碱基组成改变成在需要的宿主细胞中更优选形式。

对于本发明来说，需要修饰编码具有 $\Delta 17$ 去饱和酶活性的多肽的密码子的一部分，以便增强基因在解脂耶氏酵母中的表达。修饰了天然基因（如异丝水霉 $\Delta 17$ 去饱和酶）的核酸序列，以便使用宿主优选的密码子。技术人员可以理解的是，该优化方法同样可应用于 ω -3/ ω -6脂肪酸生物合成途径中的其他基因（例如，参见共同未决的美国临时申请NO.60/468718，在此全文收作本文参考）。此外，对异丝水霉 $\Delta 17$ 去饱和酶的调节仅仅是示例的。

基因突变

用于合成序列并且将这些序列组合在一起的方法在文献中业已完善。例如，体外诱变和选择，定点诱变，易错PCR（Melnikov等，Nucleic Acids Research, 27 (4) : 1056-1062 (February 15, 1999)），"基因改组"或其他方法可用于获得天然存在的或密码子优化的去饱和酶或延伸酶基因的突变。这样使得能够生产去饱和酶或延伸酶多肽，它们分别具有，例如更长的半衰期或需要的PUFA在体内生产的更高速度。

如果需要，对酶促活性重要的感兴趣多肽（即去饱和酶或延伸酶）的区可以通过常规诱变，所得到的突变多肽的表达以及确定它们的活性而确定。突变体可以包括缺失，插入或点突变，或它们的组合。典型的功能性分析始于缺失诱变，以便确定功能所需要的蛋白的N-末端和C-末端限制，然后进行内部缺失，进行插入或点突变，以便进一步确定功能所需要的区。还可以使用其他技术，如盒诱变或完全合成。例如，缺失诱变是通过使用核酸外切酶依次去掉5'或3'编码区实现的。可以获得用于这种技术的试剂盒。在缺失之后，在5'或3'缺失之后通

过将包括起始或终止密码子的寡核苷酸分别连接在缺失的编码区上获得所述编码区。另外，通过各种方法将编码起始或终止密码子的寡核苷酸插入所述编码区，包括定点诱变，诱变 PCR 或通过连接到在现有限制位点上消化过的 DNA 上。内部缺失同样可以通过多种方法实现，包括使用 DNA 上的现有限制位点，通过使用诱变引物进行定点诱变或诱变 PCR。插入是通过诸如接头-扫描诱变，定点诱变或诱变 PCR 的方法完成的。点突变是通过诸如定点诱变或诱变 PCR 的技术完成的。

还可将化学诱变用于鉴定对活性重要的去饱和酶或延伸酶多肽区。表达突变的构建体，并且测定所得到的改变了的蛋白作为去饱和酶或延伸酶的能力。这种结构-功能分析，可以确定哪些区可以缺失，哪些区能够承受插入，以及哪些点突变使得突变的蛋白能够以基本上与天然的去饱和酶或天然的延伸酶相同的方式起作用。所有这样的突变蛋白及其来源于本文所述密码子优化的基因的编码核苷酸序列均属于本发明范畴。

对于本发明来说，需要修饰编码具有 $\Delta 17$ 去饱和酶活性的多肽的密码子的一部分，以便增强基因在解脂耶氏酵母中的表达。修饰了天然基因（如异丝水霉 $\Delta 17$ 去饱和酶）的核酸序列，以便使用宿主优选的密码子。技术人员可以理解的是，该优化方法同样可应用于 ω -3/ ω -6 脂肪酸生物合成途径中的其他基因（例如，参见共同未决的美国临时申请 NO.60/468718，在此全文收作本文参考）。此外，对异丝水霉 $\Delta 17$ 去饱和酶的调节仅仅是示例的。

ω -3 和/或 ω -6 脂肪酸的微生物生产

ω -3 和/或 ω -6 脂肪酸的微生物生产与从诸如鱼或植物的天然来源中纯化相比具有若干优点。例如：

- 1.) 已知与高等生物相比，很多微生物具有大大简化了的油类组成，使得对需要成分的纯化更容易；
- 2.) 微生物生产不容易出现由外部因素导致的波动，如天气和食物供应；
- 3.) 微生物生产的油基本上不会受到环境污染物的污染；
- 4.) 微生物能够以可能具有特殊用途的特定形式提供 PUFAs；和，
- 5.) 微生物油类生产可以通过控制培养条件操纵，特别是通过提供

用于微生物表达的酶的特定底物，或通过添加化合物或通过遗传工程手段抑制不希望的生物化学途径。

除了上述优点之外，用重组微生物生产 ω -3和/或 ω -6脂肪酸，提供了通过在宿主中提供新的生物合成途径或通过抑制不希望的途径而改变天然存在的微生物脂肪酸特性的能力，从而提高需要的 PUFAs，或它的缀合形式的水平，以及降低不需要的 PUFAs 的水平。例如，可以改变所生产的 ω -3与 ω -6脂肪酸的比例，专门生产 ω -3或 ω -6脂肪酸，同时消除其他 ω 脂肪酸的生产，或工程生产特定 PUFA，而又没有其他 PUFA 下游或上游产物的大量积累。

表达系统，盒和载体

本发明所披露的序列的密码子优化的基因和基因产物可以在异源微生物宿主细胞中生产，特别是在含油酵母（例如，解脂耶氏酵母）的细胞中生产。在重组微生物宿主中的表达，可用于生产各种 PUFA 途径中间产物，或用于调控业已存在于宿主中的 PUFA 途径，以便利利用所述宿主合成迄今为止不可能的新的产物。

包括能指导外源蛋白高水平表达的调控序列的微生物表达系统和表达载体是本领域技术人员所熟知的。它们都可用于构建嵌合基因，以便生产本发明密码子优化的序列的任何基因产物。然后通过转化将所述嵌合基因导入合适的微生物，以便提供所编码酶的高水平表达。

因此，预计导入受合适启动子控制的编码 PUFA 生物合成途径（例如，本文所披露的 Δ 5去饱和酶， Δ 6去饱和酶， Δ 17去饱和酶，和延伸酶）的嵌合基因，会导致 ω -3和/或 ω -6脂肪酸产量的增加。预计，可将它用于在宿主微生物中表达本发明密码子优化的基因的各种组合。预计可以用其在宿主微生物中一起表达这些 PUFA 去饱和酶和延伸酶基因的多种组合。本领域技术人员显而易见的是，包括在特定表达盒内的特定基因将取决于宿主细胞，其采用天然去饱和酶和延伸酶合成 PUFAs 的能力、底物和所需终产物的可获得性。例如，可能需要构建的表达盒包括编码下列一种或多种酶促活性的基因： Δ 4去饱和酶， Δ 5去饱和酶， Δ 6去饱和酶， Δ 12去饱和酶， Δ 15去饱和酶， Δ 17去饱和酶， Δ 9去饱和酶和/或延伸酶这样，本发明涉及生产 PUFAs 的方法，包括让脂肪酸底物与本文所披露的 PUFA 酶接触，以便将所述底物转化成

需要的脂肪酸产物。因此，可以将本文所披露的每一种 PUFAs 基因和相应的酶产物（如具有合适的去饱和酶或延伸酶活性的野生型、密码子优化的、合成的和/或突变的酶）直接或间接用于生产 PUFAs。PUFAs 的直接生产是这样进行的，将脂肪酸底物直接转化成需要的脂肪酸产物，而没有任何中间步骤或中间途径。例如，ARA 生产能够在能产生或提供了 DLGA 的宿主细胞中进行，包括向所述细胞中添加或导入能提供 $\Delta 5$ 去饱和酶活性的表达盒。

相反，编码 PUFA 生物合成途径的多个基因可以组合使用，以便发生一系列反应，以便产生需要的 PUFA。例如，编码延伸酶， $\Delta 5$ 去饱和酶， $\Delta 17$ 去饱和酶和 $\Delta 4$ 去饱和酶活性的表达盒使得宿主细胞天然产生 GLA，而不是产生 DHA（以便通过延伸酶将 GLA 转化成 DGLA；然后可以通过 $\Delta 5$ 去饱和酶将 DGLA 转化成 ARA；然后通过 $\Delta 17$ 去饱和酶将 ARA 转化成 EPA，再通过延伸酶将后者转化成 DPA；和通过 $\Delta 4$ 去饱和酶将 DPA 转化成 DHA）。在优选实施方案中，其中，宿主细胞是含油酵母，编码 PUFA 生物合成所必需的每一种酶的表达盒都需要导入所述生物，因为在这些生物中天然产生的 PUFAs 局限于 18:2 脂肪酸（即，LA），并且更不常见的是 18:3 脂肪酸（即，ALA）。另外，可能需要底物补给。

用于转化合适的宿主细胞的载体或 DNA 盒为本领域所公知。存在于所述构建体中的序列的具体选择取决于需要的表达产物（同上）、宿主细胞的性质以及推荐的分离转化细胞与未转化细胞的方式。不过，通常，所述载体或盒包括指导相关基因转录和翻译的序列，选择标记，和使得它能够自主复制或染色体整合的序列。合适的载体包括所述基因的 5' 区，它控制转录起始，和 DNA 片段的 3' 区，它控制转录终止。最优选的是，当两个控制区都来自转化过的宿主细胞的基因时是最优选的，不过，可以理解的是，所述控制区不一定来自被选择用作生产宿主的特定物种的天然基因。

可用于在需要的宿主细胞中驱动去饱和酶和/或延伸酶 ORFs 表达的起始控制区或启动子是多种多样的，并且为本领域技术人员所熟知。实际上，能够指导这些基因在选定的宿主细胞中表达的任何启动子都适合本发明。在宿主细胞中的表达能够以瞬时或稳定的方式进行。瞬时表达可以通过诱导可操作地连接在感兴趣的基因上的可调控

启动子的活性而实现。稳定表达可以通过使用可操作地与感兴趣的基因连接的组成型启动子而实现。例如，当宿主细胞是酵母时，提供能够在酵母细胞中起作用的转录和翻译区，特别是来自宿主物种的那些。例如，转录起始调控区可以从以下渠道获得：1.) 糖酵解途径上的基因，如醇脱氢酶，甘油醛-3-磷酸脱氢酶（参见美国专利申请号 60/482263），磷酸甘油变位酶（参见美国专利申请号 60/482263），果糖二磷酸醛缩酶（参见美国专利申请号 60/519971），磷酸葡萄糖异构酶，磷酸甘油酸激酶等；或 2.) 可调控的基因，如酸性磷酸酶，乳糖酶，金属硫蛋白，葡糖淀粉酶，翻译延伸因子 EF1- α (TEF) 蛋白 (U. S. 6, 265, 185)，核糖体蛋白 S7 (U. S. 6, 265, 185) 等。可以使用多种调控序列中的任意一种，这取决于是需要组成型转录还是诱导型转录，启动子在表达感兴趣的 ORF 方面的效率，以及构建的方便性等。

通过修饰外源基因的翻译起始密码子 'ATG' 周围的核苷酸序列，以便它们包括有效的酵母翻译起始序列可以在酵母中获得最佳基因表达。具体地讲，可以通过定点诱变增加低效表达的基因的表达，其中，所述低效表达的基因与内源酵母基因框内融合，优选与高表达的基因融合。另外，正如本发明中在解脂耶氏酵母中所证实的，可以确定宿主中的共有翻译起始序列，并且通过工程手段将该序列结合到异源基因上，以便它们在感兴趣的宿主中实现最佳表达。

终止区可源于所述基因的 3' 区，所述起始区是从该基因上获得的或者从不同的基因上获得的。大量终止区是已知的，并且能在多种宿主中发挥令人满意的作用（当在与产生它们的相同和不同的属和种中使用）。终止区通常在更大程度上是根据方便性选择的，而不是因为任何特殊特性。优选的是，终止区源于酵母基因，特别是糖酵母属，裂殖酵母属，假丝酵母属，耶氏酵母属或克罗维酵母属。同样已知编码 γ -干扰素和 α -2 干扰素的哺乳动物基因的 3'-区在酵母中起作用。终止控制区还可以来自优选宿主天然具有的各种基因。终止位点可以任选是必需的；不过，如果具有，它是最优选的。

正如技术人员所了解的，仅仅将基因插入克隆载体，不能确保它能够在需要的水平上成功地表达。根据对高表达速度的需要，业已通过操纵多种不同的遗传因子产生了多种特化的表达载体，这些遗传因子能控制来自宿主细胞的转录，翻译，蛋白稳定性，含氧极限，以及

从宿主的分泌。更具体地讲，业已对某些分子特征进行了操作，以便控制基因表达，包括：1.) 相关转录启动子和终止子序列的性质；2.) 克隆基因的拷贝数量以及所述基因是质粒携带的或者是整合到宿主细胞的基因组上的；3.) 合成的外源蛋白的最终细胞定位；4.) 宿主生物的翻译效率；和 5.) 克隆的基因蛋白在宿主细胞中的内在稳定性。上述每一种类型的修饰包括在本发明中，作为进一步优化密码子优化的 PUFA 生物合成途径酶表达的手段。

微生物宿主的转化

一旦获得了适合在含油酵母中表达的编码密码子优化的去饱和酶或延伸酶多肽的 DNA，将它放入能够在宿主细胞中自主复制的质粒载体上，或者将它直接整合到宿主细胞的基因组上。表达盒的整合可以在宿主基因组中随机地进行，或者可以通过使用含有与宿主基因组具有足够同源性以靶定宿主基因座内的重组的区域的构建体靶定。当所述构建体被靶定于内源基因座时，转录和翻译调控区的全部或某些可以通过内源基因座提供。

当两个或两个以上的基因由独立的复制载体表达时，希望每一个载体具有不同的选择方式，并且应当缺少与其他构建体的同源性，以便保持稳定表达，并且防止构建体之间的元件的重新分类。对调控区，选择方式和导入的构建体的繁殖方法的明智的选择，可以通过实验确定，以便所有导入的基因在需要的水平上表达，以便提供需要产物的合成。

可以通过任何标准技术将包括感兴趣的基因的构建体导入宿主细胞。所述技术包括转化（例如，乙酸锂转化 [Methods in Enzymology, 194: 186-187 (1991)]），原生质体融合，bolistic 冲击，电穿孔，显微注射，或能够将感兴趣的基因导入宿主细胞的任何其他方法。可用于含油酵母（即，解脂耶氏酵母）的更具体的技术包括美国专利号 4,880,741 和 5,071,764 和 Chen, D. C.等 (Appl Microbiol Biotechnol. 48 (2): 232-235 (1997))。

为了方便起见，业已通过任何方法操作以摄入 DNA 序列（例如，表达盒）的宿主细胞在本文中被称作“转化过的”或“重组的”。转化过的宿主具有至少一个拷贝的表达构建体，并且可以具有两个或两个以上拷贝，这取决于所述基因是整合在所述基因组上，扩增的，或者存

在于具有多个拷贝的染色体外元件上。转化过的宿主细胞可以通过选择包含在导入的构建体上的标记而鉴定。另外，可以将独立的标记构建体与需要的构建体共同转化，因为很多转化技术能将很多 DNA 分子导入宿主细胞。通常，选择转化过的宿主在选择性培养基上生长的能力。选择性培养基可以掺入抗生素或缺少未转化的宿主生长所需要的因子，如养分或生长因子。导入的标记基因可以产生抗生素抗性，或者编码必须的生长因子或酶，以便当它在转化过的宿主中表达时能够在选择培养基上生长。当所表达的标记蛋白可以直接或间接检测时，还可以进行转化宿主的选择。所述标记蛋白可以单独表达，或者作为与另一种蛋白的融合体形式表达。标记蛋白可以通过以下方法检测：1.) 它的酶促活性（例如， β -半乳糖苷酶可以将底物 X-gal[5-溴-4-氯-3-吲哚基--D-半乳吡喃糖苷]转化成有颜色的产物；荧光素酶可以将荧光素转化成发光产物）；或 2.) 它的产生光或修饰特征（例如，*Aequorea victoria* 的绿色荧光蛋白在用兰色光线照射时会发出荧光）。另外，例如，可将抗生素用于检测存在于感兴趣的蛋白上的标记蛋白或分子量标记。例如，能表达标记蛋白或标记的细胞可以肉眼筛选，或者通过诸如 FACS 的技术或使用抗生素淘选。为了选择酵母转化体，可以使用能够在酵母中起作用的任何标记。理想的是，对卡那霉素，潮霉素和氨基糖苷 G418 的抗性是感兴趣的，以及在缺少尿嘧啶或亮氨酸的培养基上生产的能力。

在转化之后，将适合本发明的密码子优化的序列的基因产物（并且任选包括在宿主细胞中表达的其他 PUFA 酶）可以由所述宿主天然地或转基因地生产，或者它们可以外源提供。

ω -3 和/或 ω -6 脂肪酸在微生物中的生物合成的代谢工程

对本发明基因的密码子优化的序列以及优化其他 PUFA 基因以便在含油酵母中表达的方法的了解，可用于操作在含油酵母中的 ω -3 和/或 ω -6 脂肪酸生物合成；特别是在解脂耶氏酵母中。这可能需要在 PUFA 生物合成途径上的直接的代谢工程或途径的其他操作，这些操作可以将碳贡献给 PUFA 生物合成途径。可用于操作生物化学途径的方法为本领域技术人员所公知。

在 PUFA 生物合成途径内操作的情况下，可能需要增加 LA 的生产，以便增加 ω -3 和/或 ω -6 脂肪酸的生产。这可以通过导入和/或扩增

编码 $\Delta 9$ 和/或 $\Delta 12$ 去饱和酶的基因而实现。

为了使 ω -6 不饱和脂肪酸，如 ARA 的生产最大化，本领域技术人员公知，该生产在基本上不含 ALA 的宿主微生物中是有利的。因此，优选地，宿主是通过去除或抑制使得 LA 转化为 ALA 的 $\Delta 15$ 或 ω -3 型去饱和酶活性而选择或获得的。可以通过以下方法减少或消除内源去饱和酶活性，例如：1) 提供用于将反义序列转录到 $\Delta 15$ 去饱和酶转录产物的盒；2) 通过插入、取代和/或缺失全部或部分靶基因而破坏 $\Delta 15$ 去饱和酶基因；或 3) 使用天然具有或经过突变而具有低 $\Delta 15$ 去饱和酶活性或无 $\Delta 15$ 去饱和酶基因的宿主细胞。对不需要的去饱和酶途径的抑制也可以通过使用如 U. S. 4,778, 630 中描述的特定去饱和酶抑制剂而实现。

或者，可能需要使 ω -3 脂肪酸的生产最大化（并且使 ω -6 脂肪酸的合成最小化）。因此，可以利用宿主微生物，其中利用上述任何一种方法（也参见，例如，共同未决的美国临时申请 No. 60/484209，在此全文引入作为参考）去除或抑制了将油酸转化为 LA 的 $\Delta 12$ 去饱和酶活性。随后，可以将合适的表达盒与合适的底物（如 ALA）一起导入宿主，用于转化为 ALA 的 ω -3 脂肪酸衍生物（如 STA, ETA, EPA, DPA, DHA）。

除了目前的 PUFA 生物合成途径，预期操作导致前体脂肪酸生物合成的一些其它酶促途径，可以导致特定 PUFAs 的总体净生物合成。这些相关途径的鉴定和途径在将来是有用的。

上调需要的生物合成途径的技术

可以将额外拷贝的去饱和酶和延伸酶基因导入宿主，以便提高 ω -3 和/或 ω -6 脂肪酸生物合成途径的产量。去饱和酶或延伸酶基因的表达还可以通过使用较强的启动子（调控型的或组成型的）而在转录水平上加强，以便导致增强了的表达，包括从 mRNA 或编码的蛋白上去掉/缺失去稳定化序列，或者向 mRNA 增加稳定化序列（U. S. 4,910,141）。正如本发明所证实的，增强去饱和酶或延伸酶基因表达的另一个途径是提高所编码的 mRNAs 的翻译效率，包括用能在选定的宿主中优化基因表达的密码子取代天然基因上的密码子。

下调不希望的生物合成途径的技术

相反，可以通过基因破坏消除或通过其他方式（例如，反义

mRNA) 下调与 ω -3 和/或 ω -6 脂肪酸生物合成途径竞争能量或碳的生物化学途径, 或干扰特定 PUFA 最终产物产生的天然 PUFA 生物合成途径酶。为了进行基因破坏, 将外源 DNA 片段 (通常是选择标记基因) 插入要破坏的结构基因, 以便破坏它的编码序列, 并因此功能性灭活所述基因。将所述破坏盒转化入宿主细胞, 导致了非通过同源重组用无功能的破坏的基因替代功能性天然基因 (例如, 参见: Hamilton 等 *J. Bacteriol.* 171: 4617-4622 (1989); Balbas 等 *Gene* 136: 211-213 (1993); Gueldener 等 *Nucleic Acids Res.* 24: 2519-2524 (1996); 和 Smith 等 *Methods Mol. Cell. Biol.* 5: 270-277 (1996))。

当靶基因的序列已知时, 反义是下调基因的另一种方法。为了实现这一目的, 克隆了来自所需基因的核酸片段, 并且可操作地与启动子连接, 以便 RNA 的反义链能够转录。然后将该构建体导入宿主细胞, 并且生产 RNA 的反义链。反义 RNA 能通过抑制编码感兴趣的蛋白的 mRNA 的积累, 抑制基因表达。本领域技术人员可以理解的是, 特殊考虑与使用反义技术相关, 以便减弱特定基因的表达。例如, 反义基因的适当水平的表达可能需要使用不同的嵌合基因, 其中采用本领域技术人员所公知的不同的调控元件。

尽管当序列已知时, 定向基因破坏和反义技术提供了下调基因的有效方法, 业已开发了不是基于序列的其他特异性更低的方法, 例如, 细胞可以暴露于 UV 辐射, 然后筛选需要的表型。用化学试剂进行的诱变同样能有效产生突变体, 并且通常使用的物质包括能影响非复制 DNA 的化学物质 (例如, HNO_2 和 NH_2OH), 以及能影响复制 DNA 的试剂 (例如, 吡啶染料, 特别是导致移码突变)。使用辐射或化学试剂产生突变体的特定方法在现有技术有大量报导。例如, 参见: Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology*, 2nd ed. (1989) Sinauer Associates: Sunderland, MA; 或 Deshpande, Mukund V., *Appl. Biochem. Biotechnol.*, 36: 227(1992)。

基因破坏的另一种非特异性方法是使用可转座元件或转座子。转座子是遗传元件, 它能随机地插入 DNA, 但是随后可以根据序列回收, 以便确定是否发生了插入。体内和体外转座转座方法是公知的。这两种方法涉及组合使用可转座元件和转座酶。当可转座元件或转座子在存在转座酶的条件下与核酸片段接触时, 可转座元件能随机地插入所

述核酸片段。这种技术可用于随机诱变，并且用于基因分离，因为受到破坏的基因可以根据可转座元件的序列鉴定。用于体外转座的试剂盒可以通过商业渠道获得[例如，参见：1.) The Primer Island Transposition Kit, 可以从 Perkin Elmer Applied Biosystems 获得, Branchburg, NJ, 基于酵母 Ty1 元件; 2.) The Genome Priming System, 可以从 New England Biolabs 获得, Beverly, MA, 基于细菌转座子 Tn7; 和 3.) EZ: TN 转座子插入系统, 可以从 Epicentre Technologies 获得, Madison, WI, 基于 Tn5 细菌转座因子]。

在本发明的范围内，可以通过上述另一种方法将它用于调控脂肪酸生物合成途径的表达。例如，本发明提供了将编码生物合成途径上的关键酶的基因导入含油酵母以生产 ω -3 和/或 ω -6 脂肪酸的方法。这些基因编码以下一种或多种： Δ 6 去饱和酶， Δ 5 去饱和酶， Δ 12 去饱和酶， Δ 15 去饱和酶， Δ 4 去饱和酶， Δ 17 去饱和酶， Δ 9 去饱和酶，和 PUFA 延伸酶。特别适用于在含油酵母中表达这些基因，所述酵母天然不具有 ω -3 和/或 ω -6 脂肪酸生物合成途径，并且调控这些和其他 PUFA 生物合成基因的表达，以便使用各种方法使优选的 PUFA 产物的产量最大化，用于宿主生物的代谢工程。

用于重组生产 ω -3 和/或 ω -6 脂肪酸的优选的微生物宿主

用于生产 ω 脂肪酸的宿主细胞可以包括包括在大范围的温度和 pH 值下，在多种饲料上生长的微生物宿主，所述饲料包括简单的或复杂的碳水化合物，有机酸和醇，和/或烃。

但是，优选的微生物宿主是含油酵母。这些生物可天然地进行油类合成和积累，所述油包括超过大约 25% 的细胞干重，更优选超过大约 30% 的细胞干重，最优选超过大约 40% 的细胞干重。通常被确定为含油酵母的属包括，但不局限于：耶氏酵母属，假丝酵母属，红酵母属，红冬孢属，隐球酵母属，丝孢酵母属和油脂酵母属。更具体地讲，典型的油类合成酵母包括：类酵母红冬孢，斯达氏油脂酵母，产油油脂酵母，拉可夫氏假丝酵母，铁红假丝酵母，热带假丝酵母，产朊假丝酵母，茁芽丝孢酵母，皮状丝孢酵母，胶粘红酵母，禾木科红酵母，和解脂耶氏酵母（以前被划分为解脂假丝酵母）。

最优选的是含油酵母解脂耶氏酵母；而在其他实施方案中，最优选的是解脂耶氏酵母菌株，被命名为 ATCC #20362, ATCC #8862,

ATCC #18944, ATCC #76982 和/或 LGAM S (7) 1 (Papanikolaou S., 和 Aggelis G., Bioresour. Technol. 82 (1) : 43-9 (2002))。

过去, 已经用解脂耶氏酵母的多种菌株制备和生产: 异柠檬酸裂合酶 (DD259637); 脂酶(SU1454852, W02001083773, DD279267); 多羟基链烷酸(W02001088144); 柠檬酸(RU2096461, RU2090611, DD285372, DD285370, DD275480, DD227448, PL160027); 赤藓糖醇(EP770683); 2-氧戊二酸 (DD267999); γ -癸内酯(U. S. 6,451, 565, FR2734843); γ -十二内酯 (EP578388)和丙酮酸(JP09252790)。

PUFA 生产的发酵方法

转化过的微生物宿主细胞是在能优化去饱和酶和延伸酶的活性并且产生最大的和最经济的优选 PUFAs 的产量的条件下生长的。一般, 可以优化的培养基条件包括碳源的类型和用量, 氮源的类型和用量, 碳-氮的比例, 氧气含量, 生长温度, pH, 生物量生产期的长度, 油积累期的长度, 以及细胞收获时间。感兴趣的微生物, 如含油酵母在复杂培养基(例如, 酵母提取物-蛋白胨-葡萄糖培养液(YPD))或缺少生长所必需的成分, 以便强制选择需要的表达盒的特定基本培养基(例如, Yeast Nitrogen Base (DIFCO Laboratories, Detroit, MI))中生长。

本发明的发酵培养基必须包括合适的碳源。合适的碳源包括, 但不局限于: 单糖(例如, 葡萄糖, 果糖), 二糖(例如, 乳糖, 蔗糖), 寡糖, 多糖(例如, 淀粉, 纤维素或它们的混合物), 糖醇(例如, 甘油)或可更新饲料的混合物(例如, 干酪乳清渗透物, 玉米浆, 甜菜糖浆, 大麦芽)。另外, 碳源可以包括链烷, 脂肪酸, 脂肪酸酯, 甘油单酯, 甘油二酯, 甘油三酯, 磷脂, 和脂肪酸的各种商业来源, 包括植物油(例如, 大豆油)和动物脂肪。另外, 碳源可以包括一-碳源(例如, 二氧化碳, 甲醇, 甲醛, 甲酸和含碳胺), 业已证实了其转化成关键生物化学中间产物的代谢转化。因此, 预计, 用于本发明的碳源包括可以包括多种含碳源, 并且仅仅局限于宿主生物的选择。尽管上述所有碳源及其混合物预计都适合本发明, 优选的碳源是糖和/或脂肪酸。最优选的是葡萄糖和/或含有 10-22 个碳的脂肪酸。

氮可以由无机(例如, $(\text{NH}_4)_2\text{SO}_4$)或有机来源(例如, 尿素或谷氨酸)提供。除了合适的碳和氮源之外, 发酵培养基还必须还包括合

适的矿物质，盐，辅因子，缓冲液，维生素，和本领域技术人员已知的适合微生物生长，并且促进 PUFA 生产所需要的酶促途径的其他成分。特别关注若干金属离子（例如， Mn^{+2} ， Co^{+2} ， Zn^{+2} ， Mg^{+2} ），它们能促进脂类和 PUFAs 的合成（Nakahara, T.等, Ind.Appl. Single Cell Oils, D. J. Kyle and R. Colin, eds. pp 61-97 (1992)）。

本发明的优选的生长培养基是常见的商业制备的培养基，如酵母 Nitrogen Base (DIFCO Laboratories, Detroit, MI)。其他确定的或合成的生长培养基也可以使用，并且适合特定微生物生长的培养基是微生物学或发酵科学领域的技术人员所公知的。用于发酵的合适的 pH 范围通常为大约 pH 4.0-pH 8.0，其中 pH 5.5-pH 7.0 是起始生长条件的优选范围。发酵可以在需氧或厌氧条件下进行，其中，微需氧条件是优选的。

通常，高水平 PUFAs 在含油酵母细胞中的积累，需要两个阶段的过程，因为在脂肪的生长和合成/储存之间的代谢状态必须是“平衡的”。因此，最优选的是，需要两个阶段的发酵过程在含油酵母中生产 PUFAs。在该方法中，发酵的第一阶段被用于制备和积累细胞物质，并且以快速细胞生长和细胞分裂为特征。在发酵的第二阶段，优选在培养物中建立氮剥夺条件，以便促进高水平的脂类积累。这种氮剥夺的作用是降低细胞中 AMP 的有效浓度，以便减弱线粒体的 NAD-依赖型异柠檬酸脱氢酶的活性。当出现这种情况时，将积累柠檬酸，因此在细胞质中形成了乙酰-CoA 的丰富的资源，并且启动了脂肪酸合成。因此，这一阶段以细胞分裂终止，随后是脂肪酸的合成和油的积累为特征。

尽管细胞通常是在大约 30℃ 下生长的，某些研究业已在较低温度下增加了不饱和脂肪酸的合成（Yongmanitchai 和 Ward, Appl. Environ. Microbiol. 57: 419-25 (1991)）。根据工艺的经济学，这种温度偏移可能在两个阶段的发酵的第一阶段之后发生，此时，业已出现了大量的微生物生长。

预计，可以采用多种发酵工艺设计，其中，使用本发明的密码子优化的基因商业化生产^①脂肪酸是理想的。例如，用重组微生物宿主商业生产 PUFAs，可以通过分批发酵，补料-分批发酵活连续发酵工艺生产。

分批发酵工艺是封闭的系统，其中，培养基组成是在工艺开始时确定的，并且不再进行超出在工艺过程中维持 pH 和氧气含量需要的进一步补充。因此，在培养过程开始时，用需要的微生物接种所述培养基，并且可以在不向培养基中添加其他来源（即，碳和氮源）条件下具有生长或代谢活性。在分批发酵工艺中，该系统的代谢物和生物量组成经常地改变，直到培养终止。在典型的分批发酵工艺中，细胞从停滞阶段发展到高速生长的对数期，并最终达到静止期，在这里，生长速度减弱或终止。在不处理的条件下，静止阶段的细胞会最终死亡。标准的分批发酵工艺的变化是补料-分批发酵工艺，其中，在发酵过程中将碳源连续地添加到发酵罐中。补料-分批发酵工艺同样适用于本发明。当代谢物抑制倾向于抑制细胞代谢或希望在任何时间在培养基中具有有限数量的来源时，补料-分批工艺是有用的。测定补料-分批系统中的碳源是困难的，因此，可以根据可测定因子的变化估算，如 pH，溶解氧和废气（例如，CO₂）的分压。分批和补料-分批培养方法是本领域技术人员所公知的和熟悉的，并且其例子可以参见以下文献：Thomas D. Brock, *Biotechnology: A Textbook of Industrial Microbiology*, 2nd ed., (1989) Sinauer Associates: Sunderland, MA; 或 Deshpande, Mukund V., *Appl. Biochem. Biotechnol.*, 36: 227 (1992)，在此收作本文参考。

利用本发明的密码子优化的基因商业化生产 ω 脂肪酸的方法还可以通过连续发酵工艺实现，其中，将确定成分的培养基连续添加到生物反应器中，同时取出等量的培养物体积，用于产物回收。连续培养通常将细胞保持在恒定细胞密度的对数生长期。连续或半连续培养方法能够调节会影响细胞生长或最终产物浓度的一种因素或多种因素中的任意一种。例如，一种方法可以限制碳源，并且允许所有其他参数适当地代谢。在其他系统中，可以连续地改变影响生长的多种因素，同时保持通过培养基浊度测定的细胞浓度恒定。连续系统试图保持稳态生长，因此，细胞生长速度必须与由于将培养基从培养物中排出而导致的细胞减少。调节连续培养工艺的营养成分和生长因子的方法，以及用于增加生产形成速度的技术为工业微生物学领域所熟知，并且上文提到的 Brock 详细披露了多种方法。

PUFAs 的纯化

PUFAs 可以作为游离脂肪酸形式出现在宿主微生物中，或者以诸如酰基甘油，磷脂，硫脂或醯脂的脂化形式出现，并且可以通过本领域所熟知的多种方法从宿主细胞中提取。酵母脂类的提取技术，品质分析和可接受性标准的一种概述是以下文献：Z. Jacobs (Critical Reviews in Biotechnology 12 (5/6) : 463-491 (1992))。有关下游加工的简单概述还可以参见 A. Singh 和 O. Ward (Adv. Appl. Microbiol. 45: 271-312 (1997))。

一般，纯化 PUFAs 的方法可以包括用有机溶剂萃取，超声波处理，超临界流体萃取（例如，使用二氧化碳），皂化和物理方法，如压力，或它们的组合。特别感兴趣的是用甲醇和氯仿在存在水的条件下萃取（E.G., Bligh & W. J. Dyer, Can. J. Biochem. Physiol. 37: 911-917 (1959)）。如果需要，可以将含水层酸化，以便使带负电荷的部分质子化，并因此增加了想得到的产物向有机层中的分配。在萃取之后，可以通过在氮气流下蒸发除去有机溶剂。当以缀合形式分离时，可以对所述产物进行酶促或化学裂解，以便释放游离的脂肪酸或感兴趣的更低复杂性的缀合物，并且然后可以进行进一步的操作，以便生产需要的最终产物。理想的是，用氢氧化钾裂解脂肪酸的缀合形式。

如果需要进一步纯化，可以采用标准方法。所述方法可以包括萃取，用尿素处理，分级结晶化，HPLC，分级蒸馏，硅胶层析，高速离心或蒸馏，或这些技术的组合。对诸如酸或烯基的活性基团的保护，可以在任何步骤通过已知技术（例如，烷基化，离子化）完成。所使用的方法包括脂肪酸甲基化，以便产生甲酯。类似地，可以在任何步骤除去保护基团。理想的是，纯化含有 GLA, STA, ARA, DHA 和 EPA 的级份可以通过用尿素处理和/或分级蒸馏实现。

优选实施方案的说明

本发明证明了将 ω -3 和/或 ω -6 生物合成途径导入含油酵母以生产 PUFAs 的可行性。为此，选择 ARA (代表 ω -6 脂肪酸) 和 EPA (代表 ω -3 脂肪酸) 作为在含油酵母-解脂耶氏酵母中生产的需要的产物。因此，ARA 的合成需要将编码 Δ 6 去饱和酶、延伸酶和 Δ 5 去饱和酶活性的基因导入耶氏酵母，而 EPA 的合成需要将编码 Δ 6 去饱和酶、延伸酶、 Δ 5 去饱和酶和 Δ 17 去饱和酶活性的基因导入耶氏酵母。

在解脂耶氏酵母中表达来自多种生物的多重公共可获得的，具有将 DGAL 转化为 ARA 和将 ETA 转化为 EPA 的能力的 $\Delta 5$ 去饱和酶，并且筛选其活性，以便鉴定在替代的宿主中显示最高水平活性的基因。在此基础上，选择高山被孢霉 $\Delta 5$ 去饱和酶 (SEQ ID NO : 4) 作为在含油酵母中表达的优选基因，这是基于其将底物补料试验中大约 30% 细胞内 DGLA 转化为 ARA 的能力而选择的。

进行了其它底物补料试验，以证明由以下基因编码的酶活性：

高山被孢霉 $\Delta 6$ 去饱和酶 (SEQ ID NO : 2) 将 LA 转化为 GLA，并且将 ALA 转化为 STA (其中解脂耶氏酵母中 LA 至 GLA 的底物转化百分比为约 30%)；

异丝水霉 $\Delta 17$ 去饱和酶 (SEQ ID NO : 6) 将 DGLA 转化为 ETA，并且将 ARA 转化为 EPA (其中解脂耶氏酵母中 ARA 至 EPA 的底物转化百分比为约 23%)；

高山被孢霉高亲和力 PUFA 延伸酶 (SEQ ID NO : 8) 将 GLA 转化为 DGLA，将 STA 转化为 ETA，并且将 EPA 转化为 DPA (其中解脂耶氏酵母中 GLA 至 DGLA 的底物转化百分比为约 30%)。

根据异丝水霉 $\Delta 17$ 去饱和酶更低的底物转化百分比 (相对于 $\Delta 5$ 和 $\Delta 6$ 去饱和酶和延伸酶)，对该特定基因进行密码子优化以便增强其在耶氏酵母属中的表达。这是通过确定解脂耶氏酵母中的密码子选择和结构基因标志，指定密码子优化的 $\Delta 17$ 去饱和酶基因，然后体外合成该基因，并且使其在替代的宿主中的效率增加 (相对于野生型基因) 而实现的。

为了合成 ARA 或 EPA (由此证明含油宿主经过工程改造以生产 ω -6 和 ω -3 脂肪酸 (即 ARA 和 EPA) 的能力)，随后制备了两种不同的 DNA 表达构建体：1) 第一种含 $\Delta 6$ 去饱和酶、 $\Delta 5$ 去饱和酶和高亲和力 PUFA 延伸酶；和 2) 第二种构建体含 $\Delta 6$ 去饱和酶、 $\Delta 5$ 去饱和酶、高亲和力 PUFA 延伸酶和密码子优化的 $\Delta 17$ 去饱和酶。将两种构建体分别转化到解脂耶氏酵母中，并且整合到编码乳清苷-5'-磷酸脱羧酶的染色体 URA3 基因中 (EC 4.1.1.23)。用合适的底物补料的宿主细胞的 GC 分析检测到了 ARA (实施例 5) 和 EPA (实施例 6) 的产生。因此，这是第一次证实了含油酵母中的 PUFA 生物合成，从而将 ω -3 和/或 ω -6 生物合成途径导入了含油酵母中。

根据此处描述的教导和结果，预计本领域技术人员将认识到通过采用含油酵母作为合成多种 ω -3和/或 ω -6 PUFA的生产平台而建立的可行性和商业用途。

实施例

本发明还确定了以下实施例。应当理解的是，在这些实施方案中，尽管指明了是本发明的优选实施方案，但是，是以说明形式提供的。通过上述讨论和这些实施例，本领域技术人员可以确定本发明的实质性特征，并且，在不超出本发明的构思和范围的前提下，可以对本发明进行各种改变和改进，以便适应各种用途和条件。

一般方法

用于实施例中的标准重组 DNA 和分子克隆技术为本领域所熟知的，并且披露于以下文献中：1.) Sambrook, J., Fritsch, E. F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989) (Maniatis); 2.) T. J. Silhavy, M. L. Bannan, and L. W. Enquist, *Experiments with Gene Fusions*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1984); 和 3.) Ausubel, F. M.等, *Current Protocols in Molecular Biology*, published by Greene Publishing Assoc. and Wiley-interscience (1987)。

适合微生物培养物维持和生长的材料和方法为本领域所公知。适合在以下实施例中使用的技术可以参见以下文献：*Manual of Methods for General Bacteriology* (Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Philips, Eds), American Society for Microbiology: Washington, D. C. (1994); 或参见 Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology*, 2nd ed., Sinauer Associates: Sunderland, MA (1989)。用于生长和保持微生物细胞的所有试剂，限制酶和材料都是从 Aldrich Chemicals (Milwaukee, WI), DIFCO Laboratories (Detroit, MI), GIBCO/BRL (Gaithersburg, MD), 或 Sigma Chemical Company (St. Louis, MO) 获得的，除非另有说明。

大肠杆菌 (XL1-Blue) 感受态细胞是从 Stratagene 公司 (San

Diego, CA) 购买的。大肠杆菌菌株通常是在 37℃ 下在 Luria Bertani (LB) 平板上生长的。

按照标准方法 (Sambrook 等, 同上) 进行一般分子克隆。寡核苷酸是由 Sigma-Genosys (Spring, TX) 合成的。用 Stratagene's QuickChange™ 定点诱变试剂盒进行定点诱变, 按照生产商的说明进行。当聚合酶链反应 (PCR) 或定点诱变与亚克隆相关时, 对所述构建体进行测序, 以便证实没有将误差导入所述序列。将 PCR 产物克隆到 Promega's pGEM-T-easy 载体 (Madison, WI) 上。

DNA 序列是利用染料终止技术 (U.S. 5,366,860; EP 272,007) 和载体与插入片段特异性引物的组合在 ABI 自动测序仪上生成的。序列编辑在 Sequencher (Gene Codes Corporation, Ann Arbor, MI) 中进行。所有序列在两个方向上均表现了至少两次的覆盖。遗传序列的比较是利用 DNASTAR 软件 (DNA Star, Inc.) 实现的。或者, 遗传序列的操作是使用可以从 Genetics Computer Group Inc. (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI) 获得的程序组进行的。使用了所述 GCG 程序 "Pileup", 空位形成预设值为 12, 空位延伸预设值为 4。使用了 GCG "Gap" 或 "Bestfit" 程序, 预设的空位形成罚分为 50, 预设的空位延伸罚分为 3。除非另有说明, 在所有其他场合下, 都使用 GCG 程序的预设参数。

缩写的含义如下: "sec" 表示秒, "min" 表示分钟, "h" 表示小时, "d" 表示天, "μL" 表示微升, "mL" 表示毫升, "L" 表示升, "μM" 表示微摩尔, "mM" 表示毫摩尔, "M" 表示摩尔, "mmol" 毫摩尔, "μmole" 表示微摩尔, "g" 表示克, "μg" 表示微克, "ng" 表示纳克, "U" 表示单位, "bp" 表示碱基对, 而 "kB" 表示千碱基。

解脂耶氏酵母的培养

解脂耶氏酵母菌株 ATCC#76982 和 ATCC#90812 购自美国典型培养物保藏中心 (Rockville, MD)。解脂耶氏酵母菌株通常在 28℃ 条件下生长在 YPD 琼脂 (1% 酵母提取物、2% 细菌用蛋白胨、2% 葡萄糖、2% 琼脂) 上。为选择转化体, 采用的是基本培养基 (不含硫酸铵或氨基酸的 0.17% 酵母氮源 (NIFCO Laboratories, Detroit, MI)、2% 葡萄糖、0.1% 脯氨酸、pH6.1)。作为补充的腺嘌呤、亮氨酸、赖氨酸和/或尿嘧啶是被适当地加入直至终浓度 0.01%。

解脂耶氏酵母的脂肪酸分析

为了分析脂肪酸,先通过离心收集细胞,并如 Bligh, E. G. & Dyer, W. J. (Can. J. Biochem. Physiol. 37:911-917 (1959)) 所述方法提取脂质。脂肪酸甲酯是通过下述方法制备的,即用甲氧基钠使脂质提取物转酯 (Roughan, G.和 Nishida I. Arch Biochem Biophys. 276 (1):38-46 (1990)), 并接着用装配有 30-m X 0.25mm (内径) HP-INNOWAX (Hewlett-Packard) 柱的 Hewlett-Packard 6890 GC 对其进行分析。烤箱温度以 3.5℃/分钟的速度从 170℃ (保持 25 分钟) 上升至 185℃。

为了直接进行碱基转酯,先收获耶氏酵母培养物 (3mL), 在蒸馏水中洗涤一次,并在 Speed-Vac 内的真空条件下干燥 5-10 分钟。向样品中加入甲氧基钠 (100µl, 1%), 接着涡旋并振摇样品 20 分钟。加入 3 滴 1M NaCl 和 400µl 己烷后, 涡旋并离心样品。移出上层并如上所述通过 GC 对其进行分析。

实施例 1

构建适合在解脂耶氏酵母中表达异源基因的质粒

构建了质粒 pY5, 即 pINA532 (由 Dr. Claude Gaillardin 馈赠, Institut National Agronomique, Centre de biotechnologie Agro-Industrielle, laboratoire de Genetique Moleculaire et Cellularie INRA-CNRS, F-78850 Thiverval-Grignon, France) 衍生物, 用于在解脂耶氏酵母中表达异源基因, 如图 2 所示。

首先, 将包括 pINA532 的 ARS18 序列和 LEU2 基因的部分消化的 3598 bp EcoRI 片段亚克隆到 pBluescript (Stratagene, San Diego, CA) 的 EcoRI 位点上, 以便制备 pY2。通过 PCR 扩增来自解脂耶氏酵母基因组 DNA 的 TEF 启动子 (Muller S., 等, Yeast, 14: 1267-1283 (1998)), 用 TEF5' (SEQ ID NO: 38) 和 TEF3' (SEQ ID NO: 39) 作引物。PCR 扩增是在 50µl 的总体积中进行的, 它包括: 100 ng 耶氏酵母属基因组 DNA, 含有 10 mM KCl 的 PCR 缓冲液, 10 mM (NH₄)₂SO₄, 20 mM Tris-HCl (pH 8.75), 2 mM MgSO₄, 0.1% Triton X-100, 100 µg/mL BSA (最终浓度), 200µM 每一种脱氧核糖核苷三磷酸, 10 皮摩尔每一种引物和 1µl 的 PfuTurbo DNA 聚合酶 (Stratagene, San Diego, CA)。扩增是按以下方法进行的: 首先在 95℃ 下变性 3 分钟, 然后进行 35 轮

以下循环：95℃1分钟，56℃30秒，72℃1分钟。在72℃下进行了10分钟的最终延伸，然后在4℃下终止反应。将418的PCR产物连接到pCR-Blunt上，以便制备pIP-tef。将pIP-tef的BamHI/EcoRV片段亚克隆到pY2的BamHI/SmaI位点上，以便制备pY4。

使用pINA532作模板，用XPR5'(SEQ ID NO: 40)和XPR3'(SEQ ID NO: 41)作引物，通过PCR扩增XPR2转录终止子。PCR扩增是在50μl的总体积中进行的，使用上述成分和条件。用SacII消化179bp的PCR产物，然后连接到pY4的SacII位点上，以便制备pY5。因此，pY5(参见图3和4)可用作耶氏酵母属-大肠杆菌穿梭质粒，它包括：

- 1.) 耶氏酵母属自主复制序列(ARS18)；
- 2.) ColE1质粒复制起点；
- 3.) 氨苄青霉素-抗性基因(Amp^R)，用于在大肠杆菌中选择；
- 4.) 耶氏酵母属LEU2基因(E.C.4.2.1.33,编码异丙基苹果酸异构酶，用于在耶氏酵母属中选择)；
- 5.) 翻译延伸启动子(TEFP)，用于在耶氏酵母属中表达异源基因；和
- 6.) 细胞外蛋白酶基因终止子(XPR2)，用于在耶氏酵母属中表达的异源基因的转录终止。

作为pY5的衍生物构建了pY5-13(图4)，以便促进在解脂耶氏酵母中的亚克隆和异源基因表达。用pY5作模板，通过6轮定点诱变构建pY5-13。通过用寡核苷酸YL5和YL6(SEQ ID NOs: 106和107)进行定点诱变将SaII和ClaI位点从pY5上消除，以便制备pY5-5。通过用寡核苷酸YL9和YL10(SEQ ID NOs: 110和111)进行定点诱变将SaII位点导入pY5-5的LEU2基因和TEF启动子之间，以便制备pY5-6。用寡核苷酸YL7和YL8(SEQ ID NOs: 108和109)将PacI位点导入pY5-6的LEU2基因和ARS18之间，以便制备pY5-8。使用寡核苷酸YL3和YL4(SEQ ID NOs: 104和105)将NcoI位点导入pY5-8的TEF启动子的翻译起始密码子周围，以便制备pY5-9。使用YL1和YL2寡核苷酸(SEQ ID NOs: 102和103)将pY5-9的LEU2基因的NcoI位点消除，以便制备pY5-12。最后，用寡核苷酸YL61和YL62(SEQ ID NOs: 88和89)将BsiWI位点导入pY5-12的ColE1和XPR2区之间，以便制备pY5-13。

为了促进亚克隆, 构建了质粒 pY5 的第二种衍生物。具体地讲, 用 pY5 作模板, 通过三轮定点诱变构建了 pY5-4 (图 4)。用寡核苷酸 YL1 和 YL2 (SEQ ID NOs: 102 和 103) 将位于 Leu2 报道基因上的 NcoI 位点从 pY5 除去, 以便制备 pY5-1。通过用寡核苷酸 YL3 和 YL4 (SEQ ID NOs: 104 和 105) 进行定点诱变将 NcoI 位点导入 pY5-1 的 TEF 启动子和 XPR2 转录终止子之间, 以便制备 pY5-2。然后使用寡核苷酸 YL23 和 YL24 (SEQ ID NOs: 112 和 113) 将 PacI 位点导入 pY5-2 的 TEF 启动子和 XPR2 转录终止子之间, 以便制备 pY5-4。

实施例 2

对用于在解脂耶氏酵母内表达的 $\Delta 6$ 去饱和酶、 $\Delta 5$ 去饱和酶、 $\Delta 17$ 去饱和酶和高亲和力 PUFA 延伸酶基因的选择

将编码 ω -3 和/或 ω -6 生物合成途径的特异性基因导入含有已破坏的 $\Delta 12$ 去饱和酶的解脂耶氏酵母内之前, 有必要证实在耶氏酵母属中表达的异源 $\Delta 6$ 去饱和酶、延伸酶、 $\Delta 5$ 去饱和酶和 $\Delta 17$ 去饱和酶基因的功能性。这是通过测定各野生型蛋白在可选宿主内的转化效率而得以实现的。具体而言, 即分别表达高山被孢霉 $\Delta 5$ 去饱和酶、高山被孢霉 $\Delta 6$ 去饱和酶、异丝水霉 $\Delta 17$ 去饱和酶和高山被孢霉高亲和力 PUFA 延伸酶, 并在底物补料试验中进行活性筛选。根据这些结果, 选择高山被孢霉基因与 $\Delta 6$ 和 $\Delta 17$ 去饱和酶以及高亲和力 PUFA 延伸酶组合。

表达质粒的构建

通常是通过限制酶切消化分离野生型去饱和酶或延伸酶基因, 或通过 PCR 扩增, 并将其插入合适的载体中进行表达。各 PCR 扩增均在总体积为 50 μ l 并包括 PCR 缓冲液的反应混合物中进行, 所述 PCR 缓冲液含有: 10ng 模板、10mM KCl、10mM (NH₄)₂SO₄、20mM Tris-HCl (pH8.75)、2mM MgSO₄、0.1% Triton X-100、100 μ g/mL BSA (终浓度)、均为 200 μ M 的各种脱氧核苷三磷酸、均为 10pmole 的各种引物和 1 μ l Pfu Turbo DNA 聚合酶 (Stratagene, San Diego, CA)。扩增步骤如下 (除非另外指明): 95 $^{\circ}$ C 初始变性 3 分钟, 接着进行 35 个循环, 每个循环包括: 95 $^{\circ}$ C 1 分钟、56 $^{\circ}$ C 30 秒、72 $^{\circ}$ C 1 分钟。最终延伸循环于 72 $^{\circ}$ C 进行 10 分钟, 接着于 4 $^{\circ}$ C 终止反应。

野生型高山被孢霉 (登录号 #AF465281) $\Delta 6$ 去饱和酶

将包括高山被孢霉 $\Delta 6$ 去饱和酶基因 (SEQ ID NO: 1) 的 1384 bp 的 pCGR5 的 NcoI/NotI 片段 (U. S. 5,968,809) 插入 pY5-2 的 NcoI/NotI 位点 (实施例 1) 上, 以便制备 pY54.

野生型高山被孢霉 (登录号#AF067654) $\Delta 5$ 去饱和酶

通过以寡核苷酸 YL11 和 YL12 (SEQ ID NOs:72 和 73) 为引物、以质粒 pCGR-4 (U.S. 6,075,183) 为模板进行 PCR 扩增高山被孢霉 $\Delta 5$ 去饱和酶基因 (SEQ ID NO:3)。PCR 扩增步骤如上所述, 但延伸步骤延长至 1.5 分钟 (对循环 1-35 而言)。用 NcoI/NotI 消化 1357bp 的 PCR 产物, 并与 NcoI/NotI 消化的 pY5-13 (如实施例 1 所述) 连接生成 pYMA5bp (图 5)。

野生型异丝水霉 (ATCC #56851) $\Delta 5$ 去饱和酶

通过 PCR, 用寡核苷酸 YL13A 和 YL14 (SEQ ID NOs: 116 和 117) 作引物, 用质粒 pRSP3 (WO02/081668) 作模板, 扩增异丝水霉 $\Delta 5$ 去饱和酶基因 (SEQ ID NO: 114)。PCR 扩增步骤如上所述, 但延伸步骤延长至 1.5 分钟 (对循环 1-35 而言)。用 NcoI/PacI 消化 1.4kB 的 PCR 产物, 然后连接到 NcoI/PacI-消化过的 pY5-4 上 (图 4; 参见实施例 1), 以便制备 pYSD5。

野生型 *Isochrysis galbana* CCMP1323 $\Delta 5$ 去饱和酶

通过 PCR, 用寡核苷酸 YL19A 和 YL20 (SEQ ID NOs: 120 和 121) 作引物, 用质粒 pRIG-1 (WO02/081668) 作模板, 扩增 *I. galbana* $\Delta 5$ 去饱和酶基因 (SEQ ID NO: 118)。PCR 扩增步骤如上所述, 但延伸步骤延长至 1.5 分钟 (对循环 1-35 而言)。用 BamHI/PacII 消化 1.4kB 的 PCR 产物, 然后连接到 BamHI/PacII-消化过的 pY5-4 上 (图 4; 参见实施例 1), 以便制备 pYIG5。

野生型 *Thraustochytrium aureum* (ATCC # 34304) $\Delta 5$ 去饱和酶

通过 PCR, 用寡核苷酸 YL15 和 YL16B (SEQ ID NOs: 124 和 125) 作引物, 用质粒 pRTA4 (WO02/081668) 作模板, 扩增 *T.aureum* $\Delta 5$ 去饱和酶基因 (SEQ ID NO: 122)。PCR 扩增步骤如上所述, 但延伸步骤延长至 1.5 分钟 (对循环 1-35 而言)。用 NcoI/NcoI 消化 1.4kB 的 PCR 产物, 然后连接到 NcoI/NcoI-消化过的 pY5-2 上 (图 4; 参见实施例 1), 以便制备 pYTA5。

野生型异丝水霉 (ATCC #56851) $\Delta 17$ 去饱和酶

通过 PCR, 用寡核苷酸 YL21A (SEQ ID NO: 42) 和 YL22 (SEQ ID NO: 43) 作引物, 由质粒 pRSP19 (US2003/0196217A1) 扩增异丝水霉的野生型 $\Delta 17$ 去饱和酶基因。用 NcoI/PacI 消化 PCR 产物, 然后连接到 NcoI/PacI-消化过的 pY5-4 上 (图 4; 参见实施例 1), 以便制备 pYSD17。

野生型高山被孢霉 (登录号#AX464731) 高亲和力延伸酶

将包括高山被孢霉高亲和力 PUFA 延伸酶基因的编码区 (SEQ ID NO: 7) 的 pRPB2 的 973 bp 的 NotI 片段 (WO00/12720) 插入 pY5 的 NotI 位点 (实施例 1; 图 3 和 4), 以便制备 pY58。

解脂耶氏酵母的转化

按照 Chen, D. C. 等 (Appl Microbiol Biotechnol. 48 (2): 232-235- (1997)) 的方法将质粒 pY54, pYMA5pb, pYSD5, pYIG5, PYTA5, pYSD17 和 pY58 分别转化到解脂耶氏酵母 ATCC #76982 中。

简单地讲, 将耶氏酵母属的亮氨酸营养缺陷型在 YPD 平板上划线, 并且在 30°C 下生长大约 18 小时。将几个大环的细胞从平板上刮下来, 并且重新悬浮在 1 mL 的转化缓冲液中, 该缓冲液包括:

- 2.25 mL 的 50% PEG, 平均分子量 3350;
- 0.125 mL 的 2M 乙酸锂, pH 6.0;
- 0.125 mL 的 2M DTT; 和
- 50 μ g 的剪切的鲑鱼精子 DNA。

在 100 μ l 的重新悬浮的细胞中温育大约 500 ng 的质粒 DNA, 并且在 39°C 下保持 1 小时, 以 15 分钟的间隔进行涡旋混合。将细胞铺平板到缺少亮氨酸的基本培养基平板上, 并且在 30°C 下保持 2-3 天。

底物转化百分比的测定

含有 pY54, pYMA5pb, pYSD5, pYIG5, PYTA5, pYSD17 或 pY58 的转化解脂耶氏酵母的单个菌落分别在 3 mL 基本培养基 (20 g/L 葡萄糖, 1.7 g/L 酵母氮基质, 不含氨基酸, 1 g/L L-脯氨酸, 0.1 g/L L-腺嘌呤, 0.1 g/L L-赖氨酸, pH 6.1) 中, 在 30°C 下生长到 OD₆₀₀ 大约为 1.0。为了进行底物补料, 将 100 μ l 的细胞放在含有 10 μ g 底物的 3 mL 基本培养基中在 30°C 下继代培养大约 24 小时。然后通过离心收集细胞, 并提取脂类。脂肪酸甲酯是通过转酯制备的, 随后进行 GC 分析 (如一般方法中的描述)。底物转化百分比按以下公式测定: ([产物]/[底物

+ 产物]) *100)。

野生型高山被孢霉 $\Delta 6$ 去饱和酶的底物转化百分比

高山被孢霉 $\Delta 6$ 去饱和酶能将 LA 转化成 GLA 和/或将 ALA 转化成 STA。含有 pY54 的解脂耶氏酵母菌株按上述方法生长 (不需要补充底物), 并且分析脂类。结果表明, 具有 pY54 的耶氏酵母属菌株能将大约 30% LA 转化成 GLA。

高山被孢霉、异丝水霉、*I. galbana* 和 *T. aureum* $\Delta 5$ 去饱和酶的底物转化百分比

高山被孢霉、异丝水霉、*I. galbana* 和 *T. aureum* 的 $\Delta 5$ 去饱和酶都将 DGLA 转化为 ARA, 并且将 ETA 转化为 EPA。分别从单个菌落生长含 pYMA5pb, pYSD5, pYIG5 或 pYTA5 的解脂耶氏酵母菌株, 在含有 10 μ g 的 ARA 的基本培养基中继代培养, 并且按上述方法进行脂类分析。含有 pYMA5pb (高山被孢霉) 的耶氏酵母菌株将大约 30% 的细胞内 DGLA 转化为 ARA, 含有 pYSD5 (异丝水霉) 的耶氏酵母菌转化大约 12%, 含有 pYIG5 (*I. galbana*) 的耶氏酵母菌转化大约 7%, 含有 pYTA5 (*T. Aureum*) 耶氏酵母菌转化大约 23% 的细胞内 DGLA 转化为 ARA。

野生型异丝水霉 $\Delta 17$ 去饱和酶的底物转化百分比

异丝水霉 $\Delta 17$ 去饱和酶能将 ARA 转化成 EPA 和/或将 DGLA 转化成 ETA。含有 pYSD17 的解脂耶氏酵母菌株是从单菌落生长的, 在含有 10 μ g 的 ARA 的基本培养基中继代培养, 并且按上述方法进行脂类分析。ARA 补料实验的结果表明具有 pYSD17 的耶氏酵母属菌株能将大约 23% 的细胞内 ARA 转化成 EPA。

野生型高山被孢霉高亲和力延伸酶的底物转化百分比

高山被孢霉高亲和力 PUFA 延伸酶能将 GLA 转化成 DGLA, 将 STA 转化成 ETA 和/或将 EPA 转化成 DPA。包括 pY58 的解脂耶氏酵母菌株是从单菌落生长的, 在含有 10 μ g 的 GLA 基本培养基中继代培养, 并且按上述方法进行脂类分析。GLA 补料实验的结果表明, 具有 pY58 的耶氏酵母属菌株能将大约 30% 的细胞内 GLA 转化成 DGLA。

实施例 3

密码子优化 $\Delta 17$ 去饱和酶基因在解脂耶氏酵母内的合成和表达

基于实施例 2 的结果, 可利用的基因是编码 $\Delta 6$ 去饱和酶、延伸酶

和 $\Delta 5$ 去饱和酶活性的基因，因为这些基因在解脂耶氏酵母内均能够实现约 30% 的底物转化百分率。但异丝水霉来源的 $\Delta 17$ 去饱和酶的最大转化效率仅为 23%。因此，根据耶氏酵母属的密码子选择模式、ATG 翻译起始密码子周围的共有序列和 RNA 稳定性的普遍规律，并基于异丝水霉 DNA 序列 (SEQ ID NO:5) 设计了密码子优化 $\Delta 17$ 去饱和酶基因 (Guhaniyogi, G. 和 J. Brewer, Gene 265 (1-2):11-23 (2001))。

除了对翻译起始位点的修饰外，还对 1077bp 编码区中的 127bp 片段 (包括了 117 个密码子) 进行了密码子优化。该密码子优化 DNA 序列 (SEQ ID NO:9) 与异丝水霉 $\Delta 17$ 去饱和酶基因 DNA 序列 (SEQ ID NO:5) 之间的比较如图 6 所示，其中粗体所示核苷酸对应于密码子优化基因中被改动的核苷酸。该密码子优化基因中的所有变更均未改变编码蛋白的氨基酸序列 (SEQ ID NO:6)。

在解脂耶氏酵母中测定优选的密码子选择

在 National Center for Biotechnology Information 公共数据库中发现了大约 100 个解脂耶氏酵母的基因。通过 DNASTar 的 Editseq 程序将包括 121,167 bp 的这些基因的编码区翻译成相应的 40,389 个氨基酸，并且制成表格，以便确定由表 4 所示的解脂耶氏酵母密码子选择谱。标题为 "No." 的栏表示编码特定氨基酸的特定密码子在 40,389 个氨基酸的样品中出现的次数。标题为 "%" 的栏表示编码特定氨基酸的特定氨基酸的频率。用粗体字符表示的代码表示解脂耶氏酵母中优选的密码子。

表 3
解脂耶氏酵母的密码子选择

密码子	氨基酸	No.	%
GCA	Ala (A)	359	11.4
GCC	Ala (A)	1523	48.1
GCG	Ala (A)	256	8.1
GCU	Ala (A)	1023	32.3
AGA	Arg (R)	263	13.2
AGG	Arg (R)	91	4.6
CGA	Arg (R)	1133	56.8
CGC	Arg (R)	108	5.4
CGG	Arg (R)	209	1.0
CGU	Arg (R)	189	9.5
AAC	Ans (N)	1336	84.0
AAU	Ans (N)	255	16.0
GAC	Asp (D)	1602	66.8
GAU	Asp (D)	795	33.2
UGC	Cys (C)	268	53.2
UGU	Cys (C)	236	46.8
CAA	Gln (Q)	307	17.0
CAG	Gln (Q)	1490	83.0
GAA	Glu (E)	566	23.0
GAG	Glu (E)	1893	77.0
GGA	Gly (G)	856	29.7
GGC	Gly (G)	986	34.2
GGG	Gly (G)	148	5.1
GGU	Gly (G)	893	31.0
CAC	His (H)	618	65.5
CAU	His (H)	326	34.5
AUA	Ile (I)	42	2.1
AUC	Ile (I)	1106	53.7
AUU	Ile (I)	910	44.2
CUA	Leu (L)	166	4.7
CUC	Leu (L)	1029	29.1
CUG	Leu (L)	1379	38.9
CUU	Leu (L)	591	16.7
UUA	Leu (L)	54	1.5
UUG	Leu (L)	323	9.1

密码子	氨基酸	No.	%
AAA	Lys (K)	344	14.8
AAG	Lys (K)	1987	85.2
AUG	Met (M)	1002	100
UUC	Phe (F)	996	61.1
UUU	Phe (F)	621	38.9
CCA	Pro (P)	207	9.6
CCC	Pro (P)	1125	52.0
CCG	Pro (P)	176	8.2
CCU	Pro (P)	655	30.2
AGC	Ser (S)	335	11.3
AGU	Ser (S)	201	6.8
UCA	Ser (S)	221	7.5
UCC	Ser (S)	930	31.5
UCG	Ser (S)	488	16.5
UCU	Ser (S)	779	26.4
UAA	Term	38	46.9
UAG	Term	30	37.0
UGA	Term	13	16.1
ACA	Thr (T)	306	12.7
ACC	Thr (T)	1245	51.6
ACG	Thr (T)	269	11.1
ACU	Thr (T)	595	24.6
UGG	Trp (W)	488	100
UAC	Tyr (Y)	988	83.2
UAU	Tyr (Y)	200	16.8
GUA	Val (V)	118	4.2
GUC	Val (V)	1052	37.3
GUG	Val (V)	948	33.6
GUU	Val (V)	703	24.9

为了进一步优化在解脂耶氏酵母中的基因表达，检查了 79 个基因的 'ATG' 起始密码子周围的共有序列。在图 7 中，加下划线的 ATG 翻译起始密码子的第一个 'A' 被设为 +1。分析过的基因的 77% 在 -3 号位置上出现 "A"，表明了 "A" 在该位置上的强烈偏好。还存在 'A' 或 'C' 在 -4, -2 和 -1 号位置上的偏好，'A', 'C' 或 'T' 在 +5 号位置上的偏好，以及 'G' 或 'C' 在 +6 号位置上的偏好。因此，用于基因在解脂耶氏酵母中优化表达的密码子优化的翻译起始位点的优选的共有序列是 'MAMMATGNHS' (SEQ ID NO: 126)，其中使用了如下的核酸简并性密码：M=A/C；S=C/G；H=A/C/T；和 N=A/C/G/T。

密码子优化基因的体外合成

被用于合成密码子优化 $\Delta 17$ 去饱和酶基因的方法如图 8 所示。首先，设计 11 对寡核苷酸，以延伸异丝水霉 $\Delta 17$ 去饱和酶基因的密码子优化编码区的全长（例如 D17-1A、D17-1B、D17-2A、D17-2B、D17-3A、D17-3B、D17-4A、D17-4B、D17-5A、D17-5B、D17-6A、D17-6B、D17-7A、D17-7B、D17-8A、D17-8B、D17-9A、D17-9B、D17-10A、D17-10B、D17-11A 和 D17-11B，对应于 SEQ ID NOs:10-31）。每一对有义 (A) 和反义 (B) 寡核苷酸除 5'-末端具有 4bp 突出端外，其它部分均互补。此外，引物 D17-1A、D17-4B、D17-5A、D17-8A 和 D17-8B 还分别导入了用于后续亚克隆的 NcoI、BglII 和 Sall 限制位点。

于 37℃、体积为 20 μ l 并含有 50mM Tris-HCl (pH7.5)、10mM MgCl₂、10mM DTT、0.5mM 亚精胺、0.5mM ATP 和 10U T4 多核苷酸激酶的混合液中，使均为 100ng 的各种寡核苷酸磷酸化。将各有义和反义寡核苷酸对混合，并在采用下列参数的热循环仪中退火：95℃ (2 分钟)、85℃ (2 分钟)、65℃ (15 分钟)、37℃ (15 分钟)、24℃ (15 分钟) 和 4℃ (15 分钟)。因此，D17-1A (SEQ ID NO:10) 退火至 D17-1B (SEQ ID NO:11) 可生成双链产物 "D17-1AB"。同样，D17-2A (SEQ ID NO:12) 退火至 D17-2B (SEQ ID NO:13) 可生成双链产物 "D17-2AB"，依次类推。

接着将退火双链寡核苷酸分别混合所获得的三个集合连接在一起，如下所示：

- 集合 1: 包括 D17-1AB、D17-2AB、D17-3AB 和 D17-4AB;
- 集合 2: 包括 D17-5AB、D17-6AB、D17-7AB 和 D17-8AB; 和

●集合3: 包括 D17-9AB、D17-10AB 和 D17-11AB。

将各退火寡核苷酸集合混合在体积为 20 μ l 并含有 10U T4 DNA 连接酶的混合物中, 并于 16 $^{\circ}$ C 温育过夜。

接着通过 PCR 扩增各连接反应的产物。具体而言, 是通过以连接的“集合1”混合物(即 D17-1AB、D17-2AB、D17-3AB 和 D17-4AB)为模板, 以寡核苷酸 D17-1 (SEQ ID NO:32) 和 D17-4R (SEQ ID NO:33) 为引物进行 PCR, 扩增密码子优化 Δ 17 去饱和酶基因的第一个部分。该 PCR 扩增是在总体积为 50 μ l 并包括 PCR 缓冲液的混合物中进行, 其中的 PCR 缓冲液含有 10mM KCl、10mM(NH₄)₂SO₄、20mM Tris-HCl (pH 8.75)、2mM MgSO₄、0.1% Triton X-100、100 μ g/mL BSA (终浓度)、均为 200 μ M 的各种脱氧核苷三磷酸、均为 10 pmole 的各种引物和 1 μ l Pfu Turbo DNA 聚合酶 (Stratagene, San Diego, CA)。扩增步骤如下: 于 95 $^{\circ}$ C 初始变性 3 分钟, 接着进行 35 个循环, 每个循环包括: 95 $^{\circ}$ C 1 分钟、56 $^{\circ}$ C 30 秒、72 $^{\circ}$ C 40 秒。最终延伸循环于 72 $^{\circ}$ C 进行 10 分钟, 接着于 4 $^{\circ}$ C 终止反应。将 430bp PCR 片段亚克隆进 pGEM-T easy 载体 (Promega) 中, 可生成 pT17 (1-4)。

同样, 通过以连接的“集合2”混合物(即 D17-5AB、D17-6AB、D17-7AB 和 D17-8AB)为模板, 以寡核苷酸 D17-5 (SEQ ID NO:34) 和 D17-8D (SEQ ID NO:35) 为引物进行 PCR, 扩增密码子优化 Δ 17 去饱和酶基因的第二个部分, 并将其克隆进 pGEM-T-easy 载体中, 可生成 pT17(5-8)。最后, 同样通过以“集合3”连接混合物(即 D17-9AB、D17-10AB 和 D17-11AB)为模板, 以寡核苷酸 D17-8U (SEQ ID NO:36) 和 D17-11 (SEQ ID NO:37) 为引物进行 PCR, 扩增密码子优化 Δ 17 去饱和酶基因的第三个部分, 并将其克隆进 pGEM-T-easy 载体中, 可生成 pT17 (9-11)。

分别用 pT17 (1-4)、pT17 (5-8) 和 pT17 (9-11) 转化大肠杆菌, 并从氨苄青霉素抗性转化体中分离质粒 DNA。纯化质粒 DNA, 并用合适的限制性内切核酸酶消化, 以释放出 pT17 (1-4) 的 420bp NcoI/BglII 片段、pT17 (5-8) 的 400bp BglII/SalI 片段和 pT17 (9-11) 的 300bp SalI/NotI 片段。接着将这些片段组合、连接在一起作为模板, 并以 D17-1 (SEQ ID NO:32) 和 D17-11 (SEQ ID NO:37) 为引物扩增完整的合成密码子优化 Δ 17 去饱和酶基因。对 Δ 17 去饱和酶基因的每一个部分均

采用上述条件，在总体积为 50 μ l 的混合物中进行 PCR 扩增，热循环程序如下：于 95 $^{\circ}$ C 初始变性 3 分钟，接着进行 35 个循环，每个循环包括：95 $^{\circ}$ C 1 分钟、56 $^{\circ}$ C 30 秒、72 $^{\circ}$ C 1.1 分钟。最终延伸循环于 72 $^{\circ}$ C 进行 10 分钟，接着于 4 $^{\circ}$ C 终止反应。生成 1.1kB PCR 产物。

含密码子优化的 Δ 17 去饱和酶的质粒 pYSD17S 的构建

用 NcoI/NotI 消化包括完整合成 Δ 17 去饱和酶的上述 1.1kB PCR 产物，并将其亚克隆进 NcoI/NotI 消化的 pY5-13（实施例 1）中，可生成 pYSD17S（图 9A）。

为比较野生型和合成基因在耶氏酵母属内的效率，另一个“对照”是通过以 YL53（SEQ ID NO:44）和 YL54（SEQ ID NO:45）为引物进行定点诱变，消除 pYSD17（包括所述野生型基因；如实施例 2 所述）中富含 AT 的 PaeI 位点，可生成 pYSD17M（图 9B）。

用密码子优化的 Δ 17 去饱和酶基因对解脂耶氏酵母的转化

根据实施例 2 所述方法，分别将含有野生型和密码子优化 Δ 17 去饱和酶的质粒转化进解脂耶氏酵母 ATCC #76982 中。利用该技术，获得含有下述质粒的转化体：

表 4

转化体耶氏酵母中的质粒一览表

质粒	描述
pYSD17	野生型 Δ 17 去饱和酶
pYSD17M	野生型 Δ 17 去饱和酶，无富含 AT 的 PaeI 位点
pYSD17S	密码子优化 Δ 17 去饱和酶

密码子优化的 Δ 17 去饱和酶基因的底物转化百分率

Δ 17 去饱和酶可将 ARA 转化为 EPA（参见图 2）。野生型和密码子优化 Δ 17 去饱和酶基因的底物转化百分率（ $[\text{产物}]/[\text{底物} + \text{产物}] * 100$ ）是通过利用一般方法中所述方法，在含有各可选质粒构建体的解脂耶氏酵母中测定的。

ARA 补料试验结果显示具有对照质粒 pYSD17 或 pYSD17M 的耶氏酵母菌株可将大约 23% 的胞内 ARA 转化为 EPA（图 10A），而含有

pYSD17S 中的密码子优化 $\Delta 17$ 去饱和酶基因的菌株则可将大约 45% 的胞内 ARA 转化为 EPA (图 10B)。因此, 含有密码子优化 $\Delta 17$ 去饱和酶的耶氏酵母可转化比含有野生型异丝水霉基因的菌株多大约 2 倍的 ARA。

适合解脂耶氏酵母内多种 ω 脂肪酸生物合成基因同步表达的质粒的构建

本实施例描述了需要合成多种表达质粒, 以便构建: 1) 适合于被整合进解脂耶氏酵母属基因组内以表达 $\Delta 6$ 去饱和酶、PUFA 延伸酶、和 $\Delta 5$ 去饱和酶 (用于生产 ARA) 的 DNA 片段; 和 2) 适合于被整合进解脂耶氏酵母属基因组内以表达 $\Delta 6$ 去饱和酶、PUFA 延伸酶、 $\Delta 5$ 去饱和酶和 $\Delta 17$ 去饱和酶 (用于生产 EPA) 的 DNA 片段。

质粒 pY24 的构建

质粒 pY24 (图 11) 是用于构建适合被整合进解脂耶氏酵母基因组内的表达盒的亲本载体。pY24 的构建方法如下:

以寡核苷酸 KU5 和 KU3 (SEQ ID NOs:46 和 47) 为引物, 以耶氏酵母基因组 DNA 为模板, 通过 PCR 扩增含有耶氏酵母 URA3 基因的 1.7kB DNA 片段 (SEQ ID NO:48)。该 PCR 扩增是在总体积为 50 μ l 的混合物中进行, 该混合物包括: 100ng 耶氏酵母基因组 DNA 和含有 10mM KCl、10mM (NH₄)₂SO₄、20mM Tris-HCl (pH8.75)、2mM MgSO₄、0.1% Triton X-100、100 μ g/mL BSA (终浓度)、均为 200 μ M 的各种脱氧核苷三磷酸、均为 10 pmole 的各种引物和 1 μ l Pfu Turbo DNA 聚合酶 (Stratagene, San Diego, CA) 的 PCR 缓冲液。扩增步骤如下: 于 95 $^{\circ}$ C 初始变性 3 分钟, 接着进行 35 个循环, 每个循环包括: 95 $^{\circ}$ C 1 分钟、56 $^{\circ}$ C 30 秒、72 $^{\circ}$ C 2 分钟。最终延伸循环于 72 $^{\circ}$ C 进行 10 分钟, 接着于 4 $^{\circ}$ C 终止反应。将该 PCR 产物插入 pGEM-T easy 载体 (Promega, Madison, WI) 中, 可生成 pGYUM。

利用寡核苷酸 KI5 和 KI3 (SEQ ID NOs:50 和 51) 进行 PCR, 扩增含有 *Impatiens balsama* 的接合酶基因 (或 “imp H8”) (clone ids. pk0001.h8; E. I. du Pont de Nemours and Company, Inc., Wilmington, DE) 的 1.1kB DNA 片段 (SEQ ID NO:52)。该 PCR 扩增是利用上述组成进行, 但所用模板是 ids.pk0001.h8 的 10ng 质粒 DNA。扩增步骤如下: 于 95 $^{\circ}$ C 初始变性 3 分钟, 接着进行 35 个循环, 每个循环包括:

95℃1.5分钟、56℃30秒、72℃1.2分钟。最终延伸循环于72℃进行10分钟，接着于4℃终止反应。用NotI消化该PCR产物，并接着将其插入pY5（图3）的NotI位点，可生成pY9。

利用寡核苷酸KTI5和KTI3（SEQ ID NOs:54和55）进行PCR，扩增含有pY9的TEF::IMP H8::XPR嵌合基因的1.7kB DNA片段（SEQ ID NO:56）。如上所述进行该PCR扩增，但所用模板是pGYUM的10ng质粒DNA。扩增步骤如下：于95℃初始变性3分钟，接着进行35个循环，每个循环包括：95℃1分钟、56℃30秒、72℃2分钟。最终延伸循环于72℃进行10分钟，接着于4℃终止反应。将该PCR产物插入PCR-Script（Stratagene）中，可生成pY9R。用pGYUM的XhoI/EcoRV片段置换pY9R的1.7kB XhoI/EcoRV片段，可生成pY21。

以寡核苷酸KH5和KH3（SEQ ID NOs:58和59）为引物，以KS65的基因组DNA为模板进行PCR，扩增含有大肠杆菌潮霉素抗性基因（“HPT”；Kaster, K. R., et al., *Nucleic Acids Res.* 11:6895-6911（1983））的1kB DNA片段（SEQ ID NO:60）。该PCR扩增是利用上述组分在总体积为50μl的混合物中进行，但所用模板是ids.pk0001.h8的10ng质粒DNA。扩增步骤如下：于95℃初始变性3分钟，接着进行35个循环，每个循环包括：95℃1分钟、56℃30秒、72℃1.2分钟。最终延伸循环于72℃进行10分钟，接着于4℃终止反应。用NotI消化该PCR产物，并接着将其插入pY5的NotI位点（图3），可生成pTHPT-1。

以寡核苷酸KTH5和KTH3（SEQ ID NOs:62和63）为引物，以pTHPT-1质粒DNA为模板，如上所述扩增含有TEF::HPT::XPR融合基因的1.6kB DNA片段（SEQ ID NO:64）。用BglII消化该PCR产物，并接着将其插入pY21中，可生成pY24。

pY24-4的构建

质粒pY24（图11）被用于构建适合被整合进解脂耶氏酵母基因组的表达盒。利用pY24质粒中解脂耶氏酵母URA3基因的401bp 5'序列（SEQ ID NO:66）和568bp 3'-序列（SEQ ID NO:67）将表达盒直接整合进耶氏酵母基因组的Ura位点。通过BamHI消化，首先从pY24中移出两个嵌合基因（TEF::HPT::XPR和TEF::IMP H8::XPR），使它们自连接生成pY24-1。通过分别利用YL63和YL64（SEQ ID NOs:68

和 69) 和 YL65 和 YL66 (SEQ ID NOs:70 和 71) 引物对进行定点诱变, 将 Pacl 和 BsiWI 位点导入 pY24-1 中, 可生成 pY24-4.

用于 $\Delta 5$ 去饱和酶表达的整合载体的构建

将 pYMA5pb 的 4261bp Pacl/BsiWI 片段 (包括高山被孢霉 $\Delta 5$ 去饱和酶基因; 如实施例 2 所述) 连接进 pY24-4 (图 11) 的 Pacl/BsiWI 位点, 可生成 pYZM5 (图 5). 通过分别利用引物对 YL81 和 YL82 (SEQ ID NOs:74 和 75) 和 YL83 和 YL84 (SEQ ID NOs:76 和 77) 进行定点诱变, 将 HindIII 和 ClaI 位点导入 pYZM5, 可生成 pYZM5CH. 通过以 YL105 和 YL106 (SEQ ID NO:78 和 79) 为引物进行定点诱变, 将 PmeI 位点导入 pYZM5CH, 可生成 pYZM5CHPP. 通过以 YL119 和 YL120 (SEQ ID NO:80 和 81) 为引物进行定点诱变, 将 AscI 位点导入 pYZM5CHPP, 可生成 pYZM5CHPPA (图 5).

为优化该整合载体, 以 YL121 和 YL122 (SEQ ID NOs:82 和 83) 为引物进行 PCR, 扩增解脂耶氏酵母 URA3 基因 (SEQ ID NO:84) 上游的 440bp 5' 非编码 DNA 序列. 用 AscI 和 BsiWI 消化该 PCR 产物, 并接着用 pYZM5CHPPA 的 AscI/BsiWI 片段置换 (图 5 和 12), 可生成 pYZM5UPA (图 12). 通过利用寡核苷酸 YL114 和 YL115 (SEQ ID NOs:85 和 86) 进行定点诱变, 将 AscI 位点导入 pYZM5UPA, 可生成 pYZV5. 为了减小 pYZV5 中 URA3 基因的 3' 非编码区的大小, 通过利用寡核苷酸 YL114 和 YL115 (如上所述) 进行的定点诱变, 将第二个 Pacl 位点导入该区域的中部, 可生成 pYZV5P. 通过 Pacl 消化切除 pYZV5P 的 Pacl 片段, 再次连接后生成 pYZV16 (图 12). 用 AscI 消化 pYZV16, 可释放适用于 $\Delta 5$ 去饱和酶基因 (“MAD5”) 在解脂耶氏酵母基因组内的整合和表达的 5.2kB DNA 片段 (SEQ ID NO:87).

适合高亲和力延伸酶和 $\Delta 5$ 去饱和酶的表达的整合载体的构建

通过分别利用引物对 YL61 和 YL62 (SEQ ID NOs:88 和 89) 和 YL69 和 YL70 (SEQ ID NOs:90 和 91) 进行定点诱变, 将 BsiWI 和 HindIII 位点导入 pY58 (含有高山被孢霉高亲和力 PUFA 延伸酶的编码区; 如实施例 2 所述), 可生成 pY58BH (图 13; 延伸酶基因被标记为 “EL”). 将 pY58BH 中含有 TEF::EL::XPR 嵌合基因的 1.7kB BsiWI/HindIII 片段连接进 pYZM5CHPP (构建方法如图 5 所示) 的 BsiWI/HindIII 位点, 可生成 pYZM5EL (图 13). 该质粒适用于高山

被孢霉 $\Delta 5$ 去饱和酶和高亲和力 PUFA 延伸酶基因在解脂耶氏酵母中的整合和同步表达。

适合 $\Delta 6$ 去饱和酶、高亲和力延伸酶和 $\Delta 5$ 去饱和酶的表达的整合载体的构建

通过分别利用引物对 YL77 和 YL78 (SEQ ID NOs:92 和 93) 和 YL79A 和 YL80A (SEQ ID NOs:94 和 95) 进行的定点诱变, 将 Pacl 和 Clal 位点导入 pY54 (含有高山被孢霉 $\Delta 6$ 去饱和酶; 如实施例 2 所述), 可生成 pY54PC (图 13; $\Delta 6$ 去饱和酶基因被标记为“MAD6”)。将 pY54PC 中含有 TEF::MAD6::XPR 嵌合基因的 2kB Clal/Pacl DNA 片段连接进 pYZM5EL 的 Clal/Pacl 位点, 可生成 pYZM5EL6 (图 13)。该质粒适用于高山被孢霉 $\Delta 6$ 去饱和酶、 $\Delta 5$ 去饱和酶和高亲和力 PUFA 延伸酶基因在解脂耶氏酵母基因组内的整合和同步表达。

适合被整合进入耶氏酵母基因组、适合 $\Delta 6$ 去饱和酶、PUFA 延伸酶和 $\Delta 5$ 去饱和酶表达的 DNA 片段的构建

采用质粒 pYZV16 (构建方法如图 12 所示) 构建含有多种表达盒的质粒。

首先, 将 pYZV16 的 3.5kB BsiWI/Pacl 片段与 pYZM5EL6 (构建方法如图 13 所示) 的 7.9kB BsiWI/Pacl 片段连接, 可生成 pYZV5EL6 (图 14)。用 Ascl 消化 pYZV5EL6, 可释放适合 $\Delta 6$ 去饱和酶、PUFA 延伸酶和 $\Delta 5$ 去饱和酶基因在解脂耶氏酵母基因组内的整合和同步表达的 8.9kB DNA 片段 (SEQ ID NO:96)。

适合被整合进耶氏酵母基因组、适合 $\Delta 6$ 去饱和酶、PUFA 延伸酶、 $\Delta 5$ 去饱和酶和 $\Delta 17$ 去饱和酶表达的 DNA 片段的构建

如实施例 3 的描述, 将合成的异丝水霉 $\Delta 17$ 去饱和酶基因插入 pY5-13 的 NcoI/NotI 位点, 可生成 pYSD17S (图 9A)。通过分别利用引物对 YL101 和 YL102 (SEQ ID NOs:97 和 98) 和 YL103 和 YL104 (SEQ ID NOs:99 和 100) 进行定点诱变, 将 Clal 和 PmeI 位点导入 pYSD17S 中, 可生成 pYSD17SPC (图 14)。

用来源于 pYSD17SPC 并含有 $\Delta 17$ 去饱和酶表达盒的 1760bp Clal/PmeI 片段置换 pYZV5EL6 (图 14) 的 347bp Clal/PmeI 片段, 可生成 pYZV5E6/17。用 Ascl 消化 pYZV5E6/17, 可释放适合 $\Delta 6$ 去饱和酶、PUFA 延伸酶、 $\Delta 5$ 去饱和酶和 $\Delta 17$ 去饱和酶基因在解脂耶氏酵母

基因组内的整合和同步表达的 10.3kB DNA 片段 (SEQ ID NO:101)。

实施例 5

解脂耶氏酵母转化体中 ω -6 脂肪酸的生物合成

用 AscI 限制性内切核酸酶消化 pYZV5EL6 (来自实施例 4, 含有 Δ 6 去饱和酶、PUFA 延伸酶和 Δ 5 去饱和酶基因) 并且根据实施例 2 中描述的方法转化到解脂耶氏酵母中。

在缺乏亮氨酸的基本培养基中选择的 52 种转化体中, 34 种不能在也缺乏尿嘧啶的培养基上生长, 表明 65% 的转化体含有整合到靶定的解脂耶氏酵母 URA3 基因座中的 8.9 kB 的多基因表达盒。将来自单个菌落的转化体接种到缺乏亮氨酸的基本培养基中, 在 30°C 下温育最多 48 小时。

通过离心收集细胞, 提取脂质, 通过转酯制备脂肪酸甲酯, 随后用 Hewlett-Packard 6890 GC 进行分析(根据一般方法中描述的方法)。

GC 分析表明在含有 3 种嵌合基因的转化体中存在花生四烯酸 (ARA) (图 15), 但在野生型耶氏酵母对照菌株中不存在。工程改造解脂耶氏酵母, 使其生产 ARA, 即一种 ω -6 脂肪酸。

实施例 6

解脂耶氏酵母转化体中 ω -3 脂肪酸的生物合成

按照与实施例 5 相似的方式, 用 AscI 限制性内切核酸酶消化 pYZV5E6/17 (来自实施例 4, 含有 Δ 6 去饱和酶、PUFA 延伸酶、 Δ 5 去饱和酶和 Δ 17 去饱和酶基因) 并且转化到解脂耶氏酵母 (ATCC #76982) 中。在缺乏亮氨酸的基本培养基中选择的 133 种转化体中, 89 种不能在也缺乏尿嘧啶的培养基上生长, 表明 67% 的转化体含有整合到靶定的解脂耶氏酵母 URA3 基因座中的 10.3 kB 的多基因表达盒。

GC 分析(根据一般方法中描述的方法)表明在含有 4 种嵌合基因的转化体中存在二十碳五烯酸 (EPA) (图 16), 但在野生型耶氏酵母对照菌株中不存在。这些数据表明经过工程改造解脂耶氏酵母, 生产了 EPA, 即一种 ω -3 脂肪酸。

序列表

- <110> E. I. du Pont de Nemours and Company, Inc.
- <120> 在含油酵母中生产多不饱和脂肪酸
- <130> CL2233 PCT
- <150> US 60/468677
- <151> 2003-05-07
- <160> 126
- <170> PatentIn version 3.2
- <210> 1
- <211> 1374
- <212> DNA
- <213> 高山被孢霉 AF465281
- <400> 1
- atggctgctg ctcccagtgt gaggacgttt actcgggccc aggttttgaa tgecgaggct 60
- ctgaatgagg gcaagaagga tgecgaggca cccttcttga tgatcatga caacaagggtg 120
- tacgatgtcc gogagttcgt ccctgatcat cccggtggaa gtgtgattct cacgcacgtt 180
- ggcaaggacg gcactgacgt ctttgaoact ttccacccc aggctgcttg ggagactctt 240
- gccaactttt acgttggtga tattgaogag agcgaccgag atatcaagaa tgatgacttt 300
- goggccgagg tcogcaagct gcgtacctg ttccagtctc ttggttacta cgattcttcc 360
- aaggcatact acgccttcaa ggtctcgttc aacctctgca tctggggttt gtcgaoggtc 420
- attgtggcca agtggggcca gacctgacc ctccccaacg tgctctcggc tgegttttg 480
- ggtctgttct ggcagcagtg oggatggtg gctcacgact ttttgcata ccaggtcttc 540
- caggaccgtt tctggggtga tcttttgggc gccttcttg gaggtgtctg ccagggttc 600
- tgtcctcgt ggtggaagga caagcacaac actcaccacg cggccccaa cgtccacggc 660
- gaggatccc acattgacac ccacctctg ttgacctgga gtgagcatgc gttggagatg 720
- ttctoggatg tccagatga ggagctgacc ogcatgtggt ogcgtttcat ggtcctgaac 780
- cagacctggt tttacttccc cattctctcg tttgcccgtc tctctggtg cctccagtc 840

attctctttg tgetgcctaa cggtcaggcc cacaagccct ogggcgcgcg tgtgcccato 900
 tcgttggtcg agcagetgtc gottgogatg cactggacot ggtacctcgc caccatgttc 960
 ctgttcatca aggatecccg caacatgctg gtgtactttt tgggtgtcga ggoggtgtgc 1020
 ggaaacttgt tggcgatcgt gttctcgcgc aaccacaacg gtatgcctgt gatctcgaag 1080
 gaggaggcgg tcgatatgga tttcttcacg aagcagatca tcacgggtcg tgatgtccac 1140
 cgggtctat ttgccaactg gttcacgggt ggattgaact atcagatcga gcaccacttg 1200
 ttcccttcga tgcctcgcca caacttttca aagatccagc ctgctgtcga gaccctgtgc 1260
 aaaaagtaca atgtccgata ccacaccacc ggtatgatcg agggaactgc agaggtcttt 1320
 agccgtctga acgaggtctc caaggtacc tccaagatgg gtaaggcgcga gtaa 1374

<210> 2

<211> 457

<212> PRT

<213> 高山被孢霉 AF465281

<400> 2

Met Ala Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala Glu Val Leu
 1 5 10 15

Asn Ala Glu Ala Leu Asn Glu Gly Lys Lys Asp Ala Glu Ala Pro Phe
 20 25 30

Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Glu Phe Val Pro
 35 40 45

Asp His Pro Gly Gly Ser Val Ile Leu Thr His Val Gly Lys Asp Gly
 50 55 60

Thr Asp Val Phe Asp Thr Phe His Pro Glu Ala Ala Trp Glu Thr Leu
 65 70 75 80

Ala Asn Phe Tyr Val Gly Asp Ile Asp Glu Ser Asp Arg Asp Ile Lys

	85		90		95
Asn Asp Asp Phe Ala Ala Glu Val Arg Lys Leu Arg Thr Leu Phe Gln	100		105		110
Ser Leu Gly Tyr Tyr Asp Ser Ser Lys Ala Tyr Tyr Ala Phe Lys Val	115		120		125
Ser Phe Asn Leu Cys Ile Trp Gly Leu Ser Thr Val Ile Val Ala Lys	130		135		140
Trp Gly Gln Thr Ser Thr Leu Ala Asn Val Leu Ser Ala Ala Leu Leu	145		150		155
					160
Gly Leu Phe Trp Gln Gln Cys Gly Trp Leu Ala His Asp Phe Leu His	165		170		175
His Gln Val Phe Gln Asp Arg Phe Trp Gly Asp Leu Phe Gly Ala Phe	180		185		190
Leu Gly Gly Val Cys Gln Gly Phe Ser Ser Ser Trp Trp Lys Asp Lys	195		200		205
His Asn Thr His His Ala Ala Pro Asn Val His Gly Glu Asp Pro Asp	210		215		220
Ile Asp Thr His Pro Leu Leu Thr Trp Ser Glu His Ala Leu Glu Met	225		230		235
					240
Phe Ser Asp Val Pro Asp Glu Glu Leu Thr Arg Met Trp Ser Arg Phe	245		250		255
Met Val Leu Asn Gln Thr Trp Phe Tyr Phe Pro Ile Leu Ser Phe Ala	260		265		270
Arg Leu Ser Trp Cys Leu Gln Ser Ile Leu Phe Val Leu Pro Asn Gly					

275	280	285
Gln Ala His Lys Pro Ser Gly Ala Arg Val Pro Ile Ser Leu Val Glu 290	295	300
Gln Leu Ser Leu Ala Met His Trp Thr Trp Tyr Leu Ala Thr Met Phe 305	310	315 320
Leu Phe Ile Lys Asp Pro Val Asn Met Leu Val Tyr Phe Leu Val Ser 325	330	335
Gln Ala Val Cys Gly Asn Leu Leu Ala Ile Val Phe Ser Leu Asn His 340	345	350
Asn Gly Met Pro Val Ile Ser Lys Glu Glu Ala Val Asp Met Asp Phe 355	360	365
Phe Thr Lys Gln Ile Ile Thr Gly Arg Asp Val His Pro Gly Leu Phe 370	375	380
Ala Asn Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu 385	390	395 400
Phe Pro Ser Met Pro Arg His Asn Phe Ser Lys Ile Gln Pro Ala Val 405	410	415
Glu Thr Leu Cys Lys Lys Tyr Asn Val Arg Tyr His Thr Thr Gly Met 420	425	430
Ile Glu Gly Thr Ala Glu Val Phe Ser Arg Leu Asn Glu Val Ser Lys 435	440	445
Ala Thr Ser Lys Met Gly Lys Ala Gln 450	455	

<210> 3

<211> 1341
 <212> DNA
 <213> 高山被孢霉 AF067654

<400> 3

atgggaacgg accaagga aaacctcacc tgggaagagc tggcggccca taacaccaag 60
 gacgacctac tcttggccat ccgcggcagg gtgtacgatg tcacaaagtt cttgagccgc 120
 catcctggtg gaggggacac tctcctgctc ggagctggcc gagatgttac tccggtcttt 180
 gagatgtatc acgcgttttg ggetgcagat gccattatga agaagtaacta tgtcgggtaca 240
 ctggtctcga atgagctgcc catcttcccg gagccaacgg tgttcacaaa aaccatcaag 300
 acgagagtcg agggctactt tacggatogg aacattgatc ccaagaatag accagagatc 360
 tggggacgat acgctcttat ctttggatcc ttgatcgtt octactacgc gcagctcttt 420
 gtgcctttcg ttgtcgaacg cacatggctt cagggtgtgt ttgcaatcat catgggattt 480
 gcgtgcgcac aagtcggact caaccctctt catgatgcgt ctcacttttc agtgaccac 540
 aaccccactg tctggaagat tctgggagcc acgcacgact ttttcaacgg agcatcgtac 600
 ctggtgtgga tgtaccaaca tatgctcggc catcaccctt acaccaacat tgcctggagca 660
 gatcccgcag tctcgcagtc tgagcccgat gttcgtcgtc tcaagcccaa ccaaaagtgg 720
 tttgtcaacc acatcaacca gcacatgttt gttcctttcc tgtacggact gctggcgctc 780
 aaggtgcgca ttcaggacat caacattttg tactttgtca agaccaatga cgctattcgt 840
 gtcaatecca tctogacatg gcacactgtg atgtttctgg gcggcaaggc tttctttgtc 900
 tggatcggc tgattgttcc cctgcagtat ctgccctgg gcaaggtgct gctcttgttc 960
 acggtcggg acatggtgtc gtcttactgg ctggcgtga ccttcagge gaaccagtt 1020
 gttgaggaag ttcagtggcc gttgcctgac gagaacggga tcatccaaaa ggactgggca 1080
 gctatgcagg tcgagactac gcaggattac gcacacgatt cgcacctctg gaccagcatc 1140
 actggcagct tgaactacea ggctgtgcac catctgttcc ccaacgtgtc gcagcaccat 1200
 tatcccgata ttctggccat catcaagaac acctgcagcg agtacaaggt tccatacctt 1260
 gtcaaggata cgttttggca agcatttctt tcacatttgg agcacttggc tgttcttggc 1320

ctccgtccca aggaagagta g

1341

<210> 4

<211> 446

<212> PRT

<213> 高山被孢霉 AF067654

<400> 4

Met Gly Thr Asp Gln Gly Lys Thr Phe Thr Trp Glu Glu Leu Ala Ala
 1 5 10 15

His Asn Thr Lys Asp Asp Leu Leu Leu Ala Ile Arg Gly Arg Val Tyr
 20 25 30

Asp Val Thr Lys Phe Leu Ser Arg His Pro Gly Gly Val Asp Thr Leu
 35 40 45

Leu Leu Gly Ala Gly Arg Asp Val Thr Pro Val Phe Glu Met Tyr His
 50 55 60

Ala Phe Gly Ala Ala Asp Ala Ile Met Lys Lys Tyr Tyr Val Gly Thr
 65 70 75 80

Leu Val Ser Asn Glu Leu Pro Ile Phe Pro Glu Pro Thr Val Phe His
 85 90 95

Lys Thr Ile Lys Thr Arg Val Glu Gly Tyr Phe Thr Asp Arg Asn Ile
 100 105 110

Asp Pro Lys Asn Arg Pro Glu Ile Trp Gly Arg Tyr Ala Leu Ile Phe
 115 120 125

Gly Ser Leu Ile Ala Ser Tyr Tyr Ala Gln Leu Phe Val Pro Phe Val
 130 135 140

Val Glu Arg Thr Trp Leu Gln Val Val Phe Ala Ile Ile Met Gly Phe

145	150	155	160
Ala Cys Ala Gln Val Gly Leu Asn Pro Leu His Asp Ala Ser His Phe			
	165	170	175
Ser Val Thr His Asn Pro Thr Val Trp Lys Ile Leu Gly Ala Thr His			
	180	185	190
Asp Phe Phe Asn Gly Ala Ser Tyr Leu Val Trp Met Tyr Gln His Met			
	195	200	205
Leu Gly His His Pro Tyr Thr Asn Ile Ala Gly Ala Asp Pro Asp Val			
	210	215	220
Ser Thr Ser Glu Pro Asp Val Arg Arg Ile Lys Pro Asn Gln Lys Trp			
225	230	235	240
Phe Val Asn His Ile Asn Gln His Met Phe Val Pro Phe Leu Tyr Gly			
	245	250	255
Leu Leu Ala Phe Lys Val Arg Ile Gln Asp Ile Asn Ile Leu Tyr Phe			
	260	265	270
Val Lys Thr Asn Asp Ala Ile Arg Val Asn Pro Ile Ser Thr Trp His			
	275	280	285
Thr Val Met Phe Trp Gly Gly Lys Ala Phe Phe Val Trp Tyr Arg Leu			
	290	295	300
Ile Val Pro Leu Gln Tyr Leu Pro Leu Gly Lys Val Leu Leu Leu Phe			
305	310	315	320
Thr Val Ala Asp Met Val Ser Ser Tyr Trp Leu Ala Leu Thr Phe Gln			
	325	330	335
Ala Asn His Val Val Glu Glu Val Gln Trp Pro Leu Pro Asp Glu Asn			

340	345	350
Gly Ile Ile Gln Lys Asp Trp Ala Ala Met Gln Val Glu Thr Thr Gln		
355	360	365
Asp Tyr Ala His Asp Ser His Leu Trp Thr Ser Ile Thr Gly Ser Leu		
370	375	380
Asn Tyr Gln Ala Val His His Leu Phe Pro Asn Val Ser Gln His His		
385	390	395
400		
Tyr Pro Asp Ile Leu Ala Ile Ile Lys Asn Thr Cys Ser Glu Tyr Lys		
405	410	415
Val Pro Tyr Leu Val Lys Asp Thr Phe Trp Gln Ala Phe Ala Ser His		
420	425	430
Leu Glu His Leu Arg Val Leu Gly Leu Arg Pro Lys Glu Glu		
435	440	445

<210> 5
 <211> 1077
 <212> DNA
 <213> 异丝水霉(ATCC #56851)

<400> 5
 atgactgagg ataagacgaa ggtcagagttc ccgacgctca cggagctcaa gcaactcgatc 60
 ccgaacgcgt gctttgagtc gaacctcggc ctctcgtctet actacacggc ccgcgcgatc 120
 ttcaacgcgt cggcctcggc ggcgctgctc tacggggcgc gctcgcgcgc gttcattgcc 180
 gataacgttc tgctccacgc gctcgtttgc gccacctaca tctacgtgca gggcgtcctc 240
 ttctggggct tcttcacggt cggccacgac tgcggccact cggccttctc gcgctaccac 300
 agcgtcaact ttatcatcgg ctgcatcatg cactctgoga tttgacgcc gttcgagagc 360
 tggcgcgtga cgcaccgcca ccaccacaag aacacgggca acattgataa ggacgagatc 420
 ttttaccgc accggtcggc caaggacctc caggacgtgc gccaatgggt ctacacgctc 480

ggcggtgCGT ggtttgtcta cttgaaggTC gggtatgccc cgcgcacgat gagccacttt 540
 gaccCGTggg acccGctcct ccttcgcgcg gcgtcgccg tcatcgtgtc gctcggcgTC 600
 tggcgccgct tcttcgcgcg gtacgcgtac ctcacatact cgtcggctt tgccgtcatg 660
 ggctctact actatgcgcg gctctttgtc tttgcttCGT tcctcgtcat taGacccttc 720
 ttgcaccaca acgaCGaagc gacgcCGTgg tacggcgact cggagTggac gtacgtcaag 780
 ggcaacctct cGagcgtcga ccGctcgtac ggcgCGttcg tggacaacct gagccaccac 840
 attggcaCGc accaggtcca ccacttgTtc ccgatcattc cgcactacaa gctcaacgaa 900
 gccaccaagc actttgcgcg cgcgtaccCG cacctcgtgc gcaggaacga cGagcccatc 960
 atcagcgctt tcttcaagac cgcgcacctc tttgtcaact acggcgctgt gcccGagacg 1020
 gcgcagatct tcacgctcaa agagtcggcc gggcgcccca agccaagtc ggactaa 1077

<210> 6

<211> 358

<212> PRT

<213> 异丝水霉(ATCC #56851)

<400> 6

Met Ala Glu Asp Lys Thr Lys Val Glu Phe Pro Thr Leu Thr Glu Leu
 1 5 10 15

Lys His Ser Ile Pro Asn Ala Cys Phe Glu Ser Asn Leu Gly Leu Ser
 20 25 30

Leu Tyr Tyr Thr Ala Arg Ala Ile Phe Asn Ala Ser Ala Ser Ala Ala
 35 40 45

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

Leu His Ala Leu Val Cys Ala Thr Tyr Ile Tyr Val Gln Gly Val Ile
 65 70 75 80

Phe Trp Gly Phe Phe Thr Val Gly His Asp Cys Gly His Ser Ala Phe
85 90 95

Ser Arg Tyr His Ser Val Asn Phe Ile Ile Gly Cys Ile Met His Ser
100 105 110

Ala Ile Leu Thr Pro Phe Glu Ser Trp Arg Val Thr His Arg His His
115 120 125

His Lys Asn Thr Gly Asn Ile Asp Lys Asp Glu Ile Phe Tyr Pro His
130 135 140

Arg Ser Val Lys Asp Leu Gln Asp Val Arg Gln Trp Val Tyr Thr Leu
145 150 155 160

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

Met Ser His Phe Asp Pro Trp Asp Pro Leu Leu Leu Arg Arg Ala Ser
180 185 190

Ala Val Ile Val Ser Leu Gly Val Trp Ala Ala Phe Phe Ala Ala Tyr
195 200 205

Ala Tyr Leu Thr Tyr Ser Leu Gly Phe Ala Val Met Gly Leu Tyr Tyr
210 215 220

Tyr Ala Pro Leu Phe Val Phe Ala Ser Phe Leu Val Ile Thr Thr Phe
225 230 235 240

Leu His His Asn Asp Glu Ala Thr Pro Trp Tyr Gly Asp Ser Glu Trp
245 250 255

Thr Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr Gly Ala
260 265 270

Phe Val Asp Asn Leu Ser His His Ile Gly Thr His Gln Val His His
 275 280 285

Leu Phe Pro Ile Ile Pro His Tyr Lys Leu Asn Glu Ala Thr Lys His
 290 295 300

Phe Ala Ala Ala Tyr Pro His Leu Val Arg Arg Asn Asp Glu Pro Ile
 305 310 315 320

Ile Thr Ala Phe Phe Lys Thr Ala His Leu Phe Val Asn Tyr Gly Ala
 325 330 335

Val Pro Glu Thr Ala Gln Ile Phe Thr Leu Lys Glu Ser Ala Ala Ala
 340 345 350

Ala Lys Ala Lys Ser Asp
 355

<210> 7
 <211> 957
 <212> DNA
 <213> 高山被孢霉 AX464731

<400> 7
 atggagtoga ttgceccatt cctcccatca aagatgcogc aagatctggt tatggacct 60
 gccaccgcta tcggtgtccg ggcogcgccc tatgtogate ctctogaggc cgcgctggtg 120
 gccaggcog agaagtacat ccccacgatt gtccatcaca cgcgtgggtt cctggtcogc 180
 gtggagtgc ctttgccccg tgagctgccg ttgatgaacc cgttccacgt getgttgatc 240
 gtgctcgett atttggtcac ggtctttgtg ggcatgcaga tcatgaagaa ctttgagcgg 300
 ttcgaggta agacgttttc gctcctgca aacttttgc tggtctogat cagcgcctac 360
 atgtcggtg ggatcctgta cgaggettat caggecaact atggactggt tgagaacgt 420
 gctgatcata cttcaaggg tcttctatg gccaatga tctggctctt ctacttctcc 480
 aagatcatgg agtttgcga caccatgate atggtcctca agaagaaca cgcacgatc 540

tccttcttgc acgtttacca ccacagctcc atcttcacca tctggtggtt ggtcaccttt 600
 gttgcaacca acggtgaagc ctactttctct gctgogttga actogttcat ccatgtgato 660
 atgtacggct actactttctt gtcggccttg ggettcaage aggtgtcgtt catcaagttc 720
 tacatcacgc gctcgcagat gacacagttc tgcattgatgt oggtccagtc ttcttgggac 780
 atgtacgcca tgaaggtcct tggccgcccc ggataccctt tcttcatcac ggetctgctt 840
 tggttctaca tgtggacct gctcggctctc ttctacaact tttacagaaa gaacgccaag 900
 ttggccaagc aggccaaggc cgacgctgcc aaggagaagg caaggaagtt gcagtaa 957

<210> 8
 <211> 318
 <212> PRT
 <213> 高山被孢霉 AX464731

<400> 8

Met Glu Ser Ile Ala Pro Phe Leu Pro Ser Lys Met Pro Gln Asp Leu
 1 5 10 15

Phe Met Asp Leu Ala Thr Ala Ile Gly Val Arg Ala Ala Pro Tyr Val
 20 25 30

Asp Pro Leu Glu Ala Ala Leu Val Ala Gln Ala Glu Lys Tyr Ile Pro
 35 40 45

Thr Ile Val His His Thr Arg Gly Phe Leu Val Ala Val Glu Ser Pro
 50 55 60

Leu Ala Arg Glu Leu Pro Leu Met Asn Pro Phe His Val Leu Leu Ile
 65 70 75 80

Val Leu Ala Tyr Leu Val Thr Val Phe Val Gly Met Gln Ile Met Lys
 85 90 95

Asn Phe Glu Arg Phe Glu Val Lys Thr Phe Ser Leu Leu His Asn Phe

100	105	110
Cys Leu Val Ser Ile Ser Ala Tyr Met Cys Gly Gly Ile Leu Tyr Glu		
115	120	125
Ala Tyr Gln Ala Asn Tyr Gly Leu Phe Glu Asn Ala Ala Asp His Thr		
130	135	140
Phe Lys Gly Leu Pro Met Ala Lys Met Ile Trp Leu Phe Tyr Phe Ser		
145	150	155
Lys Ile Met Glu Phe Val Asp Thr Met Ile Met Val Leu Lys Lys Asn		
165	170	175
Asn Arg Gln Ile Ser Phe Leu His Val Tyr His His Ser Ser Ile Phe		
180	185	190
Thr Ile Trp Trp Leu Val Thr Phe Val Ala Pro Asn Gly Glu Ala Tyr		
195	200	205
Phe Ser Ala Ala Leu Asn Ser Phe Ile His Val Ile Met Tyr Gly Tyr		
210	215	220
Tyr Phe Leu Ser Ala Leu Gly Phe Lys Gln Val Ser Phe Ile Lys Phe		
225	230	235
Tyr Ile Thr Arg Ser Gln Met Thr Gln Phe Cys Met Met Ser Val Gln		
245	250	255
Ser Ser Trp Asp Met Tyr Ala Met Lys Val Leu Gly Arg Pro Gly Tyr		
260	265	270
Pro Phe Phe Ile Thr Ala Leu Leu Trp Phe Tyr Met Trp Thr Met Leu		
275	280	285
Gly Leu Phe Tyr Asn Phe Tyr Arg Lys Asn Ala Lys Leu Ala Lys Gln		

290	295	300	
Ala Lys Ala Asp Ala Ala Lys Glu Lys Ala Arg Lys Leu Gln			
305	310	315	
<210>	9		
<211>	1077		
<212>	DNA		
<213>	异丝水霉		
<400>	9		
atggetgagg ataagaccaa ggtcgagttc cctaccctga ctgagctgaa gcactctatc			60
cctaacgctt gctttgagtc caacctogga ctctegctct actaacctgc ccgagcgatc			120
ttcaacgeat ctgcctctgc tgctctgctc tacgctgccc gatctactcc cttcattgcc			180
gataacgttc tgetccaogc tctggtttgc gccacctaca tctacgtgca ggggtgctac			240
ttctgggggt tctttaccgt cggtaacgac tgtggtaact ctgccttctc ccgataccac			300
tccgtcaact tcatcattgg ctgcatcatg cactctgcca ttctgactcc cttegagtcc			360
tggcgagtga cccaccgaca ccatcacaag aacctggca acattgataa ggacgagatc			420
ttctaccctc atcggtccgt caaggacctc caggacgtgc gacaatgggt ctacaccctc			480
ggaggtgctt ggtttgtcta cctgaagtc ggatatgctc ctgaacct gtcccattt			540
gaccctggg acctctctct gcttegacga gcctccgtg tcatctgtgc cctcggagtc			600
tgggctgcct tcttcgtgc ctacgcctac ctacatact cgctcggctt tgccgtcatg			660
ggcctctact actatgctcc tctctttgtc tttgcttctg tctctgcat tactacctc			720
ttgcatcaca acgacgaagc tactccctgg tacggtgact cggagtggac ctacgtcaag			780
ggcaacctga gctcctgca ccgatcgtac ggagctttcg tggacaacct gtctcaccac			840
attgcaacc accaggtcca tcaattgttc cctateattc cccactacaa gctcaacgaa			900
gccaccaagc actttgctgc cgcttacct cactcgtga gacgtaacga cgagcccatc			960
attactgctt tctcaagac cgctcacctc tttgtcaact acggagctgt gcccgagact			1020
gctcagattt tcacctcaa agagtctgcc gctgcagcca aggccaagag cgactaa			1077

<210> 10
 <211> 105
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 D17-1A

<400> 10
 catggctgag gataagacca aggtcgagtt ccctaccctg actgagctga agcactctat 60

ccctaacgct tgctttgagt ccaacctcgg actctcgtctc tacta 105

<210> 11
 <211> 106
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 D17-1B

<400> 11
 cagtgtagta gagcgagagt ccgaggttgg actcaaagca agcgttaggg atagagtgct 60

tcagctcagt cagggttaggg aactcgacct tggctctatc ctcage 106

<210> 12
 <211> 106
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 D17-2A

<400> 12
 cactgcccca gogatcttca acgcatctgc ctctgctget ctgetctacg ctgcccgatc 60

tactcccttc attgcegata acgttctget ccacgctctg gtttgc 106

<210> 13
 <211> 106
 <212> DNA
 <213> 人工序列

<220>

<223> 引物 D17-2B

<400> 13

gtggcgcaaa ccagagcgtg gagcagaacg ttatcggcaa tgaaggagat agatcgggca 60

gcgtagagca gagcagcaga ggcagatgog ttgaagatcg ctcggg 106

<210> 14

<211> 105

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-3A

<400> 14

gccacctaca tctacgtgca ggggtgcatc ttctgggggtt tctttaccgt cggtcacgac 60

tgtgggtcaact ctgccttctc ccgataccac tccgtcaact tcate 105

<210> 15

<211> 105

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-3B

<400> 15

ccaatgatga agttgacgga gtggtatcgg gagaaggcag agtgaccaca gtcgtgaccg 60

acggtaaaga aaccccagaa gatgacaccc tgcacgtaga tgtag 105

<210> 16

<211> 105

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-4A

<400> 16

attggtgca tcatgactc tgccattctg actcccttcg agtcctggcg agtgaccac 60

cgacaccatc acaagaacac tggcaacatt gataaggacg agatc 105

<210> 17

<211> 105

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-4B

<400> 17

tagaagatct cgtccttacc aatgttgcca gtgttcttgt gatgggtcgc gtgggtcact 60

cgccaggact cgaaggagat cagaatggca gattgcatga tgcag 105

<210> 18

<211> 105

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-5A

<400> 18

acgagatctt ctaccctcat cggtcogtca aggacctcca ggacgtcga caatgggtct 60

acaccctcgg aggtgcttgg tttgtctacc tgaaggctgg atatg 105

<210> 19

<211> 107

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-5B

<400> 19

aggagcatat ccgaccttca ggtagacaaa ccaagcacct ccgagggtgt agaccattg 60

tgcacgtcc tggaggctct tgaaggaccg atgagggtag aagatct 107

<210> 20

<211> 105

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-6A

<400> 20

ctcctcgaac catgtcccac ttgaccctt gggaccctct cctgcttoga cgagcctcgg 60

ctgtcatcgt gtcctcggga gtctgggctg ccttcttcgc tgctt 105

<210> 21

<211> 106

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-6B

<400> 21

aggcgtaggc agcgaagaag gcagcccaga ctccgaggga cacgatgaca gcggaggctc 60

gtcgaagcag gagagggtcc caggggtcaa agtgggacat ggcttcg 106

<210> 22

<211> 104

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-7A

<400> 22

acgcctacct cacatactcg ctccgctttg ccgtcatggg cctctactac tatgtctctc 60

tctttgtctt tgcttcgttc ctcttcatta ctacctctt geat 104

<210> 23

<211> 103

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-7B

<400> 23
ttgtgatgca agaaggtagt aatgacgagg aacgaagcaa agacaaagag aggagcatag 60

tagtagaggc ccatgacggc aaagccgagc gagtatgtga ggt 103

<210> 24

<211> 106

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-8A

<400> 24

cacaacgacg aagctactcc ctggtacggt gactcggagt ggacctactg caagggcaac 60

ctgagctccg togaccgatc gtacggagct ttcgtggaca acctgt 106

<210> 25

<211> 106

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-8B

<400> 25

gtgagacagg ttgtccacga aagctccgta cgatcggteg acggagctca ggttgcctt 60

gacgtaggtc cactccgagt cacogtacca gggagtagct tcgtcg 106

<210> 26

<211> 102

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-9A

<400> 26

ctcaccacat tggcaccac caggtccatc acttgttccc tatcattccc cactacaagc 60

tcaacgaagc caccaagcac tttgctgccg ettaccctca cc 102

- <210> 27
 <211> 102
 <212> DNA
 <213> 人工序列
- <220>
 <223> 引物 D17-9B
- <400> 27
 cacgaggtga ggtaagcgg cagcaaagtg cttggtgct tcggtgagct tgtagtgggg 60
 aatgataggg aacaagtgat ggacctggtg ggtgccaatg tg 102
- <210> 28
 <211> 76
 <212> DNA
 <213> 人工序列
- <220>
 <223> 引物 D17-10A
- <400> 28
 tcgtgagacg taacgacgag cccatcatta ctgccttctt caagaccgct cacctctttg 60
 tcaactacgg agctgt 76
- <210> 29
 <211> 76
 <212> DNA
 <213> 人工序列
- <220>
 <223> 引物 D17-10B
- <400> 29
 cgggcacagc tcgtagttg acaaagaggt gagcggctctt gaagaaggca gtaatgatgg 60
 gctcgtcgtt acgtct 76
- <210> 30
 <211> 67
 <212> DNA
 <213> 人工序列
- <220>

<223> 引物 D17-11A

<400> 30

gcccgagact gctcagattt tcacctcaa agagtctgcc gctgcagcca aggccaagag 60

cgactaa 67

<210> 31

<211> 62

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-11B

<400> 31

ttagtcgctc ttggccttgg ctgcagcgge agactctttg agggtgaaaa tetgagcagt 60

ct 62

<210> 32

<211> 32

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-1

<400> 32

tttccatggc tgaggataag accaaggtcg ag 32

<210> 33

<211> 34

<212> DNA

<213> 人工序列

<220>

<223> 引物 D17-4R

<400> 33

ccctagaaga tctcgtcctt atcaatgttg ccag 34

<210> 34

<211> 27

- <212> DNA
 <213> 人工序列
- <220>
 <223> 引物 D17-5
- <400> 34
 cccacgagat cttctaccct catcggt 27
- <210> 35
 <211> 24
 <212> DNA
 <213> 人工序列
- <220>
 <223> 引物 D17-8D
- <400> 35
 gaaagctccg tacgatcggt cgac 24
- <210> 36
 <211> 24
 <212> DNA
 <213> 人工序列
- <220>
 <223> 引物 D17-8U
- <400> 36
 gtcgaccgat cgtacggagc ttte 24
- <210> 37
 <211> 34
 <212> DNA
 <213> 人工序列
- <220>
 <223> 引物 D17-11
- <400> 37
 aaagcgccg cttagtoget cttggccttg gctg 34
- <210> 38
 <211> 19

- <212> DNA
 <213> 人工序列
- <220>
 <223> 引物 TEF5'
- <400> 38
 agagaccggg ttggcggcg 19
- <210> 39
 <211> 30
 <212> DNA
 <213> 人工序列
- <220>
 <223> 引物 TEF3'
- <400> 39
 ttggatcctt tgaatgattc ttatactcag 30
- <210> 40
 <211> 29
 <212> DNA
 <213> 人工序列
- <220>
 <223> 引物 XPR5'
- <400> 40
 tttccgcggc cegagattcc ggcctcttc 29
- <210> 41
 <211> 31
 <212> DNA
 <213> 人工序列
- <220>
 <223> 引物 XPR3'
- <400> 41
 tttccgcgga cacaatatct ggtcaaattt c 31
- <210> 42
 <211> 33

- <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 YL21A

 <400> 42
 ttccatggc tgaggataag acgaaggtcg agt 33
- <210> 43
 <211> 36
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 YL22

 <400> 43
 cccttaatta attagtccga cttggccttg gcggcc 36
- <210> 44
 <211> 36
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 YL53

 <400> 44
 gcccaagtcgg actaagctgc taactagagc ggccgc 36
- <210> 45
 <211> 36
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 YL54

 <400> 45
 gcggccgctc tagttagcag cttagtccga cttggc 36
- <210> 46
 <211> 35

<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 KU5	
<400> 46	
tttgcccggg cgagtatctg tctgactcgt cattg	35
<210> 47	
<211> 33	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 KU3	
<400> 47	
aaagcccggg caaaggcctg tttctcgggtg tac	33
<210> 48	
<211> 1710	
<212> DNA	
<213> 解脂耶氏酵母	
<400> 48	
gtcgacgagt atctgtctga ctctcattg ccgcctttgg agtacgactc caactatgag	60
tgtgcttgga tcactttgac gatacattct tcgttgagg ctgtgggtct gacagctgcg	120
ttttcggcgc ggttgccga caacaatctc agctgcaacg tcattgctgg ctttcatcat	180
gatcacattt ttgtggcaa aggcgacgcc cagagagcca ttgaogttct ttctaatttg	240
gaccgatagc cgtatagtc agtctatcta taagttcaac taactcgtaa ctattaccat	300
aacatatact tcactgcccc agataagggt ccgataaaaa gttctgcaga ctaaatttat	360
ttcagtetcc tcttaccac caaaatgcc tctacgaag ctcgagctaa cgtccacaag	420
tccgccttg ccgctcaggt gctcaagctc gtggcagcca agaaaaccaa cctgtgtgct	480
tctctggatg ttaccaccac caaggagctc attgagcttg ccgataaggt cggaccttat	540
gtgtgcatga tcaagaccca tctcgacatc attgacgact tcacctacgc cggcactgtg	600

ctccccctca aggaacttgc tcttaagcac ggtttcttcc tgttcgagga cagaaagtte 660
 gcagatattg gcaacctgt caagcaccag tacaagaacg gtgtctaccg aatcgccgag 720
 tggtcgata tcaccaacgc ccaoggtgta ccggaacog gaatcattgc tggcctgcga 780
 gctggtgcog aggaaactgt ctctgaacag aagaaggagg acgtctctga ctacgagaac 840
 tcccagtaca aggagtctct ggtccccctct cccaacgaga agctggccag aggtctgctc 900
 atgctggcog agctgtcttg caaggctct ctggccactg gcgagtactc caagcagacc 960
 attgagcttg cccgatcoga ccccgagttt gtggttgct tcattgcca gaaccgacct 1020
 aagggcgaact ctgaggactg gcttattctg acccccggg tgggtcttga egacaaggga 1080
 gacgctctcg gacagcagta ccgaactgtt gaggatgtca tgtctaccgg aacggatata 1140
 ataattgtcg gccgaggtct gtaogccag aaccgagata ctattgagga ggccaagcga 1200
 taccagaagg ctggctggga ggcttaccag aagattaact gttagaggtt agactatgga 1260
 tatgtcattt aactgtgtat atagagagcg tgcaagtatg gagegcttgt tcagcttgta 1320
 tgatggtcag acgacotgtc tgatcgagta tgtatgatac tgcacaacct gtgtatcgc 1380
 atgatctgtc caatgggca tgttgttgtg tttctcgata cggagatgct ggttacaagt 1440
 agctaatacg attgaaactac ttatacttat atgagcttg aagaaagctg acttgtgtat 1500
 gacttattct caactacatc cccagtcaca ataccaccac tgcactacca ctacaccaa 1560
 accatgatca aaccaccat ggacttctg gaggcagaag aactgttat ggaaaagctc 1620
 aagagagaga agccaagata ctatcaagac atgtgtogca acttcaagga ggaccaagct 1680
 ctgtacaccg agaaacagc ctttctcgac 1710

<210> 49
 <211> 286
 <212> PRT
 <213> 解脂耶氏酵母

<400> 49

Met Pro Ser Tyr Glu Ala Arg Ala Asn Val His Lys Ser Ala Phe Ala
 1 5 10 15

Ala Arg Val Leu Lys Leu Val Ala Ala Lys Lys Thr Asn Leu Cys Ala
 20 25 30

Ser Leu Asp Val Thr Thr Thr Lys Glu Leu Ile Glu Leu Ala Asp Lys
 35 40 45

Val Gly Pro Tyr Val Cys Met Ile Lys Thr His Ile Asp Ile Ile Asp
 50 55 60

Asp Phe Thr Tyr Ala Gly Thr Val Leu Pro Leu Lys Glu Leu Ala Leu
 65 70 75 80

Lys His Gly Phe Phe Leu Phe Glu Asp Arg Lys Phe Ala Asp Ile Gly
 85 90 95

Asn Thr Val Lys His Gln Tyr Lys Asn Gly Val Tyr Arg Ile Ala Glu
 100 105 110

Trp Ser Asp Ile Thr Asn Ala His Gly Val Pro Gly Thr Gly Ile Ile
 115 120 125

Ala Gly Leu Arg Ala Gly Ala Glu Glu Thr Val Ser Glu Gln Lys Lys
 130 135 140

Glu Asp Val Ser Asp Tyr Glu Asn Ser Gln Tyr Lys Glu Phe Leu Val
 145 150 155 160

Pro Ser Pro Asn Glu Lys Leu Ala Arg Gly Leu Leu Met Leu Ala Glu
 165 170 175

Leu Ser Cys Lys Gly Ser Leu Ala Thr Gly Glu Tyr Ser Lys Gln Thr
 180 185 190

Ile Glu Leu Ala Arg Ser Asp Pro Glu Phe Val Val Gly Phe Ile Ala
 195 200 205

Gln Asn Arg Pro Lys Gly Asp Ser Glu Asp Trp Leu Ile Leu Thr Pro
 210 215 220

Gly Val Gly Leu Asp Asp Lys Gly Asp Ala Leu Gly Gln Gln Tyr Arg
 225 230 235 240

Thr Val Glu Asp Val Met Ser Thr Gly Thr Asp Ile Ile Ile Val Gly
 245 250 255

Arg Gly Leu Tyr Gly Gln Asn Arg Asp Pro Ile Glu Glu Ala Lys Arg
 260 265 270

Tyr Gln Lys Ala Gly Trp Glu Ala Tyr Gln Lys Ile Asn Cys
 275 280 285

<210> 50
 <211> 35
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 KI5

<400> 50
 agagcggcog catgggagaa gtgggaccca caaac

35

<210> 51
 <211> 38
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 KI3

<400> 51
 gtggcggcog ctcaaatgtc gttattgtac caataaac

38

<210> 52
 <211> 1152

<212> DNA

<213> Impatiens balsama

<400> 52

```

atgggagaag tgggacccac aaacegaacc aaaaccaagt tggacaagca acaagaatcc      60
gaaaacaggg ttcctcaoga gccacctcca ttoacactaa gtgacctaa gaaagccatc      120
ccaccccatt gcttcgagcg ctccctcgtg aaatcattct accacgtgat tcacgacatt      180
atcatcctgt octttttcta ctatgtcgcc gccaatata tccccatgct accccaaaac      240
ctccgttacg ttgcatggcc aatttattgg gccatccaag gctgtgtcca acttggata      300
ttggtcttag gccatgaatg eggccaccac gccttcagcg actaccaatg ggtagacgac      360
atggtegggt tcgtctcca ctgtcccaa ttgattccct acttctcatg gaaacatagc      420
cacgctgcc accactcaa cacggcctcc atcgagcgcg acgaggctta cccgcccgcg      480
tacaaaaacg acctgccgtg gttcgccaaa tacctacgca accccgtcgg tcgtttctc      540
atgattttcg gggcgtact gttcggtgg ccgctgtacc ttctgtcaa cgogaacggc      600
cgtctotacg accgcttgc ttcccactac gaccgcgat ccccgatctt caacaaccgc      660
gagaggctgc aagtgatcgc gtcgcagctc gggctcgtct tcgcgtactt tgctctgtac      720
aagatcgcgc tggccaaggg atttgtgtgg ttaatttgtg tgtatggcgt cccgtacgtg      780
atcctcaacg ggcttatcgt cttgatcacg ttctacagc acacgcaccg gaatctgcc      840
cgttacgacc tttccgagtg ggactggctt aggggagccc tgctgactgt ggaccgcgat      900
tacgggatgt tgaataaggt gttccataac gtgacggaca cgcacttggg gcatcatttg      960
ttcagacca tgccacatta tcgogccaag gaggcgaccg aggtgattaa accgatattg      1020
ggagactact ataagtttga cgacactccg tttctcaaag cgttgtggaa ggacatggga      1080
aagtgtattt atgtggagtc ggacgtgcct ggcaagaaca agggagttta ttggtacaat      1140
aacgacattt ga                                                                1152

```

<210> 53

<211> 383

<212> PRT

<213> Impatiens balsama

<400> 53

Met Gly Glu Val Gly Pro Thr Asn Arg Thr Lys Thr Lys Leu Asp Lys
1 5 10 15

Gln Gln Glu Ser Glu Asn Arg Val Pro His Glu Pro Pro Pro Phe Thr
 20 25 30

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

Leu Val Lys Ser Phe Tyr His Val Ile His Asp Ile Ile Ile Leu Ser
 50 55 60

Phe Phe Tyr Tyr Val Ala Ala Asn Tyr Ile Pro Met Leu Pro Gln Asn
65 70 75 80

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

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

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

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

His Ser Asn Thr Ala Ser Ile Glu Arg Asp Glu Val Tyr Pro Pro Ala
145 150 155 160

Tyr Lys Asn Asp Leu Pro Trp Phe Ala Lys Tyr Leu Arg Asn Pro Val
 165 170 175

Gly Arg Phe Leu Met Ile Phe Gly Ala Leu Leu Phe Gly Trp Pro Ser
 180 185 190

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

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

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

Lys Ile Ala Leu Ala Lys Gly Phe Val Trp Leu Ile Cys Val Tyr Gly
 245 250 255

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

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

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

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

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

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

Lys Ala Leu Trp Lys Asp Met Gly Lys Cys Ile Tyr Val Glu Ser Asp
 355 360 365

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

<210> 54

<211> 34

<212> DNA

<213> 人工序列

<220>

<223> 引物 KTI5

<400> 54

aagctcgaga ccgggttggc ggcgtatttg tgtc 34

<210> 55

<211> 38

<212> DNA

<213> 人工序列

<220>

<223> 引物 KTI3

<400> 55

ggtctcgaga tctccaccgc ggacacaata tctggtca 38

<210> 56

<211> 1756

<212> DNA

<213> 人工序列

<220>

<223> TEF/接合酶/XPR 嵌合基因

<400> 56

gaccgggttg gcgcgctatt tgtgtcccaa aaaacagccc caattgcccc aattgacccc 60

aaattgacc agtagcgggc ccaaceccgg cgagagcccc cttcacccca catatcaaac 120

ctcccccgg tcccacactt gccgttaagg gcgtaggta ctgcagtctg gaatctacgc 180

ttgttcagac tttgtactag tttctttgtc tggccatcog ggtaaccat gccggacgca 240

aaatagacta ctgaaaattt ttttgctttg tggttgggac tttagccaag ggtataaaag 300

accaccgtcc ccgaattacc tttctctctc ttttctctct ctcttgtca actcacacce 360

gaaatcgtta agcatttcct tctgagtata agaatcattc aaaggatcca ctagtctag 420
 agcggccgca tgggagaagt gggaccacaca aaccgaacca aaaccaagtt ggacaagcaa 480
 caagaatccg aaaacagggt tctcaccgag ccacctccat tcacactaag tgaccttaag 540
 aaagccatcc caccctattg ctccgagcgc tcctctgtga aatcatteta ccactgtatt 600
 caccgacatta tcctctgtc cttttctac tatgtcggc ccaattacat cccatgcta 660
 ccccaaac tccgttacgt tgcatggcca atttattggg ccatccaagg ctgtgtcaa 720
 cttggtatat tggctctagg ccatgaatgc ggcaccacg ccttcagca ctaccaatgg 780
 gtagacgaca tggctgggtt cgtcctccac tctcccaat tgattcceta cttctcatgg 840
 aaacatagcc accgtcgcca ccactccaac aaggcctcca tcgagcgca cgaggtctac 900
 ccgcccgt acaaaaacga cctgccgtgg ttcgccaat acctacgca cccgtcgtt 960
 cgtttctca tgatttcgg ggcctactg ttcggctggc cgtctacct tctgttcaac 1020
 ggaacggcc gtctctaca ccctctcct tccactacg acccgcaat ccgatctc 1080
 aacaaccgag agaggctgca agtgatecgc tccgaactc ggctcgtctt cgcgtactt 1140
 gtctgtaca agatcgcct ggcacaggga tttgtgtgt taatttgtgt gtatggcgc 1200
 ccgtactga tctcaacgg gcttatcgc ttgatcact tctacagca caagcaccg 1260
 aatctcccc gttacacct tccgagtg gactggctta ggggagcct gtcactgtg 1320
 gaccgcatc accggatggt gaataagtg ttcataacg tgacggcac gcacttggtg 1380
 catcattgt tcaagacct gccacattat cgcgccaagg aggcgaccga ggtgattaaa 1440
 ccgatattg gagaactacta taagtttgac gacactcgt ttctcaaagc gttgtggaag 1500
 gacatgggaa agtgtattta tgtggagtc gactgcctg gcaagaaca gggagtatt 1560
 tggtaacaata acgacattg agcggccgc accgcccgc gagattccg cctcttcggc 1620
 ccgaagoga cccgggtgga cgtctagagg tacctagca ttaacagata gtttccggt 1680
 gataattctc ttaacctccc aactcctt gacataacga tttatgtaac gaaactgaaa 1740
 tttgaccaga tattgt 1756

<210> 57
 <211> 383
 <212> PRT
 <213> 人工序列

<220>
 <223> TEF/接合酶/XPR 嵌合蛋白

<400> 57

Met Gly Glu Val Gly Pro Thr Asn Arg Thr Lys Thr Lys Leu Asp Lys
 1 5 10 15

Gln Gln Glu Ser Glu Asn Arg Val Pro His Glu Pro Pro Pro Phe Thr
 20 25 30

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

Leu Val Lys Ser Phe Tyr His Val Ile His Asp Ile Ile Ile Leu Ser
 50 55 60

Phe Phe Tyr Tyr Val Ala Ala Asn Tyr Ile Pro Met Leu Pro Gln Asn
 65 70 75 80

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

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

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

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

His Ser Asn Thr Ala Ser Ile Glu Arg Asp Glu Val Tyr Pro Pro Ala
145 150 155 160

Tyr Lys Asn Asp Leu Pro Trp Phe Ala Lys Tyr Leu Arg Asn Pro Val
165 170 175

Gly Arg Phe Leu Met Ile Phe Gly Ala Leu Leu Phe Gly Trp Pro Ser
180 185 190

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

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

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

Lys Ile Ala Leu Ala Lys Gly Phe Val Trp Leu Ile Cys Val Tyr Gly
245 250 255

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

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

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

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

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

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

Lys Ala Leu Trp Lys Asp Met Gly Lys Cys Ile Tyr Val Glu Ser Asp
 355 360 365

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

<210> 58

<211> 32

<212> DNA

<213> 人工序列

<220>

<223> 引物 KH5

<400> 58

tagagcggcc gcttaaacca tgaaaaagcc tg 32

<210> 59

<211> 33

<212> DNA

<213> 人工序列

<220>

<223> 引物 KH3

<400> 59

gtggcgcccg ctttaggtac ctcactatc ctt 33

<210> 60

<211> 1026

<212> DNA

<213> 大肠杆菌

<400> 60

atgaaaaagc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga aaagttcgac 60

agcgtctccg acctgatgca gctctcggag ggogaagaat ctctgtcttt cagcttcgat 120

gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt ctacaaagat 180

cgttatgttt atcggcactt tgcattcgcc gogctcccga ttcggaagt gcttgacatt 240
 ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg tgtcacgttg 300
 caagacctgc ctgaaaccga actgcccget gttctgcagc cggtcgogga ggccatggat 360
 gogatcgctg cggccgatct tagccagacg agcgggttcg gccattcgg accgcaagga 420
 atcgggtcaat aactacatg gogtgatttc atatgogga ttgctgatcc ccattgtgat 480
 cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgocagge tctgatgag 540
 ctgatgcttt gggccgagga ctgcccga gtcggcacc tcgtgcacgc ggatttcggc 600
 tccaacaatg tctgacgga caatggccgc ataacagcgg tcattgactg gagcgaggcg 660
 atgttcgggg attccaata cgaggtcgcc aacatctct tctggaggcc gtggttggt 720
 tgtatggagc agcagacgcg ctacttcgag oggaggcacc cggagcttgc aggatcgccg 780
 cggctccggg cgtatatgct ccgattggt cttgaccaac tctatcagag cttggttgac 840
 ggcaatttgg atgatgcagc ttgggcgcag ggctgatgog acgcaatcgt ccgatccgga 900
 gccgggactg tcgggcgtac acaaatgcc cgcagaagcg cggccgtctg gaccgatgce 960
 tgtgtagaag tactcgccga tagtggaac cgcgcccga gcactcgtcc gagggcaaaag 1020
 gaatag 1026

<210> 61
 <211> 341
 <212> PRT
 <213> 大肠杆菌
 <400> 61

Met Lys Lys Pro Glu Leu Thr Ala Thr Ser Val Glu Lys Phe Leu Ile
 1 5 10 15

Glu Lys Phe Asp Ser Val Ser Asp Leu Met Gln Leu Ser Glu Gly Glu
 20 25 30

Glu Ser Arg Ala Phe Ser Phe Asp Val Gly Gly Arg Gly Tyr Val Leu
 35 40 45

Arg Val Asn Ser Cys Ala Asp Gly Phe Tyr Lys Asp Arg Tyr Val Tyr
50 55 60

Arg His Phe Ala Ser Ala Ala Leu Pro Ile Pro Glu Val Leu Asp Ile
65 70 75 80

Gly Glu Phe Ser Glu Ser Leu Thr Tyr Cys Ile Ser Arg Arg Ala Gln
85 90 95

Gly Val Thr Leu Gln Asp Leu Pro Glu Thr Glu Leu Pro Ala Val Leu
100 105 110

Gln Pro Val Ala Glu Ala Met Asp Ala Ile Ala Ala Ala Asp Leu Ser
115 120 125

Gln Thr Ser Gly Phe Gly Pro Phe Gly Pro Gln Gly Ile Gly Gln Tyr
130 135 140

Thr Thr Trp Arg Asp Phe Ile Cys Ala Ile Ala Asp Pro His Val Tyr
145 150 155 160

His Trp Gln Thr Val Met Asp Asp Thr Val Ser Ala Ser Val Ala Gln
165 170 175

Ala Leu Asp Glu Leu Met Leu Trp Ala Glu Asp Cys Pro Glu Val Arg
180 185 190

His Leu Val His Ala Asp Phe Gly Ser Asn Asn Val Leu Thr Asp Asn
195 200 205

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

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

Cys Met Glu Gln Gln Thr Arg Tyr Phe Glu Arg Arg His Pro Glu Leu
 245 250 255

Ala Gly Ser Pro Arg Leu Arg Ala Tyr Met Leu Arg Ile Gly Leu Asp
 260 265 270

Gln Leu Tyr Gln Ser Leu Val Asp Gly Asn Phe Asp Asp Ala Ala Trp
 275 280 285

Ala Gln Gly Arg Cys Asp Ala Ile Val Arg Ser Gly Ala Gly Thr Val
 290 295 300

Gly Arg Thr Gln Ile Ala Arg Arg Ser Ala Ala Val Trp Thr Asp Gly
 305 310 315 320

Cys Val Glu Val Leu Ala Asp Ser Gly Asn Arg Arg Pro Ser Thr Arg
 325 330 335

Pro Arg Ala Lys Glu
 340

<210> 62
 <211> 34
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 KTH5

<400> 62
 ttttagatctc gagaccgggt tggcggcgta tttg

34

<210> 63
 <211> 31
 <212> DNA
 <213> 人工序列

<220>

<223> 引物 KTH3

<400> 63

tttagatctc caccgcgac acaatatctg g 31

<210> 64

<211> 1650

<212> DNA

<213> 人工序列

<220>

<223> TEF::HPT::XPR 融合体

<400> 64

gaccgggttg gcgcggtatt tgtgtccaa aaaacagccc caattgccc aattgacccc 60
 aaattgaccc agtagcgggc ccaaccccg cgagagccc cttcaccca catatcaaac 120
 ctccccgggt tcccacaact gccgtaagg gogtaggta ctgcagtctg gaatctacgc 180
 ttgttcagac tttgtactag tttctttgtc tggccatcc ggtaaccat gccggacgca 240
 aaatagacta ctgaaaattt tttgctttg tggttggac tttagccaag ggtataaaag 300
 accaccgtcc ccgaattacc tttctcttc tttctctct ctcttgtca actcacacc 360
 gaaatcgta agcatttct tctgagtata agaatcttc aaaggatcca ctagttctag 420
 agcggccgct taaacctga aaaagcctga actcacgog acgtctgtcg agaagtttct 480
 gatcgaaaag ttcgacagc tctccgacct gatgcagctc tcggaggcgc aagaatctcg 540
 tgctttcagc ttcgatgtag gagggcgtgg atatgtcctg cgggtaaata gctgcgccga 600
 tggtttctac aaagatcgtt atgtttatcg gcactttgca tcggccgcgc tcccattcc 660
 ggaagtgett gacattgggg aattcagcga gacccagacc tattgcatct ccgcctgtc 720
 acaggggtgc acgttgcaag acctgcctga aaccgaactg cccgtgttc tgcagccggt 780
 cgcggaggcc atggatgca toctgcggc cgatcttagc cagacgagcg ggttcggccc 840
 attoggaccg caaggaatcg gteaatacac tacatggcgt gatttcatat gcgcgattgc 900
 tgatcccat gtgtatcact ggcaactgt gatggacgac accgtcagtg cgtccgtcgc 960
 gcaggctctc gatgagctga tgctttgggc cgaggactgc cccgaagtcc ggcacctcgt 1020

gcacgoggat ttcggctcca acaatgtcct gacggacaat ggccgcataa cagcggatcat 1080
 tgactggagc gaggegatgt tcggggattc ccaatacagag gtcgccaaca tcttctctg 1140
 gaggcctgg ttggcttgta tggagcagca gacgcgctac ttcgagcgga ggcatccgga 1200
 gcttgcagga tcgccggcgc tccgggcgta tatgtctcgc attggtcttg accaactcta 1260
 tcagagcttg gttgacggca atttcgatga tgcagcttgg gcgcagggtc gatgcgacgc 1320
 aatcgtccga tccggagccg ggactgtcgg gcgtacacaa atcgcccgca gaagcgcggc 1380
 cgtctggacc gatggctgtg tagaagtact cgccgatagt ggaaaccgac gccccagcac 1440
 tcgtccgagg gcaaaggaat agtgaggtac ctaaagcggc cgccaccgcg gcccgagatt 1500
 ccggcctctt cggccgcca ggcaccggg tggacgteta gaggtaccta gcaattaaca 1560
 gatagtttgc cgggtataat tctcttaacc tcccacaetc ctttgacata acgatttatg 1620
 taacgaaact gaaatttgac cagatattgt 1650

<210> 65

<211> 341

<212> PRT

<213> 人工序列

<220>

<223> TEF::HPT::XPR 融合体

<400> 65

Met Lys Lys Pro Glu Leu Thr Ala Thr Ser Val Glu Lys Phe Leu Ile
 1 5 10 15

Glu Lys Phe Asp Ser Val Ser Asp Leu Met Gln Leu Ser Glu Gly Glu
 20 25 30

Glu Ser Arg Ala Phe Ser Phe Asp Val Gly Gly Arg Gly Tyr Val Leu
 35 40 45

Arg Val Asn Ser Cys Ala Asp Gly Phe Tyr Lys Asp Arg Tyr Val Tyr
 50 55 60

Ala Gly Ser Pro Arg Leu Arg Ala Tyr Met Leu Arg Ile Gly Leu Asp
 260 265 270

Gln Leu Tyr Gln Ser Leu Val Asp Gly Asn Phe Asp Asp Ala Ala Trp
 275 280 285

Ala Gln Gly Arg Cys Asp Ala Ile Val Arg Ser Gly Ala Gly Thr Val
 290 295 300

Gly Arg Thr Gln Ile Ala Arg Arg Ser Ala Ala Val Trp Thr Asp Gly
 305 310 315 320

Cys Val Glu Val Leu Ala Asp Ser Gly Asn Arg Arg Pro Ser Thr Arg
 325 330 335

Pro Arg Ala Lys Glu
 340

<210> 66
 <211> 401
 <212> DNA
 <213> 解脂耶氏酵母

<400> 66
 cgagtatctg tctgactcgt cattgcccgc tttggagtac gactccaact atgagtgtgc 60
 ttggataact ttgacgatac attottcgtt ggaggctgtg ggtctgacag ctgcgttttc 120
 ggcgcggttg gccgacaaca atatcagctg caacgtcatt gctggctttc atcatgatca 180
 catttttgtc ggcaaaggcg acgccagag agccattgac gttctttcta atttgaccg 240
 atagccgtat agtccagtct atctataagt tcaactaact cgtaactatt accataacat 300
 ataactcaact gccccagata aggttccgat aaaaagttct gcagactaaa tttatttcag 360
 tctcctcttc accaccaaaa tgcctccta cgaagctcga g 401

<210> 67

<211> 568
 <212> DNA
 <213> 解脂耶氏酵母

<400> 67
 atcataattg tcggccgagg tctgtacggc cagaaccgag atcctattga ggaggccaag 60
 cgataccaga aggctggctg ggaggcttac cagaagatta actgttagag gttagactat 120
 ggatatgtca ttttaactgtg tatatagaga gogtgcaagt atggagcgt tgttcagctt 180
 gtatgatggt cagacgacct gtctgatcga gtatgtatga tactgcacaa cctgtgtatc 240
 cgcgatgatct gtccaatggg gcatgttggt gtgtttctcg atacggagat gctgggtaca 300
 agtagctaat acgattgaac tacttatact tatatgaggc ttgaagaaag ctgacttgtg 360
 tatgacttat tctcaactac atccccagtc acaataccac cactgcacta ccactacacc 420
 aaaaccatga tcaaaccacc catggacttc ctggaggcag aagaacttgt tatggaaaag 480
 ctcaagagag agaagccaag atactatcaa gacatgtgto gcaacttcaa ggaggaccaa 540
 gctctgtaca ccgagaaaca ggcctttg 568

<210> 68
 <211> 36
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 YL63

<400> 68
 ttatgatatc gaattaatta acctgcagcc cggggg 36

<210> 69
 <211> 36
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 YL64

<400> 69
 cccccgggct gcaggttaat taattogata tcataa 36

<210> 70
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL65

<400> 70
tacgccgccca acccgtaagt ctcgagcttc gta 33

<210> 71
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL66

<400> 71
tacgaagctc gagacgtacg ggttgccggc gta 33

<210> 72
<211> 30
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL11

<400> 72
ttttccatgg gaacggacca aggaaaaacc 30

<210> 73
<211> 30
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL12

<400> 73
tttgccgccc cctactcttc cttgggacgg 30

<210> 74
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL81

<400> 74
gttatcgcgt cacaagcttc cacacaacgt acg 33

<210> 75
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL82

<400> 75
cgtacgttgt gtggaagctt gtgagcggat aac 33

<210> 76
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL83

<400> 76
atttgaatcg aatcgatgag cctaaaatga acc 33

<210> 77
<211> 33
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL84

<400> 77
ggttcatttt aggctcatcg attcgattca aat 33

<210> 78
<211> 38
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL105

<400> 78
ccaagcacta acctaccggt taaacaccac taaaaccc 38

<210> 79
<211> 38
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL106

<400> 79
gggttttagt ggtgtttaa cggtaggta gtgcttgg 38

<210> 80
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL119

<400> 80
cgggaaacct gtcgtggcgc gccagctgca ttaatg 36

<210> 81
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL120

<400> 81
cattaatgca gctggcgcgc cacgacaggt ttcccg 36

<210> 82
 <211> 34
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 YL121

<400> 82
 tttagcgcgc ctatcacatc acgctctcat caag 34

<210> 83
 <211> 34
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 YL122

<400> 83
 tttcgtacga accaccaccg tcagcccttc tgac 34

<210> 84
 <211> 440
 <212> DNA
 <213> 解脂耶氏酵母

<400> 84
 aaccaccacc gtcagccctt ctgactcagc tattgtagcc accgacacag gcaacagtcc 60
 gtggatagca gaatatgtct tgtecggtcca ttctcacca actttaggcg tcaagtgaat 120
 gttgcagaag aagtatgtgc cttcattgag aatcggtggt gctgatttca ataaagtctt 180
 gagatcagtt tggccagtca tgttgtgggg ggtaattgga ttgagttatc gctacagtc 240
 tgtacaggta tactcgtgc ccactttata ctttttgatt ccgctgcact tgaagcaatg 300
 tcgtttacca aaagtgagaa tgctccacag aacacacccc agggtatggt tgagcaaaaa 360
 ataaacactc cgatacgggg aatcgaaccc cggctctcac ggttctcaag aagtattctt 420
 gatgagagcg tgatgtgata 440

<210> 85	
<211> 35	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 YL114	
<400> 85	
tgatagtatc ttggcgcgcc ttctctctct tgage	35
<210> 86	
<211> 35	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 YL115	
<400> 86	
gctcaagaga gagaaggcgc gccaaagatac tatca	35
<210> 87	
<211> 5218	
<212> DNA	
<213> 人工序列	
<220>	
<223> 用于整合和表达 Δ -5 去饱和酶基因的 5218 bp 片段	
<400> 87	
tatcacatca cgctctcacc aagaatactt cttgagaacc gtggagaccg gggttcgatt	60
ccccgtatcg gagtgtttat tttttgctca accataccct ggggtgtggt ctgtggagca	120
ttctcacttt tggtaaacga cattgcttca agtgcagcgg aatcaaaaag tataaagtgg	180
gcagcgagta tacctgtaca gactgtaggc gataactcaa tccaattacc cccacaaca	240
tgactggcca aactgatctc aagactttat tgaaatcagc aacaccgatt ctcaatgaag	300
gcacatactt cttctgcaac attcacttga cgcctaaagt tggtagaaaa tggaccgaca	360
agacatattc tgctatccac ggactgttgc ctgtgtcggg ggctacaata cgtgagtcag	420

aagggtgac ggtggtggtt cgtacgttgt gtggaagctt gtgagcggat aacaatttca 480
cacaggaaac agctatgacc atgattacgc caagctcgaa attaacccctc actaaagga 540
acaaaagetg gagctccacc goggacacaa tatctgttca aatttcagtt tcgttacata 600
aatcgttatg tcaaaggagt gtgggagggtt aagagaatta tcacoggcaa actatctgtt 660
aattgctagg tacctctaga cgtccacccg ggctcgttgg cggccgaaga ggccggaatc 720
tcgggcccgc gtggcggccg cctactcttc cttgggacgg agtccaagaa cacgcaagtg 780
ctccaatgt gaagcaaatg cttgccaaaa cgtatccttg acaaggatg gaaccttgta 840
ctcgtgcag gtgttcttga tgatggccag aatateggga taatggtgct gogacacgtt 900
ggggaacaga tgggtcacag ccggtagtcc aagctgccag tgatgctggt ccagaggtgc 960
gaatcgtgtg cgtaatcctg cgtagtctcg acctgcatag ctgccagtc cttttggatg 1020
atcccgttct cgtcaggcaa cggccactga acttcctcaa caacgtggtt cgcctggaag 1080
gtcagcgcga gccagtaaga cgacaccatg tccgcgaccg tgaacaagag cagcaccttg 1140
cccaggggca gatactgcag gggaacaatc aggogatacc agacaaagaa agccttgccg 1200
cccagaaca tcacagtgtg ccatgtcgag atgggattga cacgaatage gtcattggtc 1260
ttgacaaagt acaaatgtt gatgtcctga atgcgcacct tgaacgccag cagtccgtac 1320
aggaaaggaa caaacatgtg ctggttgatg tggttgacaa accacttttg gttgggcttg 1380
atacagcgaa catogggctc agacgtcgac acgtcgggat ctgctccagc aatgttggtg 1440
taggggtgat ggccgagcat atgttggtac atccacacca ggtacgatgc tcogttgaaa 1500
aagtcgtgag ttgctcccag aatcttcag acagtggggt tgtgggtcac tgaaaagtga 1560
gacgcacat gaagagggtt gagtccgact tgtgocacg caaatcccat gatgattgca 1620
aacaccacct gaagccatgt gcgttcgaca acgaaaggca caaagagctg cgcgtagtag 1680
gaagcgatca aggatccaaa gataagagcg tatcgtcccc agatctctgg tctattcttg 1740
ggatcaatgt tccgatccgt aaagtagccc togactctcg tcttgatggt tttgtggaac 1800
accgttggtc ccgggaagat gggcagctca ttcgagacca gtgtaccgac atagtacttc 1860

ttcataatgg catctgcage cccaaaocgg tgatacatet caaagaccgg agtaacatct 1920
 eggccagetc egagcaggag agtgtocact ccaccaggat ggoggtcaa gaactttgtg 1980
 acatcgtaca ccotgcogog gatggccaag agtagtctgt ccttggtgtt atggccgccc 2040
 agctcttccc aggtgaaggt ttttccttgg tccgttccca tggatgaatga ttcttatact 2100
 cagaaggaaa tgcttaacga tttcgggtgt gagttgacaa ggagagagag aaaagaagag 2160
 gaaaggtaat tcggggacgg tggcttttta tacccttggc taaagtccca accacaaage 2220
 aaaaaaattt tcagtagtct attttgcgtc cggcatgggt taccoggatg gccagacaaa 2280
 gaaactagta caaagtctga acaagectag attccagact gcagtaccct acgcccttaa 2340
 cggcaagtgt gggaaccogg ggaggtttga tatgtgggtt gaaggggct ctcgccgggg 2400
 ttgggcccgc tactgggtca atttggggtc aattggggca attggggctg tttttggga 2460
 cacaatacgc ccgccaacc ggtctctct gaattctgca gatgggctgc aggaattccg 2520
 tcgtcgcctg agtcgacatc atttatttac cagttggcca caaaccttg acgatctctg 2580
 atgtcccctc cgacatactc ccggccggct ggggtaogtt cgatagcct atoggcctgc 2640
 acaaggtttg ggtccctagc cgataccgca ctacctgagt cacaatcttc ggaggtttag 2700
 tcttccacat agcaogggca aaagtgcgta tatatacaag agcgtttgcc agccacagat 2760
 tttactcca cacaccat cacacataca accacacaca tccacaatgg aaccgaaaac 2820
 taagaagacc aagactgact ccaagaagat tgttcttctc ggoggcgact totgtggccc 2880
 cgaggtgatt gccgaggccg tcaaggtget caagtctgtt gctgaggcct ccggcacoga 2940
 gtttgtgttt gaggacogac tcattggagg agctgccatt gagaaggagg gcgagcccat 3000
 caccgacgt actctcgaca totgccgaaa ggctgactct attatgctcg gtgctgtcgg 3060
 aggcctgcc aacacogtat ggaccctcc cgacggacga accgacgtgc gaccogagca 3120
 gggctctctc aagctcgaa aggacctgaa cctgtacgcc aacctgcgac cctgccagct 3180
 gctgtcggcc aagctcggc atctctcccc catcogaaaac gttgaggga ccgactteat 3240
 cattgtccga gagctcgtcg gaggtatcta ctttggagag cgaaaggagg atgacggatc 3300

tggcgtcgtc tccgacaccg agacctactc cgttcctgag gttgagcgaa ttgcccgaat 3360
 ggcgccttc ctggcccttc agcacaacce cccctctccc gtgtggtctc ttgacaaggc 3420
 caacgtgctg gcctcctctc gactttggcg aaagactgtc actcgagtcc tcaaggacga 3480
 attccccag ctcgagctca accaccagct gatcgactcg gccgccatga tcctcatcaa 3540
 gcagccctcc aagatgaatg gtatcatcat caccaccaac atgtttggcg atatcatctc 3600
 cgacgaggcc tccgtcatcc cggttctct gggctctgctg cctccgcct ctctggettc 3660
 tctgcccagc accaacgagg cgttcggtct gtacgagccc tgtaacggat ctgccccga 3720
 tctcggcaag cagaaggtea accccattgc caccattctg tctgcccca tgatgctcaa 3780
 gttctctott aacatgaagc ccgccggtga cgctgttgag gctgocgtca aggagtccgt 3840
 cgaggtggt atcaactaccg ccgatatcgg aggtctctcc tccacctccg aggtcggaga 3900
 cttgttgcca acaaggtcaa ggagctgctc aagaaggagt aagtcgttcc tacgacgcat 3960
 tgatggaagg agcaaaactga cgcgcctcgc ggttggtcta ccggcagggt ccgctagtgt 4020
 ataagactct ataaaaaggg cctgcctcgt ctaatgaaat gatgatttat aatttacgg 4080
 tntagcaacc ttgactagaa gaagcagatt ggggtgtgtt gtagtgagg acagtggtag 4140
 gttttgaaa cagtcttctt gaaagtgtct tgtctacagt atattcactc ataacctcaa 4200
 tagccaaggg ttagtcggt ttattaaagg aaggagttg tggctgatgt ggatagatat 4260
 ctttaagctg gcgactgcac ccaacgagtg tggtgtagc ttgttactgt atattcggt 4320
 agatatattt tgtggggttt tagtggtgtt taaacggtag gttagtgtt ggtatatgag 4380
 ttgtaggcat gacaatttg aaaggggtg actttgggaa tattgtggga tttcaatacc 4440
 ttagttgta cagggttaatt gttacaaatg atacaaagaa ctgtattctt ttcatttgt 4500
 ttttaattgt tgtatatcaa gtccgttaga cgagctcagt gccttggtt ttggcactgt 4560
 attcatttt tagaggtaca ctacattcag tgaggtatgg taaggttgag ggcataatga 4620
 aggcacctg tactgacagt cacagacctc tcaccagaaa ttttatgaga tatactoggg 4680
 ttcattttag gctcatcgat tcgattcaaa ttaattaatt cgatatcata attgtcgcc 4740

gaggtctgta cggccagaac cgagatccta ttgaggagge caagegatac cagaaggetg 4800
 gctgggagge ttaccagaag attaactggt agaggtaga ctatggatat gtcatttaac 4860
 tgtgtatata gagagcgtgc aagtatggag cgcttgttca gcttgtatga tggtcagaac 4920
 acctgtctga tcgagtatgt atgatactgc acaacctgtg tatccogcatg atctgtccaa 4980
 tggggcatgt tgttgtgttt ctogatacgg agatgctggg tacaagtagc taatacgatt 5040
 gaactactta tacttatatg aggcttgaag aaagctgact tgtgtatgac ttattctcaa 5100
 ctacateccc agtcacaata ccaccaactgc actaccacta caccaaaacc atgatcaaac 5160
 cacccatgga cttcctggag gcagaagaac ttgttatgga aaagctcaag agagagaa 5218

<210> 88
 <211> 33
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 YL61

<400> 88
 acaattccac acaacgtacg agccggaagc ata 33

<210> 89
 <211> 33
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 YL62

<400> 89
 tatgcttccg gctcgtacgt tgtgtggaat tgt 33

<210> 90
 <211> 33
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 YL69

- <400> 90
agcccatctg cagaagcttc aggagagacc ggg 33
- <210> 91
<211> 33
<212> DNA
<213> 人工序列
- <220>
<223> 引物 YL70
- <400> 91
cccggctctct cctgaagctt ctgcagatgg gct 33
- <210> 92
<211> 33
<212> DNA
<213> 人工序列
- <220>
<223> 引物 YL77
- <400> 92
tagtgagggt taattaatcg agcttggcgt aat 33
- <210> 93
<211> 33
<212> DNA
<213> 人工序列
- <220>
<223> 引物 YL78
- <400> 93
attaogccaa gctogattaa ttaaccctca cta 33
- <210> 94
<211> 34
<212> DNA
<213> 人工序列
- <220>
<223> 引物 YL79A

<400> 94		
attcctgcag cccatcgatg cagaattcag gaga		34
<210> 95		
<211> 34		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引物 YL80A		
<400> 95		
tctcctgaat tctgcatcga tgggctgcag gaat		34
<210> 96		
<211> 8894		
<212> DNA		
<213> 人工序列		
<220>		
<223> 用于整合和表达 Δ -6 和 Δ -5 去饱和酶基因和延伸酶基因的 8894 bp 片段		
<400> 96		
tatcacatca cgtctcctc aagaatactt cttgagaacc gtggagaccg gggttcgatt	60	
ccccgtatcg gagtgtttat tttttgctca accatacct ggggtgtgtt ctgtggagca	120	
ttctcacttt tggtaaacga cattgcttca agtgcagcgg aatcaaaaag tataaagtgg	180	
gcagcgagta tacctgtaca gactgtaggc gataactcaa tccaattacc ccccacaaca	240	
tgactggcca aactgatctc aagactttat tgaatcagc aacaccgatt ctcaatgaag	300	
gcacatactt cttctgcaac attcacttga cgcctaaagt tggtagagaaa tggaccgaca	360	
agacatattc tgctatccac ggactgttgc ctgtgtcggg ggctacaata cgtgagtcag	420	
aagggtgac ggtgggtggt cgtacgttgt gtggaattgt gagcggataa caatttcaca	480	
caggaaacag ctatgacat gattacgcca agctcgaat taacctcac taaaggaac	540	
aaaagctgga gctccacgc ggacacaata tetggtcaaa tttcagtttc gttacataaa	600	
tcgttatgtc aaaggagtgt gggagggttaa gagaattatc accggcaaac tatctgttaa	660	

ttgctaggta cctctagaag tccaccggg togttggcg gccgaagagg cgggaatctc 720
 gggccggcgt ggccggcgcg tactgcaact tccttgcctt ctcttggca gcgtcggcct 780
 tggcctgctt ggccaacttg gcgttcttct tgtaaaagt gtagaagaga ccgagcatgg 840
 tccacatgta gaaccaaagc agagccgtga tgaagaagg gtatccgggg cggccaagga 900
 cttcatggc gtacatgtcc caggaagact ggaccgacat catgcagaac tgtgtcatct 960
 gcgagcgcgt gatgtagaac ttgatgaac acacctgctt gaagcccaag gccgacaaga 1020
 agtagtagcc gtacatgac acatggatga acgagttcaa cgcagcagag aagtaggctt 1080
 caccgttggg tgcaacaag gtgaccaacc accagatggg gaagatggag ctgtggtggt 1140
 aaacgtgcaa gaaggagac tggcggttgt tcttcttgag gaccatgac atggtgtoga 1200
 caaactccat gatcttggag aagtagaaga gccagatcat cttggccata ggaagaccct 1260
 tgaaggtatg atcagcagcg ttctcaaaca gtccatagtt ggctgataa gcctcgtaca 1320
 ggatcccacc gcacatgtag gcgctgatcg agaccagaca aaagttgtgc aggagcgaag 1380
 acgtcttgac ctogaaccgc tcaaagttct tcatgatctg catgccaca aagaccgtga 1440
 ccaaataagc gagcacgac aacagcacgt ggaacgggtt catcaacggc agctcacggg 1500
 ccaaaggcga ctccaccgcg accaggaacc cacgcgtgtg atggacaatc gtggggatgt 1560
 acttctcggc ctgggccacc agcgcggcct cgagaggatc gacatagggc gcggcccgga 1620
 caccgatagc ggtggcaagg tcataaaca gatcttggg catctttgat gggaggaatg 1680
 gcgcaatoga ctccatgagg ccgctctaga actagtggat cctttgaatg attcttatac 1740
 tcagaaggaa atgcttaacg atttcgggtg tgagttgaca aggagagaga gaaaagaaga 1800
 ggaaaggtaa ttcggggacg gtggtctttt ataccttgg ctaaagtccc aaccacaaag 1860
 caaaaaatt ttcagtagtc tattttgcgt ccggcatggg ttaccggat ggccagacaa 1920
 agaaactagt acaaagtctg aacaagcgtg gattccagac tgcagtacce tacgccetta 1980
 acggcaagtg tgggaaccgg gggaggttg atatgtggg tgaagggggc tctcggcggg 2040
 gttgggcccg ctactgggtc aatttgggtt caattggggc aattggggct gtttttggg 2100

acacaaatac gccgccaacc cggctctctc tgaagcttgt gagcggataa caatttcaca 2160
caggaaacag ctatgacat gattacgcca agctcgaat taacctcac taaaggaac 2220
aaaagctgga gctccaccgc ggacacaata tctggtcaaa ttcagtttc gttacataaa 2280
tcgttatgtc aaaggagtgt gggaggttaa gagaattatc accggcaaac tatctgttaa 2340
ttgctaggta cctctagacg tccaccggg tcgcttgcg gcogaagagg ccggaatctc 2400
ggcgcgggt ggcggccgcc tactcttctc tgggacggag tccaagaaca cgcaagtgtc 2460
ccaatgtga agcaaatgtc tgccaaaacg taccctgac aaggtatgga acctgtact 2520
cgctgcaggt gttcttgatg atggccagaa tatcgggata atggtgctgc gacacgttg 2580
ggaacagatg gtgcacagcc tggtagttca agctgccagt gatgctggtc cagaggtgcg 2640
aatcgtgtgc gtaactctgc gtagtctcga cctgcatagc tgccagtc ttttgatga 2700
tcccgttctc gtcaggcaac ggccactgaa ctctctcaac aacgtggttc gcctggaagg 2760
tcagcgcag ccagtaagac gacaccatgt ccgcgacct gaacaagagc agcacctgc 2820
ccaggggcag ataactgcagg ggaacaatca ggcgatacca gacaaagaaa gccttgccgc 2880
cccagaacat cacagtgtgc catgtcgaga tgggattgac acgaatagc tcattgtct 2940
tgacaaagta caaatgttg atgtctgaa tgcgcacct gaacgccagc agtccgtaca 3000
ggaaggaac aaacatgtgc tggttgatgt ggttgacaaa ccacttttg ttggcttga 3060
tacgaogaac atcgggctca gacgtcgaca cgtcgggac tgetccagca atgttggtgt 3120
aggggtgatg gccgagcata tgttggtaca tccacaccag gtaogatgct ccgttgaaaa 3180
agtcgtcgt ggtcccaga atctccaga cagtggggtt gtgggtcact gaaaagtgag 3240
acgcatcatg aagagggtg agtccgactt gtgcgcacgc aatcccatg atgattgcaa 3300
acaccacctg aagccatgtg cgttcgacaa cgaaaggcac aaagagctgc gcgtagtag 3360
aagcgatcaa ggatecaaag ataagagcgt atcgtccca gatctctggt ctattcttg 3420
gatcaatgtt ccgatccgta aagtagccct cgactctcgt cttgatggtt ttgtggaaca 3480
ccgttgctc eggaagatg ggcagctcat tcgagaccag tgtaccgaca tagtactct 3540

tcataatggc atctgcagcc ccaaacgcgt gatacatctc aaagaccgga gtaacatctc 3600
ggccagctcc gagcaggaga gtgtccactc caccaggatg gcggetcaag aactttgtga 3660
catcgtacac cctgcoggg atggccaaga gtaggtogtc cttggtgta tgggcccga 3720
gctcttccca ggtgaaggtt tttccttggc ccgttccat ggtgaatgat tcttatactc 3780
agaaggaaat gcttaacgat ttoggggtg agttgacaag gagagagaga aaagaagagg 3840
aaagtaatt cggggacggt ggtctttat acccttggct aaagtccaa ccacaaagca 3900
aaaaaattt cagtagtcta ttttgcgtcc ggcatgggtt acccgatgg ccagacaaag 3960
aaactagtac aaagtotgaa caagcgtaga ttccagactg cagtacceta cgccttaac 4020
ggcaagtgtg ggaaccgggg gaggtttgat atgtggggtg aagggggctc tcgccggggt 4080
tggccccgt actgggtcaa tttgggtca attggggcaa ttggggctgt tttttggac 4140
acaaatacgc cgccaaccg gtctctcctg aattctgcag atgggctgca ggaattcogt 4200
cgtcgctga gtgcacatca tttatttacc agttggccac aaacccttga cgatctcgta 4260
tgtccctcc gacatactcc cggccggctg gggtacgttc gatagccta tcggcatcga 4320
caaggtttg gtcctagcc gataccgcaac tacctgagtc acaatcttcg gaggtttagt 4380
cttccacata gcaaggcga aagtgcgtat atatacaaga gcgtttgcca gccacagatt 4440
ttcactccac acaccacatc acacatacaa ccacacacat ccacaatgga acccgaaact 4500
aagaagacca agactgactc caagaagatt gttcttctcg gcggcgactt ctgtggcccc 4560
gaggtgattg ccgaggccgt caaggtgctc aagtctgtt ctgaggcctc cggcaccgag 4620
tttgtgttg aggaccgact cattggagga gctgccattg agaaggagg cgagcccatac 4680
accgacgcta ctctcgacat ctgccgaaag gctgactcta ttatgctcg tgctgtcgga 4740
ggcgtgcca acaccgtatg gaccactccc gacggacgaa ccgacgtgcg acccgagcag 4800
ggtctctca agctgcgaaa ggacctgaac ctgtacgcca acctgcgacc ctgccagctg 4860
ctgtgcccga agctgcgca tctctcccc atccgaaaog ttgagggcac cgaattcactc 4920
attgtccgag agctcgtcgg aggtatctac tttggagagc gaaaggagga tgaaggatct 4980

ggcgtegett ccgacaccga gacctactcc gttectgagg ttgagogaat tgcccgaatg 5040
 gcegeettcc tggceettca gcacaacccc cctetteccg tgtggtctct tgacaaggcc 5100
 aacgtgetgg cctcctctcg actttggcga aagactgtca ctcgagtcct caaggacgaa 5160
 ttccccagc tcgagctcaa ccaccagctg atcgactcgg ccgccatgat cctcatcaag 5220
 cagccctcca agatgaatgg tatcatcacc accaccaaca tgtttggcga tatcatctcc 5280
 gacgaggcct ccgtcatccc cggttctctg ggtctgctgc cctccgcctc tctggettct 5340
 ctgcccagca ccaacgaggc gttcggctg tacgagccct gtcacggacc tgccccgat 5400
 cteggcaagc agaaggctca cccattgcc accattctgt ctgcgccat gatgctcaag 5460
 ttctctotta acatgaagcc cgcgggtgac gctgttgagg ctgcctcaa ggagtcctc 5520
 gaggtggta tcactaccgc cgatatcgga ggctcttctt caacctcga ggtcggagac 5580
 ttgttgccaa caaggtcaag gagctgtca agaaggagta agtcgtttct acgacgcatt 5640
 gatggaagga gaaaactgac gcgcctcggg gttggtctac cggcagggtc cgctagtgt 5700
 taagactcta taaaaagggc cctgccctgc taatgaaatg atgatttata atttacoggt 5760
 gtagoacct tgactagaag aagcagattg ggtgtgtttg tagtgaggga cagtgttacg 5820
 ttttgaaac agtcttcttg aaagtgtctt gtctacagta tattcactca taacctcaat 5880
 agccaagggt gtagtcggtt tattaaagga agggagttgt ggctgatgtg gatagatacc 5940
 ttttaagctgg cgactgcacc caacgagtgt ggtggtagct tgttactgta tattcggtaa 6000
 gatataatgt gtggggtttt agtgggtttt aaacggtagg ttagtgcttg gtatatgagt 6060
 tgtaggcatg acaatttgga aaggggtgga cttgggaat attgtgggat ttcaatacct 6120
 tagttgtac agggtaattg ttacaaatga tacaagaac tgtatttctt ttcatttgtt 6180
 ttaattggtt gtatatcaag tccgtagac gagctcagtg ccttggtctt tggcactgta 6240
 tttcattttt agaggtacac tacattcagt gaggtatggt aaggttgagg gcataatgaa 6300
 ggcaocctgt actgacagtc acagacctct caocgagaat tttatgagat atactcgggt 6360
 tcattttagg ctcacatgat cagaattcag gagagaccgg gttggggcgg tatttgtgtc 6420

ccaaaaaaca gccccaattg ccccaattga ccccaattg acccagtagc gggcccaacc 6480
 ccggcgagag ccccttcac cccacatata aaacctccc cggttccac acttgcggt 6540
 aaggcgtag ggtactgcag tctggaatct acgcttggtc agactttgta ctagtttctt 6600
 tgtctggcca tccgggtaac ccatgccgga cgcaaatag actactgaaa attttttgc 6660
 tttgtggtt ggactttagc caagggtata aaagaccacc gtccccgaat tacotttcct 6720
 cttcttttct ctctctcctt gtcaactcac acccgaaatc gttaagcatt tcottctgag 6780
 tataagaatc attcaccatg gctgtgtctc ccagtgtgag gacgtttact cgggccgagg 6840
 ttttgaatgc cgaggtctg aatgagggca agaaggatgc cgaggcacc ttcttgatga 6900
 tcatcgacaa caaggtgtac gatgtccgag agttcgtccc tgatcatccc ggtggaagt 6960
 tgattctcac gcacgttggc aaggacggca ctgacgtctt tgacactttt caccocgagg 7020
 ctgcttggga gactcttgc aacttttacg ttggtgatat tgacgagagc gaccgcgata 7080
 tcaagaatga tgactttgag gcgagggtcc gcaagctggg taccttggtc cagtctctt 7140
 gttactaaga ttctccaag gcatactacg cctcaaggt ctggttcaac ctctgcatct 7200
 ggggtttgct gacggtcatt gtggccaagt gggccagac ctgaccctc gccaacgtgc 7260
 tctcggtgct gcttttgggt ctgttctggc agcagtggg atggttggt caccactttt 7320
 tgcatacca ggtcttcag gaccgtttct ggggtgatct ttctgggccc ttcttgggag 7380
 gtgtctgcca gggcttctg tctctgtggt ggaaggacaa gcacaacct caccacgccc 7440
 ccccaacgt ccaaggcgag gatcccgaca ttgacacca cctctgttg acctggagt 7500
 agcatgcgtt ggagatgttc tcggatgtcc cagatgagga gctgaccgc atgtggtgccc 7560
 gtttcatggt cctgaaccag acctggtttt acttccccat tctctggtt gccctctct 7620
 cctggtgcct ccagtcatt ctctttgtgc tgctaacgg tcaggccac aagccctgg 7680
 gcgcgctgt gccatctcg ttggtcagc agctgtcgtc tgcatgac tggacctggt 7740
 acctgcccac catgttctg tcatcaagg atccgtcaa catgctggtg tactttttgg 7800
 tgctgcagge ggtgtcggga aacttgttgg ccatcgtgtt ctgctcaac cacaaggta 7860

tgctgtgat ctgaggagg aggcggtcga tatgatttc ttcacgaagc agatcatcac 7920
 gggctgtgat gtccaccogg gtctatttgc caactggttc acgggtggat tgaactatca 7980
 gatcgagcac cacttgttcc ctctgatgcc tcgccacaac ttttcaaaga tccagcctgc 8040
 tgtcgagacc ctgtgcaaaa agtacaatgt cogataccac accaccggta tgategaggg 8100
 aactgcagag gtctttagcc gtctgaacga ggtctccaag gctacctcca agatgggtaa 8160
 ggcgcagtaa gggcccgcca ccggggcccg agattccggc ctcttcggcc gccaaagcac 8220
 ccgggtggac gtctagaggt acctagcaat taacagatag tttgccggtg ataattctct 8280
 taacctcca cactcctttg acataacgat ttatgtaac aaactgaaat ttgaccagat 8340
 attgtgtccg cgggtggagct ccagcttttg ttcccttag tgagggttaa ttaattcgat 8400
 atcataattg tcggccgagg tctgtacggc cagaaccgag atcctattga ggaggccaag 8460
 cgataccaga aggctggctg ggaggcttac cagaagatta actgtagag gtagactat 8520
 ggatatgtca ttttaactgtg tatatagaga gcgtgcaagt atggagcgt tgttcagctt 8580
 gtatgatggc cagaagacct gtctgatcga gtatgtatga tactgcacaa cctgtgtatc 8640
 cgcattgatc gtccaatggg gcattgtgtt gtgtttctcg ataaggagat gctgggtaca 8700
 agtagctaat acgattgaac tacttatact tatatgagc ttgaagaaag ctgacttgtg 8760
 tatgacttat tctcaactac atccccagtc acaataccac cactgcacta ccactacacc 8820
 aaaacatga tcaaaccacc catggacttc ctggaggcag aagaacttgt tatggaaaag 8880
 ctcaagagag agaa 8894

<210> 97

<211> 31

<212> DNA

<213> 人工序列

<220>

<223> 引物 YL101

<400> 97

gagcttggcg taatcgatgg tcatagctgt t

31

<210> 98
 <211> 31
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 YL102

<400> 98
 aacagctatg accatcgatt acgccaagct c 31

<210> 99
 <211> 36
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 YL103

<400> 99
 atgatgactc aggcgtttaa acgaoggaat tctctg 36

<210> 100
 <211> 36
 <212> DNA
 <213> 人工序列

<220>
 <223> 引物 YL104

<400> 100
 gcaggaattc cgtcgtttaa acgcctgagt catcat 36

<210> 101
 <211> 10328
 <212> DNA
 <213> 人工序列

<220>
 <223> 用于整合和表达 Δ -6, Δ -5 和 Δ -17 去饱和酶基因和延伸酶基因的 10328 bp 片段

<400> 101
 tatcacatca cgtctcctc aagaatactt cttgagaacc gtggagaccg gggttcgatt 60

ccccgtatcg gagtgtttat tttttgctca accatacct ggggtgtggt ctgtggagca 120
ttctcacttt tggtaaacga cattgcttca agtgcagcgg aatcaaaaag tataaagtgg 180
gcagcgagta tacctgtaca gactgtaggc gataactcaa tccaattacc cccacaaca 240
tgactggcca aactgatctc aagactttat tgaaatcagc aacaccgatt ctcaatgaag 300
gcacatactt cttctgcaac attcacttga cgocataaagt tggtagaaaa tggaccgaca 360
agacatattc tgctatccac ggactgttgc ctgtgtcggg ggetacaata cgtgagtcag 420
aagggtgac ggtggtggtt cgtacgttgt gtggaattgt gagcggataa caatttcaca 480
caggaaacag ctatgacct gattacgcca agctcgaat taacctcac taaagggaac 540
aaaagctgga gctccaccgc ggacacaata tctggtcaaa tttcagtttc gttacataaa 600
togttatgtc aaaggagtgt gggaggttaa gagaattatc accggcaaac tatctgttaa 660
ttgctaggta cctctagacg tccaccgggg tcgcttggcg gcogaagagg ccggaatctc 720
gggcccggtt ggcggcgcct tactgcaact tecttgcctt ctcttggca gctcggcct 780
tgccctgctt ggccaacttg cgtttcttcc tgtaaaagt gtagaagaga ccgagcatgg 840
tccacatgta gaaccaaagc agagccgtga tgaagaagg gtatccgggg cggccaagga 900
ccttcatggt gtacatgtcc caggaagact ggaccgacat catgcagaac tgtgtcatct 960
gcgagcgcgt gatgtagaac ttgatgaacg acacctgctt gaagcccaag gccgacaaga 1020
agtagtagcc gtacatgate acatggatga acgagttcaa cgcagcagag aagtaggctt 1080
caccgttggg tgcaacaaag gtgaccaacc accagatggt gaagatggag ctgtggtggt 1140
aaactgcaa gaaggagatc tggcggttgt tcttcttgag gacctgatc atggtgtcga 1200
caactocat gatcttggag aagtagaaga gccagatcat cttggccata ggaagacct 1260
tgaaggtatg atcagcagcg ttctcaaaca gtccatagtt ggcctgataa gcctcgtaca 1320
ggatcccacc gcacatgtag gcctgatcg agaccagaca aaagttgtgc aggagcgaag 1380
acgtcttgac ctgaaccgc tcaagttct tcatgatctg catgccaca aagaccgtga 1440
ccaaataagc gagcacgatc aacagcacgt ggaacgggtt catcaacggc agctcacggg 1500

ccaaagcga ctcaccgcg accaggaacc cacgcgtgtg atggacaatc gtggggatgt 1560
acttctcggc ctgggccacc agcgcggcct cgagaggatc gacatagggc ggggcccgga 1620
cacgatage ggtggcaagg tccataaaca gatcttgogg catctttgat gggaggaatg 1680
gogcaatoga ctccatgcgg cgcctctaga actagtggat cctttgaatg attcttatac 1740
tcagaaggaa atgcttaacg atttcgggtg tgagttgaca aggagagaga gaaaagaaga 1800
ggaaagtaa ttcggggacg gtggtctttt atacccttgg ctaaagccc aaccacaaag 1860
caaaaaatt ttcagtagtc tattttgcgt ccggcatggg ttaccggat ggccagacaa 1920
agaaactagt acaaaagtctg aacaagcgtg gattccagac tgcagtacc tacgccctta 1980
acggcaagtg tgggaaccgg gggaggttg atatgtgggg tgaagggggc tctcggggg 2040
gttgggcccg ctactgggct aatttgggt caattggggc aattggggct gtttttggg 2100
acacaaatac gccccaacc cggctctctc tgaagcttgt gagcggataa caattcaca 2160
caggaaacag ctatgacat gattacgcca agctogaaat taaccctcac taaaggaac 2220
aaaagctgga gctccaccg ggacacaata tctggtcaaa tttcagttc gttacataaa 2280
tcgttatgtc aaaggagtgt gggaggttaa gagaattatc accggcaaac tatctgttaa 2340
ttgctagta cctctagacg tccaccggg tcgcttggcg gccgaagagg ccggaatctc 2400
ggcccggggt ggoggcgcc tactcttctc tgggacggag tccaagaaca cgcaagtgtc 2460
ccaaatgta agcaaatgtc tgccaaaacg tatccttgac aaggtatgga acctgtact 2520
cgctgcaggt gttcttgatg atggccagaa tatcgggata atggtgctgc gacacgttg 2580
ggaacagatg gtgcacagcc tggtagttca agctgccagt gatgctggtc cagaggtgcg 2640
aatcgtgtgc gtaatctgc gtagtctoga cctgcatagc tgcccagtc ttttggatga 2700
tccgttctc gtcaggcaac ggccactgaa ctctctcaac aacgtggttc gcctggaagg 2760
tcagcgcag ccagtaagac gacaccatgt ccgcgaccgt gaacaagagc agcaccttgc 2820
ccaggggcag atactgcagg ggaacaatca ggcgatacca gacaagaaa gccttgccgc 2880
cccagaacat cacagtgtgc catgtcgaga tgggattgac acgaatagcg tcattgtct 2940

tgacaaagta caaaatggtg atgtcctgaa tgcgcacott gaaogccagc agtccgtaca 3000
 ggaaaggaac aaacatgtgc tggttgatgt ggttgacaaa ccacttttgg ttgggcttga 3060
 tacgacgaac atcgggctca gacgtcgaca cgtcgggatc tgctccagca atgttggtgt 3120
 aggggtgatg gccgagcata tgttggtaca tccacaccag gtaogatgct ccgttgaaaa 3180
 agtccgtcgt ggtcccaga atcttccaga cagtggggtt gtgggtcaact gaaaagtgag 3240
 acgcatcatg aagagggtt agtccgactt gtgcgcacgc aaatccatg atgattgcaa 3300
 acaccacctg aagccatgtg cgttcgacaa cgaaaggcac aaagagctgc gcgtagtagg 3360
 aagcgatcaa ggatccaaag ataagagcgt atcgtcccca gatctctggt ctattcttgg 3420
 gatcaatggt ccgatccgta aagtagccct cgaactctgt cttgatggtt ttgtggaaca 3480
 ccgttggctc cgggaagatg ggcagctcat tcgagaccag tgtaccgaca tagtacttct 3540
 tcataatggc atctgcagcc ccaaacoggt gatacatctc aaagaccgga gtaacatctc 3600
 ggccagctcc gagcaggaga gtgtccactc caccaggatg gcggctcaag aactttgtga 3660
 catcgtacac cctgccggg atggccaaga gtaggtctc cttggtgta tgggccgcca 3720
 gctcttccca ggtgaaggt tttccttggc ccgttccat ggtgaatgat tcttatactc 3780
 agaaggaaat gcttaacgat ttcgggtgtg agttgacaag gagagagaga aaagaagagg 3840
 aaagtaatt cggggacggt ggtcttttat acccttggct aaagtccaa ccacaaagca 3900
 aaaaaat ttt cagtagtcta ttttgcgtcc ggcattgggt acccgatgg ccagacaaag 3960
 aaactagtac aaagtctgaa caagcgtaga ttccagactg cagtacccta cgccttaac 4020
 ggcaagtgtg ggaaccggg gaggtttgat atgtgggtg aaggggctc tcgcccgggt 4080
 tggcccctc actgggtcaa tttgggtca attggggcaa ttgggctgt ttttgggac 4140
 acaaatacgc cgccaaccg gtctctctg aattctgcag atgggctgca ggaattcgt 4200
 cgtcgcctga gtcgacatca tttatttacc agttggccac aaaccctga cgtctcgt 4260
 gtcccctcc gacatactcc cggccggctg gggtacgtc gatagccta tcggcatcga 4320
 caaggtttg gtcctagcc gataccgcac tacctgagtc acaatcttc gaggtttagt 4380

ctccacata gcacgggcaa aagtgcgtat atatacaaga gcgtttgcca gccacagatt 4440
 ttcactccac acaccacatc acacatacaa ccacacacat ccacaatgga acccgaact 4500
 aagaagacca agactgactc caagaagatt gttctctctcg gcggcgactt ctgtggcccc 4560
 gaggtgattg ccgaggcogt caaggtgctc aagtctgttg ctgaggcctc cggcaccgag 4620
 tttgtgtttg aggaccgact cattggagga gctgccattg agaaggaggg cgagcccatc 4680
 accgacgcta ctctcgacat ctgcgaaag gctgactcta ttatgotcgg tgctgtcgga 4740
 ggcgctgcca acaccgatg gaccactccc gacggacgaa ccgaogtgcg acccgagcag 4800
 ggtctctca agctgcgaaa ggacctgaac ctgtacgcca acctgcgacc ctgccagctg 4860
 ctgtcgccca agctcgccga tctctcccc atccgaaaag ttgagggcac cgacttcac 4920
 attgtccgag agctcgtcgg aggtatctac tttggagagc gaaaggagga tgaccgatct 4980
 ggcgtcgtt cgcacaccga gacctactcc gttcctgagg ttgagogaat tgcccgaatg 5040
 gccgccttc tggcccttca gcacaacccc cctcttcccg tgttgtctct tgacaaggcc 5100
 aacgtgctgg cctcctctcg actttggcga aagactgtca ctcgagtcct caaggacgaa 5160
 ttccccage tcgagctcaa ccaccagctg atcgactcgg ccgccatgat cctcatcaag 5220
 cagccctcca agatgaatgg tatcatcacc accaccaaca tgtttgcca tatcatctcc 5280
 gacgaggcct ccgctatccc cggttctctg ggtctgtctc cctccgcctc totggcttct 5340
 ctgcccgaca ccaacgagge gttcggctg tacgagccct gtoacggatc tgccccgat 5400
 ctcgcaagc agaaggtcaa cccattgcc accattctgt ctgcgccat gatgetcaag 5460
 ttctctctta acatgaagcc cgcgggtgac gctgttgagg ctgcogtcaa ggagtccgtc 5520
 gaggetggta tcactaccgc cgatatcgga ggctcttctt ccacctccga ggtcggagac 5580
 ttgttgccaa caaggtcaag gagctgtca agaaggagta agtcgtttct acgacgcatt 5640
 gatggaagga gcaaacgtgac gcgcctgcgg gttgtctac cggcagggtc cgctagtgtg 5700
 taagactcta taaaaagggc cctgcctgc taatgaaatg atgatttata atttaccggt 5760
 gtagcaacct tgactagaag aagcagattg ggtgtgtttg tagtggagga cagtgttacg 5820

ttttgaaac agtcttcttg aaagtgtctt gtctacagta tattcaactca taacctcaat 5880
 agccaagggt gtagtcgggt tattaagga agggagttgt ggctgatgtg gatagatatac 5940
 ttaagctgg cgactgcacc caacgagtgt ggtggtagct tgttactgta tattcggtaa 6000
 gatataatgt gtgggtttt agtgggtttt aaacgacgga attcctgcag cccatctgca 6060
 gaattcagga gagaccgggt tggcgcgta tttgtgtccc aaaaaacagc cccaattgcc 6120
 ccaattgacc ccaaattgac ccagtagcgg gcccaacccc ggcgagagcc ccttcaccc 6180
 cacatatcaa acctccccg gttcccacac ttgcggttaa ggcgtaggg tactgcagtc 6240
 tggaatctac gcttgttcag actttgtact agtttctttg tctggccatc cgggtaacc 6300
 atgccggacg caaaatagac tactgaaat tttttgtctt tgtggttggg actttagcca 6360
 aggtataaaa agaccaccgt cccogaatta cctttcctct tctttctct ctctcctgt 6420
 caactcacac ccgaaatcgt taagcatttc cttctgagta taagaatcat tcaccatggc 6480
 tgaggataag accaaggctg agttccctac cctgactgag ctgaagcaact ctatccctaa 6540
 cgcttgettt gagtccaacc toggactctc gctctactac actgcccagag ogatcttcaa 6600
 cgcactctgcc tctgctctc tgctctacgc tgcccgatct actccctca ttgccgataa 6660
 cgttctctc cacgctctgg tttgcgccac ctacatctac gtgcagggtg tcatctctg 6720
 gggtttcttt accgtcggtc acgactgtgg tcaactctgcc ttctccgat accactccgt 6780
 caacttcactc attggctgca tcatgcactc tgccattctg actccctctg agtctcggcg 6840
 agtgaccac cgacaccatc acaagaacac tggcaacatt gataaggacg agatcttcta 6900
 cctcatcgg tccgtcaagg acctccagga cgtgcgacaa tgggtctaca cctcggagg 6960
 tgcttggttt gtctacctga aggtcggata tgctctcga accatgtccc actttgacct 7020
 ctgggaacct ctctgcttc gacgagcctc cgctgtcact gtgtccctcg gactctgggc 7080
 tgccctcttc gctgcctacg cctacctcac atactcctc ggctttgcg tcatgggct 7140
 ctactactat gctcctctct ttgtcttgc ttogttctc gtcattacta ccttcttga 7200
 tcacaacgac gaagctactc cctggtacgg tgactcggag tggacctacg tcaagggcaa 7260

cctgagctcc gtcgaccgat cgtacggagc tttcgtggac aaactgtctc accacattgg 7320
caccaccag gtccatcact tgttcctat cattccccac tacaagctca acgaagccac 7380
caagcacttt gctgcegett accctcacct cgtgagacgt aacgacgagc ccatcattac 7440
tgctttcttc aagaccgctc acctctttgt caactacgga gctgtgcccg agactgctca 7500
gattttcacc ctcaaagagt ctgccgtgc agccaaggcc aagagcgacc accacatca 7560
ccaccattaa gcggccgcca ccgcccgcg agattccggc ctcttcggcc gccaaagcagc 7620
ccgggtggac gtctagaggt acctagcaat taacagatag tttgcggtg ataattctct 7680
taacctocca cactccttg acataacgat ttatgtaacg aaactgaaat ttgaccagat 7740
attgtgtccg cgggtggagct ccagctttg ttcctttag tgagggttaa tttcgagctt 7800
ggcgtaatcg atgcagaatt caggagagac cgggttggcg gcgtatttgt gtoccaaaaa 7860
acagcccaaa ttgcccacat tgaccocaaa ttgaccagc agcgggcccac accccggcga 7920
gagccccctt caccocacat atcaaacctc ccccggttc cacacttgc gttaagggcg 7980
tagggtactg cagtctggaa tctacgctt ttcagacttt gtactagttt ctttgtctgg 8040
ccatccgggt aacctatgcc ggacgcaaaa tagactactg aaaattttt tgetttgtgg 8100
ttgggacttt agccaagggt ataaaagacc accgtccccg aattacctt cctctcttt 8160
tctctctctc cttgtcaact cacaccogaa atogttaagc atttctctc gagtataaga 8220
atcattcacc atggctgctg ctcccagtg gaggacgtt actcgggcg aggttttgaa 8280
tgccgaggct ctgaatgagg gcaagaagga tgccaggca cccttctga tgatcatcga 8340
caacaaggct tacgatgtcc gcgagttcgt cctgatcat ccgggtgaa gtgtgattct 8400
cacgcacgtt ggcaaggacg gcactgacgt cttgacact tttcaccocg aggtgcttg 8460
ggagactctt gccactttt acgttggta tattgacgag agcgcgcgcg atatcaagaa 8520
tgatgacttt gcggccgagg tcccaagct gogtacctg ttcagctctc ttggttacta 8580
cgattcttcc aaggcatact acgccttcaa ggtctcgttc aacctctgca tctggggttt 8640
gtcgacggtc attgtggcca agtggggcca gacctcgacc ctgcacaacg tgetctcggc 8700

tgcgcttttg ggtctgttct ggcagcagtg cggatggttg gotcagact tttgcatca 8760
 ccaggcttcc caggaccgtt tctggggatga tcttttcggc gccttcttgg gaggtgtctg 8820
 ccagggcttc tegtctctgt ggtggaagga caagcacaac actcaccacg cggcccccaa 8880
 cgtccacggc gaggatcccg acattgacac ccacctctg ttgacctgga gtgagcatgc 8940
 gttggagatg ttctcggatg tcccagatga ggagctgacc cgcattgtgt cgcgtttcat 9000
 ggtcctgaac cagacctgtt tttacttccc cattctctcg tttgccctgc tctcctggtg 9060
 cctccagtc attctctttg tgetgcctaa cggtcaggcc cacaagcct cgggcgcgcg 9120
 tgtgccatc tegtgtgtcg agcagctgc gcttgcatg cactggacct ggtacctgc 9180
 caccatgtc ctgttcatca aggatcccgt caacatgctg gtgtactttt tgggtctgca 9240
 ggcggtgtgc ggaaacttgt tggccatcgt gttctcctc aaccacaacg gtatgcctgt 9300
 gatctogaag gaggaggcgg tcgatatgga tttcttcacg aagcagatca tcacgggtcg 9360
 tgatgtccac cgggtctat ttgccaaactg gttcacgggt ggattgaact atcagatcga 9420
 gcaccaactg ttcctctga tgctctgcca caactttca aagatccagc ctgctgtcga 9480
 gacctgtgc aaaaagtaca atgtccgata ccacaccacc ggtatgatcg agggaactgc 9540
 agaggtcttt agcctctga acgaggtctc caaggctacc tccaagatgg gtaaggcga 9600
 gtaagcggcc gccaccgagg cccgagatc cggcotcttc ggccgccaag cgaccoggt 9660
 ggacgtctag aggtacctag caattaacag atagtttgc ggtgataatt ctcttaacct 9720
 cccacactcc tttgacataa cgatttatgt aacgaaactg aaatttgacc agatattgtg 9780
 tccgggtgg agctccagct tttgttccct ttagtgaggg ttaattaatt cgatataata 9840
 attgtogcc gaggtctgta cggccagaac cgatctcta ttgaggagc caagcatac 9900
 cagaaggctg gotgggagc ttaccagaag attaactgtt agaggttaga ctatgatata 9960
 gtcatttaac tgtgtatata gagagcgtgc aagtatggag cgcttgttca gcttgtatga 10020
 tggtcagacg acctgtctga tcagtatgt atgatactgc acaacctgtg tatccgatg 10080
 atctgtccaa tggggcatgt tgttgtgtt ctcgatacgg agatgtggg tacaagtagc 10140

taatacgatt gaactactta tacttatatg aggottgaag aaagctgact tgtgtatgac 10200

ttattctcaa ctacatcccc agtcacaata ccaccactge actaccacta caccaaaacc 10260

atgatcaaac cacccatgga cttcotggag gcagaagaac ttgttatgga aaagctcaag 10320

agagagaa 10328

<210> 102

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 引物 YL1

<400> 102

cagtgcctaaa agccaaggca ctgagctcgt 30

<210> 103

<211> 31

<212> DNA

<213> 人工序列

<220>

<223> 引物 YL2

<400> 103

gacgagctca gtgccttggc ttttggcact g 31

<210> 104

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> 引物 YL3

<400> 104

gtataagaat cattcacat ggatccacta gttota 36

<210> 105

<211> 36

<212> DNA

- <213> 人工序列
- <220>
- <223> 引物 YL4
- <400> 105
tagaactagt g gatccatgg tgaatgattc ttatac 36
- <210> 106
<211> 39
<212> DNA
<213> 人工序列
- <220>
- <223> 引物 YL5
- <400> 106
ccccctoga ggtc gatggt gtcgataagc ttgatatcg 39
- <210> 107
<211> 39
<212> DNA
<213> 人工序列
- <220>
- <223> 引物 YL6
- <400> 107
cgatatcaag cttatcgaca ccatcgacct cgagggggg 39
- <210> 108
<211> 37
<212> DNA
<213> 人工序列
- <220>
- <223> 引物 YL7
- <400> 108
caaccgattt cgacagttaa ttaataattt gaatoga 37
- <210> 109
<211> 37
<212> DNA

<213> 人工序列

<220>

<223> 引物 YL8

<400> 109

togattcaaa ttattaatta actgtcgaaa tcggttg

37

<210> 110

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 引物 YL9

<400> 110

tggtaaataa atgatgtoga ctcaggcgac gacgg

35

<210> 111

<211> 35

<212> DNA

<213> 人工序列

<220>

<223> 引物 YL10

<400> 111

cctgtctgc ctgagtcgac atcatttatt tacca

35

<210> 112

<211> 36

<212> DNA

<213> 人工序列

<220>

<223> 引物 YL23

<400> 112

atgatccac tagttaatta actagagcgg ccgcca

36

<210> 113

<211> 36

<212> DNA

<213> 人工序列	
<220>	
<223> 引物 YL24	
<400> 113	
tgggcgccgc tctagttaat taactagtgg atccat	36
<210> 114	
<211> 1413	
<212> DNA	
<213> 异丝水霉(ATCC #56851)	
<400> 114	
atggccccgc agacggagct ccgccagcgc cagccgcgc tcgccgagac gccggtggcc	60
ggcaagaagg cctttacatg gcaggagtc ggcagcaca acaaggcggc ctoggcctgg	120
atcattatcc gggcaaggt ctacgacgtg accgagtggg ccaacaagca ccccgcggc	180
cgcgagatgg tgctgctgca cggcggtcgc gaggccaccg acacgttcga ctcgtaccac	240
ccgttcagcg acaaggccga gtcgatcttg aacaagtatg agattggcac gttcacgggc	300
ccgtccagat tccgacett caagccggac acgggcttct acaaggagtg ccgcaagcgc	360
gttggcgagt acttcaagaa gaacaacctc catccgcagg acggcttccc gggcctctgg	420
cgcatgatgg tcgtgtttgc ggtcgccggc ctgccttgt acggcatgca cttttcgact	480
atctttgcgc tgcagctcgc ggccgcggcg ctctttggcg tctgccaggc gctgcccgtg	540
ctccaogtca tgcacgactc gtgcacgcg tcgtacacca acatgccgtt cttccattac	600
gtcgtcggcc gctttgccat ggactggttt gccggcggtc cgtatggtgc atggtcaac	660
cagcaogtgc tgggccacca catctacag aacgtcggg gctcggacc ccgatcttccg	720
gtcaacatgg aaggcgacat ccgccgcac gtgaaccgcc aggtgttcca gcccatgtac	780
gcattocagc acatctacct tccgcgcgc tatggcgtgc ttggcctcaa gttccgcac	840
caggacttca ccgacacgtt cggctcgcac acgaacggcc cgtatccggt caaccgcac	900
gcgctctcga cgtggatggc catgatcagc tccaagtctg tctgggcctt ctaccgctg	960
taccttccgc ttgcctgct ccagatgccc atcaagacgt accttgcgat cttcttctc	1020

gccgagtttg tcacgggctg gtacctgcg tcaacttcc aagtaagcca tgtctogacc 1080
 gagtgeggct acccatggcg cgacgaggcc aagatggcgc tccaggacga gtgggcagtc 1140
 tcgcaggtea agacgtcggc cgactacgcc catggctcgt ggatgacgac gttccttgcc 1200
 ggcgcgctca actaccaggt cgtgcaccac ttgttcccca gcgtgtcgca gtaccactac 1260
 ccggcgatcg cgcctcatcat cgtcgacgtc tgcaaggagt acaacatcaa gtaaccatc 1320
 ttgcggact ttacggggc gttcgttgcc cacttgaagc acctccgcaa catgggccag 1380
 caggcatcg ccgccacgat ccacatgggc taa 1413

<210> 115

<211> 470

<212> PRT

<213> 异丝水霉(ATCC #56851)

<400> 115

Met Ala Pro Gln Thr Glu Leu Arg Gln Arg His Ala Ala Val Ala Glu
 1 5 10 15

Thr Pro Val Ala Gly Lys Lys Ala Phe Thr Trp Gln Glu Val Ala Gln
 20 25 30

His Asn Thr Ala Ala Ser Ala Trp Ile Ile Ile Arg Gly Lys Val Tyr
 35 40 45

Asp Val Thr Glu Trp Ala Asn Lys His Pro Gly Gly Arg Glu Met Val
 50 55 60

Leu Leu His Ala Gly Arg Glu Ala Thr Asp Thr Phe Asp Ser Tyr His
 65 70 75 80

Pro Phe Ser Asp Lys Ala Glu Ser Ile Leu Asn Lys Tyr Glu Ile Gly
 85 90 95

Thr Phe Thr Gly Pro Ser Glu Phe Pro Thr Phe Lys Pro Asp Thr Gly

100	105	110
Phe Tyr Lys Glu Cys Arg Lys Arg Val Gly Glu Tyr Phe Lys Lys Asn		
115	120	125
Asn Leu His Pro Gln Asp Gly Phe Pro Gly Leu Trp Arg Met Met Val		
130	135	140
Val Phe Ala Val Ala Gly Leu Ala Leu Tyr Gly Met His Phe Ser Thr		
145	150	155
Ile Phe Ala Leu Gln Leu Ala Ala Ala Ala Leu Phe Gly Val Cys Gln		
165	170	175
Ala Leu Pro Leu Leu His Val Met His Asp Ser Ser His Ala Ser Tyr		
180	185	190
Thr Asn Met Pro Phe Phe His Tyr Val Val Gly Arg Phe Ala Met Asp		
195	200	205
Trp Phe Ala Gly Gly Ser Met Val Ser Trp Leu Asn Gln His Val Val		
210	215	220
Gly His His Ile Tyr Thr Asn Val Ala Gly Ser Asp Pro Asp Leu Pro		
225	230	235
Val Asn Met Asp Gly Asp Ile Arg Arg Ile Val Asn Arg Gln Val Phe		
245	250	255
Gln Pro Met Tyr Ala Phe Gln His Ile Tyr Leu Pro Pro Leu Tyr Gly		
260	265	270
Val Leu Gly Leu Lys Phe Arg Ile Gln Asp Phe Thr Asp Thr Phe Gly		
275	280	285
Ser His Thr Asn Gly Pro Ile Arg Val Asn Pro His Ala Leu Ser Thr		

290	295	300
Trp Met Ala Met Ile Ser Ser Lys Ser Phe Trp Ala Phe Tyr Arg Val		
305	310	315 320
Tyr Leu Pro Leu Ala Val Leu Gln Met Pro Ile Lys Thr Tyr Leu Ala		
	325	330 335
Ile Phe Phe Leu Ala Glu Phe Val Thr Gly Trp Tyr Leu Ala Phe Asn		
	340	345 350
Phe Gln Val Ser His Val Ser Thr Glu Cys Gly Tyr Pro Cys Gly Asp		
	355	360 365
Glu Ala Lys Met Ala Leu Gln Asp Glu Trp Ala Val Ser Gln Val Lys		
	370	375 380
Thr Ser Val Asp Tyr Ala His Gly Ser Trp Met Thr Thr Phe Leu Ala		
	385	390 395 400
Gly Ala Leu Asn Tyr Gln Val Val His His Leu Phe Pro Ser Val Ser		
	405	410 415
Gln Tyr His Tyr Pro Ala Ile Ala Pro Ile Ile Val Asp Val Cys Lys		
	420	425 430
Glu Tyr Asn Ile Lys Tyr Ala Ile Leu Pro Asp Phe Thr Ala Ala Phe		
	435	440 445
Val Ala His Leu Lys His Leu Arg Asn Met Gly Gln Gln Gly Ile Ala		
	450	455 460
Ala Thr Ile His Met Gly		
465	470	

<210> 116

<211> 30	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 YL13A	
<400> 116	
ttggatcgc agacggagct ccgccagcgc	30
<210> 117	
<211> 36	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 YL14	
<400> 117	
cccttaatta aattagccca tgtggatcgt ggcggc	36
<210> 118	
<211> 1329	
<212> DNA	
<213> <i>Isochrysis galbana</i> CCMP1323	
<400> 118	
atggtggcag gcaaatcagg cgctgcggcg cacgtgactc acagctcgac attgccccgt	60
gagtaccatg gcgcgaccaa cgactcgcgc tctgaggcgg ccgacgtcac cgtctctagc	120
atcgatgctg aaaaggagat gatcatcaac ggccgcgtgt atgacgtgtc gtcatttgtg	180
aagcggcacc caggtggctc ggtgatcaag ttccagctgg gcgccgacgc gagcgcgcgc	240
tacaacaact ttcagtcgc ctccaagaag gcggacaaga tgctgtatto gctcccgtcc	300
cgcccgcccg aggcgggcta cgcacaggac gacatctccc gcgactttga gaagctgcgc	360
ctcgagctga aggaggaggc ctacttcgag cccaacctgg tgcacgtgag ctacaggtgt	420
gtggaggttc ttgcatgta ctgggtggc gtccagctca tctggtccgg gtactggttc	480
ctcggcgcga togtggccgg cattgcgcag ggccgctgcg gctggctcca gcatgagggt	540
gggcactact cgctcaccgg caacatcaag atogaccgac atctgcagat ggccatctat	600

gggotttgct gggcatgtc gggetgetac tggcgcaacc agcacaaca gcaccacgcc 660
 acgccgcaga agctcgggac cgaccccgac ctgcagaoga tgcogetggt ggccttccac 720
 aagatcgtcg gcgccaaggc gcgaggcaag ggcaaggcgt ggetggogtg gcaggcgcgc 780
 ctcttctttg gcgggatcat ctgetcgetc gtctctttog gctggcagtt cgtgctccac 840
 cccaaccacg cgetcgcgct gcacaatcac ctggagctcg cgtacatggg cotgcggtac 900
 gtgctgtggc acctggcett tggccacctc gggctgetga gctcgetccg cctgtacgcc 960
 ttttacgtgg ccgtgggagg cacctacatc ttcaccaact tcgcoctctc gcacaccoac 1020
 aaggacgtcg tccgcccac caagcacatc tcgtgggcac tctactcggc caaccacacg 1080
 accaactgct cagactcgc ctttgtcaac tggatgatgg cctacctcaa cttccagatc 1140
 gagcaccacc tcttccgctc gatgccgcag tacaaccacc ccaagatcgc cccgagggtg 1200
 cgcgcgtctc tcgagaagca cggggctcag tatgacgtcc ggccatacct ggagtgtttt 1260
 cgggtcacgt acgtcaacct gctcgcgta ggcaaccgg agcactccta ccacgagcac 1320
 acgcaactag 1329

<210> 119
 <211> 442
 <212> PRT
 <213> *Isochrysis galbana* CCMP1323

<400> 119

Met Val Ala Gly Lys Ser Gly Ala Ala Ala His Val Thr His Ser Ser
 1 5 10 15

Thr Leu Pro Arg Glu Tyr His Gly Ala Thr Asn Asp Ser Arg Ser Glu
 20 25 30

Ala Ala Asp Val Thr Val Ser Ser Ile Asp Ala Glu Lys Glu Met Ile
 35 40 45

Ile Asn Gly Arg Val Tyr Asp Val Ser Ser Phe Val Lys Arg His Pro

245	250	255
Trp Gln Ala Pro Leu Phe Phe Gly Gly Ile Ile Cys Ser Leu Val Ser 260	265	270
Phe Gly Trp Gln Phe Val Leu His Pro Asn His Ala Leu Arg Val His 275	280	285
Asn His Leu Glu Leu Ala Tyr Met Gly Leu Arg Tyr Val Leu Trp His 290	295	300
Leu Ala Phe Gly His Leu Gly Leu Leu Ser Ser Leu Arg Leu Tyr Ala 305	310	315
Phe Tyr Val Ala Val Gly Gly Thr Tyr Ile Phe Thr Asn Phe Ala Val 325	330	335
Ser His Thr His Lys Asp Val Val Pro Pro Thr Lys His Ile Ser Trp 340	345	350
Ala Leu Tyr Ser Ala Asn His Thr Thr Asn Cys Ser Asp Ser Pro Phe 355	360	365
Val Asn Trp Trp Met Ala Tyr Leu Asn Phe Gln Ile Glu His His Leu 370	375	380
Phe Pro Ser Met Pro Gln Tyr Asn His Pro Lys Ile Ala Pro Arg Val 385	390	395
Arg Ala Leu Phe Glu Lys His Gly Val Glu Tyr Asp Val Arg Pro Tyr 405	410	415
Leu Glu Cys Phe Arg Val Thr Tyr Val Asn Leu Leu Ala Val Gly Asn 420	425	430
Pro Glu His Ser Tyr His Glu His Thr His		

435	440	
<210> 120		
<211> 36		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引物 YL19A		
<400> 120		
tttggatccg gcaggcaaat caggcgtgc ggcgca		36
<210> 121		
<211> 36		
<212> DNA		
<213> 人工序列		
<220>		
<223> 引物 YL20		
<400> 121		
ccttaattaa ctagtgcgtg tgctcgtggt aggagt		36
<210> 122		
<211> 1320		
<212> DNA		
<213> <i>Thraustochytrium aureum</i> (ATCC #34304)		
<400> 122		
atgggacgcg gcggcgaagg tcaggtgaac agcgcgcagg tggcacaagg cggcgcggga		60
acgcgaaaga cgatcctgat cgagggcgag gtctacgatg tcaccaactt taggcacccc		120
ggcgggtcga tcatcaagtt tctcagacc gacggcaccg aggctgtgga cgcgacgaac		180
gcgtttcgcg agtttcactg ccggtcgggc aaggcggaaa agtacctcaa gacctgccc		240
aagctcggcg cgcgagcaa gatgaagttt gacgccaaagg agcaggcccg gcgcgacgcg		300
atcacgcgag actacgtcaa gctgcgcgag gagatggtgg ccgagggcct cttcaagccc		360
gcgccctcc acattgteta caggtttgcg gagatcgcag ccctgttgc ggccctcgttc		420
tacctgtttt cgatgcgcgg aaacgtgttc gccacgctcg cggccatcgc agtcgggggc		480

atcgcgcagg gccgctgcgg ctggctcatg caogagtgcg gacacttctc gatgaccggg 540
 tacatccgcg ttgacgtgcg cctgcaggag ctgggtgtacg gcgtgggggtg ctcgatgtcg 600
 gogagctggt ggccggttca gcacaacaag caccacgcga ccccgagaa actcaagcac 660
 gaogtcgacc tcgacacocct gccgctcgtt gcgttcaacg agaagatcgc cgccaaggtg 720
 cgccccggct cgttccaggc caagtggctc tcggcgcagg cgtacatttt tgcgccggtg 780
 tctgcttcc tggttggtct ettctggacc ctgtttctgc acccgogcca catgccgcgc 840
 aogagccact ttgctgagat ggccgcgcgc gcggtgcgcg tcgtgggctg ggoggcgcgc 900
 atgcactcgt tcgggtacag cgggagcgac tcgttcggtc tctacatggc cacctttggc 960
 tttggtgca cctacatctt caccaacttt gcggtcagcc acaogcacct ogagtcacc 1020
 gagccggaog agttcctgca ctgggtcgag tacgccgcgc tgcacacgac caacgtgtcc 1080
 aacgactcgt ggttcacac ctgggtgatg tcgtacctca actttcagat ogagcaccac 1140
 ctctttccgt cgtgcoccca gctcaacgcc ccgcgcgctg ccccgccgt ccgcgcctc 1200
 ttcgagaagc acggcatgac ttaogacgag cgcccgtaac ttaccgcgt tggcgacaog 1260
 tttgccaacc tgcacgcct gggccaaaac gcgggccagg cggcgcccaa ggccgcttag 1320

<210> 123

<211> 439

<212> PRT

<213> *Thraustochytrium aureum* (ATCC #34304)

<400> 123

Met Gly Arg Gly Gly Glu Gly Gln Val Asn Ser Ala Gln Val Ala Gln
 1 5 10 15

Gly Gly Ala Gly Thr Arg Lys Thr Ile Leu Ile Glu Gly Glu Val Tyr
 20 25 30

Asp Val Thr Asn Phe Arg His Pro Gly Gly Ser Ile Ile Lys Phe Leu
 35 40 45

Thr Thr Asp Gly Thr Glu Ala Val Asp Ala Thr Asn Ala Phe Arg Glu
50 55 60

Phe His Cys Arg Ser Gly Lys Ala Glu Lys Tyr Leu Lys Ser Leu Pro
65 70 75 80

Lys Leu Gly Ala Pro Ser Lys Met Lys Phe Asp Ala Lys Glu Gln Ala
85 90 95

Arg Arg Asp Ala Ile Thr Arg Asp Tyr Val Lys Leu Arg Glu Glu Met
100 105 110

Val Ala Glu Gly Leu Phe Lys Pro Ala Pro Leu His Ile Val Tyr Arg
115 120 125

Phe Ala Glu Ile Ala Ala Leu Phe Ala Ala Ser Phe Tyr Leu Phe Ser
130 135 140

Met Arg Gly Asn Val Phe Ala Thr Leu Ala Ala Ile Ala Val Gly Gly
145 150 155 160

Ile Ala Gln Gly Arg Cys Gly Trp Leu Met His Glu Cys Gly His Phe
165 170 175

Ser Met Thr Gly Tyr Ile Pro Leu Asp Val Arg Leu Gln Glu Leu Val
180 185 190

Tyr Gly Val Gly Cys Ser Met Ser Ala Ser Trp Trp Arg Val Gln His
195 200 205

Asn Lys His His Ala Thr Pro Gln Lys Leu Lys His Asp Val Asp Leu
210 215 220

Asp Thr Leu Pro Leu Val Ala Phe Asn Glu Lys Ile Ala Ala Lys Val
225 230 235 240

Arg Pro Gly Ser Phe Gln Ala Lys Trp Leu Ser Ala Gln Ala Tyr Ile
 245 250 255

Phe Ala Pro Val Ser Cys Phe Leu Val Gly Leu Phe Trp Thr Leu Phe
 260 265 270

Leu His Pro Arg His Met Pro Arg Thr Ser His Phe Ala Glu Met Ala
 275 280 285

Ala Val Ala Val Arg Val Val Gly Trp Ala Ala Leu Met His Ser Phe
 290 295 300

Gly Tyr Ser Gly Ser Asp Ser Phe Gly Leu Tyr Met Ala Thr Phe Gly
 305 310 315 320

Phe Gly Cys Thr Tyr Ile Phe Thr Asn Phe Ala Val Ser His Thr His
 325 330 335

Leu Asp Val Thr Glu Pro Asp Glu Phe Leu His Trp Val Glu Tyr Ala
 340 345 350

Ala Leu His Thr Thr Asn Val Ser Asn Asp Ser Trp Phe Ile Thr Trp
 355 360 365

Trp Met Ser Tyr Leu Asn Phe Gln Ile Glu His His Leu Phe Pro Ser
 370 375 380

Leu Pro Gln Leu Asn Ala Pro Arg Val Ala Pro Arg Val Arg Ala Leu
 385 390 395 400

Phe Glu Lys His Gly Met Ala Tyr Asp Glu Arg Pro Tyr Leu Thr Ala
 405 410 415

Leu Gly Asp Thr Phe Ala Asn Leu His Ala Val Gly Gln Asn Ala Gly
 420 425 430

Gln Ala Ala Ala Lys Ala Ala
435

<210> 124
<211> 30
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL15

<400> 124
ttttccatgg gacgcggggg cgaaggtcag 30

<210> 125
<211> 37
<212> DNA
<213> 人工序列

<220>
<223> 引物 YL16B

<400> 125
ttttgcggcc gctaagcggc cttggccgcc gcctggc 37

<210> 126
<211> 10
<212> DNA
<213> 解脂耶氏酵母

<220>
<221> misc_feature
<222> (8)..(8)
<223> n 是 a, c, g, 或 t

<400> 126
mammatgnhs 10

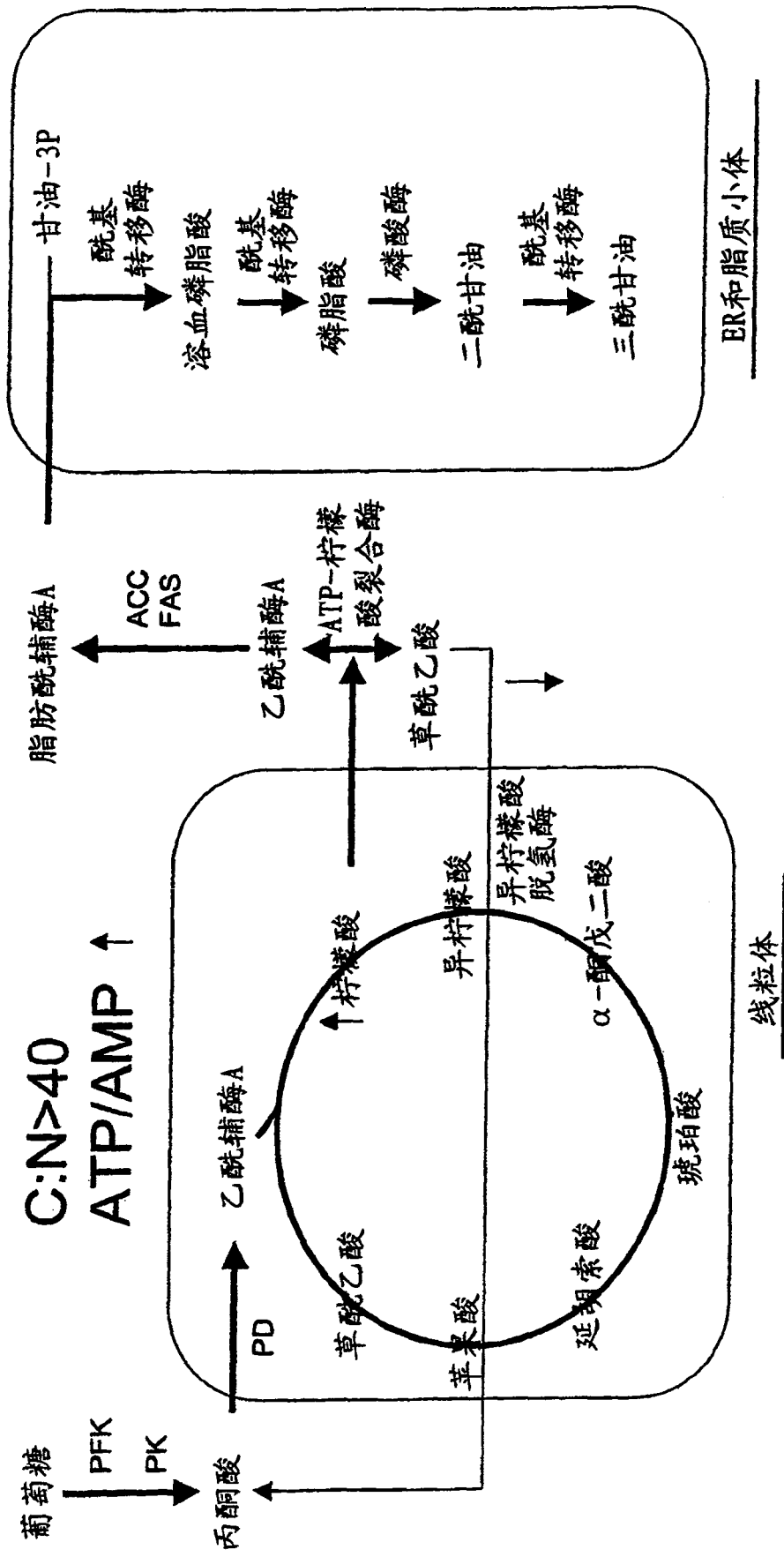


图 1

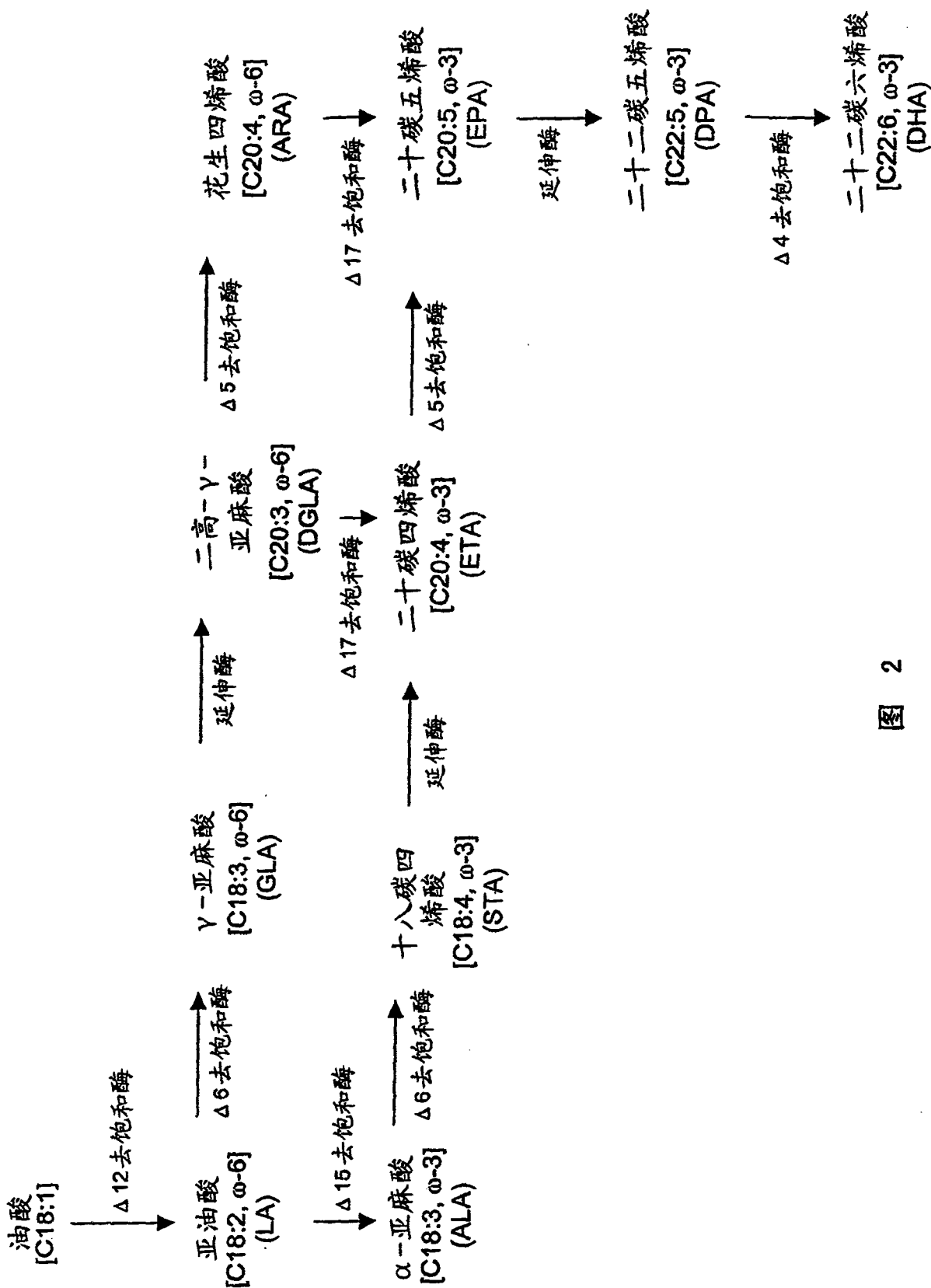


图 2

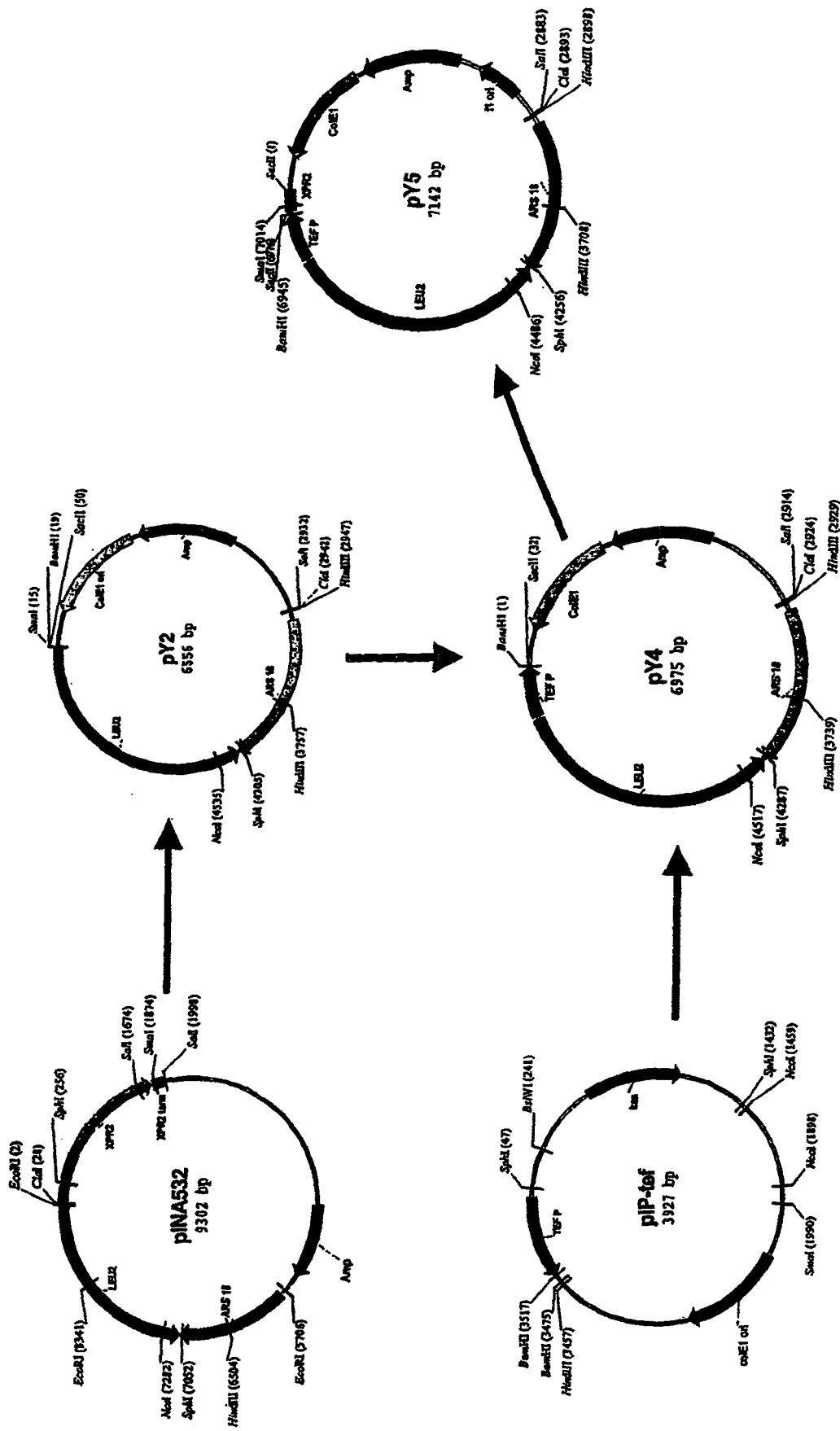


图 3

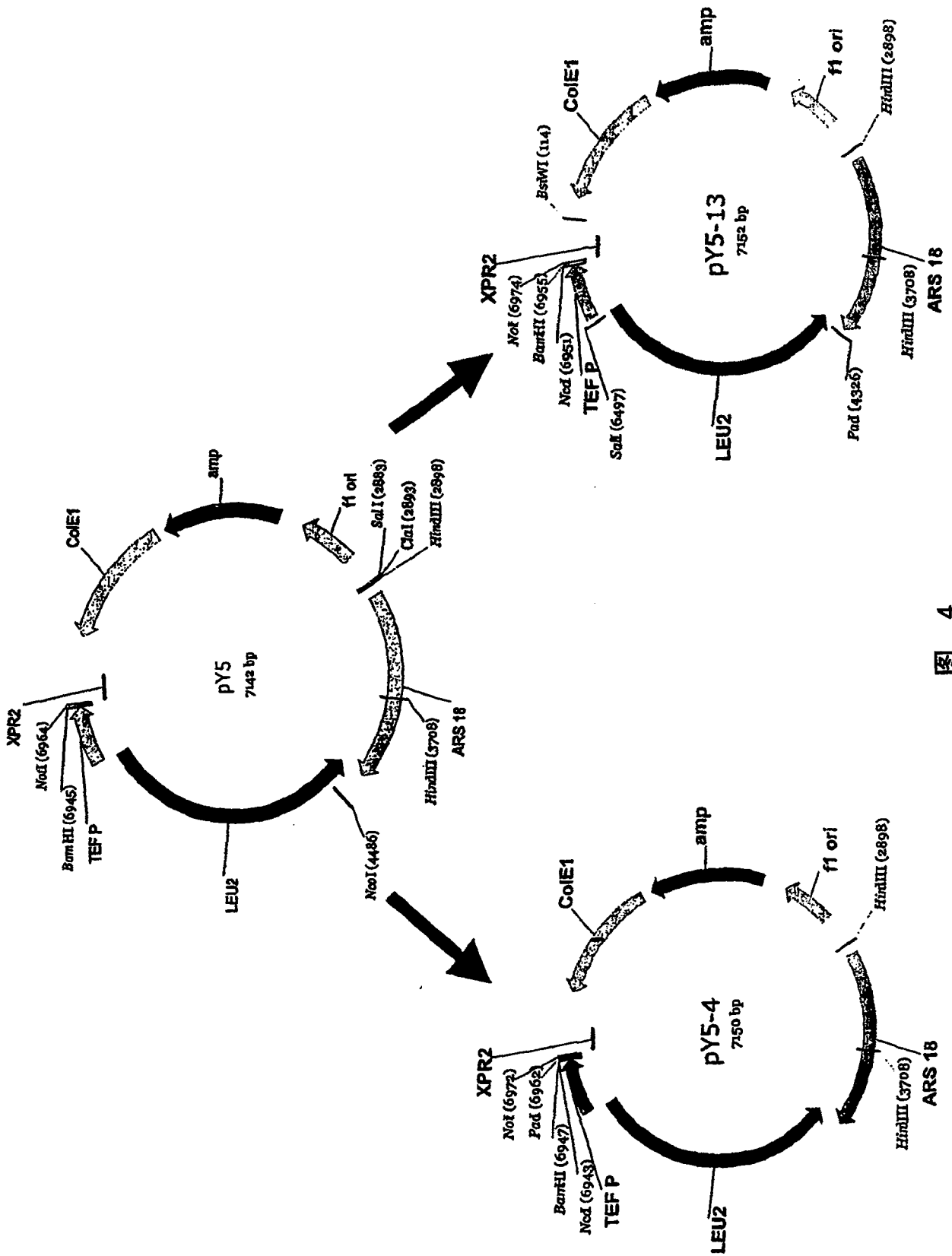


图 4

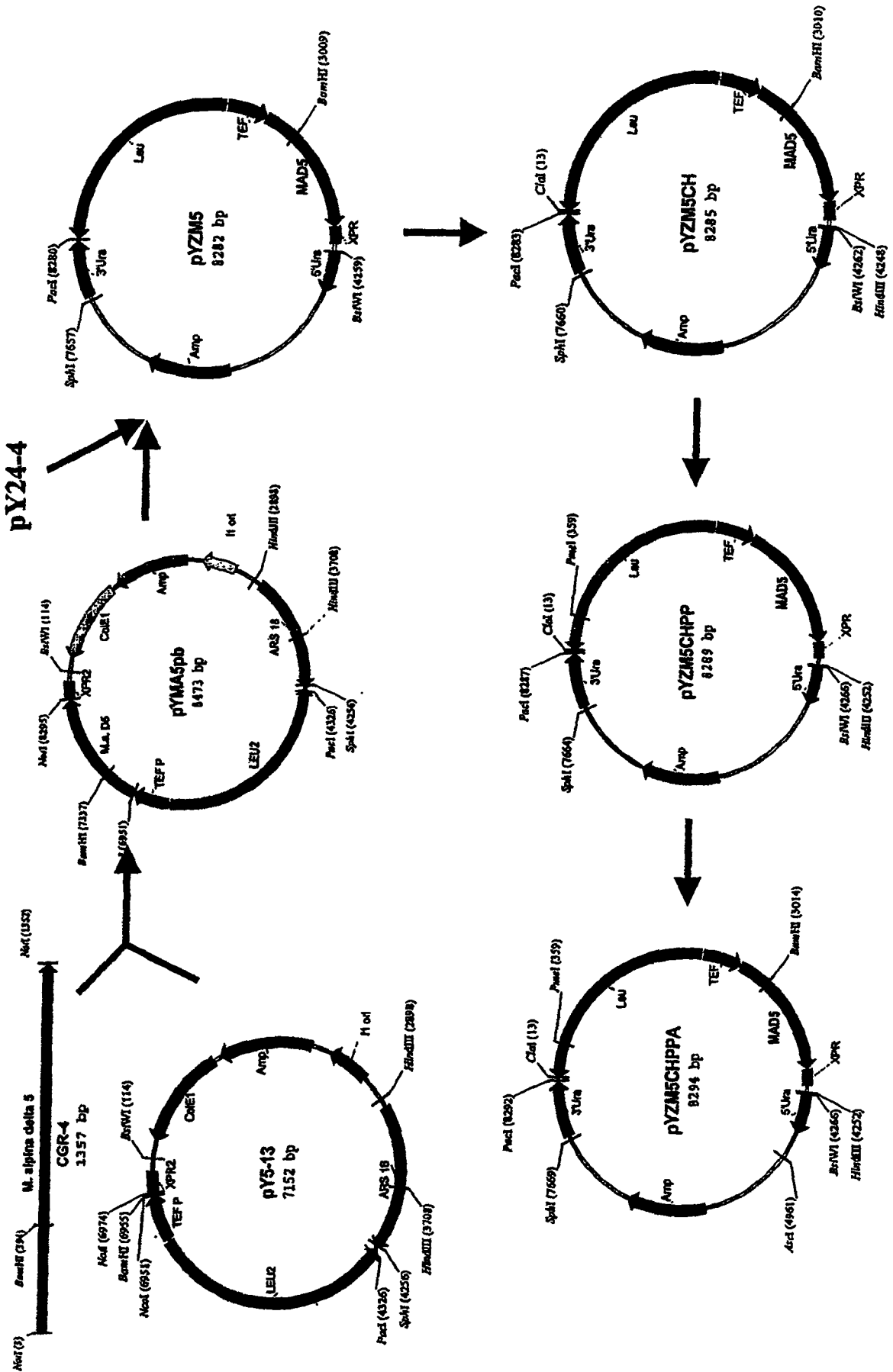


图 5

(SEQ ID NO:5) ATGACTGAGGATAAGACGAAGGTCGAGTCCCGACCGTCAAGCACTCGATCCCGAACGGCT
 (SEQ ID NO:6) M T E D K T K V E F P T L T E L K H S I P N A 70
 C T C G T C T T T
 GCTTGAGTCGAACCTCGGCCCTCTCGCTCTACTACAGGCCCGCGGATCTTCAACGGTCGGCCCTCGGC 140
 C F E S N L G L S L Y Y T A R A I F N A S A S A
 T T C A T T A T T
 GCGCTGCTCTACGGCGCGCTCGACCGCTTTCATTGCCGATAACGTTCTGCTCCACGGCTCGTTGTC 210
 A L L Y A A R S' T P F I A D N V L L H A L V C
 T T C T T G
 GCCACCTACATCTACGTCAGGGCGTCACTCTCTGGGGCTTCTTACAGGTCGGCCACGACTGGGCCACT 280
 A T Y I Y V Q G V I F W G F T V G H D C G H
 T C A T C T C T
 CGGCCTCTCGCGCTACCACAGCGTCAACTTTATCATCGGCTGCATCATGCACCTCTCGGATTTGACGCC 350
 S A F S R Y H S V N F I I G C I M H S A I L T P
 C T C A C A T T
 GTTCGAGAGCTGGCGGTGACGCACCGCCACCACCAAGAACACCGGGCAACATTGATAAGGACGAGATC 420
 F E S W R V T H R H H K N T G N I D K D E I
 C T T C A A
 TTTTACCGCACCGGTCCGGTCAAGGACCTCCAGGACGTGCGCCAATGGTCTACACGCTCGCGGTGCGT 490
 F Y P H R. S V K D L Q D V R Q W V Y T L G G A
 C A T A C T C
 GGTTCCTACTTGAAGGTGGGTATGCCCGCGCACGATGAGCCACTTTCACCGCTGGGACCCGCTCCT 560
 W F V Y L K V G Y A P R T M S H F D P W D P L L
 G A A C C T C A T T C C
 CCTTCGCGCGCTCGGCCGTCATCGTGTGCTCGGCGTCTGGGCCCTTCTTCGCGCGGTACGGGTAC 630
 L R R A S A V I V S L G V W A A F F A A Y A Y

图 6A

CTCACATACTGGCTTGGCGTCAV M G L Y Y A P L F V F A S ^{T T} 700
 L T Y S L G F A V M G L Y Y A P L F V F A S
 TCCTCGTCAATTACGACCTTCTTGGCACCACGAGCGACCGCGTGGTACGGCGACTCGGAGTGGAC ^{T T C} 770
 F L V I T T F L H H N D E A T P W Y G D S E W T
 C ^{GAGCTC} A A T ^{TCT} 840
 GTACGTCAAGGGCAACCTCTCGAGCGTCCGACCCGCTCGTACGGCGGTTCTGGACAACCTGAGCCACCAC
 Y V K G N L S S V D R S Y G A F V D N L S H H
 ATTGGCACGCACCAGGTCACCACCTTGTCCGATCATCCGCACTACAAGCTCAACGACCCACCAAGC ^{T C} 910
 I G T H Q V H H L F P I I P H Y K L N E A T K
^{T T} A A C T ^T 980
 ACTTTGGCGCGGTACCCGACCTCGTGGCAGGAAACGACGAGCCCATCATCAGGCGCTTCTTCAAGAC
 H F A A A Y P H L V R R R N D E P I I T A F F K T
 CGCGCACCTCTTTGTCAACTAGGGCGTGTGCCGAGACGGCGCAGATCTTCACGCTCAAAGAGTCGGCC ^{T T C} 1050
 A H L F V N Y G A V P E T A Q I F T L K E S A
 T A ^{AGC} 1077
 GCGCGCGCAAGGCCAAGTCGGACTAA
 A A A K A K S D .

127/1077

图 6B

[SEQ ID NO:126]

⁻⁴**C****AAA****A****T****G****N****C****G**⁺⁶
A **CC** **TC**
A

-4: A: 23/79, C: 41/79, G: 3/79, T: 12/79

-3: A: 61/79, C: 7/79, G: 7/79, T: 4/79

-2: A: 36/79, C: 30/79, G: 5/79, T: 8/79

-1: A: 32/79, C: 28/79, G: 11/79, T: 8/79

+4: A: 27/79, C: 19/79, G: 15/79, T: 17/79

+5: A: 21/79, C: 28/79, G: 6/79, T: 24/79

+6: A: 13/79, C: 28/79, G: 25/79, T: 13/79

图 7

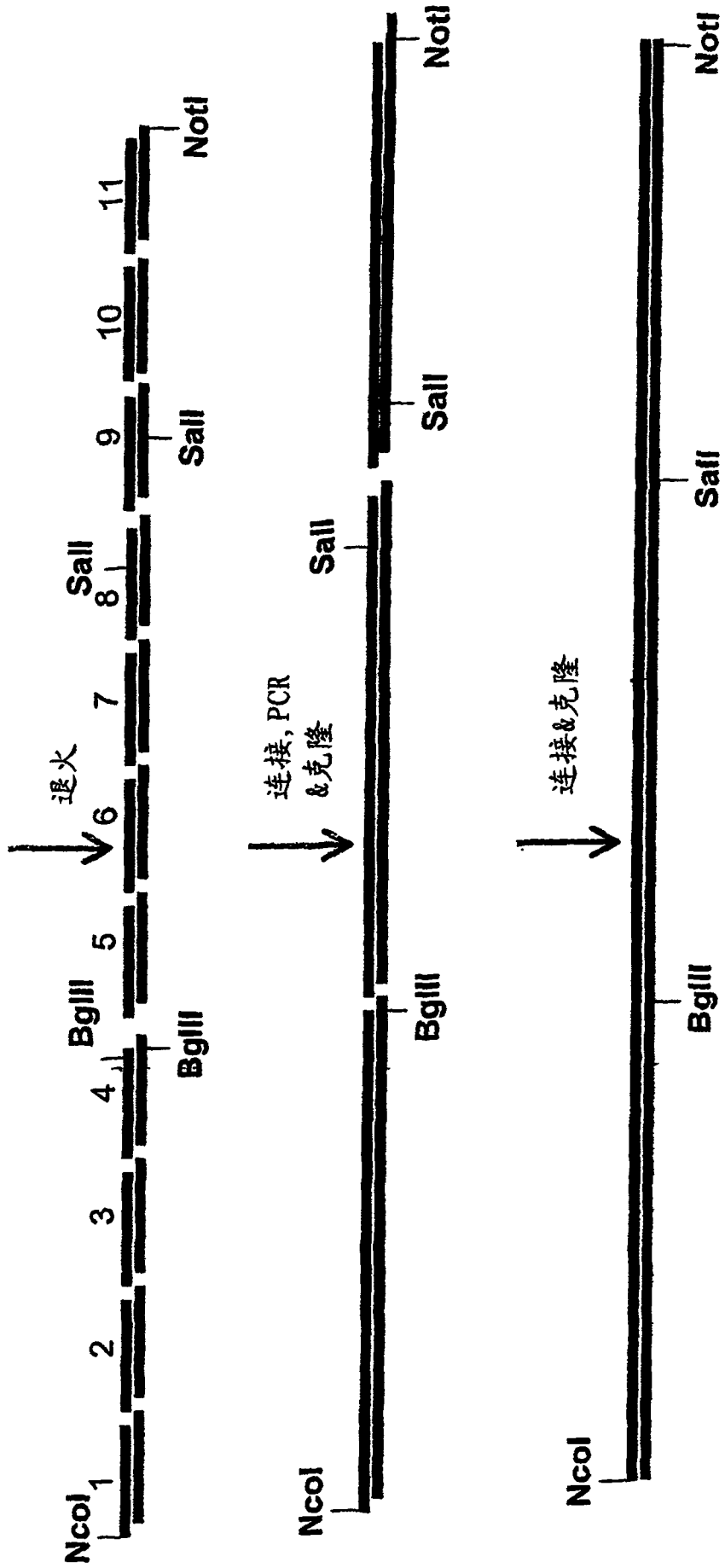


图 8

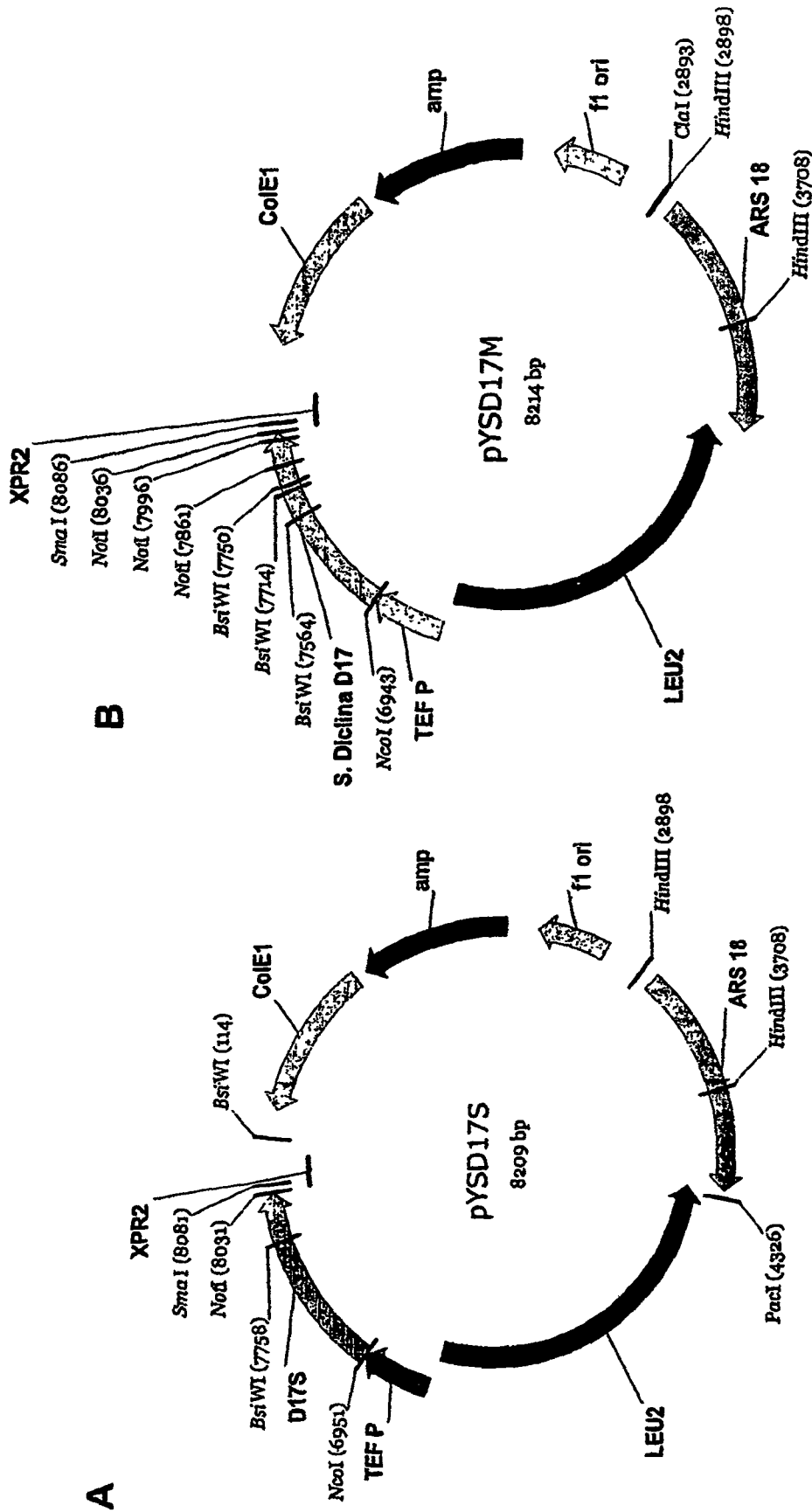


图 9

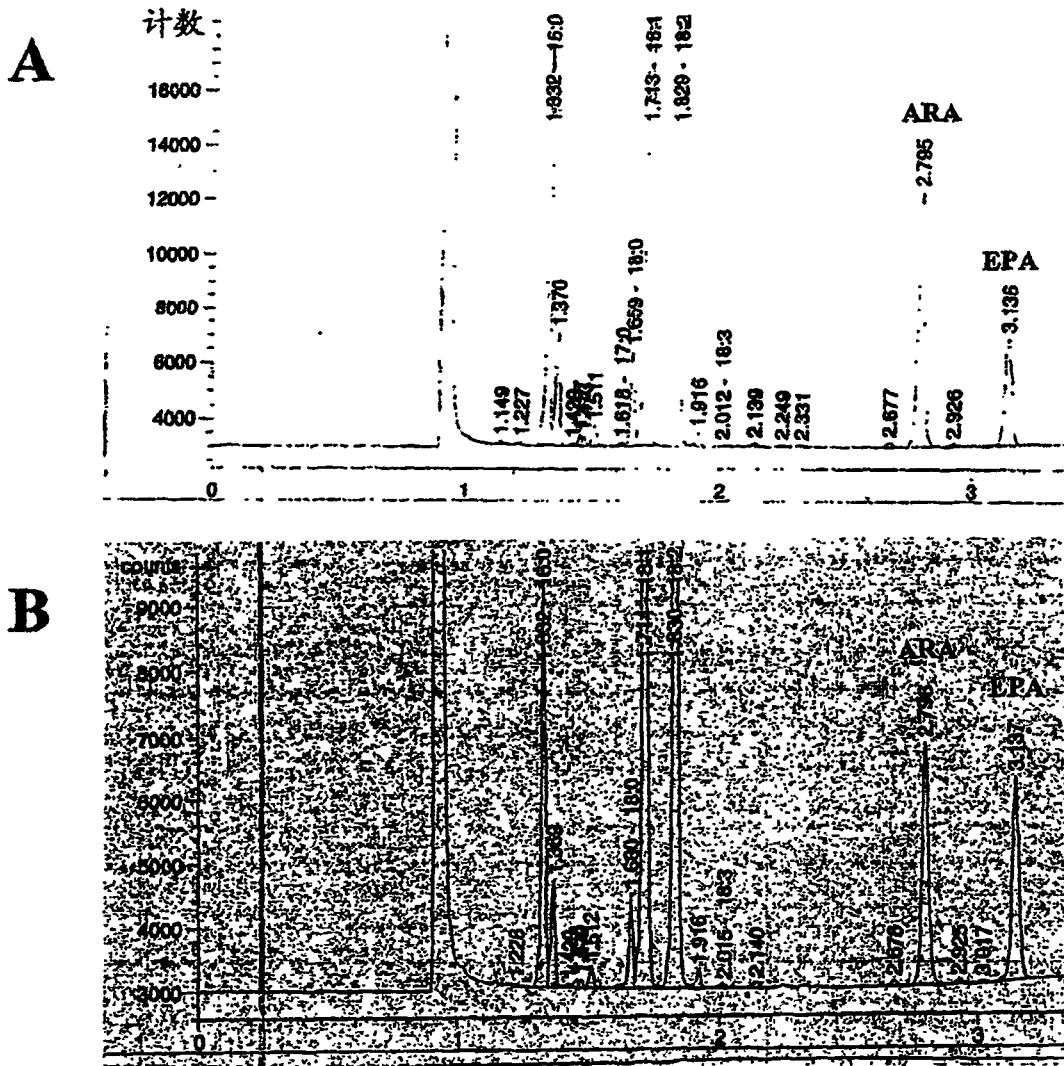


图 10

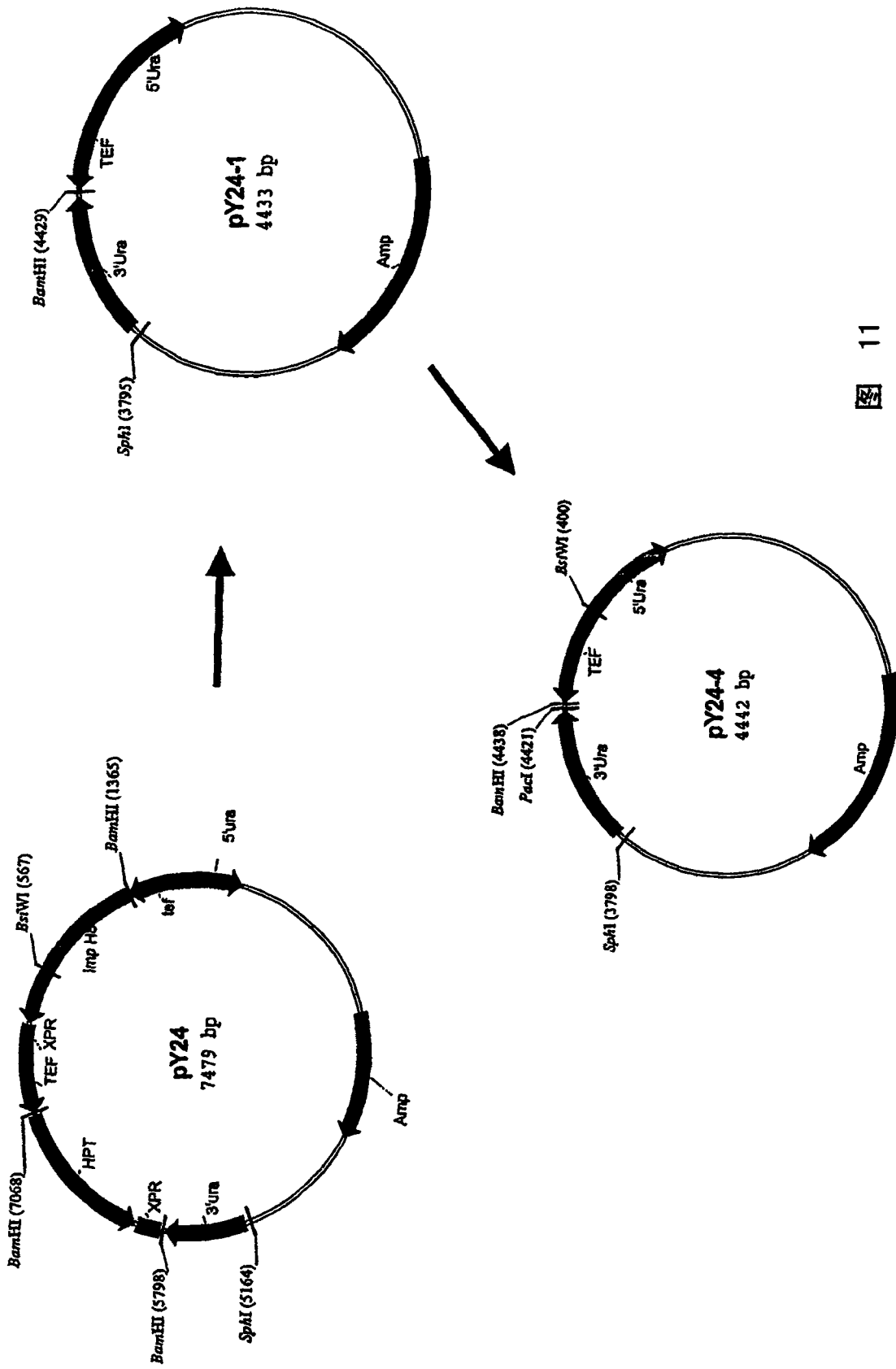


图 11

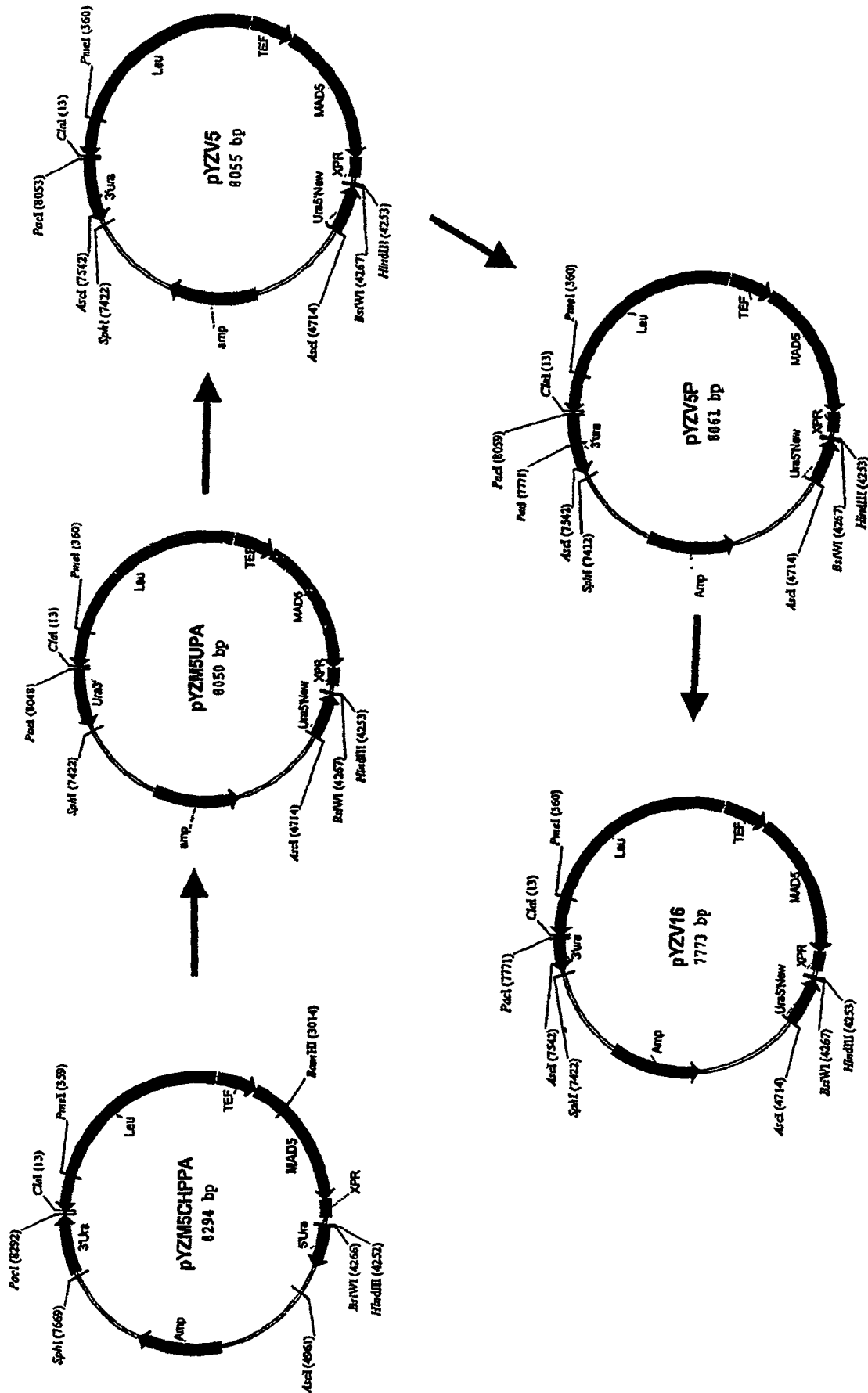


图 12

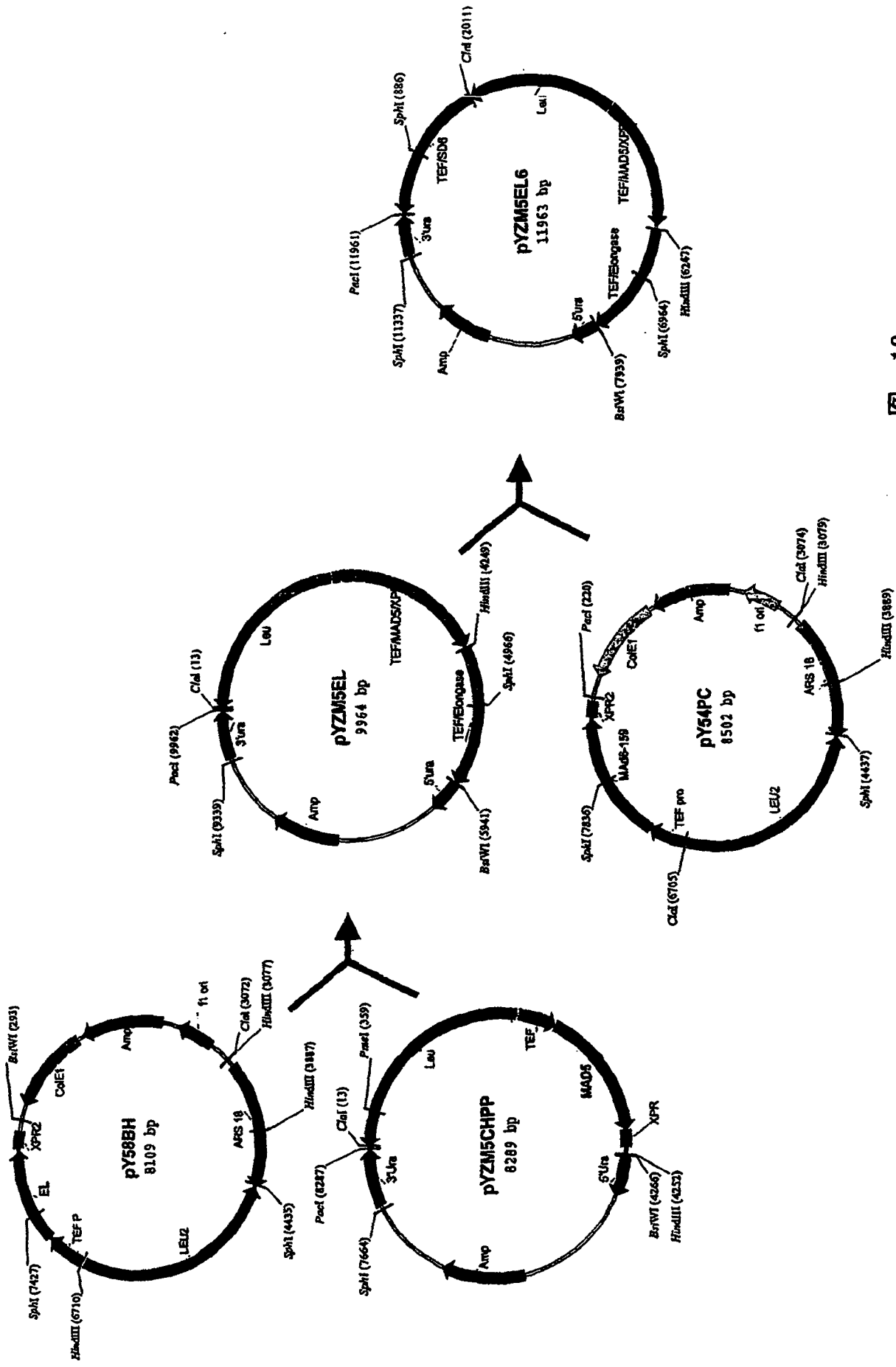


图 13

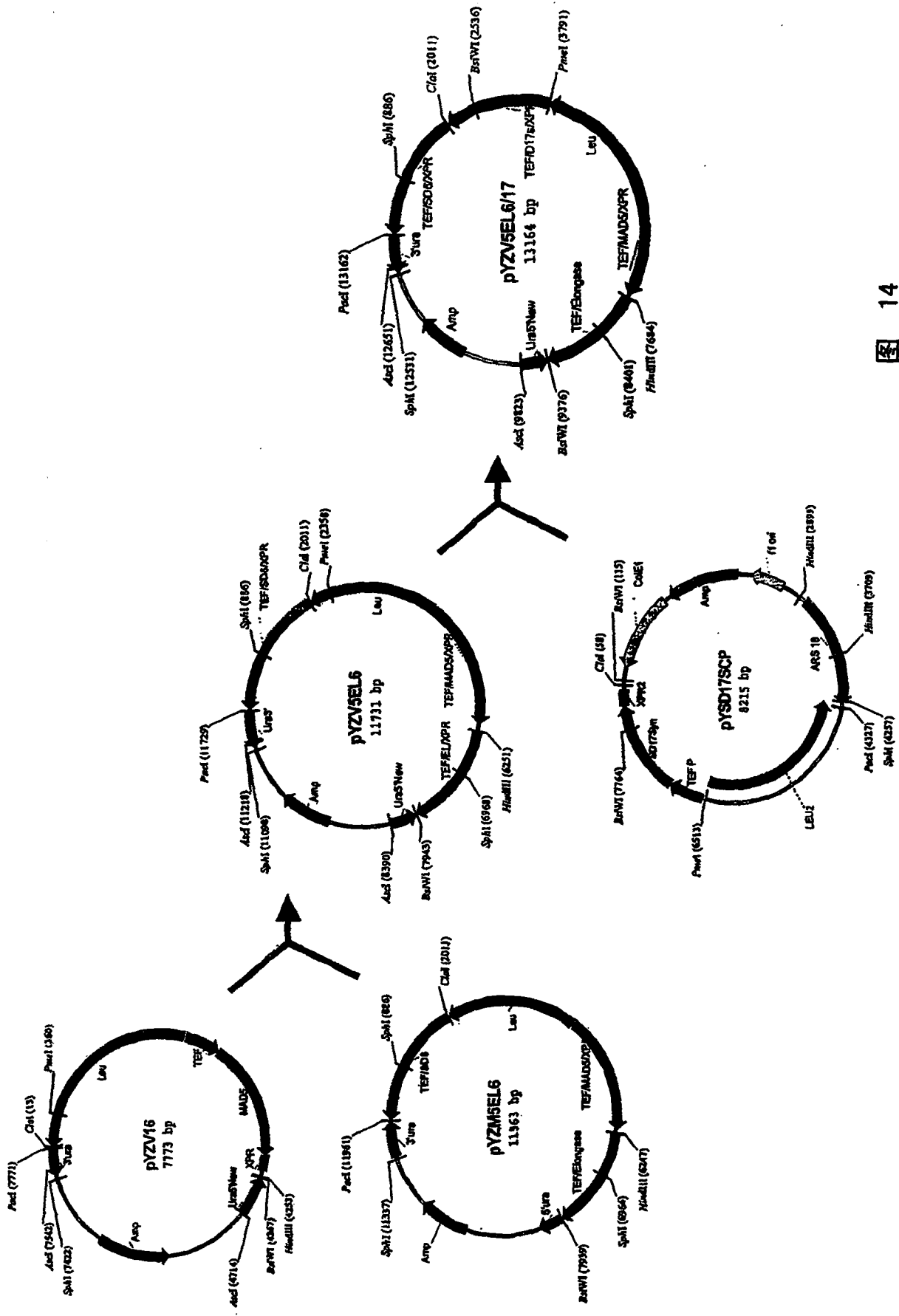


图 14

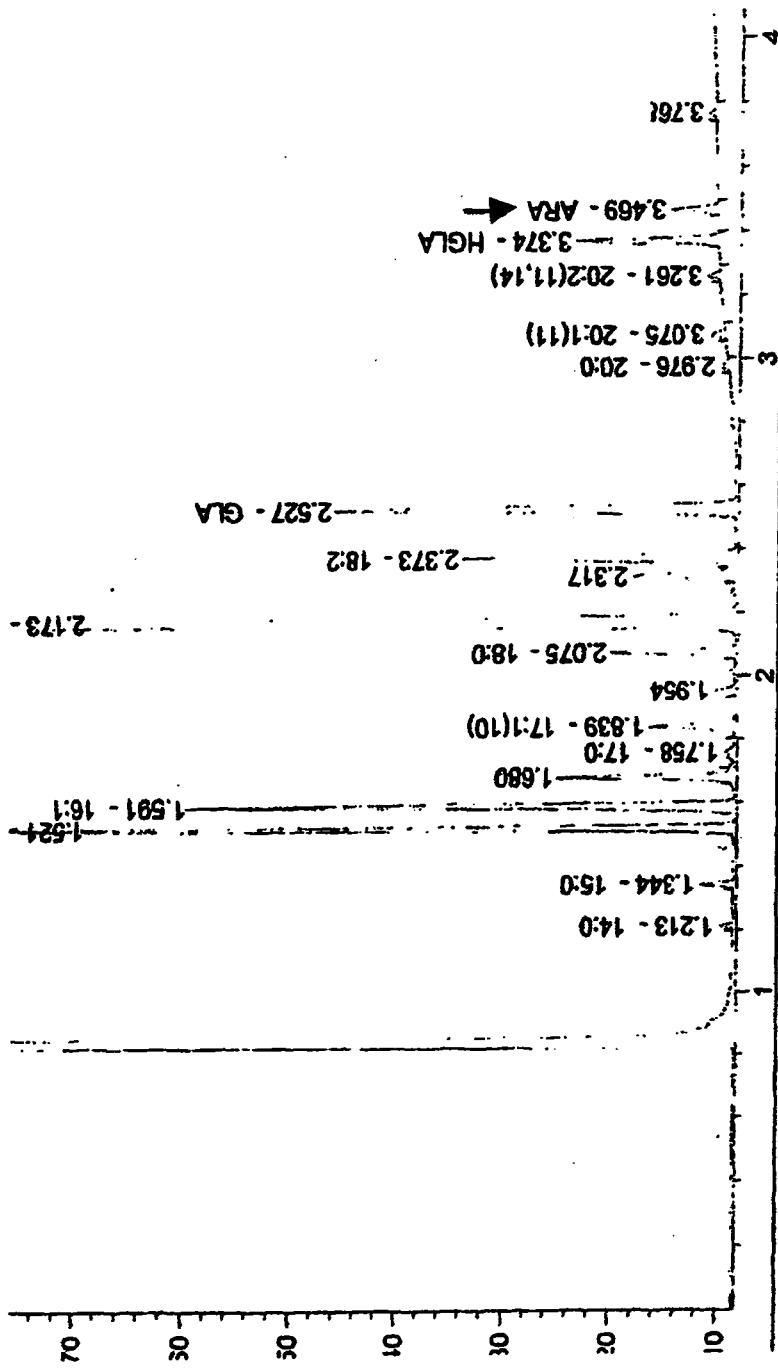


图 15

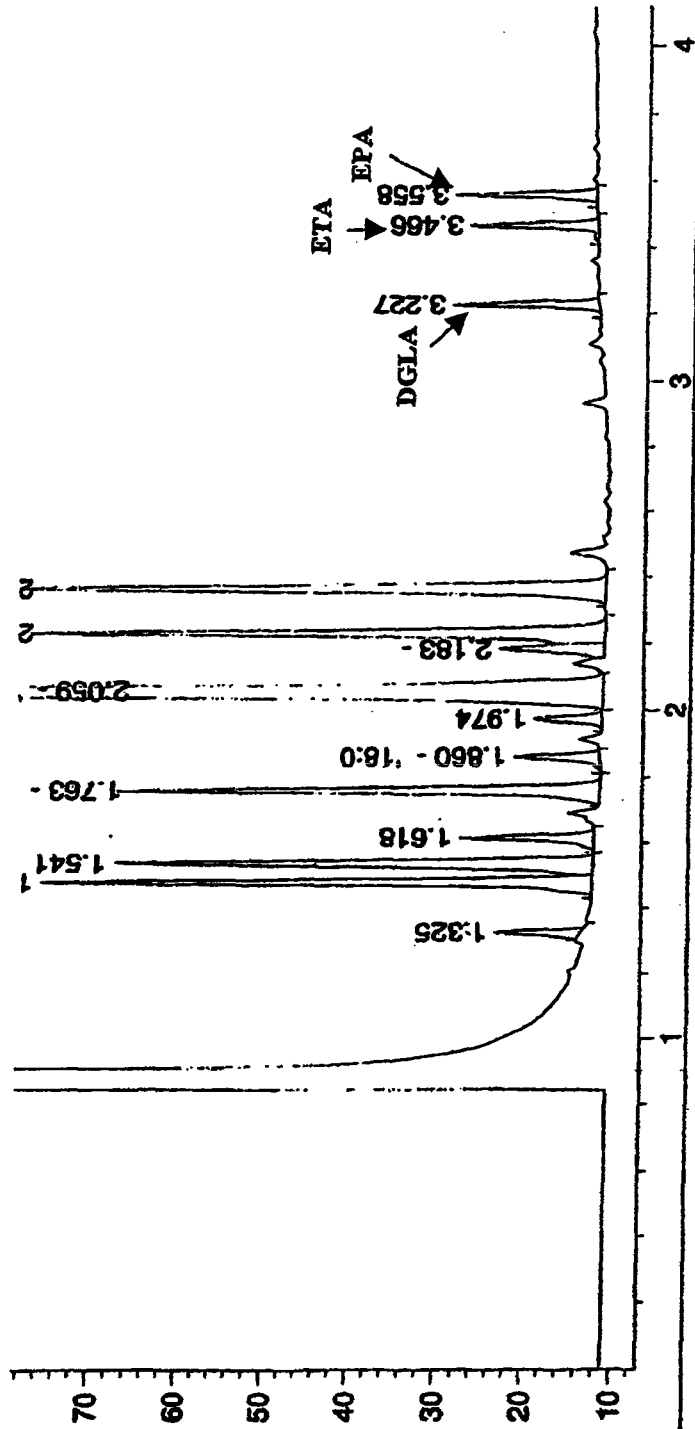


图 16