



(12) 发明专利

(10) 授权公告号 CN 101883843 B

(45) 授权公告日 2016.03.02

(21) 申请号 200880118917.4

(22) 申请日 2008.10.03

(30) 优先权数据

60/977174 2007.10.03 US

(85) PCT国际申请进入国家阶段日

2010.06.02

(86) PCT国际申请的申请数据

PCT/US2008/078671 2008.10.03

(87) PCT国际申请的公布数据

W02009/046248 EN 2009.04.09

(83) 生物保藏信息

PTA-7184 2005.10.26

PTA-7185 2005.10.26

PTA-7186 2005.10.26

PTA-7187 2005.10.26

PTA-8614 2007.08.23

PTA-8802 2007.11.29

(73) 专利权人 纳幕尔杜邦公司

地址 美国特拉华州

(72) 发明人 Z·薛 Q·Q·朱 N·S·亚达夫

P·L·夏普 S·P·洪

(74) 专利代理机构 中国专利代理(香港)有限公司 72001

代理人 李波 郭文洁

(51) Int. Cl.

C12N 1/00(2006.01)

C12P 7/64(2006.01)

C12N 15/00(2006.01)

(56) 对比文件

Gary A. Eitzen, et al. Enlarged Peroxisomes Are Present in Oleic Acid-grown. <<The Journal of Cell Biology>>. 1997, 第 137 卷(第 6 期), 1265-1278.

王显伦等. 活性豆粉复合改良剂对面粉品质影响研究.《粮食与饲料工业》. 2005, (第 09 期),

崔凯, 丁霄霖. 苏子油多不饱和脂肪酸尿素富集研究.《中国油脂》. 1996, (第 06 期),

审查员 孙瑞峰

权利要求书2页 说明书79页

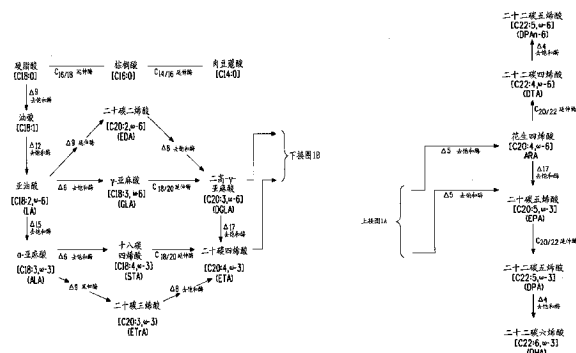
序列表175页 附图17页

(54) 发明名称

破坏过氧化物酶体生物合成因子蛋白(PEX)以改变含油真核生物中多不饱和脂肪酸和总脂质含量

(57) 摘要

通过改变过氧化物酶体生物合成因子(Pex)蛋白的活性,已经成功地找到在生产 PUFA 的含油真核生物中的总脂质级分和油级分中提高多不饱和脂肪酸(PUFA)的量的方法。在与不破坏天然 Pex 蛋白的亲本菌株比较时,在生产 PUFA 的含油真核生物菌种中破坏染色体 Pex3 基因、Pex10p 基因或 Pex16p 基因导致在菌株的总脂质级分和油级分中的 PUFA 量提高,所述 PUFA 的量以总脂肪酸百分比和干细胞重量百分比形式表示。



1. 使含油耶氏酵母属的总脂肪酸中的至少一种多不饱和脂肪酸的重量百分比提高的方法,所述方法包括:

a) 提供含油耶氏酵母属,所述含油耶氏酵母属包含:

1) 编码功能性多不饱和脂肪酸生物合成途径的基因,所述基因包含编码选自 $\Delta 15$ 去饱和酶、 $\Delta 6$ 去饱和酶和 $\Delta 9$ 延伸酶的酶的至少一种基因;和

2) 编码过氧化物酶体生物合成因子蛋白的天然基因的敲除,以及

b) 使所述耶氏酵母属在以下条件下生长:即,使得所述耶氏酵母属的总脂肪酸中的至少一种多不饱和脂肪酸的重量百分比提高,其中所述提高是相对于其中编码过氧化物酶体生物合成因子蛋白的天然基因未被敲除的耶氏酵母属的总脂肪酸中的所述至少一种多不饱和脂肪酸的重量百分比。

2. 提高含油耶氏酵母属的至少一种多不饱和脂肪酸相对于干细胞重量的重量百分比的方法,所述方法包括:

a) 提供含油耶氏酵母属,所述含油耶氏酵母属包含:

1) 编码功能性多不饱和脂肪酸生物合成途径的基因,所述基因包含编码选自 $\Delta 15$ 去饱和酶、 $\Delta 6$ 去饱和酶和 $\Delta 9$ 延伸酶的酶的至少一种基因;和

2) 编码过氧化物酶体生物合成因子蛋白的天然基因的敲除,以及

b) 使所述耶氏酵母属在以下条件下生长:即,使得耶氏酵母属的至少一种多不饱和脂肪酸相对于干细胞重量的重量百分比提高,其中所述提高是相对于其中编码过氧化物酶体生物合成因子蛋白的天然基因未被敲除的耶氏酵母属的所述至少一种多不饱和脂肪酸相对于干细胞重量的重量百分比。

3. 权利要求 1 的方法,其中所述总脂肪酸中的所述至少一种多不饱和脂肪酸的重量百分比的提高为提高至至少 1.3 倍。

4. 权利要求 1 或 2 的方法,其中所述至少一种多不饱和脂肪酸选自:

γ -亚麻酸、二高- γ -亚麻酸、花生四烯酸、二十二碳四烯酸、 ω -6 二十二碳五烯酸、 α -亚麻酸、十八碳四烯酸、二十碳四烯酸、二十碳五烯酸、 ω -3 二十二碳五烯酸、二十碳二烯酸、二十碳三烯酸、二十二碳六烯酸、这些脂肪酸的羟基化或环氧脂肪酸、 C_{18} 多不饱和脂肪酸、 C_{20} 多不饱和脂肪酸、和 C_{22} 多不饱和脂肪酸。

5. 权利要求 4 的方法,其中耶氏酵母属的总脂肪酸中的多不饱和脂肪酸的组合的重量百分比被提高。

6. 权利要求 1 或 2 的方法,其中所述耶氏酵母属中的总脂质含量被改变,其中所述改变是相对于其中编码过氧化物酶体生物合成因子蛋白的天然基因未被敲除的耶氏酵母属中的总脂质含量。

7. 权利要求 1 或 2 的方法,其中所述编码多不饱和脂肪酸生物合成途径的基因还包含编码选自下组的酶的基因:

$\Delta 9$ 去饱和酶、 $\Delta 12$ 去饱和酶、 $\Delta 5$ 去饱和酶、 $\Delta 17$ 去饱和酶、 $\Delta 8$ 去饱和酶、 $\Delta 4$ 去饱和酶、 $C_{14/16}$ 延伸酶、 $C_{16/18}$ 延伸酶、 $C_{18/20}$ 延伸酶和 $C_{20/22}$ 延伸酶。

8. 权利要求 1 或 2 的方法,其中所述过氧化物酶体生物合成因子蛋白选自:

Pex1p、Pex 2p、Pex3p、Pex3Bp、Pex4p、Pex5p、Pex5Bp、Pex5Cp、Pex5/20p、Pex6p、Pex7p、Pex8p、Pex10p、Pex12p、Pex13p、Pex14p、Pex15p、Pex16p、Pex17p、Pex14/17p、Pex18p、

Pex19p、Pex20p、Pex21p、Pex21Bp、Pex22p、Pex22p 样和 Pex26p。

9. 权利要求 8 的方法,其中所述过氧化物酶体生物合成因子蛋白是 Pex3p。

10. 解脂耶氏酵母,其具有 ATCC 命名 ATCC PTA-8614。

11. 权利要求 1 的方法,其中所述多不饱和脂肪酸选自 γ -亚麻酸、二高- γ -亚麻酸、二十碳四烯酸和二十碳五烯酸。

12. 权利要求 8 的方法,其中所述过氧化物酶体生物合成因子是 Pex 10p 或 Pex 16p。

破坏过氧化物酶体生物合成因子蛋白 (PEX) 以改变含油真核生物中多不饱和脂肪酸和总脂质含量

[0001] 本专利申请要求美国临时申请 60/977, 174 和 60/977, 177 的优先权, 所述临时申请均提交于 2007 年 10 月 3 日, 并且全文以引用方式并入本文中。

发明领域

[0002] 本发明属于生物技术领域。更具体地讲, 本发明涉及基于破坏过氧化物酶体生物合成因子 (PEX) 蛋白的、用于操纵真核生物的多不饱和脂肪酸 (PUFA) 组成和脂质含量的方法。

[0003] 发明背景

[0004] 已经有大量文献证明了与多不饱和脂肪酸 [“PUFA”], 尤其是 ω -3 和 ω -6PUFA 相关的健康有益效果。为了寻找大规模制备 ω -3 和 ω -6

[0005] PUFA 的方法, 研究者已经将他们的工作转向发现基因以及了解产生脂质和脂肪酸的编码的生物合成途径。

[0006] 一个用于生产这些 PUFA 的研究成果已经将 ω -3/ ω -6PUFA 生物合成途径导入了不会天然产生 ω -3/ ω -6PUFA 的生物体中。一种已经被广泛使用的所述生物是非含油酵母, 啤酒糖酵母 (*Saccharomyces cerevisiae*)。然而, 并无初步结果证明有限产量的亚油酸 [“LA”]、 γ -亚麻酸 [“GLA”]、 α -亚麻酸 [“ALA”]、十八碳四烯酸 [“STA”] 和 / 或二十碳五烯酸 [“EPA”] 适于商业开发。

[0007] 其它用于大规模生产 ω -3/ ω -6PUFA 的研究成果已经培育出了天然生产选定脂肪酸的微生物, 例如异养硅藻小环藻属 (*Cyclotella* sp.) 和菱形藻属 (*Nitzschia* sp.)、假单胞菌属 (*Pseudomonas*)、交替单胞菌属 (*Alteromonas*) 或希瓦氏菌 (*Shewanella*)、腐霉属 (*Pythium*) 的丝状真菌、或长被孢霉属 (*Mortierella elongata*)、*M. exigua* 或 *M. hygrophila*。

[0008] 然而所有这些研究成果都不能够充分地改善油产量或控制所生产的油的组成的特性, 因为发酵依赖微生物本身的天然能力。

[0009] 共有的美国专利 7, 238, 482 描述了使用含油酵母解脂耶氏酵母作为生产宿主用于生产 PUFA 的用途。含油酵母被定义为是那些天然能够合成并积聚油的酵母, 其中细胞干重一般大于 25%。本领域已经描述了生产宿主的优化 (参见例如国际申请公开 WO 2006/033723、美国专利申请公开 2006-0094092、美国专利申请公开 2006-0115881、和美国专利申请公开 2006-0110806)。本文所述的重组菌株包含多种嵌合基因, 所述基因表达多拷贝的异源去饱和酶、延伸酶和酰基转移酶, 并且任选地包含多个天然去饱和酶和酰基转移酶敲除以能够合成并积聚 PUFA。商业生产 PUFA 需要对宿主细胞进行进一步优化。

[0010] Lin Y. 等人提出过氧化物酶体是脂质的分解代谢和合成代谢所必需的 (*Plant Physiology*, 135:814-827 (2004))。然而, 该假说是以对拟南芥属突变体中的 Pex16p 同源物的研究为基础的, 该突变体具有异常的过氧化物酶体生物合成以及脂肪酸合成 (即, 据报道 *ssel* 种子中生产的油比野生型减少了大约 10-16%)。Binns, D. 等人 (*J. Cell Biol.*,

173(5):719-731(2006))也提供了证明过氧化物酶体和脂质体在啤酒糖酵母中的密切协作的文献。但是以前尚未在生产 PUFA 的生物中进行过对 Pex 敲除的研究。

[0011] 申请人已经通过在生产 PUFA 的生物体中破坏过氧化物酶体生物合成因子蛋白的非预知机制,解决了优化宿主细胞以用于 PUFA 商业生产的所述问题,所述机制导致在解脂耶氏酵母的重组 PUFA 生产菌株中提高 PUFA 量(以总脂肪酸百分比形式表示)的非预知结果。本文描述了在过氧化物酶体生物合成因子蛋白中存在破坏的新型菌株。

[0012] 发明概述

[0013] 本文所述的是提高在具有总脂质含量、总脂质级分和油级分的含油真核生物中至少一种多不饱和脂肪酸[“PUFA”]的重量%相对于总脂肪酸[“TFA”]的重量%的方法,所述方法包括:

[0014] a) 提供含油真核生物,所述含油真核生物包含:

[0015] 1) 编码功能性多不饱和脂肪酸生物合成途径的基因;和

[0016] 2) 在编码过氧化物酶体生物合成因子蛋白的天然基因中的破坏,由此提供 PEX 破坏的生物,以及

[0017] b) 使所述 PEX 破坏的生物在以下条件下生长:即,使得当与编码过氧化物酶体生物合成因子蛋白的天然基因未被破坏的含油真核生物中的总脂质级分或油级分中相对于总脂肪酸的重量百分比的至少一种多不饱和脂肪酸的重量百分比进行比较时,所述总脂质级分或油级分中至少一种多不饱和脂肪酸的重量百分比相对于总脂肪酸的重量百分比提高。

[0018] 该提高方法也可通过实施相同步骤(a)和(b),用于提高至少一种多不饱和脂肪酸[“PUFA”]相对于干细胞重量[“DCW”]的百分比。

[0019] 在本文所述的一些方法中,PUFA 重量%相对于 TFA 重量%提高了至少 1.3 倍。

[0020] 在一些所述方法中,PEX 破坏的生物体中的总脂质含量与天然 PEX 基因中无破坏的含油真核生物的总脂质含量相比可能提高或降低。

[0021] 在这些方法的任何一种中,提高的 PUFA 可以是单个 PUFA 或多个 PUFA 的组合。在任一种情况下,提高的 PUFA 或提高的 PUFA 组合可以包括亚油酸、共轭亚油酸、 γ -亚麻酸、二高- γ -亚麻酸、花生四烯酸、二十二碳四烯酸、 ω -6 二十二碳五烯酸、 α -亚麻酸、十八碳四烯酸、二十碳四烯酸、二十碳五烯酸、 ω -3 二十二碳五烯酸、二十碳二烯酸、二十碳三烯酸、二十二碳六烯酸、这些脂肪酸的羟基化或环氧脂肪酸、 C_{18} 多不饱和脂肪酸或这些脂肪酸的组合、 C_{20} 多不饱和脂肪酸或这些脂肪酸的组合、 $C_{20/22}$ 多不饱和脂肪酸的组合和 C_{22} 多不饱和脂肪酸或这些脂肪酸的组合。

[0022] 在这些方法的任何一种中,PEX 破坏的生物可以是以下生物的一员:耶氏酵母属(Yarrowia)、假丝酵母属(Candida)、红酵母属(Rhodotorula)、红冬孢酵母属(Rhodospiridium)、隐球酵母属(Cryptococcus)、丝孢酵母属(Trichosporon)、油脂酵母属(Lipomyces)、被孢霉属(Mortierella)、破囊壶菌属(Thraustochytrium)、裂殖壶菌属(Schizochytrium)、和糖酵母属(Saccharomyces),上述生物都具有含油特性。并且,在任一种所述方法中,PUFA 生物合成途径包括编码以下酶的任何一种或它们的组合的基因: $\Delta 9$ 去饱和酶、 $\Delta 12$ 去饱和酶、 $\Delta 6$ 去饱和酶、 $\Delta 5$ 去饱和酶、 $\Delta 17$ 去饱和酶、 $\Delta 8$ 去饱和酶、 $\Delta 15$ 去饱和酶、 $\Delta 4$ 去饱和酶、 $C_{14/16}$ 延伸酶、 $C_{16/18}$ 延伸酶、 $C_{18/20}$ 延伸酶、 $C_{20/22}$ 延伸酶和 $\Delta 9$

延伸酶。

[0023] 破坏可发生在编码以下过氧化物酶体生物合成因子蛋白的 PEX 基因中,所述蛋白包括:Pex1p、Pex 2p、Pex3p、Pex3Bp、Pex4p、Pex5p、Pex5Bp、Pex5Cp、Pex5/20p、Pex6p、Pex7p、Pex8p、Pex10p、Pex12p、Pex13p、Pex14p、Pex15p、Pex16p、Pex17p、Pex14/17p、Pex18p、Pex19p、Pex20p、Pex21p、Pex21Bp、Pex22p、Pex22p 样和 Pex26p。并且在这些方法的任何一种中,破坏可以是编码蛋白的 C-末端部分的基因的一部分中的基因敲除或缺失。在这些方法的任何一种中,缺失发生在编码蛋白的 C₃HC₄环状锌指基序的 C-末端部分的基因部分。

[0024] 本文也描述了 PEX 破坏的生物中的油级分或总脂质级分,通过如权利要求 1 所述的方法,其已经提高了至少一种 PUFA 的重量%。本文也描述了一种 PEX 破坏的解脂耶氏酵母,该酵母在编码 Pex3p 或 Pex10p 或 Pex16p 的天然基因中存在破坏。该解脂耶氏酵母的 ATCC 命名为 ATCC PTA-8614(菌株 Y4128)。

[0025] 生物保藏

[0026] 下列生物材料保藏于美国典型培养物保藏中心(American Type Culture Collection)(ATCC)(10801 University Boulevard, Manassas, VA20110-2209),并具有下列名称、保藏号和保藏日期。

[0027]

生物材料	保藏号	保藏日期
解脂耶氏酵母 (<i>Yarrowia lipolytica</i>)Y2047	ATCC PTA-7186	2005 年 10 月 26 日
解脂耶氏酵母 (<i>Yarrowia lipolytica</i>)Y2201	ATCC PTA-7185	2005 年 10 月 26 日
解脂耶氏酵母 (<i>Yarrowia lipolytica</i>)Y2096	ATCC PTA-7184	2005 年 10 月 26 日
解脂耶氏酵母 (<i>Yarrowia lipolytica</i>)Y3000	ATCC PTA-7187	2005 年 10 月 26 日
解脂耶氏酵母 (<i>Yarrowia lipolytica</i>)Y4128	ATCC PTA-8614	2007 年 8 月 23 日
解脂耶氏酵母 (<i>Yarrowia lipolytica</i>)Y4127	ATCC PTA-8802	2007 年 11 月 29 日

[0028] 上面列出的生物材料将按照国际承认的用于专利程序目的的微生物保藏的布达佩斯条约的有关条款来保存。所列出的保藏物将会被保持在指定的国际保藏机构中至少 30 年,并且一旦准予专利公开它就将会对公众开放。在由政府行为授予的专利权的部分废除中,保藏物的可用性不会构成实施主题发明的许可。

[0029] 附图简述和序列表

[0030] 图 1 由图 1A 和图 1B 组成,它们一起示出了 ω -3/ ω -6 脂肪酸生物合成途径,并且当考虑下文对该途径的描述时应被视为一起。

[0031] 图 2A 提供了解脂耶氏酵母 Pex10p (即,SEQ ID NO:10[GenBank 保藏号 CAG81606] 的氨基酸 327-364)、解脂耶氏酵母 Pex2p (即,SEQ ID NO:2[GenBank 保藏号 CAG77647] 的氨基酸 266-323) 和解脂耶氏酵母 Pex12p (即,SEQ ID NO:11[GenBank 保藏号 CAG81532] 的氨基酸 342-391) 的 C₃HC₄ 环状锌指基序的比对,它们具有星号指示的保守 C₃HC₄ 环状锌指基序的半胱氨酸和组氨酸残基。

[0032] 图 2B 图示出了解脂耶氏酵母 Pex10p C₃HC₄ 指状基序的不同氨基酸残基和它们结合的两个锌离子之间的本发明相互作用。

[0033] 图 3A 图示了解脂耶氏酵母菌株 Y4128 的发育,该菌株在总脂质级分中生产 37.6% 的 EPA。

[0034] 图 3B 提供了 pZP3-Pa777U 的质粒图谱。

[0035] 图 4 提供了下列质粒的质粒图谱:(A)pY117;和 (B)pZP2-2988。

[0036] 图 5 提供了下列质粒的质粒图谱:(A)pZKUE3S;和 (B)pFBAIN-MOD-1。

[0037] 图 6 提供了下列质粒的质粒图谱:(A)pFBAIN-PEX10;和 (B)pEXP-MOD-1。

[0038] 图 7A 提供了 pPEX10-1 的质粒图谱。图 7B 图示了解脂耶氏酵母菌株 Y4184U 的发育。

[0039] 图 8 提供了如下质粒的质粒图谱:(A)pZKL1-2SP98C;和 (B)pZKL2-5U89GC。

[0040] 图 9 提供了下列质粒的质粒图谱:(A)pYPS161;和 (B)pYRH13。

[0041] 图 10 图示了解脂耶氏酵母菌株 Y4305U3 的发育。

[0042] 图 11 提供了下列质粒的质粒图谱:(A)pZKUM;和 (B)pZKD2-5U89A2。

[0043] 图 12 提供了下列质粒的质粒图谱:(A)pY87;和 (B)pY157。

[0044] 根据下面的详细描述和附带的序列描述可以更充分地理解本发明,下面的详细描述和所附的序列描述形成了本申请的一部分。

[0045] 下面的序列遵照 37C. F. R. § 1.821-1.825 (“对包含核苷酸序列和 / 或氨基酸序列公开内容的专利申请的要求 - 序列规则”) (“Requirements for Patent Applications Containing Nucleotide Sequences and/or Amino Acid Sequence Disclosures-the Sequence Rules”), 并且符合世界知识产权组织 (World Intellectual Property Organization) (WIPO) ST. 25 标准 (1998) 以及 EPO 和 PCT 的序列表要求 (细则 5.2 和 49.5(a-bis) 以及行政指令 (Administrative Instructions) 的 208 节和附录 C)。用于核苷酸和氨基酸序列数据的符号和格式遵循在 37C. F. R. § 1.822 中示出的规则。

[0046] SEQ ID NO:1-86 是引物、编码基因或蛋白 (或它们的片段) 的 ORF、或质粒,如表 1 所述。

[0047] 表 1

[0048] 核酸和蛋白质 SEQ ID 号总汇

[0049]

描述和缩写	核酸 SEQ ID NO	蛋白质 SEQ ID NO

解脂耶氏酵母 Pex1p (GenBank 保藏号 CAG82178)	--	1 (1024AA)
解脂耶氏酵母 Pex2p (GenBank 保藏号 CAG77647)	--	2 (381AA)
解脂耶氏酵母 Pex3p (GenBank 保藏号 CAG78565)	--	3 (431AA)
解脂耶氏酵母 Pex3Bp (GenBank 保藏号 CAG83356)	--	4 (395AA)
解脂耶氏酵母 Pex4p (GenBank 保藏号 CAG79130)	--	5 (153AA)
解脂耶氏酵母 Pex5p (GenBank 保藏号 CAG78803)	--	6 (598AA)

[0050]

解脂耶氏酵母 Pex6p (GenBank 保藏号 CAG82306)	--	7 (1024AA)
解脂耶氏酵母 Pex7p (GenBank 保藏号 CAG78389)	--	8 (356AA)
解脂耶氏酵母 Pex8p (GenBank 保藏号 CAG80447)	--	9 (671AA)
解脂耶氏酵母 Pex10p (GenBank 保藏号 CAG81606)	--	10 (377AA)
解脂耶氏酵母 Pex12p (GenBank 保藏号 CAG81532)	--	11 (408AA)
解脂耶氏酵母 Pex13p (GenBank 保藏号 CAG81789)	--	12 (412AA)
解脂耶氏酵母 Pex14p (GenBank 保藏号 CAG79323)	--	13 (380AA)
解脂耶氏酵母 Pex16p (GenBank 保藏号 CAG79622)	--	14 (391AA)
解脂耶氏酵母 Pex17p (GenBank 保藏号 CAG84025)	--	15 (225AA)

解脂耶氏酵母 Pex19p (GenBank 保藏号 AAK84827)	--	16 (324AA)
解脂耶氏酵母 Pex20p (GenBank 保藏号 CAG79226)	--	17 (417AA)
解脂耶氏酵母 Pex22p (GenBank 保藏号 CAG77876)	--	18 (195AA)
解脂耶氏酵母 Pex26p (GenBank 保藏号 NC_006072, 核苷酸 117230-118387 的反义翻译)	--	19 (386AA)
包含解脂耶氏酵母 Pex10 基因的重叠群, 该基因编码过氧化物酶体生物合成因子蛋白 (Pex10p) (GenBank 保藏号 AB036770)	20 (3387bp)	--
解脂耶氏酵母 Pex10 (GenBank 保藏号 AB036770, 核苷酸 1038-2171) (蛋白序列与 SEQ ID NO:10100%相同)	21 (1134bp)	22 (377AA)

[0051]

解脂耶氏酵母 Pex10 (GenBank 保藏号 AJ012084, 它对应于 GenBank 保藏号 AB036770 的核苷酸 1107-2171) (相对于 SEQ ID NO:10 和 22 的蛋白序列截短了最初的 23 个氨基酸)	23 (1065bp)	24 (354AA)
解脂耶氏酵母 Pex10p C ₃ HC ₄ 环状锌指基序 (即, SEQ ID NO:10 的氨基酸 327-364)	--	25 (38AA)
解脂耶氏酵母截短 Pex10p (GenBank 保藏号 CAG81606 [SEQ ID NO:10], 缺失了 C-末端的 32 个氨基酸)	--	26 (345AA)
解脂耶氏酵母乙酰羟酸合酶 (AHAS) 突变基因, 该基因包含 W497L 突变	27 (2987bp)	--
质粒 pZP3-Pa777U	28 (13,066bp)	--
质粒 pY117	29 (9570bp)	--
质粒 pZP2-2988	30 (15,743bp)	--
质粒 pZKUE3S	31 (6303bp)	--
引物 pZP-GW-5-1	32	--

引物 pZP-GW-5-2	33	--
引物 pZP-GW-5-3	34	--
引物 pZP-GW-5-4	35	--
引物 pZP-GW-3-1	36	--
引物 pZP-GW-3-2	37	--
引物 pZP-GW-3-3	38	--
引物 pZP-GW-3-4	39	--
Genome Walker 衔接子 [顶链]	40	--
Genome Walker 衔接子 [底链]	41	--
巢式衔接子引物	42	--
引物 Per10F1	43	--
引物 ZPGW-5-5	44	--
引物 Per10R	45	--

[0052]

质粒 pFBAIN-MOD-1	46 (7222bp)	--
质粒 pFBAln-PEX10	47 (8133bp)	--
引物 PEX10-R-BsiWI	48	--
引物 PEX10-F1-SaII	49	--
引物 PEX10-F2-SaII	50	--
质粒 pEXP-MOD1	51 (7277bp)	--
质粒 pPEX10-1	52 (7559bp)	--

质粒 pPEX10-2	53 (8051bp)	---
质粒 pZKL1-2SP98C	54 (15, 877bp)	---
质粒 pZKL2-5U89GC	55 (15, 812bp)	---
质粒 pYPS161	56 (7966bp)	---
引物 Pex-10de11 3' . 正向	57	---
引物 Pex-10de12 5' . 反向	58	---
质粒 pYRH13	59 (8673bp)	---
引物 PEX16Fii	60	---
引物 PEX16Rii	61	---
引物 3UTR-URA3	62	---
引物 Pex16-conf	63	---
实时 PCR 引物 ef-324F	64	---
实时 PCR 引物 ef-392R	65	---
实时 PCR 引物 Pex16-741F	66	---
实时 PCR 引物 Pex16-802R	67	---

[0053]

TaqMan 探针 ef-345T 的核苷酸部分	68	---
TaqMan 探针 PEX16-760T 的核苷酸部分	69	---
质粒 pZKUM	70 (4313bp)	---
质粒 pZKD2-5U89A2	71 (15, 966bp)	---

解脂耶氏酵母二酰基甘油酰基转移酶 (DGAT2) (美国专利公开 7, 267, 976)	72 (2119bp)	73 (514AA)
合成 $\Delta 12$ 去饱和酶, 该酶来源于串珠镰刀菌, 其密码子经优化以在解脂耶氏酵母 (“FmD12S”) 中表达	74 (1434bp)	75 (477AA)
合成突变型 $\Delta 8$ 去饱和酶 (“EgD8M”), 该酶来源于小眼虫 (“EgD8S”; 美国专利公开 7, 256, 033)	76 (1272bp)	77 (422AA)
合成 $\Delta 9$ 延伸酶, 该酶来源于小型绿藻属 CCMP389, 其密码子经优化以在解脂耶氏酵母 (“E389D9eS”) 中表达	78 (792bp)	79 (263AA)
合成 $\Delta 5$ 去饱和酶, 该酶来源于小眼虫, 其密码子经优化以在解脂耶氏酵母 (“EgD5S”) 中表达	80 (1350bp)	81 (449AA)
质粒 pY157	82 (6356bp)	---
质粒 pY87	83 (5910bp)	---
被 Cre 重组酶识别的大肠杆菌 (<i>Escherichia coli</i>) LoxP 重组位点	84 (34bp)	---
引物 UP 768	85	---
引物 LP 769	86	---

[0054] 发明详述

[0055] 本文所述的是用于操纵在生产 PUFA 的真核生物中的长链多不饱和脂肪酸 [“LC-PUFA”] 的浓度 (以总脂肪酸百分比形式表示) 和含量 (以干细胞重量百分比形式表示) 的通用方法。这些方法依赖于破坏在宿主中的天然过氧化物酶体生物合成因子 [“Pex”] 蛋白并将对具有天然的或遗传工程的 PUFA 生产能力的多种真核生物具有广泛适用性, 所述真核生物包括藻类、真菌、卵菌、酵母、类眼虫、原生藻菌、植物和一些哺乳动物系统。

[0056] PUFA 或其衍生物可用作膳食替代品或补充剂, 尤其是婴儿代乳品 (formula), 用于接受静脉内营养法的患者或用来预防或治疗营养不良。例如, PUFA 可掺入烹饪油、脂肪或人造黄油并作为消费者的一般饮食部分被摄入, 从而给予消费者所需的饮食补充剂。此外, PUFA 还可以掺入婴儿代乳品、营养补充剂或其它食品中, 并且可用作抗炎或降胆固醇剂。任选地, 该组合物可以用于药用 (人药或兽药)。

[0057] 定义

[0058] 在本公开中, 使用了大量术语和缩写。给出了如下定义。

[0059] “开放阅读框”缩写为 ORF。

[0060] “聚合酶链反应”缩写为 PCR。

[0061] “美国典型培养物保藏中心”缩写为“ATCC”。

[0062] “多不饱和脂肪酸”缩写为“PUFA”。

[0063] “三酰基甘油”缩写为“TAG”。

[0064] “总脂肪酸”缩写为“TFA”。

[0065] “脂肪酸甲酯”缩写为“FAME”。

[0066] “干细胞重量”缩写为“DCW”。

[0067] 如本文所用,术语“发明”或“本发明”不旨在限制而是通常适用于权利要求中的或本文所述的任何发明。

[0068] 术语“过氧化物酶体”指普遍存在于所有真核细胞中的细胞器。它们具有单独的脂双层膜,该膜将它们的内容物与细胞溶质分离开来,并且包含多种下文所述功能必需的膜蛋白。过氧化物酶体经由“扩展的穿梭机制 (extended shuttle mechanism)”选择性输入蛋白。更具体地讲,存在至少 32 个已知的过氧化物酶体蛋白 (也称为 peroxin),它们参与通过 ATP 水解穿过过氧化物酶体膜而输入蛋白的过程。一些过氧化物酶体蛋白在其 N- 末端或 C- 末端包含特异性蛋白信号,即,过氧化物酶体靶信号或“PTS”,它们引导蛋白通过过氧化物酶体膜。一旦细胞蛋白进入过氧化物酶体,它们通常通过一些方法降解。例如,过氧化物酶体包含氧化酶,例如过氧化氢酶、D- 氨基酸氧化酶和尿酸氧化酶,这些酶能够降解对细胞有毒的物质。作为另外一种选择,过氧化物酶体降解脂肪酸分子以生成乙酰 -CoA 游离分子,它们被回输到细胞溶质中,该过程称为 β - 氧化。

[0069] 术语“过氧化物酶体生物合成因子蛋白”、“过氧化物酶体蛋白”和“Pex 蛋白”是可互换的,并且指涉及过氧化物酶体生物合成的和 / 或参与通过 ATP 水解使细胞蛋白穿过过氧化物酶体膜的过程的蛋白。编码任何这些蛋白的基因的首字母缩写是“Pex 基因”。Pex 基因的命名体系描述于 Distel 等人, *J. Cell Biol.*, 135 :1-3 (1996)。迄今已经在多种真核生物中鉴定了至少 32 个不同的 Pex 基因。已经从对突变体的分析中分离出了多种 Pex 基因,所述突变体显示具有异常的过氧化物酶体功能或结构。根据 Kiel, J. A. K. W. 等人的文献 (Traffic, 7 :1291-1303 (2006)), 其中对 17 个不同真菌菌种中的基因组序列进行了电脑模拟的信息学分析 (silico analysis), 鉴定了以下 Pex 蛋白 :Pex1p、Pex2p、Pex3p、Pex3Bp、Pex4p、Pex5p、Pex5Bp、Pex5Cp、Pex5/20p、Pex6p、Pex7p、Pex8p、Pex10p、Pex12p、Pex13p、Pex14p、Pex15p、Pex16p、Pex17p、Pex14/17p、Pex18p、Pex19p、Pex20p、Pex21p、Pex21Bp、Pex22p、Pex22p 样和 Pex26p。因此,本文将这些蛋白中的每一个称为“Pex 蛋白”、“过氧化物酶体蛋白”或“过氧化物酶体生物合成因子蛋白”,并且每个蛋白由至少一个“Pex 基因”编码。

[0070] 术语“保守结构域”或“基序”指进化上相关的蛋白质的比对序列中在特定位置处保守的一组氨基酸。虽然同源蛋白质之间在其它位置的氨基酸可以发生变化,但在特定位置高度保守的氨基酸表明为是对蛋白质的结构、稳定性或活性来说必需的氨基酸。因为它们可通过它们在蛋白质同系物家族的比对序列中的高度保守来鉴定,所以它们可用作识别标签或“签名”来确定具有新的测定序列的蛋白质是否属于以前鉴定的蛋白质家族。参考本文相关部分,Pex2p、Pex10p 和 Pex12p 在它们的羧基末端都有一个富含半胱氨酸的基序,已知该基序是 C_3H_4 环状锌指基序。该基序是它们的活性所必需的,涉及蛋白停靠并转移到过氧化物酶体中 (Kiel, J. A. K. W., 等人, Traffic, 7 :1291-1303 (2006))。

[0071] 术语“ C_3HC_4 环状锌指基序”或“ C_3HC_4 基序”一般指结合两个锌离子的保守半胱氨酸富含基序,它们通过如式 I 所示的氨基酸序列进行鉴定:

[0072] 式 I : $CX_2CX_9\text{ }_{27}CX_1\text{ }_3HX_2CX_2CX_4\text{ }_{48}CX_2C$

[0073] 在解脂耶氏酵母的编码过氧化物酶体生物合成因子 10 蛋白(即 Y1Pex10p)的基因中的 C_3HC_4 环状锌指基序,位于 SEQ ID NO:10 的氨基酸 327-364 之间,并且由 $CX_2CX_{11}CX_1HX_2CX_2CX_{10}CX_2C$ 基序(SEQ IDNO:25)限定。在解脂耶氏酵母的编码过氧化物酶体生物合成因子 2 蛋白(即 Y1Pex2p)的基因中的 C_3HC_4 环状锌指基序,位于 SEQ ID NO:2 的氨基酸 266-323 之间。解脂耶氏酵母过氧化物酶体生物合成因子 12 蛋白,即 Y1Pex12p,包含位于 SEQ ID NO:11 的氨基酸 342-391 之间的不完全 C_3HC_4 环指基序。图 2A 列出了对应于 Y1Pex10、Y1Pex2 和 Y1Pex12 的 C_3HC_4 环状锌指基序的蛋白序列;星号表示基序的保守半胱氨酸或组氨酸残基。

[0074] 认为 Y1Pex10、Y1Pex2 和 Y1Pex12 通过蛋白-蛋白相互作用形成环指复合物。本发明的 Y1Pex10p C_3HC_4 指状基序(具有两个锌残基)的胱氨酸和组氨酸残基间相互作用在图 2B 中用图表示出。

[0075] 术语“Pex10”指编码过氧化物酶体生物合成因子 10 蛋白或过氧化物酶体装配蛋白 Peroxin 10 的基因,其中所述 peroxin 蛋白在下文中称为“Pex10p”。Pex10p 的功能尚未被清楚的阐明,虽然对其它生物的研究已经揭示 Pex10 产物位于过氧化物酶体膜内并且是细胞器正常功能所必需的。 C_3HC_4 环状锌指基序在 Pex10p 的 C-末端区域是保守的(Kalish, J. E. 等人, Mol. Cell Biol., 15:6406-6419(1995); Tan, X. 等人, J. Cell Biol., 128:307-319(1995); Warren, D. S., 等人, Am. J. Hum. Genet., 63:347-359(1998)), 并且是酶活性所必需的。

[0076] 术语“Y1Pex10”指解脂耶氏酵母的编码过氧化物酶体生物合成因子 10 蛋白的基因,其中所述蛋白在下文中称为“Y1Pex10p”。该特定的 peroxin 近来由 Sumita 等人(FEMS Microbiol. Lett., 214:31-38(2002))进行了研究。Y1Pex10 的核苷酸序列在 GenBank 中注册了多个保藏号,包括 GenBank 保藏号 CAG81606(SEQ ID NO:10)、AB036770(SEQ IDNOs: 20、21 和 22)以及 AJ012084(SEQ ID NO:23 和 24)。如在 SEQ IDNO:24 中所示的 Y1Pex10p 的序列长度为 354 个氨基酸。相比之下,如在 SEQ ID NO:10 和 SEQ ID NO:22 中所示的 Y1Pex10p 序列每个长度为 377 个氨基酸,它们 100%相同的序列在蛋白 N-末端具有附加地 23 个氨基酸(对应于与 GenBank 保藏号 AJ012084(SEQ ID NO:24)中鉴定的起始密码子不同的起始密码子)。

[0077] 术语“Pex3”指编码过氧化物酶体生物合成因子 3 蛋白或过氧化物酶体装配蛋白 Peroxin 3 的基因,其中 peroxin 蛋白在下文中称为“Pex3p”。虽然尚未清楚地了解关于 Pex3p 的机制细节,但是清楚的是 Pex3p 是在早期的过氧化物酶体生物合成中所需的过氧化物酶体整合膜蛋白,该蛋白用于形成过氧化物酶体膜(参见例如 Baerends, R. J. 等人, J. Biol. Chem., 271:8887-8894(1996); Bascom, R. A. 等人, Mol. Biol. Cell, 14:939-957(2003))。

[0078] 术语“Y1Pex3”指解脂耶氏酵母的编码过氧化物酶体生物合成因子 3 蛋白的基因,其中所述蛋白在下文中称为“Y1Pex3p”。Y1Pex3 的核苷酸序列在 GenBank 中登记为保藏号 CAG78565(SEQ ID NO:3)。

[0079] 术术语“Pex16”指编码过氧化物酶体生物合成因子 16 蛋白或过氧化物酶体装配蛋白 Peroxin 16 的基因,其中 peroxin 蛋白在下文中称为“Pex16p”。虽然对多种生物的研究已经揭示 Pex16 产物在形成过氧化物酶体膜和调节过氧化物酶体增殖中起到作用,但是 Pex16p 的功能仍未被清楚地阐明 (Platta, H. W. 和 R. Erdmann, Trends Cell Biol., 17(10):474-484(2007))。

[0080] 术语“YlPex16”指解脂耶氏酵母的编码过氧化物酶体生物合成因子 16 蛋白的基因,其中所述蛋白在下文中称为“YlPex16p”。该特定 peroxin 描述于 Elizen G. A. 等人 (J. Cell Biol., 137:1265-1278(1997)) 和 Titorenko, V. I. 等人 (Mol. Cell Biol., 17:5210-5226(1997))。YlPex16 的核苷酸序列在 GenBank 中登记为保藏号 CAG79622 (SEQ ID NO:14)。

[0081] 在天然 Pex 基因中的或与之相关的术语“破坏”指在一部分该基因中的插入、缺失、或定向突变,所述破坏导致全部基因敲除使得该基因从基因组中缺失并且不翻译蛋白,或者导致翻译的 Pex 蛋白具有插入、缺失、氨基酸取代或其它定向突变。蛋白中的破坏位置可以是在例如蛋白的 N-末端部分中或在蛋白的 C-末端部分中。破坏的 Pex 蛋白相对于未破坏的 Pex 蛋白将具有减弱的活性,并且可能是无功能的。在编码 Pex 蛋白的天然基因中的破坏也包括导致 Pex 蛋白低表达或缺乏表达的替代方法,例如能够经由操纵调控序列、转录和翻译因子和 / 或信号转导途径或使用有义、反义或 RNAi 技术等方法导致破坏。

[0082] 如本文所用,术语“PEX 破坏的生物”指包含以下基因的任何含油真核生物,所述基因编码功能性多不饱和脂肪酸生物合成途径并且在编码过氧化物酶体生物合成因子蛋白的天然基因中具有上述破坏。

[0083] 术语“脂质”指任何脂溶性的 (即,亲脂的)、天然存在的分子。脂质是具有多种关键生物学功能的化合物的不同组,所述功能例如细胞膜的结构组分、能量贮存来源和信号途径的中间体。可以将脂质广泛定义为疏水性或两亲性小分子,它们完全地或部分地起源于酮脂酰或异戊二烯基团。表 2 显示了对脂质的综述,它基于 Lipid Metabolites and Pathways Strategy (LIPID MAPS) 分类系统 (National Institute of General Medical Sciences, Bethesda, MD)。

[0084] 表 2

[0085] 脂质类别综述

[0086]

结构构件	脂质类别	脂质类别实例
来源于酮脂酰子单元的缩合	脂肪酰	包括脂肪酸、类二十烷酸、脂肪族酯和脂酰胺
	甘油脂	主要包括单、双、和三取代甘油脂，最为人们熟知的是脂肪酸甘油酯[“三酰基甘油”]
	甘油磷脂或磷脂	包括磷脂酰胆碱、磷酸酰乙醇胺、磷脂酰丝氨酸、磷脂酰肌醇和磷脂酸
	鞘脂类	包括神经酰胺、磷酸鞘脂类（例如鞘磷脂）、糖鞘脂（例如神经节苷脂）、鞘氨醇、脑苷脂类
	糖脂	包括糖酰胺、糖酰胺多糖、酰基海藻糖、酰基海藻糖多糖
	聚酮化合物	包括卤化内酯、聚烯、直链四环素、聚醚抗生素、类黄酮，芳族聚酮化合物
来源于异戊二烯子单元缩合	甾醇脂质	包括甾醇（例如胆固醇）、C18 甾体化合物（例如雌激素）、C19 甾体化合物（例如雄性激素）、C21 甾体化合物（例如孕激素、糖皮质激素和盐皮质激素）、开环甾体化合物、胆汁酸
	异戊烯醇 (Prenol Lipids)	包括类异戊二烯、类胡萝卜素、醌、氢醌、聚异戊烯醇、萜烷类化合物

[0087] 本文术语细胞的“总脂质级分”指细胞的所有酯化脂肪酸。可分离在总脂质级分中的不同亚级分,包括三酰基甘油[“油”]级分、磷脂酰胆碱级分和磷脂酰乙醇胺级分,但是这并不包括所有亚级分。

[0088] “脂质体”指通过单层磷脂以及通常通过特异性蛋白结合的脂质小滴。这些细胞器是大多数生物运输/贮存中性脂类的位点。认为脂质体来源于包含 TAG 生物合成酶的内质网的微区。它们的合成及尺寸受特异性蛋白组分的控制。

[0089] “中性脂类”指那些一般以贮存脂肪和油形式存在于脂质体中的细胞中的脂质,它们的名称是因为在细胞 pH 下,所述脂质无带电基团。它们一般是完全非极性的,对水无亲和力和力。中性脂类一般指脂肪酸的甘油单酯、二酯、和/或三酯,也分别称为单酰基甘油、二酰基甘油或三酰基甘油,或统称为酰基甘油。为了从酰基甘油中释放游离脂肪酸,必须发生水解反应。

[0090] 术语“三酰基甘油”[“TAG”]和“油”是可互换的,它们指由酰化甘油分子的三个脂肪酰残基组成的中性脂类。TAG 能够包含长链 PUFA,以及较短的饱和的和不饱和的脂肪酸以及较长链的饱和脂肪酸。细胞的 TAG 级分也称作“油级分”,并且“油的生物合成”一般指在细胞中合成 TAG。油或 TAG 级分是总脂质级分的亚级分,虽然它也是构成总脂质含量的主要部分,总脂质含量以含油生物细胞

中的总脂肪酸重量占干细胞重量 [参见下文] 的百分比表示。油 [“TAG”] 级分中的脂肪酸组成和总脂质级分的脂肪酸组成一般是相似的。因此,总脂质级分中的 PUFA 浓度的提高或降低将对应于油 [“TAG”] 级分中的 PUFA 浓度的升高或降低,反之亦然。

[0091] 术语“总脂肪酸” [“TFA”] 本文指所有细胞脂肪酸的总量,所述脂肪酸在给定实例中可通过碱酯交换方法 (本领域已知的方法) 被衍生化成脂肪酸甲酯 [“FAME”], 例如它可以是总脂质级分或油级分。因此,总脂肪酸包括来自中性和极性脂质级分的脂肪酸,包括磷脂酰胆碱级分、磷脂酰乙醇胺级分 (phosphatidyletanolamine fraction) 和二酰基甘油、单酰基甘油和三酰基甘油 [“TAG 或油”] 级分,但是不包括游离脂肪酸。

[0092] 术语细胞的“总脂质含量”是 TFA 的量度,以干细胞重量 [“DCW”] 百分比的形式表示。因此,总脂质含量 [“TFA % DCW”] 等同于例如每 100 毫克 DCW 的总脂肪酸毫克数。

[0093] 脂肪酸浓度本文一般表示为 TFA 的重量百分比 [“% TFA”], 例如每 100 毫克 TFA 的给定脂肪酸毫克数。除非在本文公开内容中另作具体说明,给定脂肪酸相对于总脂质的百分比等同于以 % TFA 表示的脂肪酸浓度 (例如总脂质的 % EPA 等同于 EPA % TFA)。

[0094] 在一些情况下,可使用细胞中给定脂肪酸占干细胞重量的百分比 [“% DCW”] 形式表达给定脂肪酸的含量。因此例如二十碳五烯酸 % DCW 将根据下式进行测定 : (二十碳五烯酸 % TFA) * (TFA % DCW) / 100。

[0095] 术语“脂质分布”和“脂质组成”是可互换的,并且指在特定脂质级分 (例如在总脂质级分或油 [“TAG”] 级分中) 中包含的单个脂肪酸的量,其中所述量用 TFA 百分比形式表示。混合物中存在的各单个脂肪酸的总量应当是 100。

[0096] 如本文所用,术语“成倍增加”指通过乘以某数获得的增加。例如,通过乘以 1.3 的数量、量、浓度、重量百分比等,提供 1.3 倍的增加。

[0097] 术语“脂肪酸”指不同链长的长链脂族酸 (链烷酸), 链长为约 C₁₂ 至 C₂₂ (尽管更长和更短链长的酸均是已知的)。主链长介于 C₁₆ 和 C₂₂ 之间。脂肪酸的结构可用简单的记号系统 “X:Y” 来表示,其中 X 表示具体脂肪酸中碳 (“C”) 原子的总数,而 Y 表示双键的数目。另外的关于“饱和脂肪酸”与“不饱和脂肪酸”、“单不饱和脂肪酸”与“多不饱和脂肪酸” (“PUFA”) 以及 “ω-6 脂肪酸” (ω-6 或 n-6) 与 “ω-3 脂肪酸” (ω-3 或 n-3) 之间的区别的详细信息在美国专利 7, 238, 482 中有所提供。

[0098] 表 3 提供了用于描述本文 PUFA 的命名。在标题为“简化符号”一栏中, ω- 指代系统用于表明碳数目、双键的数目和最接近 ω 碳的双键位置,双键位置的计数从 ω 碳开始 (为此 ω 碳的编号为 1)。该表的其余部分汇总了 ω-3 和 ω-6 脂肪酸及其前体的俗名、在整个说明书中使用的缩写以及每种化合物的化学名称。

[0099] 表 3

[0100] 多不饱和脂肪酸及其前体的命名

[0101]

俗名	缩写	化学名称	简化符号
肉豆蔻酸	--	十四酸	14:0
棕榈酸	棕榈酸	十六酸	16:0
棕榈油酸	--	9- 十六碳烯酸	16:1

硬脂酸	--	十八酸	18:0
油酸	--	顺式-9-十八碳烯酸	18:1
亚油酸	LA	顺式-9,12-十八碳二烯酸	18:2 ω -6
γ -亚麻酸	GLA	顺式-6,9,12-十八碳三烯酸	18:3 ω -6
附子脂酸	EDA	顺式-11,14-二十碳二烯酸	20:2 ω -6
二高- γ -亚麻酸	DGLA	顺式-8,11,14-二十碳三烯酸	20:3 ω -6
花生四烯酸	ARA	顺式-5,8,11,14-二十碳四烯酸	20:4 ω -6
α -亚麻酸	ALA	顺式-9,12,15-十八碳三烯酸	18:3 ω -3
硬脂艾杜糖酸	STA	顺式-6,9,12,15-十八碳四烯酸	18:4 ω -3
二十碳三烯酸	ETrA	顺式-11,14,17-二十碳三烯酸	20:3 ω -3
金松烯酸	SCI	顺式-5,11,14-二十碳三烯酸	20:3b ω -6
Juniperonic	JUP	顺式-5,17,11,14-二十碳四烯酸	20:4b ω -3
二十碳四烯酸	ETA	顺式-8,11,14,17-二十碳四烯酸	20:4 ω -3
二十碳五烯酸	EPA	顺式-5,8,11,14,17-二十碳五烯酸	20:5 ω -3

[0102]

二十二碳三烯酸	DRA	顺式-10,13,16-二十二碳三烯酸	22:3 ω -3
二十二碳四烯酸	DTA	顺式-7,10,13,16-二十二碳四烯酸	22:4 ω -3
二十二碳五烯酸	DPAn-6	顺式-4,7,10,13,16-二十二碳五烯酸	22:5 ω -6
二十二碳五烯酸	DPA	顺式-7,10,13,16,19-二十二碳五烯酸	22:5 ω -3
二十二碳六烯酸	DHA	顺式-4,7,10,13,16,19-二十二碳六烯酸	22:6 ω -3

[0103] 虽然使用本文所述方法,表 3 列出的 ω -3/ ω -6PUFA 最可能在含油酵母的油级分中积聚,但是该列表不应理解为限制性的或完全的。

[0104] 如本文所用,术语“多不饱和脂肪酸的组合”或“多不饱和脂肪酸的任何组合”指上表 3 中列出的任何两种或更多种多不饱和脂肪酸的混合物。此类组合具有能相对于细胞中的多种浓度或重量百分比进行测量,包括相对于细胞中的总脂肪酸的重量百分比进行测量的浓度和重量百分比属性。

[0105] 代谢途径或生物合成途径在生物化学意义上可以认为是发生于细胞内由酶催化的一系列化学反应,以实现细胞待使用的或待贮存的代谢产物的形成,或启动另一代谢途径(称作流量产生步骤)。很多此类途径均很精细,并涉及对起始物质的逐步修饰以使之形成具有期望的精确化学结构的产物。

[0106] 术语“PUFA 生物合成途径”指将油酸转化成诸如 LA、EDA、GLA、DGLA、ARA、DRA、DTA 和 DPAn-6 之类的 ω -6 脂肪酸和诸如 ALA、STA、ETra、ETA、EPA、DPA 和 DHA 之类的 ω -3 脂肪酸的代谢过程。文献中详细描述了该过程。参见例如 Int' . App. Pub. No. WO 2006/052870。简而言之,该过程涉及通过添加碳原子来延长碳链和通过加入双键来使分子去饱和,这通过存在于内质网膜内的一系列特异性去饱和酶和延伸酶(称为“PUFA 生物合成途径酶”)进行。更具体地讲,“PUFA 生物合成途径酶”指如下与 PUFA 生物合成相关的任何酶(以及编码它们的基因),所述酶包括: 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、 Δ 8 去饱和酶、 Δ 9 延伸酶、 $C_{14/16}$ 延伸酶、 $C_{16/18}$ 延伸酶、 $C_{18/20}$ 延伸酶和/或 $C_{20/22}$ 延伸酶。

[0107] 术语“ ω -3/ ω -6 脂肪酸生物合成途径”指在合适条件下表达时编码催化 ω -3 和 ω -6 脂肪酸二者之一或两者产生的酶的一组基因。通常,参与 ω -3/ ω -6 脂肪酸生物合成途径的基因编码 PUFA 生物合成途径酶。图 1 示出了一条代表性途径,提供了从肉豆蔻酸经过多种中间产物向 DHA 的转化,演示了 ω -3 和 ω -6 脂肪酸两者是如何可以从共同来源产生。该途径自然分成两部分,使得一部分只生成 ω -3 脂肪酸而另一部分只生成 ω -6 脂肪酸。只产生 ω -3 脂肪酸的部分在本文中称作 ω -3 脂肪酸生物合成途径,而只产生 ω -6 脂肪酸的部分在本文中称作 ω -6 脂肪酸生物合成途径。

[0108] 如本文中关于 ω -3/ ω -6 脂肪酸生物合成途径所用的,术语“功能性的”指该途径中的一些(或全部)基因表达活性酶,导致体内催化或底物转化。应当理解,“ 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”并不意味着上面段落中列出的所有基因都是必需的,因为许多脂肪酸产物将仅需要表达该途径中的一亚组基因。

[0109] 术语“ Δ 6 去饱和酶/ Δ 6 延伸酶途径”指最低程度包括至少一种 Δ 6 去饱和酶和至少一种 $C_{18/20}$ 延伸酶的 PUFA 生物合成途径,从而使得能分别从 LA 和 ALA 开始,以 GLA 和/或 STA 作为脂肪酸中间产物来生物合成 DGLA 和/或 ETA。通过其它去饱和酶和延伸酶的表达,还可以合成 ARA、EPA、DPA 和 DHA。

[0110] 术语“ Δ 9 延伸酶/ Δ 8 去饱和酶途径”指最低程度包括至少一种 Δ 9 延伸酶和至少一种 Δ 8 去饱和酶的 PUFA 生物合成途径,从而使得能分别从 LA 和 ALA 开始,以 EDA 和/或 ETra 作为脂肪酸中间产物来生物合成 DGLA 和/或 ETA。通过其它去饱和酶和延伸酶的表达,还可以合成 ARA、EPA、DPA 和 DHA。

[0111] 术语“去饱和酶”指能够通过从邻接碳原子中的一个上移除氢原子、并且从而在碳

原子之间导入双键来去饱和脂肪酸中的邻接碳原子的多肽。去饱和产生了脂肪酸或受关注的前体。尽管在整个说明书中使用 ω - 指代系统来指代特定的脂肪酸,但使用 Δ - 系统从底物的羧基端计数来表示去饱和酶的活性更方便。本文特别关注的是:1) $\Delta 5$ 去饱和酶,它催化脂肪酸底物 DGLA 转化成 ARA 和 / 或催化脂肪酸底物 ETA 转化成 EPA ;2) $\Delta 17$ 去饱和酶,它在从分子羧基末端编号为第 17 和第 18 的碳原子之间去饱和脂肪酸,并且它例如催化底物脂肪酸 ARA 转化成 EPA 和 / 或催化底物脂肪酸 DGLA 转化成 ETA ;3) $\Delta 6$ 去饱和酶,它催化脂肪酸底物 LA 转化成 GLA 和 / 或催化脂肪酸底物 ALA 转化成 STA ;4) $\Delta 12$ 去饱和酶,它催化底物脂肪酸油酸转化成 LA ;5) $\Delta 15$ 去饱和酶,它催化脂肪酸底物 LA 转化成 ALA 和 / 或催化脂肪酸底物 GLA 转化成 STA ;6) $\Delta 4$ 去饱和酶,它催化底物脂肪酸 DPA 转化成 DHA 和 / 或催化底物脂肪酸 DTA 转化成 DPA $n-6$;7) $\Delta 8$ 去饱和酶,它催化底物脂肪酸 EDA 转化成 DGLA 和 / 或催化底物脂肪酸 ETrA 转化成 ETA ;和,8) $\Delta 9$ 去饱和酶,它催化底物脂肪酸棕榈酸转化成棕榈油酸 (16:1) 和 / 或催化底物脂肪酸硬脂酸转化成油酸。基于 $\Delta 15$ 和 $\Delta 17$ 去饱和酶将 $\omega-6$ 脂肪酸转化成它们的 $\omega-3$ 对应物的能力 (如分别将 LA 转化成 ALA 以及将 ARA 转化成 EPA),偶尔也将它们称作“ $\omega-3$ 去饱和酶”、“W-3 去饱和酶”和 / 或“ $\omega-3$ 去饱和酶”。所期望的是通过用脂肪酸去饱和酶的基因来转化合适的宿主并测定它对该宿主脂肪酸分布的作用,从而经验性地测定特定脂肪酸去饱和酶的特异性。

[0112] 术语“延伸酶”指能延长脂肪酸碳链从而产生比该延伸酶作用于其上的脂肪酸底物长 2 个碳原子的酸的多肽。该延长过程在与脂肪酸合酶相关的多步骤机制中发生,如美国专利公开 2005/0132442 和国际专利公开 WO 2005/047480 所述。延伸酶体系催化反应的实例如 GLA 转化成 DGLA、STA 转化成 ETA 以及 EPA 转化成 DPA。通常,延伸酶的底物选择性有些广泛,但由链长度和不饱和的程度及类型两者来区分。例如, $C_{14/16}$ 延伸酶利用 C_{14} 底物 (如肉豆蔻酸), $C_{16/18}$ 延伸酶利用 C_{16} 底物 (如棕榈酸), $C_{18/20}$ 延伸酶 (也称为 $\Delta 6$ 延伸酶,两个术语可以互换使用) 利用 C_{18} 底物 (如 GLA、STA),而 $C_{20/22}$ 延伸酶利用 C_{20} 底物 (如 EPA)。以类似方式, $\Delta 9$ 延伸酶能够催化 LA 和 ALA 分别转化成 EDA 和 ETrA。重要的是,须注意一些延伸酶具有广泛的特异性因而单个酶可能能够催化几种延伸酶反应。例如单个酶可因此作为 $C_{16/18}$ 延伸酶和 $C_{18/20}$ 延伸酶。

[0113] 术语“转化效率”和“底物转化百分比”指特定酶 (如去饱和酶) 能够将底物转化成产物的效率。转化效率根据下面的公式测量: $([产物]/[底物+产物]) \times 100$, 其中‘产物’包括中间产物和该途径中来源于它的所有产物。

[0114] 术语“含油的”指那些倾向于以油形式贮存它们的能源的生物 (Weete, Fungal Lipid Biochemistry, 第 2 版, Plenum, 1980)。

[0115] 术语“含油酵母”指那些能制造油 (即 TAG) 的、分类为酵母的微生物。通常,含油微生物的细胞油或 TAG 含量符合 S 形曲线,其中脂质浓度增加直至在对数生长期晚期或稳定生长期早期它达到最高浓度,随后在稳定生长期晚期和死亡期期间逐渐下降 (Yongmanitchai 和 Ward, Appl. Environ. Microbiol., 57:419-25(1991))。本文所述的含油微生物通常积聚超过它们的干细胞重量约 25% 的油或 TAG。含油酵母的实例包括但不限于如下属:耶氏酵母属 (Yarrowia)、假丝酵母属 (Candida)、红酵母属 (Rhodotorula)、红冬孢酵母属 (Rhodosporidium)、隐球酵母属 (Cryptococcus)、丝孢酵母属 (Trichosporon) 和油脂酵母属 (Lipomyces)。

[0116] 如本文所用的,术语“分离的核酸片段”和“分离的核酸分子”可互换使用,并且是指单链或双链的,任选含有合成的、非天然的或改变了的核苷酸碱基的 RNA 或 DNA 聚合物。-DNA 聚合物形式的分离的核酸片段可由 cDNA、基因组 DNA 或合成 DNA 的一个或多个片段构成。

[0117] 当在合适的温度和溶液离子强度条件下单链形式的核酸片段可以退火至另一核酸片段时,核酸片段“可杂交”至另一核酸片段,例如 cDNA、基因组 DNA 或 RNA 分子。杂交条件和洗涤条件是众所周知的,并在 Sambrook, J., Fritsch, E. F. 和 Maniatis, T. Molecular Cloning: A Laboratory Manual, 第 2 版, Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989) 中得到举例说明。

[0118] 氨基酸或核苷酸序列的“基本部分”指这样的部分,该部分包括的多肽的氨基酸序列或基因的核苷酸序列足以推定鉴定所述多肽或基因,所述鉴定或者可以由本领域技术人员通过人工评价序列来完成,或者可以利用诸如 BLAST (Basic Local Alignment Search Tool; Altschul, S. F. 等人, J. Mol. Biol., 215:403-410 (1993)) 之类的算法通过计算机自动化序列比较和鉴定来完成。一般来讲,为了推测鉴定多肽或核酸序列是否与已知的蛋白质或基因同源,需要有 10 个或更多邻接氨基酸或者 30 个或更多邻接核苷酸的序列。此外,对于核苷酸序列,包含 20-30 个邻接核苷酸的基因特异性寡核苷酸探针可用于序列依赖性的基因鉴定(如 DNA 杂交)和基因分离(如细菌菌落或噬斑的原位杂交)的方法中。此外,12 至 15 个碱基的短寡核苷酸可在 PCR 中用作扩增引物,以便获得包含该引物的特定核酸片段。因此,核苷酸序列的“基本部分”所包含的序列应足以特异性地鉴定和/或分离包含该序列的核酸片段。

[0119] 术语“互补的”用于描述核苷酸碱基之间能够彼此杂交的关系。例如,对于 DNA,腺嘌呤与胸腺嘧啶互补,而胞嘧啶与鸟嘌呤互补。

[0120] 术语“同源性”或“同源”本文互换使用。它们指这样的核酸片段,即其中一个或多个核苷酸碱基改变并不会影响该核酸片段介导基因表达或产生某种表型的能力。这些术语也指本文所述的 Pex 核酸片段的修饰(例如缺失或插入一个或多个核苷酸),相对于初始的未经修饰的核酸片段,该修饰基本上不会改变所得核酸片段的功能特性。

[0121] 此外,技术人员认识到,同源核酸序列也由它们在中等严格条件(如 $0.5\times\text{SSC}$, 0.1% SDS, 60°C)下,与本文所示例的序列杂交的能力,或杂交至本文公开的核苷酸序列的任何部分以及杂交至与其功能相当的序列的能力所限定。

[0122] “密码子简并性”指允许核苷酸序列在不影响所编码的多肽的氨基酸序列的情况下发生变化的遗传密码的性质。技术人员非常了解具体宿主细胞在使用核苷酸密码子确定给定氨基酸时所表现出的“密码子偏好性”。因此,当合成基因用以改善在宿主细胞中的表达时,希望对基因进行设计,使得其密码子使用频率接近该宿主细胞优选的密码子使用频率。

[0123] “合成的基因”可由使用本领域技术人员已知的方法化学合成的寡核苷酸构件装配而成。将这些寡核苷酸基本单位构件进行退火并随后连接以形成基因节段,该基因节段随后在酶促作用下装配而构建成完整的基因。因此,基于最优化核苷酸序列以反映宿主细胞的密码子偏好性,可以定制基因用以最优化基因表达。如果密码子使用偏向于宿主偏好的那些密码子,则技术人员能预期成功的基因表达的可能。优选的密码子的确定可基于对

来源于宿主细胞的基因（其中序列信息可获得）的检测。

[0124] “基因”指表达特定蛋白的核酸片段，并且可以指单独的编码区或可以包含位于编码序列之前的调控序列（5′ 非编码区）和之后的调控序列（3′ 非编码区）。“天然基因”是指天然存在的具有其自己的调控序列的基因。“嵌合基因”是指不是天然基因的任何基因，包含在天然情况下不是一起存在的调控序列和编码序列。因此，嵌合基因可包括源于不同来源的调控序列和编码序列，或者包括源于同一来源但以不同于天然存在的方式排列的调控序列和编码序列。“内源性基因”指位于生物基因组内它的天然位置的天然基因。“外来”基因指通过基因转移导入到宿主生物内的基因。外来基因可包括插入到非天然生物内的天然基因、导入到天然宿主内的新位置的天然基因，或嵌合基因。“转基因”是已通过转化方法导入基因组内的基因。“密码子优化的基因”是其密码子使用频率经设计用以模仿宿主细胞优选的密码子使用频率的基因。

[0125] “编码序列”指编码特定氨基酸序列的 DNA 序列。“合适的调控序列”指位于编码序列的上游（5′ 非编码序列）、中间或下游（3′ 非编码序列）的核苷酸序列，其可影响相关编码序列的转录、RNA 加工或稳定性或者翻译。调控序列可包括启动子、增强子、静默子、5′ 非翻译前导序列（例如在转录起始位点和翻译启动密码子之间的序列）、内含子、多腺苷酸化识别序列、RNA 加工位点、效应子结合位点和茎-环结构。

[0126] “启动子”指能够控制编码序列或功能性 RNA 表达的 DNA 序列。一般来讲，编码序列位于启动子序列的 3′ 端。启动子可整个源于天然基因，或者由源于不同的天然存在的启动子的不同元件组成，或者甚至包含合成的 DNA 片段。本领域内的技术人员应当理解，不同的启动子可以在不同的组织或细胞类型中，或者在不同的发育阶段，或者响应不同的环境条件或生理条件而引导基因的表达。导致基因在大部分时间内在大多数细胞类型中表达的启动子通常称为“组成型启动子”。还应当进一步认识到，由于在大多数情况下调控序列的确切边界尚未完全确定，因此不同长度的 DNA 片段可能具有相同的启动子活性。

[0127] 术语“3′ 非编码序列”和“转录终止子”指位于编码序列下游的 DNA 序列。这包括多腺苷酸化识别序列和编码能影响 mRNA 加工或基因表达的调控信号的其它序列。多腺苷酸化信号通常特征在于影响多腺苷酸片添加到 mRNA 前体的 3′ 末端。3′ 区可影响相关编码序列的转录、RNA 加工或稳定性或翻译。

[0128] “RNA 转录物”指由 RNA 聚合酶催化 DNA 序列的转录所产生的产物。当 RNA 转录物是 DNA 序列的完全互补的拷贝时，它被称为初级转录物，或者它可以是源自初级转录物的转录后加工的 RNA 序列并被称作成熟 RNA。“信使 RNA”或“mRNA”指无内含子并且可以由细胞翻译成蛋白质的 RNA。“eDNA”指与 mRNA 互补并源于 mRNA 的双链 DNA。“有义”RNA 指包含 mRNA 并因而能由细胞翻译成蛋白质的 RNA 转录物。“反义 RNA”指与全部或部分靶初级转录物或 mRNA 互补，并阻止靶基因表达的 RNA 转录物（美国专利 5,107,065；国际申请公开 WO 99/28508）。反义 RNA 可以与特定基因转录物的任何部分，即 5′ 非编码序列、3′ 非编码序列或编码序列互补。“功能性 RNA”指反义 RNA、核酶 RNA 或其它不被翻译但是对细胞过程有影响作用的 RNA。

[0129] 术语“可操作地连接”指单个核酸片段上的核酸序列的关联，使得其中一个核酸序列的功能受到另一个核酸序列的影响。例如，当启动子能够影响编码序列的表达时，它可操作地连接编码序列。即，编码序列处于启动子的转录控制下。编码序列可以有义或反义

的取向可操作地连接至调控序列。

[0130] 如本文所用,术语“表达”指源于核酸片段的有义 RNA (mRNA) 或反义 RNA 的转录和稳定积聚。表达也可指将 mRNA 翻译成多肽。

[0131] “成熟”蛋白指经翻译后加工的多肽,即已经去除了存在于初始翻译产物中的任何前肽或肽原的多肽。“前体”蛋白质指 mRNA 的初级翻译产物,即前肽和肽原仍然存在。前肽和肽原可以是但不限于细胞内定位信号。

[0132] “转化”指将核酸分子转移至宿主生物中,导致在遗传上稳定遗传。例如,核酸分子可以是自主复制的质粒,或者它可以整合进宿主生物的基因组中。含有转化核酸片段的宿主生物被称为“转基因”或“重组”或“转化”生物体。

[0133] “稳定转化”指将核酸片段转移至宿主生物的基因组(包括核基因组和细胞器基因组)中,导致在遗传上稳定遗传。相反,“瞬时转化”指将核酸片段转移至宿主生物的核中或包含 DNA 的细胞器中,导致基因表达而不整合或不稳定遗传。含有转化核酸片段的宿主生物被称为“转基因”生物体。

[0134] 术语“质粒”和“载体”指通常携带有不属于细胞中心代谢部分的基因的染色体外元件,并且常常是环状双链 DNA 片段的形式。这类元件可以是源自任何来源的自主复制序列、基因组整合序列、噬菌体或单链或双链 DNA 或 RNA 的核苷酸序列(线性或环状),其中多个核苷酸序列已连接或重组为一种独特构建体,该独特构建体能够将表达盒引入细胞中。

[0135] 术语“表达盒”指包含如下编码序列的 DNA 片段:所选基因的编码序列和所选基因产物表达所需的位于编码序列之前(5' 非编码序列)和之后(3' 非编码序列)的调控序列。因此,表达盒通常由如下序列构成:(1) 启动子序列;2) 编码序列,即,开放阅读框[“ORF”]和,3) 3' 非翻译区域,即,在真核细胞中通常包含聚腺苷酸位点的终止子。表达盒通常包含于载体中以有利于克隆和转化。可以将不同表达盒转化进包括细菌、酵母、植物和哺乳动物细胞在内的不同生物体中,只要能针对每种宿主使用正确的调控序列。

[0136] 术语“同一性百分比”指两种或更多种多肽序列之间或两种或更多种多核苷酸序列之间的关系,该关系通过对序列进行比较而确定。“同一性”还表示多肽或多核苷酸序列之间的序列关联的程度,根据具体情况,它由比较序列的序列串之间的匹配百分比确定。“同一性百分比”和“相似性百分比”可容易地通过已知方法计算出来,所述的方法包括但不限于以下文献中所描述的那些:1.) Computational Molecular Biology (Lesk, A.M. 编辑) Oxford University: NY (1988); 2.) Biocomputing: Informatics and Genome Projects (Smith, D.W. 编辑) Academic: NY (1993); 3.) Computer Analysis of Sequence Data, Part I (Griffin, A.M. 和 Griffin, H.G. 编辑) Humana: NJ (1994); 4.) Sequence Analysis in Molecular Biology (von Heinje, G. 编辑) Academic (1987); 和 5.) Sequence Analysis Primer (Gribskov, M. 和 Devereux, J. 编辑) Stockton: NY (1991)。

[0137] 设定确定同一性百分比的优选方法来用于给出待测试序列之间的最佳匹配。用于测定同一性百分比和相似性百分比的方法在公开可获得的计算机程序中进行编辑。序列比对和同一性百分比计算可以用 LASERGENE 生物信息学计算机软件包 (LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI)) 中的 MegAlign™ 程序进行。序列的多重比对采用包括几种改变形式的算法在内的“Clustal 比对方法”进行,包括“Clustal V 比对方法”和“Clustal W 比对方法”(在 Higgins 和 Sharp, CABIOS, 5:

151-153(1989);Higgins, D. G. 等人, Comput. Appl. Biosci., 8:189-191(1992) 中有所描述) 和可以在 LASERGENE 生物信息学计算机软件包(DNASTAR Inc.) 中的 MegAlign™v6.1 程序中找到 的比对方法。用 Clustal 程序比对序列后,可通过查看程序中的“序列距离”表来获得“同一性百分比”。

[0138] 本领域的技术人员非常清楚,多种程度的序列同一性百分比可用于从其它物种中鉴定多肽,其中这类多肽具有相同或相似的功能或活性。同一性百分比的有用实例包括但不限于 50%、55%、60%、65%、70%、75%、80%、85%、90%、或 95%、或从 50%至 100%的任何整数百分比。实际上,从 50%至 100%的任何整数氨基酸同一性可用于描述在本文所述的方法和宿主细胞中编码多肽的合适核酸片段(分离的多核苷酸),例如 51%、52%、53%、54%、55%、56%、57%、58%、59%、60%、61%、62%、63%、64%、65%、66%、67%、68%、69%、70%、71%、72%、73%、74%、75%、76%、77%、78%、79%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或 99%。在一些情况下,合适的核酸片段(分离的多核苷酸)编码与本文报道的氨基酸序列至少约 70%相同,优选地至少约 75%相同,更优选地至少约 80%相同的多肽。优选的核酸片段编码与本文所报道的氨基酸序列具有至少约 85%同一性的氨基酸序列。更优选的核酸片段编码与本文所报道的氨基酸序列具有至少约 90%同一性的氨基酸序列。最优选的是编码与本文所报道的氨基酸序列具有至少约 95%同一性的氨基酸序列的核酸片段。

[0139] 合适的核酸片段不但具有上述同源性,而且通常编码具有至少 50 个氨基酸,优选至少 100 个氨基酸,更优选至少 150 个氨基酸,更优选至少 200 个氨基酸,最优选至少 250 个氨基酸的多肽。

[0140] 术语“序列分析软件”指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”可商购获得或独立开发。典型的序列分析软件包括但不限于:1.) GCG 程序包(Wisconsin Package Version 9.0, Genetics Computer Group(GCG), Madison, WI);2)BLASTP, BLASTN, BLASTX(Altschul 等人, J. Mol. Biol., 215:403-410(1990));3)DNASTAR(DNASTAR, Inc. Madison, WI);4)Sequencher(Gene Codes Corporation, Ann Arbor, MI);和 5.)整合了 Smith-Waterman 算法的 FASTA 程序(W. R. Pearson, Comput. Methods Genome Res., [Proc. Int. Symp.] (1994), 召开年份:1992, 111-20. 编辑:Suhai, Sandor. Plenum:New York, NY)。在这一描述中,除非另外指明,只要序列分析软件用于分析,分析结果都基于所用程序的“默认值”。在此所用的“默认值”指在首次初始化软件时软件最初加载的任何值或参数集。

[0141] 本文使用的标准重组 DNA 和分子克隆技术是本领域熟知的并且已经在如下文献中有所描述:Sambrook, J., Fritsch, E. F. 和 Maniatis, T., Molecular Cloning: A Laboratory Manual, 第 2 版, Cold Spring Harbor Laboratory: Cold Spring Harbor, NY(1989)(下文称为“Maniatis”);Silhavy, T. J.、Bennan, M. L. 和 Enquist, L. W., Experiments with Gene Fusions, Cold Spring Harbor Laboratory: Cold Spring Harbor, NY(1984);以及 Ausubel, F. M. 等人, Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience 出版, Hoboken, NJ(1987)。

[0142] 综述:脂肪酸和三酰基甘油的生物合成

[0143] 通常,含油微生物中的脂质蓄积是响应存在于生长培养基中的总的碳氮比而触

发。该过程（导致含油微生物中游离棕榈酸（16:0）的从头合成）在美国专利 7,238,482 中有详细描述。棕榈酸是长链饱和和不饱和脂肪酸衍生物的前体，这些脂肪酸衍生物通过延伸酶和去饱和酶的作用形成（图 1）。

[0144] TAG（脂肪酸的主要贮存单位）通过涉及如下反应的一系列反应形成：(1) 一分子的酰基辅酶 A 和甘油-3-磷酸通过酰基转移酶酯化而生成溶血磷脂酸；(2) 第二分子的酰基辅酶 A 通过酰基转移酶酯化而生成 1,2-二酰基甘油磷酸（通常称为磷脂酸）；(3) 通过磷脂酸磷酸酶移除磷酸以生成 1,2-二酰基甘油 [“DAG”]；和 4) 通过酰基转移酶的作用加入第三脂肪酸以形成 TAG。

[0145] 广谱的脂肪酸可被引入 TAG 中，包括饱和的和不饱和的脂肪酸以及短链和长链的脂肪酸。一些能通过酰基转移酶被引入 TAG 的脂肪酸的非限制性实例包括：癸酸（10:0）、月桂酸（12:0）、肉豆蔻酸（14:0）、棕榈酸（16:0）、棕榈油酸（16:1）、硬脂酸（18:0）、油酸（18:1）、异油酸（18:1）、LA（18:2）、桐油酸（18:3）、GLA（18:3）、ALA（18:3）、STA（18:4）、花生酸（20:0）、EDA（20:2）、DGLA（20:3）、ETra（20:3）、ARA（20:4）、ETA（20:4）、EPA（20:5）、二十二烷酸（22:0）、DPA（22:5）、DHA（22:6）、木蜡酸（24:0）、神经酸（24:1）、蜡酸（26:0）和褐煤酸（28:0）脂肪酸。在本文所述的方法和宿主细胞中，最期望将“长链”PUFA 引入 TAG 中，其中长链 PUFA 包括来源于 18:1 底物的任何脂肪酸，所述底物具有至少 18 个碳原子，即， C_{18} 或更长。其也包括羟基化脂肪酸、环氧脂肪酸和共轭亚油酸。

[0146] 虽然大多数 PUFA 以中性脂类的形式掺入 TAG 并贮存于脂质体中，但是重要的是认识到在含油生物中测量总 PUFA 应包括那些位于磷脂酰胆碱级分、磷脂酰乙醇胺级分、和三酰基甘油（也称为 TAG 或油）级分中的 PUFA。

[0147] ω 脂肪酸的生物合成

[0148] 其中油酸转化成 ω -3/ ω -6 脂肪酸的代谢过程包括通过添加碳原子使碳链延长和通过加入双键使分子去饱和。这需要存在于内质网膜内的一系列专门的去饱和酶和延伸酶。然而，如在图 1 中看到的和如下所述的，通常存在多个用于产生特定 ω -3/ ω -6 脂肪酸的替代途径。

[0149] 具体地讲，图 1 描述了下文所述的途径。所有途径都需要最初通过 $\Delta 12$ 去饱和酶将油酸转化成亚油酸 LA（第一个 ω -6 脂肪酸）。然后，利用“ $\Delta 6$ 去饱和酶/ $\Delta 6$ 延伸酶途径”并将 LA 用作底物来如下形成长链 ω -6 脂肪酸：1) 通过 $\Delta 6$ 去饱和酶将 LA 转化成 γ -亚麻酸 [“GLA”]；2) 通过 $C_{18/20}$ 延伸酶将 GLA 转化成二高- γ -亚麻酸 [“DGLA”]；3) 通过 $\Delta 5$ 去饱和酶将 DGLA 转化成花生四烯酸 [“ARA”]；4) 通过 $C_{20/22}$ 延伸酶将 ARA 转化成二十二碳四烯酸 [“DTA”]；以及，5) 通过 $\Delta 4$ 去饱和酶将 DTA 转化成二十二碳五烯酸 [“DPA ω -6”]。

[0150] 作为另外一种选择，“ $\Delta 6$ 去饱和酶/ $\Delta 6$ 延伸酶途径”能够如下使用 α -亚麻酸 [“ALA”] 作底物生产长链 ω -3 脂肪酸：(1) 通过 $\Delta 15$ 去饱和酶将 LA 转化成 ALA（第一个 ω -3 脂肪酸）；(2) 通过 $\Delta 6$ 去饱和酶将 ALA 转化成十八碳四烯酸 [“STA”]；(3) 通过 $C_{18/20}$ 延伸酶将 STA 转化成二十碳四烯酸 [“ETA”]；(4) 通过 $\Delta 5$ 去饱和酶将 ETA 转化成二十碳五烯酸 [“EPA”]；(5) 通过 $C_{20/22}$ 延伸酶将 EPA 转化成二十二碳五烯酸 [“DPA”]；以及，(6) 通过 $\Delta 4$ 去饱和酶将 DPA 转化成二十二碳六烯酸 [“DHA”]。任选地， ω -6 脂肪酸可转化成 ω -3 脂肪酸；例如，ETA 和 EPA 通过 $\Delta 17$ 去饱和酶活性分别从 DGLA 和 ARA 生成。

[0151] ω -3/ ω -6 脂肪酸生物合成的替代途径利用 Δ 9 延伸酶和 Δ 8 去饱和酶,即,“ Δ 9 延伸酶/ Δ 8 去饱和酶途径”。更具体地讲,LA 和 ALA 可通过 Δ 9 延伸酶分别转化成 EDA 和 ETrA; Δ 8 去饱和酶然后将 EDA 转化成 DGLA 和/或将 ETrA 转化成 ETA。随后如上所述形成下游 PUFA。

[0152] 本文宿主生物必须天然或经由基因工程技术具有生产 PUFA 的能力。虽然多种微生物能够在常见的细胞代谢过程中合成 PUFA(包括 ω -3/ ω -6 脂肪酸),并且它们中的一些能进行商业培养,但是这些生物很少生产或不生产具有期望的油含量的油和用于药物、膳食替代品、医疗食品、营养补充剂、其它食物产品、工业油脂化学制品或其它终端产品中的组合物。因此,越来越着重于工程化微生物以生产“设计”的脂质和油的能力,其中脂肪酸含量和组成通过基因工程进行详细指定。在此基础上,期望宿主可能包含编码功能性 PUFA 生物合成途径的异源基因,但这是非必需的。

[0153] 如果宿主生物不天然生产期望的 PUFA 或具有期望的脂质分布,本领域的技术人员熟悉将编码用于 PUFA 生物合成的合适酶的一个或多个表达盒导入所选的宿主生物所必需的构思和技术。文献为技术人员提供了多个教导,用于将此类表达盒导入多个宿主生物。将使用宿主生物解脂耶氏酵母的一些文献提供如下:美国专利 7,238,482、国际申请公布 WO 2006/033723、专利申请公布 US-2006-0094092、专利申请公布 US-2006-0115881-A1 和专利申请公布 US-2006-0110806-A1。上述列举并非穷尽性的因此不应理解为限制性的。

[0154] 简而言之,多种 ω -3/ ω -6PUFA 产品能在它们转化成 TAG 之前被生产出来,这取决于脂肪酸底物和存在于或被转化到宿主细胞中的 ω -3/ ω -6 脂肪酸生物合成途径的特定基因。同样地,期望脂肪酸产品的生产可能直接发生或间接发生。当不经过任何中间步骤或中间途径将脂肪酸底物直接转化成期望的脂肪酸产品时,直接生产发生。当编码 PUFA 生物合成途径的多个基因可被组合使用,使得一系列反应发生以生产期望的 PUFA 时,间接生产发生。具体地讲,可期望用包含 Δ 12 去饱和酶、 Δ 6 去饱和酶、 $C_{18/20}$ 延伸酶、 Δ 5 去饱和酶和 Δ 17 去饱和酶的表达盒转化含油酵母以过剩生产 EPA。参见美国专利 7,238,482 和国际申请公布 WO 2006/052870。为本领域的技术人员熟知的是可将编码 PUFA 生物合成途径酶的基因的多种其它组合用于在含油生物中表达(参见图 1)。包含于具体表达盒中的具体基因取决于宿主生物、它的 PUFA 分布和/或去饱和酶/延伸酶分布、底物的可利用性和期望的终产物。

[0155] 许多候选基因具有所需去饱和酶和/或延伸酶活性,它们能根据公开可获得的文献进行鉴定,如 GenBank、专利文献、和对具有生产 PUFA 能力的生物进行的实验分析。可用的去饱和酶和延伸酶序列可源自任何来源,例如,分离自天然来源如细菌、藻类、真菌、卵菌、酵母、植物、动物等、经由半合成途径产生或从头合成。在鉴定这些候选基因后,选择具有去饱和酶或延伸酶活性的特定多肽时的考虑事项包括:1) 多肽的底物特异性;2) 多肽或其组件是否为限速酶;3) 去饱和酶或延伸酶是否是合成期望的 PUFA 所必需的;4) 多肽所需的辅因子;和/或 5) 多肽在其产生后是否被修饰,例如,通过激酶或异戊烯转移酶修饰。

[0156] 表达的多肽优选具有与它在宿主细胞中的位置的生化环境相容的参数。参见美国专利 7,238,482。考虑每种特定的去饱和酶和/或延伸酶的转化效率也可以是有用的。更具体地讲,由于每种酶极少能以 100%的效率将底物转化成产物,宿主细胞内所产生的未纯化的油的最终脂质分布通常是由期望的 ω -3/ ω -6 脂肪酸及多种上游 PUFA 中间产物组成

的多种 PUFA 的混合物。因此,当使期望的脂肪酸的生物合成最优化时,每种酶的转换效率也是要考虑的变量。

[0157] 过氧化物酶体生物合成和 Pex 基因

[0158] 如前文所述,过氧化物酶体是普遍存在于所有真核细胞中的细胞器。它们的主要作用是降解细胞定位细胞器中的各种物质,例如毒性化合物、脂肪酸等。例如,在过氧化物酶体中能够发生 β -氧化过程,其中脂肪酸分子发生降解,最后产生乙酰-CoA 的游离分子(它被回输到细胞溶质中)。虽然线粒体中的 β -氧化过程导致 ATP 合成,在过氧化物酶体中的 β -氧化引起高电位电子转移到 O_2 中并导致形成 H_2O_2 , H_2O_2 随后被过氧化物酶体过氧化氢酶转化成水和 O_2 。特长链(如 C_{18} 至 C_{22})脂肪酸在过氧化物酶体中经过初始 β -氧化,然后经过线粒体 β -氧化。

[0159] 已知负责通过 ATP 水解使蛋白穿过过氧化物酶体膜的蛋白是过氧化物酶体生物合成因子蛋白,或“peroxins”。这些过氧化物酶体生物合成因子蛋白也包括那些涉及过氧化物酶体生物合成/装配的蛋白。过氧化物酶体生物合成因子蛋白的基因首字母缩写是 Pex;并且命名体系描述于 Distel 等人, *J. Cell Biol.*, 135:1-3(1996)。迄今已经在多种真核生物中鉴定了至少 32 个不同的 Pex 基因。然而在真菌中, Kiel 等人近来的文献(*Traffic*, 7:1291-1303(2006))提出过氧化物酶体生物合成/基质蛋白输入所需的最小数目是 17,因此仅需要 Pex1p、Pex2p、Pex3p、Pex4p、Pex5p、Pex6p、Pex7p、Pex8p、Pex10p、Pex12p、Pex13p、Pex14p、Pex17p、Pex19p、Pex20p、Pex22p 和 Pex26p。这些蛋白以协同方式作用以增生(复制)过氧化物酶体并经由转移进入过氧化物酶体以输入蛋白(参见 Waterham, H. R. 和 J. M. Cregg, *BioEssays*. 19(1):57-66(1996))。

[0160] 最初从对突变体的分析中分离了多种 Pex 基因,所述突变体显示具有异常的过氧化物酶体功能或结构。然而,由于全部基因组序列是可获得的,所以通过基于同源性的计算机序列搜索来鉴定 Pex 基因变得更加容易。Kiel 等人(*Traffic*, 7:1291-1303(2006))引用过氧化物酶体生物合成机制的强保守性,尽管它偶尔会出现低序列相似性。更具体地讲,在酵母和丝状真菌中,它们的数据指示迄今鉴定的几乎所有 Pex 蛋白都是保守的。下表 4 显示 Kiel 等人鉴定的过氧化物酶体生物合成因子蛋白(同上),它们来源于啤酒糖酵母(*Saccharomyces cerevisiae*)、光滑假丝酵母(*Candida glabrata*)、棉阿舒囊霉(*Ashbya gossypii*)、乳酸克鲁维酵母(*Kluyveromyces lactis*)、白假丝酵母(*Candida albicans*)、汉逊德巴利酵母(*Debaryomyces hansenii*)、巴斯德毕赤酵母(*Pichia pastoris*)、多形汉逊酵母(*Hansenula polymorpha*)、解脂耶氏酵母(*Yarrowia lipolytica*)、烟曲霉(*Aspergillus fumigatus*)、构巢曲霉(*Aspergillus nidulans*)、产黄青霉(*Penicillium chrysogenum*)、稻瘟病菌(*Magnaporthe grisea*)、粗糙脉孢菌(*Neurospora crassa*)、玉蜀黍赤霉(*Gibberella zeae*)、玉米黑粉菌(*Ustilago maydis*)、新生隐球菌新生变种(*Cryptococcus neoformans* var. *neoformans*)和栗酒裂殖酵母(*Schizosaccharomyces pombe*)。

[0161]

表 4

真菌过氧化氢酶体生物合成因子蛋白的 GenBank 保藏号

[从表 2 中重建, Kiel 等人, (Traffic, 7:1291-1303 (2006))]

	啤酒糖 酵母 (<i>Saccharomyces cerevisiae</i>)	光滑假丝 酵母 (<i>Candida glabrata</i>)	棉阿舒囊霉 (<i>Ashbya gossypii</i>)	乳酸克鲁维 酵母 (<i>Kluyveromyces lactis</i>)	白假丝 酵母 (<i>Candida albicans</i>)	汉逊德巴利 酵母 (<i>Debaryomyces hansenii</i>)	巴斯德毕赤 酵母 (<i>Pichia pastoris</i>)	多形汉逊 酵母 (<i>Hansenula polymorpha</i>)	解脂耶氏 酵母
Pex1p	CAA82041	CAG60131	AAS53742	CAH02218	EAL02496	CAG89689	CAA85450	AAD52811	CAG82178
Pex2p	CAA89508	CAG60461	AAS50677	CAH00186	EAK95929	CAG85956	CAA65646	AAT97412	CAG77647
Pex3p	AAB64764	CAG62379	AAS52217	CAG99801	EAK94771	CAG89890	CAA96530	AAC49471	CAG78565
Pex3Bp	—	—	—	—	—	—	na	—	CAG83356
Pex4p	CAA97146	CAG60639	AAS53685	CAG99212	EAL03336	CAG87262	AAA53634	AAC16238	CAG79130
Pex5p	CAA89730	CAG61665	AAS53824	CAH01742	EAK94251	CAG89098	AAB40613	AAC49040	CAG78803
Pex5Bp	—	CAG61076	—	—	—	—	na	—	—
Pex5Cp	CAA89120 (Ymr018wp)	—	—	—	—	—	na	—	—
Pex5/20p	—	—	—	—	—	—	na	—	—
Pex5Rp	—	—	—	—	—	—	na	—	—
Pex6p	AAA16574	CAG58438	AAS54884	CAG99125	EAK95956	CAG87108	CAA80278	AAD52812	CAG82306
Pex7p	CAA57183	CAG57936	AAS54301	CAG99215	EAK95226	CAG87150	AAC08303	ABA64462	CAG78389

[0162]

Pex8p	CAA97079	CAG61238	AAS52889	CAH01253	EAK91777, EAK91778*	CAG89446	AAC41653	CAA82928	CAG80447
Pex9p	被错误鉴定的 ORF	—	—	—	—	—	—	—	—
Pex10p	AAB64453	CAG62699	AAS53069	CAG99788	翻译 AACQ- 01000128, 核 苷酸 37281- 36306(包含内 含子)	CAG89101	AAB09086	CAA86101	CAG81606
Pex12p	CAA89129	CAG62649	AAS50837	CAG99378	EAL00707	CAG84342	AAC49402	AAM66157	CAG81532
Pex13p	AAB46885	CAG57840	AAS51456	CAG99931	EAK97421	CAG86337	AAB09087	DQ345349	CAG81789
Pex14p	AAS56829	CAG58828	AAS54871	CAG99440	EAK90926	CAG91028	AAG28574	AAB40596	CAG79323
Pex15p	CAA99046	CAG58938	AAS51506	CAG98135	—	—	na	—	—
Pex16p	—	—	—	—	—	—	na	—	CAG79622
Pex17p	CAA96116	CAG61398	AAS50595	CAH01010	EAK95385	CAG86168	AAF19606	DQ345350	CAG84025
Pex14/17p	—	—	—	—	—	—	na	—	—
Pex18p	AAB68992	—	—	—	—	—	na	—	—
Pex19p	CAA98630	CAG58359	AAS52741	CAG99258	EAK97275	CAG84799	AAD43507	AAK84070	AAK84827
Pex20p	—	—	—	—	EAK91603, EAK94766*	CAG87898	AAX11696	AAX14715	CAG79226

[0163]

Pex21p	CAA97267	CAG59241	AAS51769	CAG99735	—	—	na	—	—
Pex21Bp	—	CAG60281	—	—	—	—	na	—	—
Pex22p	AAC04978	CAG60970	AAS52329	CAG97800	EAK91040	CAG88727	AAD45664	DQ384616	CAG77876
Pex22p 样	—	—	—	—	—	na	—	—	EAL90994
Pex26p	—	—	—	—	EAK91093	CAG88929	na	DQ645588	反义翻译 NC_006072, 核苷酸 117230-118387

[0164]

	烟曲霉 (<i>Aspergillus fumigatus</i>)	构巢曲霉 (<i>Aspergillus nidulans</i>)	产黄青霉 (<i>Penicillium chrysogenum</i>)	稻瘟病菌 (<i>Magnaporthe grisea</i>)	粗梗脉孢菌 (<i>Neurospora crassa</i>)	玉蜀黍赤霉 (<i>Gibberella zeae</i>)	玉米黑粉菌 (<i>Ustilago maydis</i>)	新生隐球菌 新生变种 (<i>Cryptococcus neoformans</i> var. <i>neoformans</i>)	粟酒裂殖酵母 (<i>Schizosaccharomyces pombe</i>)
Pex1p	EAL93310	EAA57740	AAG09748	XP_364454	EAA34641	EAA76787	EAK85195	AAW43248	CAA19256
Pex2p	EAL88068	EAA58944	DQ793192	XP_368589	EAA35361	EAA70670	EAK81310	AAW40683	CAA16981
Pex3p	EAL91965	EAA64392	DQ793193	XP_369909	EAA33751	EAA76989	EAK87104	AAW42444	CAB10141
Pex3Bp	—	—	—	—	—	—	—	—	—
Pex4p	EAL87211	翻译 AACD0-1000130, 核苷酸 150195-150738 (包含内含子)	DQ793194	XP_369064	EAA34737	EAA76379	翻译 AACD0-1000006, 核苷酸 97041-96550 (包含内含子)	—	CAB91184
Pex5p	EAL85289	EAA63772	AAR12222	XP_360528	EAA36111	EAA68640	EAK83659	AAW46349	CAA22179
Pex5Bp	—	—	—	—	—	—	—	—	—
Pex5Cp	—	—	—	—	—	—	—	—	—
Pex5/20p	—	—	—	—	—	—	EAK82973	AAW41849	—

[0165]

Pex5Rp	—	—	—	—	—	—	—	—	—	—	—	—
Pex6p	EAL92776	EAA63496	AAG09749	XP_368715	EAA36040	EAA73732	EAK83459	AAW45333	CAB11501	—	—	—
Pex7p	EAL90870	EAA65909	DQ793195	XP_363555	AAN39560	EAA74171	EAK84499	AAW41119	P78798	—	—	—
Pex8p	EAL93137	EAA57947	DQ793196	XP_359449	EAA27783	EAA77627	EAK83936	AAW43468	CAB53406	—	—	—
Pex9p	—	—	—	—	—	—	—	—	—	—	—	—
Pex10p	EAL87045	EAA62774	DQ793197	XP_369099	EAA34967	EAA76761	EAK83811	AAW45079	CAB51769	—	—	—
Pex12p	EAL93972	EAA61357	DQ793198	XP_363845	EAA32773	EAA76413	EAK81282	AAW46724	CAD27496	—	—	—
Pex13p	EAL85282	EAA63824	DQ793199	XP_369087	EAA35785	EAA68396	EAK84395	AAW42381	CAB16740	—	—	—
Pex14p	EAL92562	EAA61046	DQ793200	XP_368216	EAA28304	EAA76904	EAK83123	AAW46857	CAA18656	—	—	—
Pex15p	—	—	—	—	—	—	—	—	—	—	—	—
Pex16p	EAL88469	EAA62294	DQ793201	XP_364166	EAA34648	EAA71849	EAK82801	AAW43797	CAA22819	—	—	—
Pex17p	参见 Pex14/17p	—	—	—	—	—	—	—	—	—	—	—
Pex14/17p	EAL93590	EAA58642	DQ793202	XP_368163	EAA27748	EAA73655	EAK81127	—	—	—	—	—
Pex18p	—	—	—	—	—	—	—	—	—	—	—	—
Pex19p	EAL92487	EAA60977	DQ793203	XP_368273	EAA31855	EAA70162	EAK86072	AAW42876	CAA97344	—	—	—
Pex20p	EAL90176	EAA60479	DQ793204	XP_368606	AAN39561	EAA76911	—	—	—	—	—	—
Pex21p	—	—	—	—	—	—	—	—	—	—	—	—
Pex21Bp	—	—	—	—	—	—	—	—	—	—	—	—
Pex22p	—	—	—	—	—	—	—	—	—	—	—	—

[0166]

Pex22p 样	EAL90994	EAA66006	DQ793205	XP_365689	EAA26537	翻译 AACM0- 1000080, 核苷 酸 4362-3039 (包含内含子)	—	—	—
Pex26p	EAL93994	EAA61336	DQ793206	XP_359606	EAA28582	EAA76391	—	—	—

* 在不重叠群上编码的部分 ORF。

[0167] 导致过氧化物酶体生物合成减弱的 Pex 基因突变在酵母、人类和植物中引起严重的代谢和发育紊乱 (Eckert, J. H. 和 R. Erdmann, Rev. Physiol. Biochem Pharmacol.,

147:75-121(2003);Weller, S. 等人, AnnualReview of Genomics and Human Genetics, 4:165-211(2003);Wanders, R. J., Am. J. Med. Genet., 126A:355-375(2004);Mano, S. 和 M. Nishimura, Vitam Horm., 72:111-154(2005);Wanders, J. A., 和 H. R. Waterham, Annu. Rev. Biochem., 75:295-332(2006);Fujiki, Yukio. Peroxisome Biogenesis Disorders. In, Encyclopedia of Life Sciences. JohnWiley&Sons, 2006)。例如 X-连锁肾上腺脑白质营养不良 [“X-ALD”] 和 Zellweger 综合征、以及若干个严重程度较低的疾病,它们可能由单个酶缺乏和 / 或过氧化物酶体生物合成失调引起。

[0168] 在酵母、解脂耶氏酵母中,已经分离并鉴定了多种不同的 Pex 基因,如上表 4 中鉴定的那些 Pex 基因。更具体地讲,Bascom, R. A. 等人 (Mol. Biol. Cell, 14:939-957(2003)) 描述了 YlPex3p;Szilard, R. K. 等人 (J. Cell Biol., 131:1453-1469(1995)) 描述了 YlPex5p;Nuttley, W. M. 等人 (J. Biol. Chem., 269:556-566(1994)) 描述了 YlPex6p;Elizen G. A., 等人 (J. Biol. Chem., 270:1429-1436(1995)) 描述了 YlPex9p;Elizen G. A., 等人 (J. Cell Biol., 137:1265-1278(1997)) 和 Titorenko, V. I. 等人 (Mol. Cell Biol., 17:5210-5226(1997)) 描述了 YlPex16p;Lambkin, G. R. 和 R. A. Rachubinski (Mol. Biol. Cell., 12(11):3353-3364(2001)) 描述了 YlPex19;以及 Titorenko V. I., 等人 (J. Cell Biol., 142:403-420(1998)) 和 Smith J. J. 以及 R. A. Rachubinski (J. Cell Biol., 276:1618-1625(2001)) 描述了 YlPex20p。

[0169] 本文最初关注的基因是 YlPex10p (GenBank 保藏号 CAG81606、AB036770 和 AJ012084)。它由 Sumita 等人证明 (FEMS Microbiol. Lett., 214:31-38(2002)): 1) YlPex10p 用作过氧化物酶体的组分;并且 2) YlPex10p 的 C₃HC₄ 环状锌指基序是蛋白功能必需的,如经由制造 C341S、C346S 和 H343W 点突变并随后进行生长分析来所测定的。

[0170] 已经在其它生物中完成了对 Pex10 的 C₃HC₄ 环状锌指基序的研究并具有类似结果。例如,发现改变在巴斯德毕赤酵

[0171] 母 (*Pichia pastoris*) 的 Pex10p C₃HC₄ 基序中的保守残基的点突变会使蛋白无功能 (Kalish, J. E. 等人, Mol. Cell Biol., 15:6406-6419(1995))。同样地,在对纤维原细胞系进行功能互补分析后,Warren D. S. 等人 (Hum. Mutat., 15(6):509-521(2000)) 得出结论:C₃HC₄ 基序对 Pex10p 功能是至关重要的。若干个研究结果显示 Pex10p 在拟南芥属中丧失功能会引起胚芽在心期死亡 (Hu, J. 等人, Science, 297:405-409(2002);Schmummann, U. 等人, Proc. Natl. Acad. Sci. U. S. A., 100:9626-9631(2003);Sparkes, I. A. 等人, Plant Physiol., 133:1809-1819(2003);Fan, J. 等人, Plant Physiol., 139:231-239(2005))。在后继研究中,Schemann, U. 等人 (Proc. Natl. Acad. Sci. U. S. A., 104:1069-1074(2007)) 研究了 Pex10p 在非致命的功能部分丧失拟南芥属突变体中的功能。具体地讲,在拟南芥属野生型植株中制备具有功能异常的 C₃HC₄ 基序的表达 Pex10p 的四个 T-DNA 插入序列。突变植物显示具有减少的叶片过氧化物酶体,并且作者提出 Pex10p 中环指基序的失活消除了连结过氧化物酶体与叶绿体以及在过氧化物酶体和叶绿体之间转移代谢物所需的蛋白相互作用。

[0172] 虽然研究尚未鉴定其它 Pex 蛋白中必需的结构域,但是研究认识到了多种 Pex 突变体的效应,从而认识到各种不同的生物进化出的用于装配、保持、繁殖和遗传过氧化物酶体 (一种已知在脂质代谢中起作用的细胞器) 的方法和分子机制。例如, Bascom,

R. A. 等人已经进行了解脂耶氏酵母 Pex3p 的基因敲除和超表达 (Mol. Biol. Cell, 14: 939-957 (2003))。基因敲除细胞不包含野生型过氧化物酶体, 相反具有多个小囊泡; 超表达导致细胞具有较少、较大和聚集的过氧化物酶体。他们假定 Pex3p 涉及通过固定过氧化物酶体生物合成组件启动过氧化物酶体装配, 即, 过氧化物酶体靶信号 (PTS) 1 和 2 输入机制。同样地, Guo, T. 等人, 敲除解脂耶氏酵母 Pex16p 导致过多地增殖不成熟的过氧化物酶体囊泡并显著地降低它们转化成成熟过氧化物酶体的速率和效率 (J. Cell Biol., 162: 1255-1266 (2003)), 然而超表达导致产生很少但是增大的过氧化物酶体 (Eitzen 等人, J. Cell Biol., 137: 1265-1278 (1997))。Guo 等人得出结论: Pex16p 负调节早期的过氧化物酶体前体分裂所需的膜分裂事件。

[0173] 尽管已有上文概述的研究进展, 关于不同 Pex 蛋白的作用、它们彼此相互作用和过氧化物酶体中的生物合成 / 装配机制的细节仍未被阐明。同样地, 本专利申请所述的数字是在对其它植物或动物的研究中尚未验证的新观察数据, 其中在 YlPex10p 的 C_3HC_4 基序中的突变或 YlPex3p、YlPex10p 或 YlPex16p 的敲除导致产生解脂耶氏酵母突变体, 该突变体具有提高的 PUFA 掺入能力, 尤其是掺入长链 PUFA 如 C_{20} 至 C_{22} 分子到细胞的总脂质级分和油级分中的能力。

[0174] 已经提出过氧化物酶体是脂质的分解代谢和合成代谢所必需的 (LinY. 等人, Plant Physiology, 135: 814-827 (2004)); 然而, 该假说是基于对 Pex16p 的同源物的研究。更具体地讲, Lin, Y. 等人 (同上) 报道拟南芥属 Shrunken 种子 1 (sse1) 突变体具有异常的过氧化物酶体生物合成和脂肪酸合成, 这基于 sse1 种子中的油与野生型相比减少了 10-16%。Binns, D. 等人 (J. Cell Biol., 173 (5): 719-731 (2006)) 检查了啤酒糖酵母中的过氧化物酶体 - 脂质体相互作用并测定两个细胞器之间的广泛物理接触促进脂质体内的结合脂解及过氧化物酶体脂肪酸氧化。更具体地讲, 检查不同 Pex 基因敲除菌株的游离脂肪酸与 TAG 的比率, 发现比率相对于野生型提高。很明显, 为了了解过氧化物酶体, 尤其是 Pex3p、Pex10p 和 Pex16p 蛋白的代谢作用, 进一步的研究将是必须的。

[0175] 不希望受任何具体说明或理论的限制, 假定在含油酵母细胞中的 Pex 基因破坏或敲除影响过氧化物酶体中天然存在的脂质分解代谢和合成代谢, 或受过氧化物酶体的影响。与尚未破坏其天然过氧化物酶体生物合成因子蛋白的含油酵母相比, 破坏或敲除导致总脂质级分和油级分中的 PUFA 量增加, 所述 PUFA 量以总脂肪酸百分比形式表示。在一些情况下, 也观察到在总脂质级分和油级分中以干细胞重量百分比形式表示的 PUFA 量增加, 和 / 或以干细胞重量百分比形式表示的总脂质含量增加。假定这种通用机制可应用于所有真核生物中, 如藻类、真菌、卵菌、酵母、类眼虫、原生藻菌、植物和一些哺乳动物体系, 因为所有这些生物都包含过氧化物酶体。

[0176] 鉴定并分离 Pex 同源物

[0177] 当在优选宿主生物中的特定 Pex 基因或蛋白的序列不是已知的时, 本领域技术人员认识到在调控基因编码的蛋白活性之前鉴定和分离这些基因或它们的部分基因将是最期望的, 所述调控继而促进掺入真核生物总脂质级分和油级分中的 PUFA 量的变化, 所述量以总脂肪酸百分比形式表示。对优选的宿主 Pex 基因序列的认识将有利于通过定向破坏来破坏同源染色体基因。

[0178] 表 4 中的 Pex 序列或它们的序列部分可用于在相同或其它藻类、真菌、卵菌、类眼

虫、原生藻菌、酵母或植物物种中使用序列分析软件搜索 Pex 同源物。通常,这种计算机软件通过将同源程度赋予多种置换、缺失和其它修饰来匹配相似的序列。使用软件算法,如具有低复杂度滤波器和以下参数的 BLASTP 比对方法:Expect value = 10, matrix = Blosum 62(Altschul, 等人, Nucleic Acids Res. 25 :3389-3402(1997)), 该方法是用于将表 4 中的任何 Pex 蛋白对核酸或蛋白序列数据库进行比较并从而鉴定在优选宿主生物中的相似已知序列的熟知方法。

[0179] 使用算法软件搜索已知序列数据库尤其适于分离对公开可获得的 Pex 序列具有相对低的同一性百分比的同源物,如那些表 4 中描述的序列。可预测的一点是:分离与公开可获得的 Pex 序列具有至少约 70% -85% 同一性的 Pex 同源物将是相对更容易的。此外,那些至少约 85% -90% 相同的序列将尤其适于分离,那些至少约 90% -95% 相同的序列将最容易分离。

[0180] 通过使用 Pex 酶独有的基序已经分离了一些 Pex 同源物。例如,为人熟知的是 Pex2p、Pex10p 和 Pex12p 在它们的羧基末端都有一个富含半胱氨酸的基序,已知该基序是 C_3HC_4 环状锌指基序(图 2A)。该“保守结构域”区域对应于一组在特定位点高度保守的氨基酸,并且可能提供对蛋白结构、稳定性或活性来说必需的 Pex 蛋白区域。基序通过它们在蛋白同源物家族的比对序列中的高度保守性进行鉴定。作为独有的“标记”,它们能决定具有新测定序列的蛋白是否属于以前鉴定过的蛋白家族。这些基序可用作诊断工具以分别迅速鉴定新的 Pex2、Pex10 和 / 或 Pex12 基因。

[0181] 作为另外一种选择,公开可获得的 Pex 序列或它们的基序可以是用于鉴定同源物的杂交试剂。核酸杂交试验的基本组成包括探针、怀疑含有目的基因或基因片段的样本及特定的杂交方法。探针通常是与待检测核酸序列互补的单链核酸序列。探针与待检测的核酸序列是可杂交的。尽管探针的长度可在 5 个碱基到数万个碱基之间变化,但通常约 15 个碱基到约 30 个碱基的探针长度是合适的。只需要探针分子的部分与待检测的核酸序列互补。另外,探针和靶序列之间不需要完全互补。杂交确实可以在并不完全互补的分子之间发生,结果是杂交区内的一定比率的碱基未与适当的互补碱基配对。

[0182] 杂交方法是已知的。通常探针和样品必须在允许核酸杂交的条件下混合。这涉及在适当浓度和温度条件下在存在无机或有机盐时使探针和样品接触。探针和样品核酸必须接触足够长的时间,使探针和样品核酸之间的任何可能的杂交均会发生。混合物中的探针或靶标的浓度决定杂交发生所需的时间。探针或靶标的浓度越高,所需的杂交孵育时间就越短。任选地,可以加入离液剂(如氯化胍、硫氰酸胍、硫氰酸钠、四氯乙酸锂、高氯酸钠、四氯乙酸铷、碘化钾或三氟乙酸铯)。如果需要,能将甲酰胺加入杂交混合物,通常为 30-50% (v/v) [“按体积计”]。

[0183] 可以采用多种杂交溶液。通常,这些杂交溶液包含约 20% 至 60% 体积,优选 30% 体积的极性有机溶剂。通常的杂交溶液采用约 30-50% v/v 甲酰胺、约 0.15 至 1M 氯化钠、约 0.05 至 0.1M 缓冲液(如柠檬酸钠、Tris-HCl、PIPES 或 HEPES (pH 范围约 6-9))、约 0.05 至 0.2% 的去污剂(如十二烷基硫酸钠)或 0.5-20mM 的 EDTA、FICOLL (Pharmacia Inc.) (约 300-500 千道尔顿)、聚乙烯吡咯烷酮(约 250-500 千道尔顿)和血清白蛋白。一般的杂交溶液还包含约 0.1 至 5mg/mL 未经标记的载体核酸、片段化的核酸 DNA (如小牛胸腺或蛙精 DNA 或酵母 RNA),以及任选约 0.5% 至 2% wt/vol [“重量体积比”] 的甘氨酸。可以

包含其它添加剂,例如包括极性水溶性或可膨胀试剂(如聚乙二醇)、阴离子聚合物(如聚丙烯酸酯或聚甲基丙烯酸酯)和阴离子糖类聚合物(如硫酸葡聚糖)在内的体积排阻剂。

[0184] 核酸杂交可适用于多种测定形式。最合适形式之一是夹心测定形式。夹心测定尤其适用于在非变性条件下杂交。夹心型测定的主要成分是固体支持体。固体支持体具有吸附或共价连接至其上的固定核酸探针,该探针未经标记并且与序列的一部分互补。

[0185] 任何 Pex 核酸片段或任何鉴定的同源物可用于从相同或其它藻类、真菌、卵菌、类眼虫、原生藻菌、酵母或植物物种中分离编码同源蛋白的基因。使用序列依赖性规程分离同源基因是本领域熟知的。序列依赖性规程的实例包括但不限于:1) 核酸杂交方法;2) DNA 和 RNA 扩增方法,例如核酸扩增技术的多种应用,如聚合酶链反应[“PCR”](美国专利 4,683,202);连接酶链反应[“LCR”](Tabor, S. 等人, Proc. Natl. Acad. Sci. U. S. A., 82:1074(1985));或链置换扩增[“SDA”], Walker 等人, Proc. Natl. Acad. Sci. U. S. A., 89:392(1992));和 3.) 文库构建和互补筛选方法。

[0186] 例如,编码蛋白或多肽的基因类似于公开可获得的 Pex 基因或它们的基序,可使用所有或部分那些公开可获得的核酸片段作为 DNA 杂交探针,以使用熟知的方法筛选来自任何期望生物的文库而直接分离该基因。基于公开可获得核酸序列的特异性,寡核苷酸探针可通过本领域已知的方法(Maniatis, 同上)设计并合成。而且,整个序列可直接用于通过熟练技术人员已知的方法,例如随机引物 DNA 标记、切口平移或末端标记技术,来合成 DNA 探针,或使用可获得的体外转录体系来合成 RNA 探针。此外,能设计特异性引物并用于扩增部分或全长公开可获得的序列或它们的基序。所得的扩增产物可在扩增反应过程中直接标记或在扩增反应后标记,并用作探针以在合适的严格条件下分离全长的 DNA 片段。

[0187] 通常,在 PCR 类型的扩增技术中,引物具有不同的序列而且彼此之间不互补。取决于期望的检测条件,应当设计引物序列以提供既有效又可靠的靶核酸的复制。PCR 引物设计方法是常见且熟知的(Thein 和 Wallace, “The use of oligonucleotides as specific hybridization probes in the Diagnosis of Genetic Disorders”, Human Genetic Diseases: A Practical Approach, K. E. Davis 编辑, (1986) 第 33-50 页, IRL: Herndon, VA; Rychlik, W., In Methods in Molecular Biology, White, B. A. Ed., (1993) 第 15 卷, 第 31-39 页, PCR Protocols: Current Methods and Applications. Humana: Totowa, NJ)。

[0188] 通常,可以将可获得的 Pex 序列的两个短片段在 PCR 规程中用于从 DNA 或 RNA 扩增编码同源基因的更长的核酸片段。也可以对克隆的核酸片段文库进行 PCR, 其中一个引物的序列来源于可获得的核酸片段或它们的基序。其它引物序列利用 mRNA 前体编码基因 3' 末端存在的多腺苷酸片。

[0189] 作为另一种选择,第二个引物序列可以基于来源于克隆载体的序列。例如,技术人员可以按照 RACE 规程(Frohman 等人, Proc. Natl. Acad. Sci. U. S. A., 85:8998(1988)), 通过用 PCR 扩增在转录物内单个位点与 3' 或 5' 端之间的区域的拷贝来产生 cDNA。以 3' 和 5' 方向取向的引物可以从能利用可获得的序列设计。使用可商业获得的 3' RACE 或 5' RACE 体系(例如 BRL, Gaithersburg, MD), 可以分离特异性的 3' 或 5' cDNA 片段(Ohara 等人, Proc. Natl. Acad. Sci. U. S. A., 86:5673(1989); Loh 等人, Science, 243:217(1989))。

[0190] 基于所讨论的这些熟知方法中的任何一种,在选择任何优选真核生物中鉴定和

/或分离 Pex 基因同源物将是可能的。能通过定向破坏 PUFA 生产宿主生物中的内源基因容易地确认任何推定的 Pex 基因的活性,因为总脂质级分和油级分的脂质分布相对于它们在缺乏定向 Pex 基因破坏的生物中的那些脂质分布发生了改变。

[0191] 经由破坏天然过氧化物酶体生物合成因子蛋白提高总脂质级分和油级分中的 PUFA 量

[0192] 如上所述,本发明的公开内容涉及下述提高含油真核生物中的一种 PUFA 或 PUFA 组合的重量百分比的方法,所述方法包括:

[0193] a) 提供含油真核生物,所述生物在编码过氧化物酶体生物合成因子蛋白的天然基因中包含破坏,这产生 PEX 破坏生物;和编码功能性 PUFA 生物合成途径的基因;以及

[0194] b) 在下述条件下培养 (a) 的真核生物:当与未破坏其天然过氧化物酶体生物合成因子蛋白的含油真核生物中的那些的重量%相比较时,(a) 真核生物的总脂质级分和油级分中的一种 PUFA 或 PUFA 组合的重量%相对于总脂肪酸的重量%增加。

[0195] 以总脂肪酸百分比形式提高的 PUFA 的量可能是:1) 作为功能性 PUFA 生物合成途径期望终产品形式的 PUFA,这与以中间体或副产品形式产生的 PUFA 相反;2) C_{20} 至 C_{22} 的 PUFA;和/或 3) 总 PUFA。

[0196] 除了相对于总脂肪酸重量%提高一种 PUFA 或 PUFA 组合的重量%之外,在一些情况下,还可以提高或降低细胞的总脂质含量 (TFA% DCW)。这意味着无论 PEX 基因破坏是否引起 PEX 破坏细胞中的总脂质量提高或降低,该破坏总是引起一种 PUFA 或 PUFA 组合的重量%提高。

[0197] 本文提供的另一种方法涉及在编码过氧化物酶体生物合成因子蛋白的天然基因中的破坏,其中当与亲本菌株中的该百分比进行比较时,所述破坏能够导致一种 PUFA 或 PUFA 组合相对于干细胞重量的百分比提高,所述亲本菌株的天然 Pex 蛋白未被破坏或其表达破坏天然 Pex 蛋白的“后备”拷贝。

[0198] 在上述方法的优选方面,在编码过氧化物酶体生物合成因子蛋白的天然基因中的破坏导致 PUFA 的量提高,PUFA 是功能性 PUFA 生物合成途径期望的终产品,而不是 PUFA 中间体或副产品,PUFA 量以干细胞重量百分比相对于亲本菌株的重量百分比表示,所述亲本菌株的天然 Pex 蛋白未被破坏或其表达破坏天然 Pex 蛋白的“后备”拷贝。在一些情况下,组合 PUFA 相对于干细胞重量的百分比的提高是 C_{20} 至 C_{22} PUFA 的组合或总 PUFA 的提高。

[0199] 上文也描述了通过这些方法产生的生物,所述生物包含至少一个过氧化物酶体生物合成因子蛋白的破坏。也描述了获取自这些生物的脂质和油、获取自脂质和油加工的产品、这些脂质和油用于食品、动物饲料、或工业应用的用途和/或副产品用于食品或动物饲料的用途。

[0200] 上述方法中优选的真核生物包括藻类、真菌、卵菌、酵母、类眼虫、原生藻菌、植物和一些哺乳动物系统。

[0201] 用于这些方法中任何一种的过氧化物酶体生物合成因子蛋白可选自:Pex1p、Pex2p、Pex3p、Pex3Bp、Pex4p、Pex5p、Pex5Bp、Pex5Cp、Pex5/20p、Pex6p、Pex7p、Pex8p、Pex10p、Pex12p、Pex13p、Pex14p、Pex15p、Pex16p、Pex17p、Pex14/17p、Pex18p、Pex19p、Pex20p、Pex21p、Pex21B、Pex22p、Pex22p 样和 Pex26p(以及它们的蛋白同源物)。在本文所述的一些优选方法中,被破坏的过氧化物酶体生物合成因子蛋白选自:Pex2p、Pex3p、

Pex10p、Pex12p 和 / 或 Pex16p。然而在一些更优选的方法中,被破坏的过氧化物酶体生物合成因子蛋白选自:Pex3p、Pex10p 和 / 或 Pex16p。

[0202] 在编码过氧化物酶体生物合成因子蛋白的天然基因中的破坏可能是在部分基因(例如在蛋白 N- 末端部分或在蛋白 C- 末端部分)中进行的插入、缺失、或定向突变。作为另外一种选择,该破坏能导致完全的基因敲除,使得基因从宿主细胞基因组中除去。或者该破坏可能是导致无功能蛋白的定向突变。

[0203] 破坏方法

[0204] 本发明包括在天然基因中的破坏,所述基因编码优选宿主细胞内的过氧化物酶体生物合成因子蛋白。虽然本领域的技术人员可使用多种技术获得破坏,但是特定基因的内源活性一般能通过以下技术减少或消除,例如:1) 通过插入、取代和 / 或缺失所有或部分靶基因破坏所述基因;或 2) 操纵控制所述蛋白表达的调控序列。这些技术在下文中讨论。然而,本领域的技术人员认识到这些技术在现有文献中已有详细描述,并且不受本文所述的方法、宿主细胞、和产品的限制。本领域的技术人员也认识到大多数适用的技术使用任何特定的含油酵母。

[0205] 经由插入、取代和 / 或缺失的破坏:就基因破坏而言,将外源 DNA 片段(通常是一个选择性标记基因)插入结构基因。这打断结构基因的编码序列并引起基因失活。将破坏盒转化到宿主细胞中导致非功能性破坏基因通过同源重组置换功能性天然基因。参见例如:Hamilton 等人, *J. Bacteriol.*, 171:4617-4622(1989); Balbas 等人, *Gene*, 136:211-213(1993); Guedener 等人, *Nucleic Acids Res.*, 24:2519-2524(1996); 和 Smith 等人, *Methods Mol. Cell. Biol.*, 5:270-277(1996)。本领域的技术人员了解基因定向常规方法的许多改良方法,所述方法容许有阳性选择和阴性选择、生成基因敲除、以及将外来 DNA 序列插入到哺乳动物系统、植物细胞、丝状真菌、藻类、卵菌、类眼虫、原生藻菌、酵母和 / 或微生物系统的特定基因组位点中。

[0206] 相反地,基因破坏的非特异性方法是使用转座元件或转座子。转座子是随机插入 DNA 但随后能根据序列进行检索以测定插入位点的基因元件。体内和体外转座技术是已知的,并且涉及转座元件与转座酶的组合使用。当转座元件或转座子与核酸片段在存在转座酶的情况下接触时,转座元件随机插入核酸片段中。该技术用于随机诱变和基因分离,因为被破坏的基因可基于转座元件的序列进行鉴定。用于体外转座的试剂盒是可商购获得的并包括:Primer Island Transposition Kit,得自 PerkinElmer Applied Biosystems, Branchburg, NJ, 基于酵母的 Ty1 元件;Genome Priming System,得自 New England Biolabs, Beverly, MA, 基于细菌转座子 Tn7;和 EZ::TN 转座子插入系统,得自 Epicentre Technologies, Madison, WI, 基于 Tn5 细菌转座元件。

[0207] Pex 调控序列的操纵:本领域熟知与编码序列附连的调控序列包括转录和翻译“控制”核苷酸序列,所述序列位于编码序列的上游(5' 非编码序列)、内部、或下游(3' 非编码序列),并且影响附连的编码序列的转录、RNA 加工或稳定性、或翻译。因此,操纵 Pex 基因调控序列可以指操纵特定 Pex 基因的启动子、静默子、5' 非转录前导序列(介于转录起始位点好翻译启动密码子之间)、内含子、增强子、启动控制区、多腺苷酸化识别序列、RNA 加工位点、效应子结合位点和茎-环结构。然而在所有情况下,操纵的结果是下调 Pex 基因的表达,这促使与天然过氧化物酶体生物合成因子蛋白未被破坏的含油酵母相比,总脂质

级分和油级分中以总脂肪酸百分比形式表示的 PUFA 量增加。

[0208] 例如,能缺失或破坏 Pex10 基因的启动子。作为另外一种选择,驱动 Pex10 基因表达的天然启动子可以用与天然启动子相比启动子活性减弱的异源启动子取代。用于操纵调控序列的方法是为人们熟知的。

[0209] 技术人员能够使用这些技术和其它熟知技术破坏在本文所述的优选宿主细胞中的天然过氧化物酶体生物合成因子蛋白,所述优选宿主例如哺乳动物系统、植物细胞、丝状真菌、藻类、卵菌、类眼虫、原生藻菌和酵母。

[0210] 本领域的技术人员能够识别破坏天然 Pex 基因的最佳方法,以使得与天然过氧化物酶体生物合成因子蛋白未被破坏的真核生物相比,在总脂质级分和油级分中积聚的 PUFA 量增加,所述 PUFA 量以总脂肪酸的百分比形式表示。

[0211] ω -3 和 / 或 ω -6 脂肪酸生物合成的代谢工程

[0212] 本文所述方法除了用于破坏天然过氧化物酶体生物合成因子蛋白之外,还可使用操纵 ω -3 和 / 或 ω -6 脂肪酸生物合成。这种操纵可能需要直接在 PUFA 生物合成途径中进行代谢工程改造或需要另外对为 PUFA 生物合成途径贡献碳的途径进行操纵。可用于上调期望的生化途径和下调不期望的生化途径的技术是本领域熟知的。例如,与 ω -3 和 / 或 ω -6 脂肪酸生物合成途径竞争能量或碳的生化途径或干扰特定 PUFA 终产物产生的天然 PUFA 生物合成途径中的酶可以通过基因破坏来去除或通过其它手段(如反义 mRNA 和锌指靶向技术(zinc-finger targeting technologies))来下调。

[0213] 以下讨论改变 PUFA 生物合成途径从而分别提高 GLA、ARA、EPA 或 DHA 含量的方法,以及在 TAG 生物合成途径和 TAG 降解途径中的所期望的操纵:分别是国际申请公布 WO 2006/033723,国际申请公布 WO 2006/055322[美国专利申请公布 2006-0094092-A1],国际申请公布 WO 2006/052870[美国专利申请公布 2006-0115881-A1]和国际申请公布 WO 2006/052871[美国专利申请公布 2006-0110806-A1]。

[0214] 表达系统、表达盒、载体和宿主细胞转化

[0215] 制备重组构建体并将其导入到优选真核宿主中以破坏天然过氧化物酶体生物合成因子蛋白和 / 或导入编码 PUFA 生物合成途径的基因可能是必需的,所述真核宿主例如哺乳动物系统、植物细胞、丝状真菌、藻类、卵菌、类眼虫、原生藻菌和酵母。本领域的技术人员了解的标准来源材料如下所述:1) 构建、操纵和分离大分子(例如 DNA 分子、质粒等)的特定条件和程序;2) 生成重组 DNA 片段和重组表达构建体;和 3) 克隆的筛选和分离。参见 Sambrook 等人,Molecular Cloning: A Laboratory Manual,第 2 版,Cold Spring Harbor Laboratory: Cold Spring Harbor, NY(1989); Maliga 等人,Methods in Plant Molecular Biology, Cold Spring Harbor, NY(1995); Birren 等人,Genome Analysis: Detecting Genes, v. 1, Cold Spring Harbor, NY(1998); Birren 等人,Genome Analysis: Analyzing DNA, v. 2, Cold Spring Harbor: NY(1998); Plant Molecular Biology: A Laboratory Manual, Clark, ed. Springer: NY(1997)。

[0216] 一般来讲,存在于构建体中的序列的具体选择取决于期望的表达产物、宿主细胞的性质以及提出的相对于未转化细胞分离转化细胞的手段。技术人员熟知基因元件必须存在于质粒载体上以成功地转化、选择并增殖包含嵌合基因的宿主细胞。然而,通常载体或盒含有引导相关基因的转录和翻译的序列、选择标记和允许自主复制或染色体整合的序列。

合适的载体包括控制转录起始的基因 5' 区（即启动子）和控制转录终止的 DNA 片段 3' 区（即终止子）。最优选两个控制区都来源于来自转化宿主细胞的基因。

[0217] 用于驱动在期望的宿主细胞中的异源基因或部分异源基因表达的启动控制区或启动子有多种,并且为人们所熟知。这些控制区可包含启动子、增强子、静默子、内含子序列、3' UTR 和 / 或 5' UTR 区域、以及蛋白和 / 或 RNA 稳定元件。此类元件的长度和特异性可以不同。实际上能够引导这些基因在所选宿主细胞中的表达的任何启动子（即,天然的、合成的、或嵌合的启动子）都是适用的。在宿主细胞中的表达能够以诱导型或组成型的方式发生。诱导型表达通过诱导可操作地连接至受关注的 Pex 基因的可调控启动子的活性发生。组成型表达通过利用可操作地连接至受关注基因的组成型启动子发生。

[0218] 当宿主细胞是例如酵母时,酵母细胞中的功能性转录和翻译区域尤其从宿主菌种中提供。用于在解脂耶氏酵母中使用的优选转录起始调节区参见国际申请公布 WO 2006/052870。可以使用许多调控序列的任意一种,这取决于期望组成型转录还是诱导型转录、启动子在表达所关注的 ORF 中的效率、构建的容易性等。

[0219] 必须在重组构建体中提供编码转录终止信号的 3' 非编码序列,即,“终止区”,该序列可以来自从中获取初始区的基因或不同基因的 3' 区域。大量的终止区是已知的并且当用于与它们源自的属和物种相同和不同的属和物种两者时,其在多种宿主中发挥的功能令人满意。选择终止区更多的是从方便的角度出发而不是为了任何特性。终止区也可以源于优选宿主的多种天然基因。

[0220] 尤其有用的酵母终止区是那些来源于酵母基因,尤其是糖酵母属、裂殖糖酵母属、假丝酵母属、耶氏酵母属或克鲁维酵母属的终止区。编码 γ -干扰素和 α -2 干扰素的哺乳动物基因的 3' 区域也已知能在酵母中起作用。3' - 区也可以是合成的,因为本领域的技术人员可利用可获得的信息来设计和合成用作转录终止子的 3' - 区序列。终止区可以是非必需的,但是是高度优选的。

[0221] 载体除了上述调控元件之外还可包含选择性的和 / 或可评分的标记。优选地,标记基因是抗生素抗性基因,使得用抗生素处理细胞引起未转化细胞的生长抑制或死亡,但不抑制转化细胞的生长。为了选择酵母转化体,可使用在酵母中发挥功能的任何标记,使之具有对卡那霉素、潮霉素、和氨基配糖物 G418 的抗性并能够在缺乏尿嘧啶、赖氨酸、组氨酸、或亮氨酸的培养基上生长的标记是尤其有用的。

[0222] 仅仅将基因插入克隆载体不确保它能以期望的速率、浓度、量等表达。根据对高表达率的需要,通过操作许多控制转录、RNA 稳定性、翻译、蛋白质稳定性和位置、氧限制和从宿主细胞分泌的不同遗传元件,已经建立了许多专用的表达载体。一些操纵特征包括:相关转录启动子和终止子序列的性质,所克隆基因的拷贝数目以及该基因是质粒携带的还是整合进了宿主细胞的基因组中,所合成的外来蛋白质的最终细胞定位,蛋白质在宿主生物内的翻译和正确折叠的效率,所克隆基因的 mRNA 和蛋白质在宿主细胞内的固有稳定性和所克隆基因中的密码子使用,以使得其频率与宿主细胞的优选密码子使用频率接近。这些特征的每一种可用于本文所述的方法和宿主细胞中,以进一步优化 PUFA 生物合成途径基因的表达并减少天然 Pex 基因的表达。

[0223] 在制备适于破坏或敲除天然过氧化物酶体生物合成因子蛋白和 / 或表达编码 PUFA 生物合成途径活性的基因的重组构建体（例如包含具有启动子、ORF 和终止子的嵌合

基因的构建体)后,将其置于能够在宿主细胞中自主复制的质粒载体中,或直接整合进宿主细胞基因组中。表达盒的整合可以在宿主基因组中随机地发生或者可以通过使用下述构建体来靶向,所述构建体含有足以靶向宿主基因座中的重组的与宿主基因组同源的区。如果构建体靶向内源性基因座,则转录和翻译调控区的全部或某些可以由内源性基因座提供。

[0224] 当两个或更多个基因从独立的复制载体表达时,每个载体可具有不同的选择手段并且应当缺乏与其它构建体的同源性以便维持稳定表达和防止元件在构建体之间重配(reassortment)。可用实验方法确定对调控区、选择手段和所引入的构建体的增殖方法的正确选择,以便所有导入的基因均以需要的水平表达从而提供期望的产品的合成。

[0225] 包含目的基因的构建体可以通过任何标准技术导入到宿主细胞中。这些技术包括转化(如醋酸锂转化 [Methods in Enzymology, 194:186-187(1991)])、原生质体融合、基因枪轰击(bolistic impact)、电穿孔、显微注射、真空过滤或将所关注的基因引入宿主细胞中的任何其它方法。

[0226] 为方便起见,已经通过任何方法被操纵以摄取DNA序列(如表达盒)的宿主细胞在本文中称为“转化的”或“重组的”。转化的宿主将具有至少一个表达构建体的拷贝,而且可以具有两个或更多个拷贝,这取决于该基因是整合进基因组内、被扩增还是存在于具有多拷贝数的染色体外元件上。

[0227] 转化宿主细胞能通过选择导入构建体上包含的标记进行鉴定。作为另外一种选择,分离的标记构建体可用期望的构建体进行共转化,该方法与多种将多个DNA分子导入宿主细胞的转化技术一样。通常转化宿主通过它们在选择培养基上生长的能力进行选择。选择培养基可掺入抗生素或缺乏未转化宿主必需的生长因子,如营养物质或生长因子。导入的标记基因可赋予抗生素抗性,或编码必需的生长因子或酶,从而当在转化宿主中表达时允许在选择培养基上生长。转化宿主的选择也能发生在表达标记蛋白能被直接地或间接地检测出来时。标记蛋白可单独表达或融合到另一种蛋白中进行表达。标记蛋白可通过其酶活性(例如 β -半乳糖苷酶可将X-gal[“5-溴-4-氯-3-吲哚基- β -D-半乳糖苷”]底物转化成着色产物;荧光素酶能将荧光素转化成发光产物)或其产光或改性特性(例如维多利亚多管水母(Aequorea victoria)的绿荧光蛋白,它当用蓝光照亮时发荧光)进行检测。作为另外一种选择,能使用抗体检测标记蛋白或在例如所关注蛋白上的分子标记。能对表达标记蛋白或标记物的细胞进行选择,例如通过视觉检测或通过例如使用抗体分选或批选荧光活化细胞的技术。

[0228] 无论选择哪种宿主或表达构建体,必须筛选多个转化体以获取显示期望的表达水平、调节和模式的菌株或植物品系,因为不同的独立转化事件导致不同水平和模式的表达(Jones等人,EMBO J., 4:2411-2418(1985);De Almeida等人,Mol.Gen.Genetics, 218:78-86(1989))。这种筛选可以通过DNA印迹的Southern分析(Southern, J.Mol.Biol., 98:503(1975))、mRNA表达的Northern分析(Kroczeck, J.Chromatogr.Biomed.Appl., 618(1-2):133-145(1993))、蛋白质表达的Western和/或Elisa分析、PUFA产物的表型分析或GC分析来完成。

[0229] 优选的真核宿主生物

[0230] 多种真核生物适合作为本文的宿主,因此产生在天然过氧化物酶体生物合成因子

蛋白和编码 PUFA 生物合成途径的基因中的包含破坏的转化宿主生物,与尚未破坏其天然过氧化物酶体生物合成因子蛋白的真核生物相比,其中转化的真核宿主生物的 PUFA 含量提高,所述 PUFA 被掺入到总脂质级分和油级分中,以总脂肪酸百分比形式表示。可获得的宿主可以是不同的哺乳动物体系、植物细胞、真菌、藻类、卵菌、酵母、原生藻菌和 / 或类眼虫。虽然优选含油生物,但本文也可使用非含油生物,使得当它们的一种天然 PEX 基因被破坏时,在总脂质级分或油级分中至少一种多不饱和脂肪酸的重量%相对于总脂肪酸的重量%提高,并且可导致 PUFA 提高 1.3 倍。此外,PUFA 百分比可相对于非含油生物的干细胞重量提高。在另一个实施方案中,能基因修饰非含油生物使之变成含油生物,例如酵母如啤酒糖酵母。

[0231] 含油生物天然能够合成并积聚油,其中总油含量通常大于细胞干重的约 25%。将多种藻类、藓、真菌、酵母、原生藻菌和植物被天然地归类为含油生物。

[0232] 优选的含油微生物包括那些藻类、原生藻菌和真菌生物,它们天然生产 ω -3/ ω -6PUFA。例如,ARA、EPA 和 / 或 DHA 经由小环藻属 (*Cyclotella* sp.)、菱形藻属 (*Nitzschia* sp.)、腐霉属 (*Pythium*)、破囊壶菌属 (*Thraustochytrium* sp.)、裂殖壶菌属 (*Schizochytrium* sp.) 和被孢霉属 (*Mortierella*) 生产。Mackenzie 等人 (*Appl. Environ. Microbiol.*, 66 :4655 (2000)) 描述了转化高山被孢霉 (*M. alpina*) 的方法。类似地,用于转化破囊壶菌目 (*Thraustochytriales*) 微生物 (例如破囊壶菌属 (*Thraustochytrium*)、裂殖壶菌属 (*Schizochytrium*)) 的方法在美国专利 7,001,772 中公开。

[0233] 更优选含油酵母,包括那些天然生产和那些经遗传工程化以生产 ω -3/ ω -6PUFA 的酵母。通常鉴定为含油酵母的属包括但不限于:耶氏酵母属 (*Yarrowia*)、假丝酵母属 (*Candida*)、红酵母属 (*Rhodotorula*)、红冬孢酵母属 (*Rhodospiridium*)、隐球酵母属 (*Cryptococcus*)、丝孢酵母属 (*Trichosporon*) 和油脂酵母属 (*Lipomyces*)。更具体地讲,示例性的油合成酵母包括:圆红冬孢酵母 (*Rhodospiridium toruloides*)、斯达氏油脂酵母 (*Lipomyces starkeyii*)、产油油脂酵母 (*L. lipoferus*)、拉可夫氏假丝酵母 (*Candida revkaufi*)、铁红假丝酵母 (*C. pulcherrima*)、热带假丝酵母 (*C. tropicalis*)、产朊假丝酵母 (*C. utilis*)、茁芽丝孢酵母 (*Trichosporon pullans*)、皮状丝孢酵母 (*T. cutaneum*)、胶粘红酵母 (*Rhodotorula glutinus*)、禾本科红酵母 (*R. graminis*) 和解脂耶氏酵母 (*Yarrowia lipolytica*) (以前归类为解脂假丝酵母 (*Candida lipolytica*))。

[0234] 最优选的是含油酵母解脂耶氏酵母;而且,在其它实施方案中,最优选的是命名为 ATCC#76982、ATCC#20362、ATCC#8862、ATCC#18944 和 / 或 LGAM S(7)1 的解脂耶氏酵母菌株 (Papanikolaou S. 和 Aggelis G., *Bioresour. Technol.*, 82(1) :43-9 (2002))。

[0235] 涉及转化解脂耶氏酵母的具体教导内容包括美国专利 4,880,741 和美国专利 5,071,764, 以及 Chen, D.C. 等人 (*Appl. Microbiol. Biotechnol.*, 48(2) : 232-235 (1997)), 而合适的选择技术描述于美国专利 7,238,482 和国际申请公布 WO 2005/003310 和 WO 2006/052870。

[0236] 在解脂耶氏酵母中表达基因的优选方法是通过将线性 DNA 整合进宿主基因组中。当期望基因高水平表达时,整合进基因组中的多个位置可能是尤其有用的,例如在 Ura3 基因座 (GenBank 保藏号 AJ306421)、Leu2 基因座 (GenBank 保藏号 AF260230)、Lys5 基因座 (GenBank 保藏号 M34929)、Aco2 基因座 (GenBank 保藏号 AJ001300)、Pox3 基因座 (Pox3 :

GenBank 保藏号 XP_503244 或 Aco3 :GenBank 保藏号 AJ001301)、 Δ 12 去饱和酶基因座 (美国专利 7, 214, 491)、Lip1 基因座 (GenBank 保藏号 Z50020)、Lip2 基因座 (GenBank 保藏号 AJ012632)、SCP2 基因座 (GenBank 保藏号 AJ431362)、Pex3 基因座 (GenBank 保藏号 CAG78565)、Pex16 基因座 (GenBank 保藏号 CAG79622) 和 / 或 Pex10 基因座 (GenBank 保藏号 CAG81606)。

[0237] 优选的用于解脂耶氏酵母的选择方法是对卡那霉素、潮霉素和氨基糖苷 G418 的抗性以及在缺乏尿嘧啶、亮氨酸、赖氨酸、色氨酸或组氨酸的培养基上生长的能力。5- 氟乳清酸 [5- 氟尿嘧啶 -6- 羧酸一水合物或“5-FOA”] 也可用于选择酵母 Ura- 突变体。该化合物对具有编码乳清酸苷 5' - 单磷酸脱羧酶 [OMP 脱羧酶] 的功能性 URA3 基因的酵母细胞有毒性;因此,基于这种毒性,5-FOA 特别可用于选择和鉴定 Ura- 突变型酵母菌株 (Bartel, P.L. 和 Fields, S., Yeast 2-Hybrid System, Oxford University :New York, 第 7 卷, 第 109-147 页, 1997; 也参见国际申请公开 W02006/052870, 用于 5-FOA 在耶氏酵母属中的应用)。

[0238] 用于耶氏酵母属的备选的优选选择方法依赖于用于解脂耶氏酵母的基于磺酰脲 (氯嘧磺隆 ;E. I. duPont de Nemours & Co., Inc., Wilmington, DE) 抗性的显性非抗生素标记。更具体地讲,该标记基因是具有单个氨基酸改变 (即 W497L) 的天然乙酰羟酸合酶 (“AHAS”或乙酰乳酸合酶 ;E. C. 4. 1. 3. 18), 从而赋予了磺酰脲除草剂抗性 (国际申请公布 WO 2006/052870)。AHAS 是用于生物合成支链氨基酸 (即缬氨酸、亮氨酸、异亮氨酸) 的途径中的第一常用酶,并且它是磺酰脲和咪唑啉酮除草剂的靶标。

[0239] 用于多不饱和脂肪酸生产的发酵方法

[0240] 使转化的宿主细胞在优化 PUFA 生物合成基因的表达并且产生最大并且最经济产量的期望 PUFA 的条件下生长。通常,可通过改变碳源的类型和量、氮源的类型和量、碳氮比、不同矿物离子的量、氧水平、生长温度、pH、生物质生产期的时长、油积聚期的时长以及细胞收获的时间和方法来优化培养基条件。通常使所关注的含油酵母例如解脂耶氏酵母在复合培养基 (例如酵母提取物 - 蛋白胨 - 葡萄糖肉汤培养基 (YPD)) 或缺乏生长所必需的成分并强迫选择期望表达盒的确定成分的极限培养基 (如 Yeast Nitrogen Base (DIFCO Laboratories, Detroit, MI)) 中生长。

[0241] 用于本文所述方法和宿主细胞的发酵培养基必须包含合适的碳源如在美国专利 7, 238, 482 中提出的碳源。合适的碳源涵盖多种来源,优选糖、甘油和 / 或脂肪酸。最优选的碳源是葡萄糖和 / 或包含介于 10-22 个碳之间的脂肪酸。

[0242] 氮可以由无机源 (如 $(\text{NH}_4)_2\text{SO}_4$) 或有机源 (如尿素或谷氨酸) 提供。除了合适的碳源和氮源之外,发酵培养基还必须含有合适的矿物质、盐、辅因子、缓冲液、维生素和本领域技术人员已知的适合含油酵母生长并促进 PUFA 产生的酶促途径的其它组分。要特别注意促进脂质和 PUFA 合成的几种金属离子 (如 Fe^{+2} 、 Cu^{+2} 、 Mn^{+2} 、 Co^{+2} 、 Zn^{+2} 和 Mg^{+2}) (Nakahara, T. 等人, Ind. Appl. Single Cell Oils, D. J. Kyle 和 R. Colin 编辑, 第 61-97 页 (1992))。

[0243] 用于本文所述方法和宿主细胞的优选生长培养基是常规的商业制备培养基,如酵母氮源培养基 (Yeast Nitrogen Base, DIFCO Laboratories, Detroit, MI)。也可以使用其它确定成分的生长培养基或合成的生长培养基,适于转化宿主细胞生长的培养基对微生物学或发酵科学来说是已知的。适于发酵的 pH 范围通常介于约 pH 4.0 至 pH 8.0 之间,其中

优选将 pH 5.5 至 pH 7.5 作为初始生长条件的范围。发酵可以在有氧或厌氧条件下进行，其中优选为微好氧条件。

[0244] 通常，在含油酵母细胞中积聚高水平的 PUFA 和 TAG 需要一个两阶段过程，因为代谢状态必须在生长和脂肪的合成 / 贮存之间达到“平衡”。因此，最优选的是，两阶段的发酵过程是在含油酵母中产生油所必需的。这种方法在美国专利 7, 238, 482 中有所描述，其也描述了多种合适的发酵工艺设计（即分批式、补料分批式和连续式）和生长过程中的注意事项。

[0245] PUFA 油的纯化和处理

[0246] 包括 PUFA 在内的脂肪酸可以游离脂肪酸或酯化形式如酰基甘油、磷脂、硫脂或糖脂形式存在于宿主生物中。这些脂肪酸可通过本领域熟知的多种方法从宿主细胞中提取。关于酵母脂质的提取技术、质量分析和可接受标准的一篇综述是 Z. Jacobs (Critical Reviews in Biotechnology, 12(5/6) :463-491 (1992)) 的综述。有关下游加工的简述也可由 A. Singh 和 O. Ward (Adv. Appl. Microbiol., 45 :271-312 (1997)) 提供。

[0247] 一般来讲，用于纯化脂肪酸（包括 PUFA）的手段可包括用有机溶剂萃取（例如，美国专利 6, 797, 303 和美国专利 5, 648, 564）、超声处理、超临界流体萃取（如使用二氧化碳）、皂化和物理手段（例如挤压）、或它们的组合。参见美国专利 7, 238, 482。

[0248] 油在食物、保健食品、药物和动物饲料中的用途

[0249] 市场包含许多种掺入了 ω -3 和 / 或 ω -6 脂肪酸（尤其是 ALA、GLA、ARA、EPA、DPA 和 DHA）的食物和饲料产品。可以预期，通过本文所述方法和宿主细胞制备的包含长链 PUFA 的微生物物质、包含 PUFA 的部分纯化微生物物质、包含 PUFA 的纯化微生物油、和 / 或纯化的 PUFA 赋予健康有益效果，通过加入这些物质改善了对食物或饲料的摄取。举例来说，可将这些油加到食物类似物、饮料、肉产品、谷类产品、烘焙食品、小吃食品和乳制品中。参见美国专利公开 2006/0094092。

[0250] 这些组合物通过加入到医疗食品（包括医疗营养物质、饮食补充剂、婴儿代乳品和药品）中可赋予健康有益效果。技术人员将会知道加到食品、饲料、饮食补充剂、营养物质、药品、和其它可摄取的产品中以赋予健康有益效果的油的量。本领域描述了摄取这些油获得的健康有益效果，它是技术人员已知的并进行了持续的研究。本文所指的量是“有效”量，除了其它方面的原因，它取决于摄入的包含这些油的产品的性质以及它们旨在治疗的病症。

[0251] 优选实施方案的描述

[0252] 如实施例所证明并如表 5 所概述的（下文），破坏 Y1Pex10p 的 C_3HC_4 环状锌指基序的 C- 末端部分、缺失整个染色体 Y1Pex10 基因或整个染色体 Y1Pex16 基因、缺失整个染色体 Y1Pex10 和 Y1Pex16 基因、和缺失整个染色体 Y1Pex3 基因都导致产生解脂耶氏酵母的工程化 PUFA 生产菌株，相对于不破坏天然 Pex 蛋白的亲本菌株，该菌株具有以总脂肪酸百分比形式表示的、提高的 PUFA 重量%。在解脂耶氏酵母的工程化 EPA 生产菌株中表达染色体外的 Y1Pex10p 逆转了该效应，所述菌株在基因组 Pex10p 中具有破坏和提高的总脂质级分和油级分中的 PUFA 的量。

[0253] 表 5 汇编了来自实施例 3、4、5、7、9、11 和 12 的数据，使得关于总脂质含量 [“TFA% DCW”]、以总脂肪酸重量%表示的给定脂肪酸浓度 [“TFA%”]、和干细胞重量百分比形式

表示的给定脂肪酸含量 [“DCW%”] 的趋势能基于存在 / 不存在 Pex 破坏或敲除被推论出来。“期望 PUFA% TFA”和“期望 PUFA% DCW”分别定量期望的 PUFA 产物 (即 DGLA 或 EPA) 的特定浓度或含量,所述产物通过设计的工程化 PUFA 生物合成途径进行生产。“全部 PUFA”包括 LA、ALA、EDA、DGLA、ETrA、ETA 和 EPA、然而“C20PUFA”被限定为 EDA、DGLA、ETrA、ETA 和 EPA。

[0254]

表 5

在具有突变 Pex 基因的解脂耶氏酵母菌株中的 PUFA%TFA 和 DCW%

实施例	菌株	基因组 Pex 基因	TFA %DCW	%TFA			%DCW		
				期望的 PUFA	全部 PUFA	C20 PUFA	期望的 PUFA	全部 PUFA	C20 PUFA
3, 4	Y4086	野生型 Pex10	28.6	9.8 [EPA]	60.1	25.2	2.8 [EPA]	17.2	7.2
	Y4128	突变型 * Pex10	11.2	42.8 [EPA]	79.3	57.9	4.8 [EPA]	8.9	6.4
5	Y4128U1 + pFBAIn-PEX10	突变型 * Pex10 + 质粒野生型 Pex10 在嵌合 FBAInm::Pex10::Pex20 基因内	29.2	10.8 [EPA]	60	27.3	3.1 [EPA]	17.5	8.0
	Y4128U1 + pPEX10-1	突变型 * Pex10 + 质粒野生型 Pex10, 在 Pex10-5' (500bp) :: Pex10::Pex10-3' 基因内	27.1	10.7 [EPA]	60.1	26.7	2.9 [EPA]	16.2	7.2
	Y4128U1 + pPEX10-2	突变型 * Pex10 + 质粒野生型 Pex10, 在 Pex10-5' 内 (991bp) :: Pex10::Pex10-3' 基因	28.5	10.8 [EPA]	59	26.9	3.1 [EPA]	16.8	7.7
	Y4128U1 + 对照物	突变型 * Pex10	22.8	27.7 [EPA]	62.6	42.3	6.3 [EPA]	14.2	9.6

[0255]

7	Y4184U	野生型 Pex10	11.8	20.6 [EPA]	nq ⁺	nq ⁺	2.4 [EPA]	nq ⁺	nq ⁺
			8.8	23.2 [EPA]	nq ⁺	nq ⁺	2.0 [EPA]	nq ⁺	nq ⁺
	Y4184U ΔPex10	突变型 Pex10	17.6	43.2 [EPA]	nq ⁺	nq ⁺	7.6 [EPA]	nq ⁺	nq ⁺
			13.2	46.1 [EPA]	nq ⁺	nq ⁺	6.1 [EPA]	nq ⁺	nq ⁺
9	Y4036 (avg)	野生型 Pex16	Nq ⁺	23.4 [DGLA]	61.5	33.7	nq ⁺	nq ⁺	nq ⁺
	Y4036 (ΔPex16) (avg)	突变型 Pex16	Nq ⁺	43.4 [DGLA]	69.1	49.1	nq ⁺	nq ⁺	nq ⁺
11	Y4305U (Δpex10) (avg)	突变型 Pex10 和野生型 Pex16	30	44.7 [EPA]	76.6	55.4	13.4 [EPA]	23.0	16.6
	Y4305 (ΔPex10, ΔPex16) (avg)	突变型 Pex10, 突变型 Pex16	30	48.3 [EPA]	79.0	57.7	14.5 [EPA]	23.7	17.3
12	Y4036	野生型 Pex3	4.7	19 [DGLA]	57	27	0.9 [DGLA]	2.7	1.3

[0256]

	Y4036 (Δ Pex3)	突变型 Pex3	6.1	46 [DGLA]	68	56	2.8 [DGLA]	4.4	3.4
			5.9	46 [DGLA]	68	56	2.7 [DGLA]	4.0	3.3

* 在 Y4128 中的 Pex10 破坏导致截短的蛋白，其中 C-末端的最后 32 个氨基酸(对应于 C₃HC₄ 环状锌指基序的 C 末端部分)不存在。

◆ nq = 未定量

[0257] 虽然数据不能直接地在实施例之间进行比较,但作为不同的耶氏酵母属菌株和生长条件的结果,可得出以下结论(相对于其天然 Pex 蛋白未被破坏的亲本菌株或其表达破

坏天然 Pex 蛋白的“后备”拷贝的亲本菌株)：

[0258] 1) 在生产 PUFA 的耶氏酵母属中的 Pex 破坏导致总脂质级分和油级分中的单个 PUFA 例如 EPA 或 DLGA 的重量%相对于总脂肪酸的重量% (% TFA) 提高；

[0259] 2) 在生产 PUFA 的耶氏酵母属中的 Pex 破坏导致总脂质级分和油级分中的 C20PUFA 的重量%相对于总脂肪酸的重量% (% TFA) 提高；

[0260] 3) 通过延伸点 1)，在生产 PUFA 的耶氏酵母属中的 Pex 破坏导致总脂质级分和油级分中的任何 PUFA 和所有 PUFA 组合的量相对于总脂肪酸的重量%提高；和

[0261] 4) 在生产 PUFA 的耶氏酵母属中的 Pex 破坏导致单个 PUFA 例如 EPA 或 DLGA 相对于干细胞重量的百分比提高。

[0262] 当比较在“所有 PUFA% DCW”、“C20PUFA% DCW”和 TFA% DCW 中的 Pex 破坏的效应时，观察到不同的结果。具体地讲，在一些情况下，在生产 PUFA 的耶氏酵母属中的 Pex 破坏导致总脂质级分和油级分中的 C₂₀PUFA 或所有 PUFA 的量（以干细胞重量百分比形式表示）提高（相对于天然 Pex 蛋白未被破坏的亲本菌株）。在其它情况下，在总脂质级分和油级分中存在以干细胞重量百分比形式表示的减少量的 C₂₀PUFA 或所有 PUFA（相对于天然 Pex 蛋白未被破坏的亲本菌株）。也观察到相对于总脂质含量（TFA% DCW）的类似结果，因此 Pex 破坏的效应可能导致总脂质含量提高或降低。

[0263] 虽然上文概述的每项内容都是受关注的，尤其有用的是检查 Pex 破坏对期望的 PUFA 相对于总 PUFA 量的比率的效应，对生物进行工程化以生产期望的 PUFA。

[0264] 例如，在包含 Pex10 破坏的菌株 Y4128 中 54% 的 PUFA (% TFA) 是 EPA，所述破坏导致截去 C 末端最后 32 个氨基酸，然而在亲本菌株 Y4086 中仅有 16.3% 的 PUFA (% TFA) 是 EPA。因此，所述破坏导致期望的 PUFA (% TFA) 量提高 3.3 倍（实施例 3、4）。在类似情况下，在菌株 Y4036 (Δ Pex16) 中 62.8% 的 PUFA (% TFA) 是 DGLA，然而在 Y4036- 中仅有 38.1% 的 PUFA (% TFA) 是 DGLA，提高了 1.65 倍（实施例 9）。并且在菌株 Y4036 (Δ Pex3) 中 67.7% 的 PUFA (% TFA) 是 DGLA，然而在 Y4036- 中仅有 33.3% 的 PUFA (% TFA) 是 DGLA，提高了 2.0 倍（实施例 12）。这些结果证实这一假说：Pex 破坏导致总脂质级分和油级分中的期望的 PUFA 的量 (% TFA) 选择性地提高，所述 PUFA 通过工程化生物进行生产。

[0265] 当检查 Pex 破坏对 C20PUFA 相对于总 PUFA 量的比率的效应时，观察到较低的选择性。例如，在包含 Pex10 破坏的菌株 Y4128 中，73% 的 PUFA (% TFA) 是 C20PUFA，然而在菌株 Y4086 中，仅有 42% 的 PUFA (% TFA) 是 C20PUFA。因此，破坏导致总脂质级分和油级分中积聚的 C20PUFA 相对于总 PUFA 的量提高 1.7 倍（实施例 3、4）。在类似情况下，在菌株 Y4036 (Δ Pex16) 中 71% 的 PUFA (% TFA) 是 C20PUFA，然而在 Y4036- 中仅有 54.8% 的 PUFA (% TFA) 是 C20PUFA，提高了 1.3 倍（实施例 9）。并且在菌株 Y4036 (Δ Pex3) 中 82.4% 的 PUFA (% TFA) 是 C20PUFA，然而在 Y4036- 中仅有 47.4% 的 PUFA (% TFA) 是 C20PUFA，提高了 1.7 倍（实施例 12）。

[0266] 基于教导和本文所述的结果，期望将实现以下方法的可行性和商业可用性：利用在编码过氧化物酶体生物合成因子蛋白的天然基因中的多个破坏作为提高在生产 PUFA 的真核生物中的 PUFA 量的方法。生产 PUFA 的真核生物能够合成多种 ω -3 和 / 或 ω -6PUFA，该合成利用 Δ 9 延伸酶 / Δ 8 去饱和酶途径或 Δ 6 去饱和酶 / Δ 6 延伸酶途径。

实施例

[0267] 在以下实施例中进一步描述本发明,所述实施例示出了实施本发明的简要方法,但不完全限定其所有可能变型。

[0268] 一般方法

[0269] 实施例中使用的标准重组 DNA 技术和分子克隆技术是本领域所熟知的并且在如下文献中有所描述:1) Sambrook, J., Fritsch, E.F. 和 Maniatis, T. 的 Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989) (Maniatis); 2) T. J. Silhavy, M. L. Bannan 和 L. W. Enquist, Experiments with Gene Fusions; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1984); 以及 3.) Ausubel, F.M. 等人, Current Protocols in Molecular Biology, 由 Greene Publishing Assoc. 和 Wiley-Interscience 出版, Hoboken, NJ (1987)。

[0270] 适用于微生物培养物的维持和生长的材料和方法是本领域熟知的。适用于下面实施例的技术可在 Manual of Methods for General Bacteriology (Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg 和 G. Briggs Phillips 编辑), American Society for Microbiology: Washington, D. C. (1994)); 或 Thomas D. Brock in Biotechnology: A Textbook of Industrial Microbiology, 第 2 版, Sinauer Associates: Sunderland, MA (1989) 的描述中找到。用于生长和维持微生物细胞的所有试剂、限制性酶和材料得自 Aldrich Chemicals (Milwaukee, WI)、DIFCO Laboratories (Detroit, MI)、New England Biolabs, Inc. (Beverly, MA)、GIBCO/BRL (Gaithersburg, MD)、或 Sigma Chemical Company (St. Louis, MO)。通常于 37°C 下在 Luria Bertani (LB) 平板上培养大肠杆菌菌株。

[0271] 一般的分子克隆根据标准方法 (Sambrook 等人, 同上) 来完成。DNA 序列是在 ABI 自动测序仪上采用染料终止剂技术 (美国专利 5,366,860; EP 272,007) 使用载体和插入特异性引物的组合来产生的。序列编辑是在 Sequencher (Gene Codes Corporation, Ann Arbor, MI) 中进行的。所有序列在两个方向上覆盖至少两次。除非另外指明, 使用 DNASTAR 软件 (DNASTAR Inc., Madison, WI) 比较本文的基因序列。

[0272] 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, “ μ L”表示微升, “mL”表示毫升, “L”表示升, “ μ M”表示微摩尔浓度, “mM”表示毫摩尔浓度, “M”表示摩尔浓度, “mmol”表示毫摩尔, “ μ mole”表示微摩尔, “g”表示克, “ μ g”表示微克, “ng”表示纳克, “U”表示单位, “bp”表示碱基对, 并且 “kB”表示千碱基对。

[0273] 表达盒的命名:

[0274] 表达盒的结构通过简单的记号系统 “X::Y::Z” 表示, 其中 X 描述启动子片段, Y 描述基因片段而 Z 描述终止子片段, 它们均彼此可操作地连接。

[0275] 解脂耶氏酵母的转化和培养

[0276] 解脂耶氏酵母菌株 ATCC#20362 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是在根据下面所示配方的数种培养基中于 28–30°C 下培育。根据标准方法, 需要通过将 20g/L 琼脂添加到每种液体培养基中来制备琼脂平板。

[0277] YPD 琼脂培养基 (每升): 10g 酵母提取物 [Difco]、20g 细菌用蛋白胨 [Difco]、以及 20g 葡萄糖。

[0278] 基本培养基 (MM) (每升):20g 葡萄糖、1.7g 无氨基酸酵母氮源、1.0g 脯氨酸,并且 pH6.1 (不需要调节)。

[0279] 基本培养基 + 尿嘧啶 (MM+ 尿嘧啶或 MMU) (每升):如上制备 MM 培养基并加入 0.1g 尿嘧啶和 0.1g 尿苷。

[0280] 基本培养基 + 尿嘧啶 + 磺酰脲类 (MMU+SU) (每升):如上制备 MMU 培养基并加入 280mg 磺酰脲类。

[0281] 基本培养基 + 亮氨酸 + 赖氨酸 (MMLeuLys) (每升):如上制备 MM 培养基并加入 0.1g 亮氨酸和 0.1g 赖氨酸。

[0282] 基本培养基 +5- 氟乳清酸 (MM+5-FOA) (每升):20a 葡萄糖、6.7g 酵母氮源、75mg 尿嘧啶、75mg 尿苷,并且基于对一系列从 100mg/L 至 1000mg/L 的浓度的 FOA 活性试验 (因为供应商提供的每批料中会发生改变),选用适量的 FOA (Zymo Research Corp., Orange, CA)。

[0283] 高葡萄糖培养基 (HGM) (每升):80 葡萄糖、2.58g KH_2PO_4 和 5.36g K_2HPO_4 , pH7.5 (不需要调节)。

[0284] 无酵母提取物的发酵培养基 (FM 无 YE) (每升):6.70g 酵母氮源、6.00g KH_2PO_4 , 2.00g K_2HPO_4 、1.50g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 和 20g 葡萄糖。

[0285] 发酵培养基 (FM) (每升):如上制备无 YE 培养基的 FM 并加入 5.00g 酵母提取物 (BBL)。

[0286] 合成右旋糖培养基 (SD) (每升):6.7g 酵母氮源,具有硫酸铵并且无氨基酸;和 20g 葡萄糖。

[0287] Complete Minimal Glucose Broth Minus Uracil (CSM-Ura):目录号

[0288] C8140, Teknova, Hollister, CA (0.13% 氨基酸缺陷型粉末,无尿嘧啶。0.17% 酵母氮源,0.5% $(\text{NH}_4)_2\text{SO}_4$,2.0% 葡萄糖)。

[0289] 解脂耶氏酵母的转化根据 Chen, D. C. 等人 (Appl. Microbiol Biotechnol., 48 (2) : 232-235 (1997)) 的方法完成,除非另外说明。简而言之,将耶氏酵母划线到 YPD 平板上,并在 30°C 下生长大约 18 小时。从平板上刮掉几大环量的细胞并将其重悬浮于含有下面成分的 1mL 转化缓冲液中:2.25mL 50% PEG, 平均分子量为 3350 ;0.125mL 2M 醋酸锂, pH6.0 ;以及 0.125mL 2M DTT。然后,将大约 500ng 线性化的质粒 DNA 在 100 μL 重悬浮的细胞中孵育,并将其在 39°C 下维持 1 小时,每间隔 15 分钟进行涡旋混合。将细胞铺在选择培养基平板上并在 30°C 下维持 2 至 3 天。

[0290] 解脂耶氏酵母的脂肪酸分析

[0291] 为了进行脂肪酸分析,将细胞通过离心收集并如 Bligh, E. G. & Dyer, W. J. (Can. J. Biochem. Physiol., 37 :911-917 (1959)) 中所述提取脂质。通过将脂质提取物与甲醇钠进行酯交换反应而制备脂肪酸甲酯 [“FAME” (Roughan, G. 和 Nishida I., Arch Biochem Biophys., 276 (1) :38-46 (1990)) 并随后将其用配有 30m $\times$ 0.25mm (内径) HP-INNOWAX (Hewlett-Packard) 柱的 Hewlett-Packard 6890 气相色谱仪 (GC) 进行分析。炉温以 3.5°C / 分钟从 170°C (保持 25 分钟) 升到 185°C。

[0292] 为了进行直接的碱催化酯交换反应,收获耶氏酵母培养物 (3mL),在蒸馏水中洗涤一次,并在 Speed-Vac 中在真空下干燥 5-10 分钟。将甲醇钠 (100 μL , 浓度为 1%) 加入样

本中,然后将样本涡旋振荡 20 分钟。在加入 3 滴 1M 氯化钠和 400 μ L 己烷后,涡旋并离心样本。移出上层并如上所述用 GC 进行分析。

[0293] 实施例 1

[0294] 产生解脂耶氏酵母菌株 Y4086 以通过 $\Delta 9$ 延伸酶/ $\Delta 8$ 去饱和酶途径生产占总脂质约 14% 的 EPA

[0295] 本实施例描述了来源于解脂耶氏酵母 ATCC#20362 的菌株 Y4086 的构建,该菌株能通过 $\Delta 9$ 延伸酶/ $\Delta 8$ 去饱和酶途径的表达生产相对于总脂质而言约 14% 的 EPA(图 3A)。

[0296] 菌株 Y4086 的产生需要构建菌株 Y2224(由野生型耶氏酵母菌株 ATCC#20362 的 Ura3 基因自主突变而产生的 FOA 抗性突变体)、菌株 Y4001(产生 17% EDA,具有 Leu⁻表型)、菌株 Y4001U(Leu⁻和 Ura 表型)、菌株 Y4036(产生 18% DGLA,具有 Leu⁻表型)、菌株 Y4036U(具有 Leu⁻和 Ura⁻表型)和菌株 Y4070(产生 12% ARA,具有 Ura⁻表型)。关于构建菌株 Y2224、Y4001、Y4001U、Y4036、Y4036U 和 Y4070 的更多细节描述于实施例 7,国际申请公布 WO 2008/073367。

[0297] 相对于野生型解脂耶氏酵母 ATCC#20362,菌株 Y4070 的最终基因型是 Ura3⁻、unknown 1⁻、unknown 3⁻、Leu⁺、Lys⁺、GPD::\Delta 12 去饱和酶基因[国际申请公布 WO 2005/047485];ME3S 是经密码子优化的 C_{16/18} 延伸酶基因,来源于高山被孢霉[国际申请公布 WO 2007/046817];EgD9e 是小眼虫 $\Delta 9$ 延伸酶基因[国际申请公布 WO 2007/061742];EgD9eS 是经密码子优化的 $\Delta 9$ 延伸酶基因,来源于小眼虫[国际申请公布 WO 2007/061742];EgD8M 是合成突变型 $\Delta 8$ 去饱和酶[国际申请公布 WO 2008/073271],来源于小眼虫[美国专利 7,256,033];EgD5 是小眼虫 $\Delta 5$ 去饱和酶[美国专利申请公布 US 2007-0292924-A1];EgD5S 是经密码子优化的 $\Delta 5$ 去饱和酶基因,来源于小眼虫[美国专利申请公布 2007-0292924];以及 RD5S 是经密码子优化的 $\Delta 5$ 去饱和酶,来源于多甲藻属 CCMP626[美国专利申请公布 2007-0271632])。

[0298] 产生菌株 Y4086 以生产占总脂质约 14% 的 EPA

[0299] 构建体 pZP3-Pa777U(图 3B;SEQ ID NO:28)描述于表 19,国际申请公布 WO 2008/07336,它被生成以将三个 $\Delta 17$ 去饱和酶基因整合进菌株 Y4070 的 Pox3 位点中(GenBank 保藏号 AJ001301),从而生成 EPA。 $\Delta 17$ 去饱和酶基因是 PaD17,一种瓜果腐霉菌 $\Delta 17$ 去饱和酶(国际申请公布 WO 2008/054565),和 PaD17S,一种经密码子优化的 $\Delta 17$ 去饱和酶,来源于瓜果腐霉菌(国际申请公布 WO 2008/054565)。

[0300] 将 pZP3-Pa777U 质粒用 AscI/SphI 消化,然后根据“一般方法”将其用于转化菌株 Y4070。将转化细胞铺在 MM 平板上并在 30°C 下维持 2 至 3 天。将单菌落再次划线至 MM 平板上,随后将其在 30°C 下接种至液体 MMLeuLys 中并以 250rpm/min 摇动 2 天。离心收集细胞,提取脂质,通过酯交换反应来制备 FAME,并随后用 Hewlett-Packard 6890GC 进行分析。

[0301] GC 分析显示,EPA 存在于含有 pZP3-Pa777U 的 3 种嵌合基因的转化体中,但不存在于亲本 Y4070 菌株中。所选的 96 个菌株中大部分产生占总脂质约 10-13% 的 EPA。有 2 个菌株(即 #58 和 #79)产生了占总脂质约 14.2% 和 13.8% 的 EPA。将这两个菌株分别命名

为 Y4085 和 Y4086。

[0302] 相对于野生型解脂耶氏酵母 ATCC#20362, 菌株 Y4086 的最终基因型是 Ura3⁺、Leu⁺、Lys⁺、unknown 1⁻、unknown 2⁻、YALI0F24167g⁻、GPD::FmD12::Pex20、YAT1::FmD12::OCT、YAT1::ME3S::Pex16、GPAT::EgD9e::Lip2、EXP1::EgD9eS::Lip1、FBAINm::EgD9eS::Lip2、FBAINm::EgD8M::Pex20、EXP 1::EgD8M::Pex16、FBAIN::EgD5::Aco、EXP1::EgD5S::Pex20、YAT1::RD5S::OCT、YAT1::PaD17S::Lip1、EXP1::PaD17::Pex16、FBAINm::PaD17::Aco。

[0303] 实施例 2

[0304] 产生解脂耶氏酵母菌株 Y4128 以通过 $\Delta 9$ 延伸酶 / $\Delta 8$ 去饱和酶途径生产占总脂质约 37% 的 EPA

[0305] 本实施例描述来源于解脂耶氏酵母 ATCC#20362 的菌株 Y4128 的构建, 该菌株能够产生相对于总脂质约 37.6% 的 EPA (即, 相对于 Y4086 以总脂肪酸的百分比表示的 EPA 浓度有大于 2 倍的提高; 图 3A)。

[0306] 菌株 Y4128 的产生需要构建菌株 Y2224、Y4001、Y4001U、Y4036、Y4036U、Y4070 和 Y4086 (如实施例 1 所述), 以及构建菌株 Y4086U1 (Ura⁻)。

[0307] 产生菌株 Y4086U1 (Ura⁻)

[0308] 菌株 Y4086U1 经由在菌株 Y4086 的构建体 pY117 中暂时表达 Cre 重组酶 (图 4A; SEQ ID NO:29; 描述于国际申请公布 W0 2008/073367 的表 20 中) 制备 Ura⁻ 表型。这从基因组释放出 LoxP 夹着的 Ura3 基因。突变的耶氏酵母属乙酰羟酸合酶 [“AHAS”; E. C. 4. 1. 3. 18] (即, GenBank 保藏号 XP501277, 包含 W497L 突变, 如 SEQ ID NO:27 所示; 参见国际申请公布 W0 2006/052870) 在质粒 pY117 中, 赋予其磺酰脲类抗除草剂性 (SU^R), 该抗性用作阳性筛选标记。

[0309] 根据“一般方法”将质粒 pY117 用于转化菌株 Y4086。在转化后, 将细胞置于 MMU+SU (280 μ g/mL 磺酰脲类; 也称作氯嘧磺隆, E. I. duPont de Nemours & Co., Inc., Wilmington, DE) 平板上并在 30°C 下保持 2 至 3 天。挑取在 MMU+SU 平板上生长的 SU^R 单菌落, 将其在 30°C 下划线至 YPD 液体培养基中并以 250rpm/min 摇动 1 天以固化 pY117 质粒。将生长培养基划线接种在 MMU 平板上。在 30°C 下保持两天后, 将单克隆再划线接种在 MM 和 MMU 平板上。选择可以在 MMU 平板上生长但不在 MM 平板上生长的那些菌落。将这些菌株中具有 Ura⁻ 表型的两个菌株命名为 Y4086U1 和 Y4086U2。

[0310] 产生菌株 Y4128 以生产占总脂质约 37% 的 EPA

[0311] 制备构建体 pZP2-2988 (图 4B; SEQ ID NO:30; 描述于表 21, 国际申请公布 W0 2008/073367) 以将一个 $\Delta 12$ 去饱和酶基因 (即, FmD12S, 一种经密码子优化的 $\Delta 12$ 去饱和酶基因, 来源于串珠镰刀菌 [国际申请公布 W0 2005/047485])、两个 $\Delta 8$ 去饱和酶基因 (即, EgD8M) 和一个 $\Delta 9$ 延伸酶基因 (即, EgD9eS) 整合进菌株 Y4086U1 的 Pox2 位点 (GenBank 保藏号 AJ001300) 中, 从而在较高水平上生产 EPA。将 pZP2-2988 质粒用 AscI/SphI 消化, 然后根据“一般方法”将其用于转化菌株 Y4086U1。将转化细胞铺在 MM 平板上并在 30°C 下维持 2 至 3 天。将单菌落再次划线至 MM 平板上, 随后将其在 30°C 下接种至液体 MMLeuLys 中并以 250rpm/min 摇动 2 天。通过离心收集细胞, 将其重悬浮于 HGM 中, 然后以 250rpm/min 摇动 5 天。离心收集细胞, 提取脂质, 通过酯交换反应来制备 FAME, 并随后用

Hewlett-Packard 6890GC 进行分析。

[0312] GC 分析显示所选择的 96 个菌株中的大多数生产了占总脂 12-15.6% 的 EPA。有 2 个菌株 (即, Group I 中的 #37 和 Group II 中的 #33) 生产占总脂质约 37.6% 和 16.3% 的 EPA。将这两个菌株分别命名为 Y4128 和 Y4129。

[0313] 相对于野生型解脂耶氏酵母 ATCC#20362 的菌株 Y4128 的最终基因型是: YALI0F24167g-、Pex10-、unknown 1-、unknown 2-、GPD::Fmd12::Pex20、YAT1::Fmd12::OCT、GPM/FBAIN::Fmd12S::OCT、YAT1::ME3S::Pex16、GPAT::EgD9e::Lip2、EXP1::EgD9eS::Lip1、FBAINm::EgD9eS::Lip2、FBA::EgD9eS::Pex20、FBAINm::EgD8M::Pex20、EXP1::EgD8M::Pex16、GPDIN::EgD8M::Lip1、YAT1::EgD8M::Aco、FBAIN::EgD5::Aco、EXP1::EgD5S::Pex20、YAT1::RD5S::OCT、YAT1::PaD17S::Lip1、EXP1::PaD17::Pex16、FBAINm::PaD17::Aco。

[0314] 在 2007 年 8 月 23 日将解脂耶氏酵母菌株 Y4128 保藏于美国典型培养物保藏中心, 并且命名为 ATCC PTA-8614。

[0315] 产生具有 Ura- 表型的 Y4128U 菌株

[0316] 为了破坏菌株 Y4128 中的 Ura3 基因, 制备构建体 pZKUE3S (图 5ASEQ ID NO:31; 描述于表 22, 国际申请公布 WO 2008/073367) 以将 EXP1::ME3S::Pex20 嵌合基因整合进菌株 Y4128 的 Ura3 基因中。将质粒 pZKUE3S 用 SphI/PacI 消化, 然后根据“一般方法”将其用于转化菌株 Y4128。转化后, 将细胞铺至 MM+5-FOA 选择平板上并在 30°C 下维持 2 至 3 天。

[0317] 挑取在 MM+5-FOA 选择平板上生长的总共 24 个转化体, 并且再次划线接种到新 MM+5-FOA 平板上。从板上剥离细胞, 提取脂质, 通过酯交换反应来制备 FAME, 并随后用 Hewlett-Packard 6890GC 进行分析。

[0318] GC 分析显示来自平板的具有 pZKUE3S 的所有转化体中存在 10-15% 的 EPA。鉴定为 #3、#4、#10、#12、#19 和 #21 的菌株分别生产占总脂质 12.9%、14.4%、15.2%、15.4%、14% 和 10.9% 的 EPA, 将它们分别命名为 Y4128U1、Y4128U2、Y4128U3、Y4128U4、Y4128U5 和 Y4128U6 (统称为 Y4128U)。

[0319] Y4128 (37.6%) 对 Y4128U (平均 13.8%) % EPA 定量差异是因为不同的生长条件而产生的。具体地讲, 前者的培养基是在液体培养基中生长两天后进行分析的, 而后的培养基是在琼脂平板上生长后进行分析的。申请人已经观察到当比较琼脂平板与液体培养基中的结果时, % EPA 提高了 2-3 倍。因此, 虽然不能直接比较结果, 但 Y4128 和 Y4128U 菌株都证明产生了 EPA。

[0320] 实施例 3

[0321] 测定解脂耶氏酵母菌株 Y4128 的总脂质含量

[0322] 通过 GC 分析法测定菌株 Y4128 生产的脂质总量和脂质中的每种脂肪酸物质的百分比。具体地讲, 如一般方法所述, 提取总脂质, 通过酯交换反应来制备 FAME, 并随后用 Hewlett-Packard 6890GC 进行分析。

[0323] 干细胞重量如下进行测定: 从 10mL 培养物中经由离心收集细胞, 用水洗涤细胞一次以除去残余培养基, 在 80°C 真空炉中干燥细胞过夜, 称量干细胞重量。样本中的 FAME 总量通过比较 GC 特征图中的所有峰面积与加入的已知量内部标准品 C15:0 脂肪酸的峰面积而进行测定。

[0324] 基于上述分析,测定菌株 Y4086 和 Y4128 的以干细胞重量百分比形式表示的脂质含量(DCW)和脂质组成。菌株 Y4128 与菌株 Y4086 相比脂质含量降低(11.2TFA% DCW 对 28.6TFA% DCW)。相反地,菌株 Y4128 与菌株 Y4086 相比脂质中的 EPA 浓度提高,如下表 6 所示。脂肪酸被鉴定为 18:0(硬脂酸)、18:1(油酸)、LA、ALA、EDA、DGLA、ETrA、ETA 和 EPA;脂肪酸组成表示为总脂肪酸的重量百分比(wt. %)(TFA)。

[0325] 表 6

[0326] 在解脂耶氏酵母菌株 Y4086 和 Y4128 中的脂质组成

[0327]

样本	18:0	18:1	18:2 [LA]	18:3 (n-3) [ALA]	20:2 [EDA]	20:3 (n-6) [DGLA]	20:3 (n-3) [ETrA]	20:4 (n-3) [ETA]	20:5 (n-3) [EPA]
Y4086	4.6	26.8	28.0	6.9	7.6	0.9	4.9	2.0	9.8
Y4128	1.8	6.7	19.6	1.8	4.2	3.4	1.5	6.0	42.8

[0328] 细胞中的 EPA 含量以 mg EPA/g 干细胞形式表示,并且按照下式计算:(% EPA/脂质)*(%脂质/干细胞重量)*0.1,其从菌株 Y4086 中的 28mg EPA/g DCW 提高到菌株 Y4128 中的 47.9mg EPA/g DCW。

[0329] 因此,表 6 中的结果显示与亲本菌株 Y4086 相比,菌株 Y4128 具有更低的总脂质含量(TFA% DCW)(11.2%对 28.6%),更高的 EPA% TFA(42.8%对 9.8%),和更高的 EPA% DCW(4.8%对 2.8%)。此外,菌株 Y4128 的 EPA 相对于总 PUFA 的量提高了 3.3 倍(54%的 PUFA[% TFA]对 16.3%的 PUFA[% TFA]),而 C20PUFA 相对于总 PUFA 的量提高了 1.7 倍(73%的 PUFA[% TFA]对 42%的 PUFA[% TFA])。

[0330] 实施例 4

[0331] 测定在解脂耶氏酵母菌株 Y4128 中的整合位点 pZP2-2988 发生 Pex10 整合

[0332] 菌株 Y4128 中的 pZP2-2988 基因组整合位点通过基因组步移法,使用来自 Clontech(Palo Alto, CA) 的 Universal GenomeWalker™试剂盒,按照制造商推荐的规程进行测定。基因质粒序列,设计以下引物用于基因组步移法:pZP-GW-5-1(SEQ ID NO:32)、pZP-GW-5-2(SEQ IDNO:33)、pZP-GW-5-3(SEQ ID NO:34)、pZP-GW-5-4(SEQ ID NO:35)、pZP-GW-3-1(SEQ ID NO:36)、pZP-GW-3-2(SEQ ID NO:37)、pZP-GW-3-3(SEQ ID NO:38) 和 pZP-GW-3-4(SEQ ID NO:39)。

[0333] 使用 Qiagen Miniprep 试剂盒用改进规程从菌株 Y4128 中制备基因组 DNA。从 YPD 培养基平板上刮除细胞并置于 1.5mL 微管中。将细胞沉淀物(100 μL)用 250 μL 缓冲液 P1 重悬,该缓冲液 P1 包含 0.125M β-巯基乙醇和 1mg/mL 酵母裂解酶 20T(MP Biomedicals, Inc., Solon, OH)。细胞悬浮液在 37℃ 下培养 30 分钟。然后将缓冲液 P2(250 μL)加到管中。在将管颠倒几次进行混合后,加入 350 μL 缓冲液 N3。然后将微管中的混合物在 14,000rpm 下离心 5 分钟。将上清液注入 Qiagen 小量制备离心柱中,离心 1 分钟。加入 0.75mL 缓冲液

PE 洗涤一次柱子,然后在 14,000rpm 下离心 1 分钟。在 14,000rpm 下再离心 1 分钟干燥该柱。通过加入 50 μ L 的缓冲液 EB 到柱中洗脱基因组 DNA,允许静置 1 分钟并在 14,000rpm 下离心 1 分钟。

[0334] 使用纯化的基因组 DNA 进行基因组步移法。按照 GenomeWalker 试剂盒的规程,用限制性酶 DraI、EcoRV、PvuII 和 StuI 分别消化 DNA。就每次消化而言,反应混合物包含 10 μ L 10X 限制缓冲液、10 μ L 适用的限制性酶和 8 μ g 基因组 DNA,总体积为 100 μ L。反应混合物在 37 $^{\circ}$ C 下培养 4 小时。然后使用 Qiagen PCR 纯化试剂盒,严格按照制造商规程纯化经消化的 DNA 样本。DNA 样本用 16 μ L 水洗脱。然后将纯化的经消化基因组 DNA 样本连接到 genome walker 衔接子上(下文)。每个连接混合物包含 1.9 μ L genome walker 衔接子、1.6 μ L 10X 连接缓冲液、0.5 μ L T4DNA 连接酶和 4 μ L 消化过的 DNA。反应混合物在 16 $^{\circ}$ C 下培养过夜。然后将 72 μ L 的 50mM TrisHCl、1mM EDTA, pH7.5 加入到每个连接混合物中。

[0335] 就 5' 末端基因组步移法而言,使用 1 μ L 的每个连接混合物作模板进行四个 PCR 反应。此外,每个反应混合物包含 1 μ L 的 10 μ M 引物 pZP-GW-5-1 (SEQ ID NO :32)、1 μ L 的 10 μ M 试剂盒配送的 GenomeWalker 衔接子、41 μ L 水、5 μ L 10X cDNA PCR 反应缓冲液和 1 μ L 来自 Clontech 的 Advantage cDNA 聚合酶混合物。Genome Walker 衔接子序列 (SEQ ID NOs :40 [顶链] 和 41 [底链]) 显示如下 :

[0336] 5' -GTAATACGACTCACTATAGGGCACGCGTGGTCGACGGCCCCG

[0337] GGCTGGT-3'

[0338] 3' -H2N-CCCGACCA-5'

[0339] PCR 条件如下 :95 $^{\circ}$ C 1 分钟,随后是 30 个 95 $^{\circ}$ C 20 秒和 68 $^{\circ}$ C 3 分钟的循环,最后在 68 $^{\circ}$ C 下延伸 7 分钟。将 PCR 产物每个进行 1:100 的稀释,并且将 1 μ L 稀释 PCR 产物用作模板进行第二轮 PCR。条件是完全相同的,除了用 pZP-GW-5-2 (SEQ ID NO :33) 代替 pZP-GW-5-1 (SEQ ID NO :32)。

[0340] 就 3' -末端基因组步移而言,如上文所述进行四个 PCR 反应,除了使用引物 pZP-GW-3-1 (SEQ ID NO :36) 和巢式衔接子引物 (SEQ ID NO :42)。PCR 产物同样进行稀释并用作第二轮 PCR 的模板,使用 pZP-GW-3-2 (SEQ ID NO :37) 代替 pZP-GW-3-1 (SEQ ID NO :36)。

[0341] 通过凝胶电泳分析 PCR 产物。使用 EcoRV 消化的基因组 DNA 作为模板和引物 pZP-GW-3-2 以及巢式衔接子引物,一个反应产物生成 ~ 1.6kB 的片段。分离该片段,用 Qiagen 凝胶纯化试剂盒纯化并克隆到 pCR2.1-TOPO 中。序列分析显示该片段包括两部分 :质粒 pZP2-2988 和 耶氏酵母属来自染色体 C 的基因组 DNA。它们在染色体 C 的核苷酸位点 139826 上接合。这是位于 Pex10 基因的编码区内部 (GenBank 保藏号 CAG81606 ;SEQ ID NO :10)。

[0342] 为了测定接合的 5' 末端,使用来自菌株 Y4128 的基因组 DNA 作为模板以及引物 Per10F1 (SEQ ID NO :43) 和 ZPGW-5-5 (SEQ ID NO :44) 进行 PCR 扩增。反应混合物包括 1 μ L 的每个均为 20 μ M 的引物、1 μ L 基因组 DNA、22 μ L 水和 25 μ L TaKaRa ExTaq 2X 预混物 (TaKaRa BioInc., Otsu Shiga, Japan)。热循环仪条件是 :94 $^{\circ}$ C 1 分钟,然后为 30 个 94 $^{\circ}$ C 20 秒、55 $^{\circ}$ C 20 秒和 72 $^{\circ}$ C 2 分钟的循环,之后在 72 $^{\circ}$ C 下进行 7 分钟最终延伸反应。扩增了 1.6kB

的 DNA 片段并将其克隆到 pCR2.1-TOPO 中。序列分析显示它是一个嵌合片段,该片段位于耶氏酵母属来自染色体 C 的基因组 DNA 和 pZP2-2988 之间。接合位于染色体 C 的核苷酸位点 139817。因此,染色体 C 的 10 个核苷酸的片段被来自菌株 Y4128 中的 pZP2-2988 (图 4B) AscI/SphI 片段置换。因此,菌株 Y4128 中的 Pex10 缺少编码蛋白的最后 32 个氨基酸。

[0343] 基于上述结论,实施例 2 中分离的 Y4128U 菌株 2 (同上) 后来称为 Δ pex10 菌株。为清楚起见,菌株 Y4128U1 等同于菌株 Y4128U1 (Δ pex10)。

[0344] 实施例 5

[0345] 在解脂耶氏酵母菌株 Y4128U1 (Δ pex10) 中的 Pex10 质粒表达

[0346] 构建递送解脂耶氏酵母 Pex10 基因的三个质粒:1)pFBAln-PEX10 允许 Pex10 ORF 在 FBAlNm 启动子控制下表达;和 2)pPEX10-1 和 pPEX10-2 允许在天然 Pex10 启动子控制下表达 Pex10,虽然 pPEX10-1 使用较短版本 (~ 500 bp) 的启动子而 pPEX10-2 使用较长版本 (~ 900 bp) 的启动子。在构建这些表达质粒并转化后,测定 Pex10 质粒表达对解脂耶氏酵母菌株 Y4128U1 (Δ pex10) 总油含量和 EPA 含量的效应。缺失 Pex10 导致细胞中以 TFA 百分比形式表示的 EPA 的量提高,但是以 DCW 百分比形式表示的总脂质的量降低。

[0347] 构建 pFBAln-PEX10、pPEX10-1 和 pPEX10-2

[0348] 为了构建 pFBAln-PEX10,使用引物 Per10F1 (SEQ ID NO:43) 和 Per10R (SEQ ID NO:45),使用解脂耶氏酵母基因组 DNA 作为模板扩增 Pex10 基因的编码区。PCR 反应混合物包含 1 μ L 每个均为 20 μ M 的引物、1 μ L 解脂耶氏酵母基因组 DNA (~ 100 ng)、25 μ L ExTaq 2X 预混物和 22 μ L 水。反应如下进行:94 $^{\circ}$ C 1 分钟,然后为 30 个 94 $^{\circ}$ C 20 秒、55 $^{\circ}$ C 20 秒和 72 $^{\circ}$ C 90 秒的循环,之后在 72 $^{\circ}$ C 下进行 7 分钟最终延伸反应。PCR 产物是一个 1168bp 的 DNA 片段,它用 Qiagen PCR 纯化试剂盒进行纯化,用 NcoI 和 NotI 消化,并克隆到用相同的两个限制性酶消化的 pFBAln-MOD-1 (SEQ ID NO:46;图 5B) 中。

[0349] 8 个单克隆进行序列分析,2 个具有无错的正确 Pex10 序列。pFBAln-PEX10 (SEQ ID NO:47;图 6A) 组分在下表 7 中列出。

[0350] 表 7

[0351] 质粒 pFBAln-PEX10 (SEQ ID NO:47) 的组分

[0352]

SEQ ID NO:47 内的 RE 位点和核苷酸	片段和嵌合基因组分的描述
BglIII-BsiWI (6040-318)	FBAlNm::Pex10::Pex20,它包含: ● FBAlNm:解脂耶氏酵母 FBAlNm 启动子 (美国专利 7,202,356); ● Pex10:解脂耶氏酵母 Pex10ORF (GenBank 保藏号 AB036770,核苷酸 1038-2171;SEQ ID NO:21); ● Pex20:来自耶氏酵母属 Pex20 基因 (GenBank 保藏号 AF054613) 的 Pex20 终止子序列
PacI-BglIII (4530-6040)	耶氏酵母属 URA3 (GenBank 保藏号 AJ306421)

(3123-4487)	耶氏酵母属自主复制序列 18(ARS18 ;GenBank 保藏号 A17608)
(2464-2864)	大肠杆菌 f1 复制起点
(1424-2284)	用于在大肠杆菌中选择的氨苄青霉素抗性基因 (Amp ^R)
(474-1354)	ColE1 质粒复制起点

[0353] 为了构建 pPEX10-1 和 pPEX10-2, 设计并合成引物 PEX10-R-BsiWI (SEQ ID NO : 48)、PEX10-F1-SalI (SEQ ID NO : 49) 和 PEX10-F2-SalI (SEQ ID NO : 50)。使用基因组解脂耶氏酵母 DNA 以及引物 PEX10-R-BsiWI 和 PEX10-F1-SalI 进行的 PCR 扩增生成 1873bp 的片段, 该片段包含 Pex10 ORF, Pex10 基因的 500bp 的 5' 上游区域和 215bp 的 3' 下游区域, 该基因的两端侧接 SalI 和 BsiWI 限制性位点。该片段用 Qiagen PCR 纯化试剂盒进行纯化, 用 SalI 和 BsiWI 消化, 并克隆到用相同的两个酶消化的 pEXP-MOD-1 (SEQ ID NO : 51 ; 图 6B) 中以生成 pPEX10-1 (SEQ ID NO : 52 ; 图 7A)。质粒 pEXP-MOD1 类似于 pFBAIn-MOD-1 (SEQ ID NO : 46 ; 图 5B), 除了后者的 FBAInm 启动子被 EXP1 启动子置换。表 8 列出 pPEX10-1 的组分。

[0354] 表 8

[0355] 质粒 pPEX10-1 (SEQ ID NO : 52) 的组分

[0356]

SEQ ID NO : 52 中的 RE 位点和核苷酸	片段和嵌合基因组分的描述
SalI-BsiWI (5705-1)	Pex10-5' ::Pex10::Pex10-3' , 它包含 : ● Pex10-5' : 解脂耶氏酵母 Pex10 基因的 500bp 的 5' 启动子区域 ; ● Pex10 : 解脂耶氏酵母 Pex10ORF (GenBank 保藏号 AB036770, 核苷酸 1038-2171 ; SEQ ID NO : 21) ; ● Pex10-3' : 来自耶氏酵母属 Pex10 基因的 215bp 的 Pex10 终止子序列 (GenBank 保藏号 AB036770) [注意图中的整个 Pex10-5' ::Pex10::Pex10-3' 表达盒统一标记成“PEX10”]
PacI-SalI (4216-5703)	耶氏酵母属 URA3 基因 (GenBank 保藏号 AJ306421)
(2806-4170)	耶氏酵母属自主复制序列 18(ARS18 ;GenBank 保藏号 A17608)
(2147-2547)	大肠杆菌 f1 复制起点

(1107-1967)	用于在大肠杆菌中选择的氨苄青霉素抗性基因 (Amp ^R)
(157-1037)	ColE1 质粒复制起点

[0357] 使用 PEX10-R-BsiWI (SEQ ID NO :48) 和 PEX10-F2-SalI (SEQ ID NO :50) 对解脂耶氏酵母基因组 DNA 进行 PCR 扩增生成 2365bp 的片段,该片段包含 Pex100RF, Pex10 基因的 991bp 的 5' 上游区域和 215bp 的 3' 下游区域,该基因的两端侧接 SalI 和 BsiWI 限制性位点。用 Qiagen PCR 纯化试剂盒纯化该片段,用 SalI 和 BsiWI 进行消化,然后将其克隆到进行了相似消化的 pEXP-MOD-1 中。这导致合成 pPEX10-2 (SEQ ID NO :53),其构建类似于质粒 pPEX10-1 的构建 (表 8,同上),除了在嵌合 Pex10-5' ::Pex10::Pex10-3' 基因中有更长的 Pex10-5' 启动子。

[0358] 在菌株 Y4128U1 (Δ pex10) 中表达 Pex10

[0359] 按照一般方法中的规程将质粒 pFBAIN-MOD-1 (对照物;SEQ ID NO :46)、pFBAln-PEX10 (SEQ ID NO :47)、pPEX10-1 (SEQ ID NO :52) 和 pPEX10-2 (SEQ ID NO :53) 转化到 Y4128U1 (Δ pex10) 中。将转化体置于 MM 平板上。递送上述质粒的转化体总脂质含量和脂肪酸组成按照实施例 3 所述进行分析。

[0360] 下表 9 显示以干细胞重量 (DCW) 百分比表示的脂质含量和脂质组成。具体地讲,将脂肪酸鉴定为 18:0 (硬脂酸)、18:1 (油酸)、LA、ALA、EDA、DGLA、ETrA、ETA 和 EPA ;脂肪酸组成表示为总脂肪酸的重量百分比 (wt. %)。

[0361]

表 9
在用多个 Pex10 质粒转化过的解脂耶氏酵母菌株 Y4128U1 (Δ pex10) 中的脂质组成

质粒	TFA %DCW	18:0	18:1	18:2 [LA]	18:3 (ω 3) [ALA]	20:2 [EDA]	20:3 (ω 6) [DGLA]	20:3 (ω 3) [ETrA]	20:4 (ω 3) [ETA]	20:5 (ω 3) [EPA]
pFBAIN-MOD-1	22.8	1.9	9.6	18.3	2.0	4.3	2.3	2.1	5.9	27.7
pFBAIN-PEX10	29.2	4.0	24.9	25.1	7.6	6.6	1.0	5.3	3.6	10.8
pPEX10-1	27.1	3.9	25.0	25.2	8.2	6.4	0.9	5.2	3.5	10.7
pPEX10-2	28.5	4.3	25.4	24.5	7.6	6.4	1.0	5.3	3.4	10.8

[0362] 表 9 中的结果显示在 Y4128U1 (Δ pex10) 中从天然解脂耶氏酵母 Pex10 启动子或从解脂耶氏酵母 FBAINm 启动子表达 Pex10, 将 EPA 百分比降低回 Y4086 的水平, 同时将总脂

质含量 (TFA% DCW) 提高到 Y4086 的水平 (参见表 6 数据用于比较)。每克干细胞的 EPA 含量从对照样本中 (即, 递送 pFBAln-MOD-1 的细胞) 的 63.2mg 变成递送 pFBAln-PEX10 的细胞中的 31.5mg、递送 pPEX10-1 的细胞中的 29mg 以及递送 pPEX10-2 的细胞中的 30.8mg。这些结果证明破坏 Pex10 的环指结构域提高了细胞中的 EPA 的量但降低细胞中的总脂质含量。

[0363] 因此, 表 9 的结果显示与具有对照质粒的 Y4128U1 (Δ pex10) 转化体相比, 所有具有 Pex10 表达质粒的转化体显示更高的脂质含量 (TFA% DCW) ($> 27\%$ 对 22.8%)、更低 EPA% TFA (大约 10.8% 对 27.7%)、和更低的 EPA% DCW ($< 3.1\%$ 对 6.3%)。此外, 具有对照质粒的菌株 Y4128U1 (Δ pex10) 转化体与那些具有 Pex10 表达质粒的转化体相比, EPA 相对于总 PUFA 的量提高了 2.5 倍 (44% 的 PUFA[% TFA] 对 17.5% (平均值) 的 PUFA[% TFA]), C20PUFA 相对于总 PUFA 的量提高了 1.5 倍 (67% 的 PUFA[% TFA] 对 44% (平均值) 的 PUFA[% TFA])。

[0364] 实施例 6

[0365] 生成 Y4184U 菌株用于生产 EPA

[0366] 解脂耶氏酵母菌株 Y4184U 用作实施例 7 中的宿主, 下文菌株 Y4184U 来源于解脂耶氏酵母 ATCC#20362, 并且能够经由 $\Delta 9$ 延伸酶 / $\Delta 8$ 去饱和酶途径的表达生产 EPA。该菌株具有 Ura⁻ 表型并且其构建如实施例 7 所述, 国际申请公布 WO 2008/073367。

[0367] 然而概括地说, 菌株 Y4184U 的产生需要构建菌株 Y2224、菌株 Y4001、菌株 Y4001U、菌株 Y4036、菌株 Y4036U 和菌株 Y4069 (如上文实施例 1 所述)。进一步产生菌株 Y4184U (在图 7B 用图表表示) 需要生成菌株 Y4084、菌株 Y4084U1、菌株 Y4127 (在 2007 年 11 月 29 日保藏于美国典型培养物保藏中心, 保藏号 ATCC PTA-8802)、菌株 Y4127U2、菌株 Y4158、菌株 Y4158U1 和菌株 Y4184。质粒构建体 pZKL1-2SP98C 用于转化菌株 Y4127U2, 在图 8A 中用图表表示 (SEQ IDNO:54; 描述于表 23, 国际申请公布 WO 2008/073367)。质粒 pZKL2-5U89GC 用于转化菌株 Y4158U1, 在图 8B 中显示 (SEQ IDNO:55; 描述于表 24, 国际申请公布 WO 2008/073367)。

[0368] 相对于野生型解脂耶氏酵母 ATCC#20362 的菌株 Y4184 (生产占总脂质 31% 的 EPA) 的最终基因型是 unknown 1-、unknown 2-、unknown 4-、unknown 5-、unknown 6-、unknown 7-、YAT1::ME3S::Pex16、EXP1::ME3S::Pex20 (2 拷贝)、GPAT::EgD9e::Lip2、FBAlNm::EgD9eS::Lip2、EXP1::EgD9eS::Lip1、FBA::EgD9eS::Pex20、YAT1::EgD9eS::Lip2、GPD::EgD9eS::Lip2、GPDIN::EgD8M::Lip1、YAT 1::EgD8M::Aco、EXP1::EgD8M::Pex16、FBAlNm::EgD8M::Pex20、FBAlN::EgD8M::Lip1 (2 拷贝)、GPM/FBAlN::FmD 12S::Oct、EXP1::FmD12S::Aco、YAT1::FmD12::Oct、GPD::FmD12::Pex20、EXP1::EgD5S::Pex20、YAT1::EgD5S::Aco、YAT1::Rd5S::Oct、FBAlN::EgD5::Aco、FBAlNm::PaD17::Aco、EXP1::PaD17::Pex16、YAT1::PaD17S::Lip1、YAT1::Y1CPT1::Aco、GPD::Y1CPT1::Aco (其中 FmD12 是串珠镰刀菌 $\Delta 12$ 去饱和酶基因 [国际申请公布 WO2005/047485]; FmD12S 是经密码子优化的 $\Delta 12$ 去饱和酶基因, 来源于串珠镰刀菌 [国际申请公布 WO 2005/047485]; ME3S 是经密码子优化的 C_{16/18} 延伸酶基因, 来源于高山被孢霉 [国际申请公布 WO 2007/046817]; EgD9e 是小眼虫 $\Delta 9$ 延伸酶基因 [国际申请公布 WO 2007/061742]; EgD9eS 是经密码子优化的 $\Delta 9$ 延伸酶基因, 来源于小眼虫 [国际申请公布

WO 2007/061742];EgD8M 是合成突变型 $\Delta 8$ 去饱和酶 [国际申请公布 WO 2008/073271], 来源于小眼虫 [美国专利 7, 256, 033];EgD5 是小眼虫 $\Delta 5$ 去饱和酶 [美国专利申请公布 US 2007-0292924-A1];EgD5S 是经密码子优化的 $\Delta 5$ 去饱和酶基因, 来源于小眼虫 [美国专利申请公布 2007-0292924];RD5S 是经密码子优化的 $\Delta 5$ 去饱和酶, 来源于多甲藻属 CCMP626 [美国专利申请公布 2007-0271632];PaD17 是瓜果腐霉菌 $\Delta 17$ 去饱和酶 [国际申请公布 WO 2008/054565];PaD17S 是一种经密码子优化的 $\Delta 17$ 去饱和酶, 来源于瓜果腐霉菌 [国际申请公布 WO2008/054565];和 Y1CPT1 是一种解脂耶氏酵母二酰基甘油胆碱磷酸转移酶基因 [国际申请公布 WO 2006/052870])。

[0369] 为了破坏菌株 Y4184 中的 Ura3 基因, 构建体 pZKUE3S (图 5A SEQ ID NO :31;描述于表 22, 国际申请公布 WO 2008/073367) 用于将 EXP1::ME3S::Pex20 嵌合基因整合进菌株 Y4184 的 Ura3 基因中以分别产生菌株 Y4184U1 (占总脂质 11.2% 的 EPA)、Y4184U2 (占总脂质 10.6% 的 EPA) 和 Y4184U4 (占总脂质 15.5% 的 EPA) (统称为 Y4184U)。

[0370] 实施例 7

[0371] 解脂耶氏酵母菌株 Y4184U4 中的 Pex10 染色体缺失提高了 EPA 的积聚和总脂质含量

[0372] 构建体 pYPS161 (图 9A, SEQ ID NO :56) 用于从生产 EPA 的耶氏酵母属菌株 Y4184U4 中敲除染色体 Pex10 基因 (实施例 6)。用 Pex10 敲除构建体转化解脂耶氏酵母菌株 Y4184U4 导致生成菌株 Y4184 ($\Delta pex10$)。测定并比较 Pex10 敲除对总油量和 EPA 含量的效应。具体地讲, 敲除 Pex10 导致细胞中 EPA (% TFA 和 % DCW) 百分比提高以及总脂质含量提高。

[0373] 构建体 pYSP161

[0374] 构建体 pYPS161 包含以下组分:

[0375] 表 10

[0376] 质粒 pYPS161 (SEQ ID NO :56) 的描述

[0377]

SEQ ID NO :56 中的 RE 位点和核苷酸	片段和嵌合基因组分的描述
AscI/BsiW I (1521-157)	耶氏酵母属 Pex10 基因的 1364bp 的 Pex10 敲除片段 #1 (GenBank 保藏号 AB036770)
PacI/SphI (5519-4229)	耶氏酵母属 Pex10 基因的 1290bp 的 Pex10 敲除片段 #2 (GenBank 保藏号 AB036770)
SalI/EcoRI (7170-5551)	耶氏酵母属 URA3 基因 (GenBank 保藏号 AJ306421)
2451-1571	ColE1 质粒复制起点
3369-2509	用于在大肠杆菌中选择氨苄青霉素抗性基因 (Amp ^R)

3977-3577

大肠杆菌 f1 复制起点

[0378] 生成解脂耶氏酵母基因敲除菌株 Y4184(Δ Pex10)

[0379] 用于转化解脂耶氏酵母菌株 Y4184U4(实施例 6) 的标准规程使用 Pex10 敲除构建体 pYPS161 的纯化 5.3kB AscI/SphI 片段(同上),并且还制备了细胞单独对照。在每个转化实验中有大约 200 至 250 个菌落存在,然而在细胞单独平板上无菌落存在(每期望值)。

[0380] 菌落 PCR 用于筛选具有 Pex10 缺失的细胞。具体地讲,PCR 反应使用 MasterAmp Taq 聚合酶(Epicentre Technologies, Madison, WI),使用 PCR 引物 Pex-10del1 3' . 正向 (SEQ ID NO:57) 和 Pex-10del2 5' . 反向 (SEQ ID NO:58),按照标准规程进行。PCR 反应条件为 94℃ 5 分钟,然后 30 个 94℃ 30 秒、60℃ 30 秒和 72℃ 2 分钟的循环,之后在 72℃ 下进行 6 分钟的最终延伸反应。然后将该反应保持在 4℃ 下。如果 Pex10 基因敲除构建体在 Pex10 区域整合,期望生成长度为 2.8KB 的单一 PCR 产物。相反地,如果该菌株在除 Pex10 区域之外的染色体区域内整合 Pex10 基因敲除构建体,则将生成两个 PCR 片段,即 2.8kb 和 1.1kb 的片段。在 288 个筛选的菌落中,大部分菌落具有在随机位点整合的 Pex10 基因敲除构建体。288 个菌落中仅有一个菌落包含 Pex10 敲除。将该菌株命名为 Y4184(Δ pex10)。

[0381] 评估解脂耶氏酵母菌株 Y4184 和 Y4184(Δ Pex10) 的总油和 EPA 产量

[0382] 为了评估 Pex10 基因敲除对总脂质级分中的 PUFA 百分比和细胞中的总脂质含量的效应,菌株 Y4184 和 Y4184(Δ pex10) 在可比较的含油条件下生长。具体地讲,培养物在初始 OD₆₀₀为~ 0.1 的情况下开始生长,在含有 25mL 发酵培养基 (FM) 或无酵母提取物 (FM 无 YE) 的 FM 培养基的 250mL 烧瓶中生长 48 小时。在 50mL 圆锥管中,以 8000rpm 离心 10 分钟收获细胞。弃去上清液,将细胞重悬于 25mL HGM 中并转移到新的 250mL 烧瓶中。细胞在 30℃ 下透气培养另外 120 小时。

[0383] 为了测定干细胞重量 (DCW),处理来自 5mL FM 生长培养基和 10mL 无 YE 生长培养物的 FM 的细胞。培养的细胞在 4300rpm 下离心 10 分钟。使用 10mL 生理盐水重悬沉淀物并在相同条件下二次离心。然后用 1mL 无菌 H₂O(第三次)重悬沉淀物,将其转移到预先称重的铝盘中。细胞在 80℃ 真空炉中干燥过夜。测量细胞重量。

[0384] 递送上述质粒的转化体的总脂质含量和脂肪酸组成如实施例 3 所述分析。DCW、总脂质含量 (TFA% DCW)、总 EPA% TFA、和 EPA% DCW 在下表 11 中显示。

[0385] 表 11**[0386] 在解脂耶氏酵母菌株 Y4184 和 Y4184 中的脂质组成 (Δ Pex10)****[0387]**

培养基	菌株	DCW	TFA %DCW	EPA %TFA	EPA %DCW
FM	Y4184	11.5	11.8	20.6	2.4
	Y4184 (Δ Pex10)	11.5	17.6	43.2	7.6
FM 无 YE	Y4184	4.6	8.8	23.2	2.0
	Y4184 (Δ Pex10)	4.0	13.2	46.1	6.1

[0388] 表 11 中的结果显示敲除 Y4184(Δ Pex10) 中的染色体 Pex10 基因与菌株 Y4184 中的 EPA 百分比和总油含量相比提高 EPA 百分比(% TFA 和 % DCW) 并提高总油含量, 菌株 Y4184 的天然 Pex10p 未被敲除。更具体地讲, 在 FM 培养基中, EPA(% TFA) 有约 109% 的提高, EPA 产量(% DCW) 有约 216% 的提高, 总油量(TFA% DCW) 有约 49% 的提高。在无 YE 的 FM 培养基中, EPA(% TFA) 有约 100% 的提高, EPA 产量(% DCW) 有约 205% 的提高, 总油量(TFA% DCW) 有约 50% 的提高。

[0389] 因此, 表 11 中的结果显示与亲本菌株 Y4184 相比, Y4184(Δ Pex10) 菌株在 FM 培养基中具有更高的脂质含量(TFA% DCW)(17.6% 对 11.8%), 更高的 EPA% TFA(43.2% 对 20.6%), 和更高的 EPA% DCW(7.6% 对 2.4%)。同样的, 与亲本菌株 Y4184 相比, Y4184(Δ Pex10) 菌株在无 YE 的 FM 培养基中具有更高的脂质含量(TFA% DCW)(13.2% 对 8.8%), 更高的 EPA% TFA(46.1% 对 23.2%), 和更高的 EPA% DCW(6.1% 对 2.0%)。

[0390] 实施例 8(预测的)

[0391] 在解脂耶氏酵母的 PUFA 生产菌株中的替代 Pex 基因的染色体敲除

[0392] 本实施例描述了已经进行了工程化以生产 ω -3/ ω -6PUFA 的解脂耶氏酵母的多个菌株。如果使用实施例 7 的方法同上破坏编码 Pex1p、Pex2p、Pex3p、Pex3Bp、Pex4p、Pex5p、Pex6p、Pex7p、Pex8p、Pex12p、Pex13p、Pex14p、Pex16p、Pex17p、Pex19p、Pex20p、Pex22p 或 Pex26p 的染色体基因, 预期这些解脂耶氏酵母宿主菌株中的任何一种都能进行工程化以生产在总脂质级分和油级分中含量提高的 ω -3/ ω -6PUFA。

[0393] 更具体地讲, 申请人的受让人已经工程化了多种解脂耶氏酵母菌株, 经由表达异源 Δ 6 去饱和酶/ Δ 6 延伸酶 PUFA 途径或异源 Δ 9 延伸酶/ Δ 8 去饱和酶 PUFA 途径来生产高浓度的多种 ω -3/ ω -6PUFA。

[0394] 生产 ω -3/ ω -6PUFA 的代表性解脂耶氏酵母菌株概述

[0395] 虽然在下表中概述了一些代表性菌株, 但公开的生产 ω -3/ ω -6PUFA 的解脂耶氏酵母菌株不受本文菌株的任何方式的限制。相反地, 除了以下共有的和共同未决的专利申请之外, 在本专利申请中提供的所有教导对于开发合适的工程化解脂耶氏酵母菌株以生产 ω -3/ ω -6PUFA 也是有用的。这些教导具体地讲包括以下申请人的受让人的共同未决的专利和专利申请: 美国专利 7, 125, 672、美国专利 7, 189, 559、美国专利 7, 192, 762、美国专利 7, 198, 937、美国专利 7, 202, 356、美国专利 7, 214, 491、美国专利 7, 238, 482、美国专利 7, 256, 033、美国专利 7, 259, 255、美国专利 7, 264, 949、美国专利 7, 267, 976、美国专利 7, 273, 746、美国专利申请 10/985254 和 10/985691(提交于 2004 年 11 月 10 日)、美国专利申请 11/183664(提交于 2005 年 7 月 18 日)、美国专利申请 11/185301(提交于 2005 年 7 月 20 日)、美国专利申请 11/190750(提交于 2005 年 7 月 27 日)、美国专利申请 11/198975(提交于 2005 年 8 月 8 日)、美国专利申请 11/253882(提交于 2005 年 10 月 19 日)、美国专利申请 11/264784 和 11/264737(提交于 2005 年 11 月 1 日)、美国专利申请 11/265761(提交于 2005 年 11 月 2 日)、美国专利申请 11/601563 和 11/601564(提交于 2006 年 11 月 16 日)、美国专利申请 11/635258(提交于 2006 年 12 月 7 日)、美国专利申请 11/613420(提交于 2006 年 12 月 20 日)、美国专利申请 11/787772(提交于 2007 年 4 月 18 日)、美国专利申请 11/737772(提交于 2007 年 4 月 20 日)、美国专利申请 11/740298(提交于 2007 年 4 月 26 日)、美国专利申请 12/111237(提交于 2008 年 4 月 29 日)、美国专利申请 11/748629 和

11/748637(提交于 2007 年 5 月 15 日)、美国专利申请 11/779915(提交于 2007 年 7 月 19 日)、美国专利申请 60/991266(提交于 2007 年 11 月 30 日)、美国专利申请 11/952243(提交于 2007 年 12 月 7 日)、美国专利申请 61/041716(提交于 2008 年 4 月 2 日)、美国专利申请 12/061738(提交于 2008 年 4 月 3 日)、美国专利申请 12/099811(提交于 2008 年 4 月 9 日)、美国专利申请 12/102879(提交于 2008 年 4 月 15 日)、美国专利申请 12/111237(提交于 2008 年 4 月 29 日)、美国专利申请 61/055511(提交于 2008 年 5 月 23 日)和美国专利申请 61/093007(提交于 2008 年 8 月 29 日)。

[0396]

表 12: 工程化以生产 ω -3/ ω -6 PUFA 的代表性解脂耶氏酵母菌株的脂质分布

菌株	参照序列	ATCC 保藏号	脂肪酸含量 (总脂肪酸百分比[%])														脂质 %dcw
			16:0	16:1	18:0	18:1	18:2	18:3 (ALA)	GLA	20:2	DGLA	ARA	ETA	EPA	DPA	DHA	
野生型	US	#76982	14	11	3.5	34.8	31	--	0	--	--	--	--	--	--	--	--
pDMW208	2006-0035351-A1;	--	11.9	8.6	1.5	24.4	17.8	--	25.9	--	--	--	--	--	--	--	--
pDMW208D62	WO2006/033723	--	16.2	1.5	0.1	17.8	22.2	--	34	--	--	--	--	--	--	--	--
M4	US 2006-0115881-A1; WO2006/052870	--	15	4	2	5	27	--	35	--	8	0	0	0	--	--	--
Y2034	US	--	13.1	8.1	1.7	7.4	14.8	--	25.2	--	8.3	11.2	--	--	--	--	--
Y2047	2006-0094092-A1;	PTA-7186	15.9	6.6	0.7	8.9	16.6	--	29.7	--	0	10.9	--	--	--	--	--
Y2214	WO2006/055322	--	7.9	15.3	0	13.7	37.5	--	0	--	7.9	14	--	--	--	--	--
EU	US	--	19	10.3	2.3	15.8	12	--	18.7	--	5.7	0.2	3	10.3	--	--	36
Y2072	2006-0115881-A1;	--	7.6	4.1	2.2	16.8	13.9	--	27.8	--	3.7	1.7	2.2	15	--	--	--
Y2102	WO2006/052870	--	9	3	3.5	5.6	18.6	--	29.6	--	3.8	2.8	2.3	18.4	--	--	--
Y2088		--	17	4.5	3	2.5	10	--	20	--	3	2.8	1.7	20	--	--	--
Y2089		--	7.9	3.4	2.5	9.9	14.3	--	37.5	--	2.5	1.8	1.6	17.6	--	--	--
Y2095		--	13	0	2.6	5.1	16	--	29.1	--	3.1	1.9	2.7	19.3	--	--	--
Y2090		--	6	1	6.1	7.7	12.6	--	26.4	--	6.7	2.4	3.6	26.6	--	--	22.9

[0397]

Y2096	PTA-7184	8.1	1	6.3	8.5	11.5	--	25	--	5.8	2.1	2.5	28.1	--	--	20.8
Y2201	PTA-7185	11	16.1	0.7	18.4	27	--	--	3.3	3.3	1	3.8	9	--	--	--
Y3000	US 2006-0110806-A1; WO2006/052871	5.9	1.2	5.5	7.7	11.7	--	30.1	--	2.6	1.2	1.2	4.7	18.3	5.6	--
Y4001	--	4.3	4.4	3.9	35.9	23	0	--	23.8	0	0	0	--	--	--	--
Y4036	--	7.7	3.6	1.1	14.2	32.6	0	--	15.6	18.2	0	0	--	--	--	--
Y4070	--	8	5.3	3.5	14.6	42.1	0	--	6.7	2.4	11.9	--	--	--	--	--
Y4158	--	3.2	1.2	2.7	14.5	30.4	5.3	--	6.2	3.1	0.3	3.4	20.5	--	--	27.3
Y4184	--	3.1	1.5	1.8	8.7	31.5	4.9	--	5.6	2.9	0.6	2.4	28.9	--	--	23.9
	WO2008/073367															

[0398] Pex 基因的染色体敲除[0399] 在选择优选解脂耶氏酵母菌株生产期望的 ω -3/ ω -6PUFA(或它们的 PUFA 组

合)后,本领域的技术人员能够容易地工程化合适的基因敲除构建体(类似于在实施例7中的pYPS161)以使得染色体Pex敲除基因通过转化进入解脂耶氏酵母菌株。优选的Pex基因将包括:Y1Pex1p(GenBank保藏号CAG82178;SEQ ID NO:1)、Y1Pex2p(GenBank保藏号CAG77647;SEQ ID NO:2)、Y1Pex3p(GenBank保藏号CAG78565;SEQ ID NO:3)、Y1Pex3Bp(GenBank保藏号CAG83356;SEQ ID NO:4)、Y1Pex4p(GenBank保藏号CAG79130;SEQ ID NO:5)、Y1Pex5p(GenBank保藏号CAG78803;SEQ ID NO:6)、Y1Pex6p(GenBank保藏号CAG82306;SEQ ID NO:7)、Y1Pex7p(GenBank保藏号CAG78389;SEQ ID NO:8)、Y1Pex8p(GenBank保藏号CAG80447;SEQ ID NO:9)、Y1Pex12p(GenBank保藏号CAG81532;SEQ ID NO:11)、Y1Pex13p(GenBank保藏号CAG81789;SEQ ID NO:12)、Y1Pex14p(GenBank保藏号CAG79323;SEQ ID NO:13)、Y1Pex16p(GenBank保藏号CAG79622;SEQ ID NO:14)、Y1Pex17p(GenBank保藏号CAG84025;SEQ ID NO:15)、Y1Pex19p(GenBank保藏号AAK84827;SEQ ID NO:16)、Y1Pex20p(GenBank保藏号CAG79226;SEQ ID NO:17)、Y1Pex22p(GenBank保藏号CAG77876;SEQ ID NO:18)和Y1Pex26p(GenBank保藏号NC_006072,反义翻译核苷酸117230-118387;SEQ ID NO:19)。

[0400] 期望的是染色体破坏Pex基因将导致与天然过氧化物酶体生物合成因子蛋白未被破坏的真核生物相比,在总脂质级分和油级分中积聚的以总脂肪酸百分比形式表示的PUFA量增加,其中所述PUFA量可能是:1)以功能性PUFA生物合成途径期望的终产品形式产生的PUFA,与以中间体或副产品形式产生的PUFA相反,2) C_{20} 和 C_{22} PUFA,和/或3)总PUFA。优选的结果是与天然过氧化物酶体生物合成因子蛋白未被破坏的真核生物相比,不但以总脂肪酸百分比表示的PUFA量提高,而且以干细胞重量百分比形式表示的PUFA量也提高。此外,PUFA量可能是:1)以功能性PUFA生物合成途径期望的终产品形式产生的PUFA,与以中间体或副产品形式产生的PUFA相反,2) C_{20} 和 C_{22} PUFA,和/或3)总PUFA。在一些情况下,相对于未破坏天然过氧化物酶体生物合成因子蛋白的真核生物,总脂质含量也提高。

[0401] 实施例9

[0402] 在解脂耶氏酵母菌株Y4036U中的Pex16染色体缺失提高了积聚的DGLA百分比

[0403] 本实施例描述使用构建体pYRH13(图9B;SEQ ID NO:59)敲除DGLA生产耶氏酵母属菌株Y4036U(实施例1)中的染色体Pex16基因。用Pex16敲除构建体转化解脂耶氏酵母菌株Y4036U导致生成菌株Y4036U(Dpex16)。测定并比较Pex16敲除对DGLA含量的效应。具体地讲,敲除Pex16导致细胞中以总脂肪酸百分比形式存在的DGLA百分比提高。

[0404] 构建体pYRH13

[0405] 质粒pYRH13来源于质粒pYPS161(图9A,SEQ ID NO:56;实施例7)。具体地讲,解脂耶氏酵母Pex16基因(GenBank保藏号CAG79622)的1982bp的5'启动子区域置换pYPS161的AscI/BsiWI片段,解脂耶氏酵母Pex16基因(GenBank保藏号CAG79622)的448bp的3'终止子区域置换pYPS161的PacI/SphI片段以生成pYRH13(SEQ ID NO:59;图9B)。

[0406] 生成解脂耶氏酵母基因敲除菌株Y4036(Δ Pex16)

[0407] 使用标准规程,用Pex16敲除构建体pYRH13的纯化6.0kBAscI/SphI片段转化解脂耶氏酵母的菌株Y4036U(实施例1)。

[0408] 为了筛选具有 Pex16 缺失的细胞,使用 Taq 聚合酶 (Invitrogen ;Carlsbad, CA)、PCR 引物 PEX16Fii (SEQ ID NO :60) 和 PEX16Rii (SEQ ID NO :61) 进行菌落 PCR。设计这组引物以扩增完整 Pex16 基因的 1.1kB 区域,因此 Pex16 缺失突变体 (即, Δ pex16) 将不生成该条带。设计第二组引物生成仅当 Pex16 基因被缺失时出现的条带。具体地讲,一个引物 (即 3UTR-URA3 ;SEQ ID NO :62) 结合到导入的 6.0kB AscI/SphI 破坏片段的载体序列区域上,而另一个引物 (即, PEX16-conf ;SEQ ID NO :63) 结合到在破坏片段的同源序列之外染色体 Pex16 终止子序列上。

[0409] 更具体地讲,使用反应混合物进行菌落 PCR,所述混合物包含 :20mM Tris-HCl (pH 8.4)、50mM KCl、1.5mM MgCl₂、400 μ M 各种 dGTP、dCTP、dATP、和 dTTP、2 μ M 的各个引物、20 μ L 的水和 2U Taq 聚合酶。扩增如下进行 :94°C 120 秒进行初始变性,随后在 94°C 60 秒进行 35 个循环的变性,在 55°C 下退火 60 秒,在 72°C 下延伸 120 秒。最后在 72°C 下进行 5 分钟的最终延伸循环,然后在 4°C 下终止反应。

[0410] 在 205 个筛选的菌落中,195 个菌落具有在染色体随机位点整合的 Pex16 敲除片段,因此不是 Δ pex16 突变体 (然而,由于存在 pYRH13,细胞能够在 ura⁻ 平板上生长)。这些随机整合的菌株中的三个,命名为 Y4036U-17、Y4036U-19 和 Y4036U-33,用作脂质生产实验中的对照物 (下文)。

[0411] 筛选的剩余 10 个菌落 (即,总计 205 个菌落中的剩余菌落) 包含 Pex16 基因敲除。将在 Y4036U 菌株背景中的这十个 Δ pex16 突变体命名为 RHY25 至 RHY34。

[0412] 通过定量实时 PCR 确认解脂耶氏酵母基因敲除菌株 Y4036U (Δ Pex16)

[0413] 通过定量实时 PCR 在菌株 RHY25 至 RHY34 中对 Pex16 基因敲除进行进一步确认,使用耶氏酵母属翻译延伸因子 (tef-1) 基因 (GenBank 保藏号 AF054510) 作为对照物。

[0414] 首先,设计分别靶向于 Pex16 基因和 tef-1 基因的实时 PCR 引物和 TaqMan 探针,该设计使用 Primer Express 软件 v 2.0 (Applied Biosystems, Foster City, CA)。具体地讲,设计实时 PCR 引物 ef-324F (SEQ ID NO :64)、ef-392R (SEQ ID NO :65)、PEX16-741F (SEQ ID NO :66) 和 PEX16-802R (SEQ ID NO :67) 以及 TaqMan 探针 ef-345T (即,5' 6-FAMTM-TGCTGGTGGTGTGGTGAGTT-TAMRATM,其中核苷酸序列如 SEQ ID NO :68 所示) 和 PEX16-760T (即,5' -6FAMTM-CTGTCCATTCTGCGACCCCTC-TAMRATM,其中核苷酸序列如 SEQ ID NO :69 所示)。TaqMan 荧光探针 5' 末端具有结合的 6FAMTM 荧光报告基因染料,而 3' 末端包含 TAMRATM 淬灭剂。所有引物和探针获取自 Sigma-Genosys (Woodlands, TX)。

[0415] 敲除候选 DNA 通过在 50 μ L 水中悬浮 1 个菌落进行制备。tef-1 和 PEX16 的反应分别进行,每个样本重复三次。实时 PCR 反应包括 20pmoles 各个正向和反向引物 (即, ef-324F、ef-392R、PEX16-741F 和 PEX16-802R 5' , 同上)、5pmoles TaqMan 探针 (即, ef-345T 和 PEX16-760T)、10 μ L TaqMan Universal PCR Master Mix--No AmpErase[®] Uracil-N-Glycosylase (UNG) (目录号 PN 4326614, Applied Biosystems)、1 μ L 菌落悬浮液和 8.5 μ L 去 RNase/DNase 水,每个反应总体积为 20 μ L。反应在 ABI PRISM[®] 7900 Sequence Detection System 上进行,使用以下条件 :初始在 95°C 变性 10 分钟,随后进行 95°C 变性 15 秒、60°C 退火 1 分钟的循环 40 次。在每个循环期间通过监控 6-FAMTM 荧光自动收集实时数据。使用用于数据归一化的 tef-1 基因阈值循环 (C_T) 值,按照 ABI PRISM[®] 7900 Sequence Detection System 使用说明书进行数据分析。

[0416] 基于这一分析,得出结论:所有十个 Y4036U(Δ pex16) 菌落(即,RHY25 至 RHY34)是有效的 Pex16 基因敲除菌落,其中 pYRH13 构建体已经整合进染色体 Y1Pex16 中。

[0417] 评估用于 DGLA 生产的解脂耶氏酵母菌株 Y4036U 和 Y4036U(δ Pex16)

[0418] 为了评估 Pex16 基因敲除对总脂质级分中的 PUFA 百分比和细胞中的总脂质含量的效应,使 Y4036U 和 Y4036U(Δ pex16) 菌株在可比较的含油条件下生长。更具体地讲,菌株 Y4036U-17、Y4036U-19 和 Y4036U-33 具有在染色体随机位点整合的 Pex16 基因敲除片段,上述菌株被认为是 Pex16 野生型(即,Y4036U),而菌株 RHY25 至 RHY34 是 Pex16 突变型菌株(即,Y4036U(Δ pex16))。每个菌株在起始 OD₆₀₀为~0.1 的情况下,在包含 90mg/L L-亮氨酸的 25mL MM 中,置于 125mL 烧瓶中培养 48 小时。在 50mL 圆锥管中,以 4300rpm 离心 5 分钟收获细胞。弃去上清液,将细胞重悬于 25mL HGM 中并转移到新的 125mL 烧瓶中。细胞在 30℃下透气培养另外 120 小时。

[0419] 下表 13 中显示了每个菌株的脂肪酸组成(即,LA(18:2),ALA,EDA 和 DGLA);脂肪酸组成以占总脂肪酸的重量百分比(wt.%)的形式表示。菌株 Y4036U 和 Y4036U(Δ pex16)的平均脂肪酸组成用灰色突出显示并且用“Ave”指示。在 MM+L-亮氨酸培养基中无受试菌株提供足够的细胞群,因此不能分析总脂质含量。

[0420] 表 13

[0421] 解脂耶氏酵母菌株 Y4036U 和 Y4036U(Δ pex16) 中的脂质组成

[0422]

菌株	样本	18:2	ALA	EDA	DGLA
Y4036U	Y4036U-17	26.1	2.4	9.9	24.9
	Y4036U-19	29.4	1.6	9.9	18.1
	Y4036U-33	20.7	3.1	11.2	27.3
Y4036U	AVE	25.4	2.4	10.3	23.4
Y4036U (Δ pex16)	RHY25-1	14.9	5.1	5.5	44.1
	RHY25-2	14.3	5.0	5.4	42.6
	RHY26-1	14.4	5.1	5.6	42.9
	RHY26-2	13.8	4.9	5.9	44.6
	RHY27-1	14.4	5.0	5.4	42.6
	RHY27-2	15.1	4.9	5.6	44.2
	RHY28	15.3	4.6	5.7	42.6
	RHY29	15.4	4.8	5.5	43.9
	RHY30	15.5	4.5	5.9	43.6
	RHY31	15.5	4.7	5.8	43.9
	RHY32	15.5	4.9	5.9	44.4
	RHY33	15.9	4.7	5.9	41.8
	RHY34	15.9	4.9	6.2	43.5
Y4036U (Δ pex16)	AVE	15.1	4.9	5.7	43.4

[0423] 表 13 中的结果显示敲除 Y4036U (Δ pex16) 中的染色体 Pex16 基因与天然 Pex16p 未被敲除的菌株 Y4036U 中的 DGLA % TFA 相比较, DGLA % TFA 提高了大约 85 %。然而, Y4036U (Δ pex16) 积聚的 LA (18:2) 也降低了 ~ 40 %。

[0424] 因此, 表 13 中的结果显示与亲本菌株 Y4036 相比, Y4036 (Δ Pex16) 菌株具有更高的平均 DGLA % TFA (43.4 % 对 23.4 %)。此外, 菌株 Y4036U (Δ pex16) 的 DGLA 相对于总 PUFA 的量提高了 1.65 倍 (62.8 % 的 PUFA [% TFA] 对 38.1 % 的 PUFA [% TFA]), 而 C20PUFA 相对于总 PUFA 的量提高了 1.3 倍 (71 % 的 PUFA [% TFA] 对 54.8 % 的 PUFA [% TFA])。

[0425] 实施例 10

[0426] 产生菌株 Y4305 以生产占总脂质约 53.2 % 的 EPA

[0427] 解脂耶氏酵母菌株 Y4305U 具有 Ura⁻ 表型, 它在下文的实施例 11 中被用作宿主。菌株 Y4305 (Ura⁺ 菌株, 它是 Y4305U 的亲本) 来源于解脂耶氏酵母 ATCC#20362, 该菌株能通过 Δ 9 延伸酶 / Δ 8 去饱和酶途径的表达生产相对于总脂质而言约 53.2 % 的 EPA。

[0428] 菌株 Y4305U 的产生需要构建菌株 Y2224、菌株 Y4001、菌株 Y4001U、菌株 Y4036、菌株 Y4036U、菌株 Y4070 和菌株 Y4086 (如上文实施例 1 所述)。进一步开发菌株 Y4305U 需要构建菌株 Y4086U1、菌株 Y4128 和菌株 Y4128U3 (如上文实施例 2 所述)。随后产生菌株 Y4305U (如图 10 所示) 需要构建菌株 Y4217 (生产 42 % 的 EPA)、菌株 Y4217U2 (Ura⁻)、菌株

Y4259(生产 46.5%的 EPA)、菌株 Y4259U2(Ura⁻) 和菌株 Y4305(生产 53.2%的 EPA)。

[0429] 虽然本文未详细阐述关于转化和选择 EPA 生产菌株(在菌株 Y4128U3 后开发的菌株)的细节,但是在实施例 1 和 2 中描述了分离菌株 Y4217、菌株 Y4217U2、菌株 Y4259、菌株 Y4259U2、菌株 Y4305 和菌株 Y4305U 的方法。

[0430] 简而言之,生成构建体 pZKL2-5U89GC(图 8B;SEQ ID NO:55;描述于表 24,国际申请公布 W0 2008/073367)以将一个 $\Delta 9$ 延伸酶基因(即, EgD9eS)、一个 $\Delta 8$ 去饱和酶基因(即, EgD8M)、一个 $\Delta 5$ 去饱和酶基因(即, EgD5S)、和一个解脂耶氏酵母二酰基甘油胆碱磷酸转移酶(CPT1)基因整合进菌株 Y4128U3 的 Lip2 位点(GenBank 保藏号 AJ012632)中,从而产生较高含量的 EPA。六个分别命名为 Y4215、Y4216、Y4217、Y4218、Y4219 和 Y4220 的菌株分别生产占总脂质约 41.1%、41.8%、41.7%、41.1%、41%和 41.1%的 EPA。

[0431] 菌株 Y4217U1 和 Y4217U2 经由构建体 pZKUE3S 破坏菌株 Y4217 中的 Ura3 基因进行制备(图 5A SEQ ID NO:31;描述于表 22,国际申请公布 W0 2008/073367),它包含定向于 Ura3 基因的嵌合 EXP1::ME3S::Pex20 基因。利用构建体 pZKL1-2SP98C(图 8A SEQ ID NO:54;描述于表 23,国际申请公布 W0 2008/073367)将一个 $\Delta 9$ 延伸酶基因(即, EgD9eS)、一个 $\Delta 8$ 去饱和酶基因(即, EgD8M)、一个 $\Delta 12$ 去饱和酶基因(即, FmD12S)、和一个解脂耶氏酵母 CPT1 基因整合进菌株 Y4217U2 的 Lip1 位点(GenBank 保藏号 Z50020)中,从而产生分离菌株 Y4259、Y4260、Y4261、Y4262、Y4263 和 Y4264,分别生产占总脂质约 46.5%、44.5%、44.5%、44.8%、44.5%和 44.3%的 EPA。

[0432] 然后经由用构建体 pZKUM 进行转化制备 Ura⁻ 衍生物(即, 菌株 Y4259U2)(图 11A; SEQ ID NO:70;描述于表 33,国际申请公布 W02008/073367),该构建体将 Ura3 突变基因整合进菌株 Y4259 的 Ura3 基因中,从而分别产生分离的菌株 Y4259U1、Y4259U2 和 Y4259U3(统称为 Y4259U)(分别生产占总脂质 31.4%、31%和 31.3%的 EPA)。

[0433] 最后,生成构建体 pZKD2-5U89A2(图 11B;SEQ ID NO:71)以将一个 $\Delta 9$ 延伸酶基因、一个 $\Delta 5$ 去饱和酶基因、一个 $\Delta 8$ 去饱和酶基因、和一个 $\Delta 12$ 去饱和酶基因整合进菌株 Y4259U2 的二酰基甘油酰基转移酶(DGAT2)位点中,从而使 EPA 产量提高。pZKD2-5U89A2 质粒含有如下组分:

[0434] 表 14

[0435] 描述质粒 pZKD2-5U89A2(SEQ ID NO:71)

[0436]

SEQ ID NO:71 中的 RE 位点和核苷酸	片段和嵌合基因组分的描述
AscI/BsiWI (1-736)	耶氏酵母属 DGAT2 基因的 728bp 的 5' 部分(SEQ ID NO:72)(图中标记为“YLDGAT5”;美国专利 7,267,976)
PacI/SphI (4164-3444)	耶氏酵母属 DGAT2 基因的 714bp 的 3' 部分(SEQ ID NO:72)(图中标记为“YLDGAT3”;美国专利 7,267,976)

SwaI/BsiWI (13377-1)	YAT1::FmD12S::Lip2,它包含: ● YAT1:解脂耶氏酵母 YAT1 启动子 (在图中标记为“YAT”;专利申请公布 US 2006/0094102-A1); ● FmD12S:经密码子优化的 $\Delta 12$ 延伸酶 (SEQ ID NO:74),来源于串珠镰刀菌 (图中标记为“F. D12S”;国际申请公布 WO 2005/047485); ● Lip2:来自耶氏酵母属 Lip2 基因 (GenBank 保藏号 AJ012632) 的 Lip2 终止子序列
-------------------------	---

[0437]

PmeI/SwaI (10740-13377)	FBAIN::EgD8M::Lip1,它包含: ● FBAIN:解脂耶氏酵母 FBAIN 启动子 (美国专利 7,202,356); ● EgD8M:合成突变型 $\Delta 8$ 去饱和酶 (SEQ ID NO:76;专利申请公布 US 2008-0138868A1),来源于小眼虫 (“EgD8S”;美国专利 7,256,033); ● Lip1:来自耶氏酵母属 Lip1 基因 (GenBank 保藏号 Z50020) 的 Lip1 终止子序列
ClaI/PmeI (8846-10740)	YAT1::E389D9eS::OCT,它包含: ● YAT1:解脂耶氏酵母 YAT1 启动子 (在图中标记为“YAT”;专利申请公布 US 2006/0094102-A1); ● E389D9eS:经密码子优化的 $\Delta 9$ 延伸酶 (SEQ ID NO:78),来源于小型绿藻属 CCMP389(图中标记为“D9ES-389”;国际申请公布 WO 2007/061742); ● OCT:耶氏酵母属 OCT 基因 (GenBank 保藏号 X69988) 的 OCT 终止子序列
ClaI/EcoRI (8846-6777)	耶氏酵母属 Ura3 基因 (GenBank 保藏号 AJ306421)
EcoRI/PacI (6777-4164)	EXP1::EgD5S::ACO,它包含: ● EXP1:解脂耶氏酵母输出蛋白 (EXP1) 启动子 (在图中标记为“Exp”;国际申请公布 WO 2006/052870)。 ● EgD5S:经密码子优化的 $\Delta 5$ 去饱和酶 (SEQ ID NO:80),来源于小眼虫 (专利申请公布 US 2007-0292924-A1); ● Aco:耶氏酵母属 Aco 基因 (GenBank 保藏号 AJ001300) 的 Aco 终止子序列

[0438] 将 pZKD2-5U89A2 质粒用 AscI/SphI 消化,然后根据“一般方法”将其用于转化菌株 Y4259U2。将转化细胞铺在 MM 平板上并在 30℃ 下维持 3 至 4 天。将单菌落再划线接种到 MM 平板上,所得菌落用于接种液体 MM。液体培养物以 250rpm/min,在 30℃ 下摇动 2 天。通过离心收集细胞,将其重悬浮于 HGM 中,然后以 250rpm/min 摇动 5 天。通过离心收集细

胞,并提取脂质。通过酯交换制备 FAME 并随后用 Hewlett-Packard 6890GC 进行分析。

[0439] GC 分析显示大多数选择的 96 个菌株生产占总脂质 40-46% 的 EPA。命名为 Y4305、Y4306、Y4307 和 Y4308 的四个菌株分别生产占总脂质约 53.2%、46.4%、46.8% 和 47.8% 的 EPA。Y4305 的全部脂质分布如下:16:0(2.8%)、16:1(0.7%)、18:0(1.3%)、18:1(4.9%)、18:2(17.6%)、ALA(2.3%)、EDA(3.4%)、DGLA(2.0%)、ARA(0.6%)、ETA(1.7%) 和 EPA(53.2%)。总脂质%干细胞重量是 27.5。

[0440] 相对于野生型解脂耶氏酵母 ATCC#20362 的菌株 Y4305 的最终基因型是 SCP2-(YALI0E01298g)、YALI0C18711g-、Pex10-、YALI0F24167g-, unknown 1-、unknown 3-、unknown 8-、GPD::FmD12::Pex20、YAT1::FmD12::OCT、GPM/FBAIN::FmD12S::OCT、EXP1::FmD12S::Aco、YAT1::FmD12S::Lip2、YAT1::ME3S::Pex16、EXP1::ME3S::Pex20(3copies)、GPAT::EgD9e::Lip2、EXP1::EgD9eS::Lip1、FBAINm::EgD9eS::Lip2、FBA::EgD9eS::Pex20、GPD::EgD9eS::Lip2、YAT1::EgD9eS::Lip2、YAT1::E389D9eS::OCT、FBAINm::EgD8M::Pex20、FBAIN::EgD8M::Lip1(2 个拷贝)、EXP1::EgD8M::Pex16、GPDIN::EgD8M::Lip1、YAT1::EgD8M::Aco、FBAIN::EgD5::Aco、EXP1::EgD5S::Pex20、YAT1::EgD5S::Aco、EXP1::EgD5S::ACO、YAT1::RD5S::OCT、YAT1::PaD17S::Lip1、EXP1::PaD17::Pex16、FBAINm::PaD17::Aco、YAT1::Y1CPT1::ACO、GPD::Y1CPT1::ACO。

[0441] 为了破坏菌株 Y4305 中的 Ura3 基因,使用构建体 pZKUM(图 11A;SEQ ID NO:70;描述于表 33,国际申请公布 WO 2008/073367) 以将 Ura3 突变基因整合进菌株 Y4305 的 Ura3 基因中。挑取在 MM+5-FOA 平板上生长的总共 8 个转化体,并且再次分别划线接种到 MM 平板和 MM+5-FOA 平板上。所有 8 个菌株具有 Ura- 表型(即,细胞能在 MM+5-FOA 平板上生长,但是不能在 MM 平板上生长)。从 MM+5-FOA 平板上刮除细胞并提取脂质。通过酯交换制备 FAME 并随后用 Hewlett-Packard 6890GC 进行分析。

[0442] GC 分析显示在 MM+5-FOA 平板上生长的 pZKUM 转化体 #1、#6 和 #7 中存在占总脂质 37.6%、37.3% 和 36.5% 的 EPA。这三个菌株分别命名为菌株 Y4305U1、Y4305U2 和 Y4305U3(统称为 Y4305U)。为清楚起见,在实施例 11 中菌株 Y4305U 被称为菌株 Y4305U(Δ pex10)。

[0443] 实施例 11

[0444] 在解脂耶氏酵母菌株 Y4305U(Δ pex10) 中的 Pex16 染色体缺失进一步提高了积聚的 EPA 百分比(Pex10-Pex16 双敲除)

[0445] 本实施例描述使用构建体 pYRH13(图 9B;SEQ ID NO:59) 敲除在耶氏酵母属菌株 Y4305U 中的染色体 Pex16(Δ pex10)(实施例 10),从而导致生成 Pex10-Pex16 双突变体。测定并比较 Pex10-Pex16 双敲除对总油量和 EPA 含量的效应。具体地讲,在菌株 Y4305U(Δ pex10)(Δ pex16) 中的 Pex10-Pex16 双突变体的效应导致细胞中 EPA 的量与单突变体相比(即,菌株 Y4305U(Δ pex10)) 有所提高(EPA% TFA 和 EPA% DCW)。

[0446] 生成解脂耶氏酵母基因敲除菌株 Y4305U(Δ pex10)(Δ pex16)

[0447] 采用 Pex16 敲除构建体 pYRH13(实施例 9;SEQ ID NO:59) 的纯化 6.0kB AscI/SphI 片段,使用标准规程来转化解脂耶氏酵母菌株 Y4305U(Δ pex10)(实施例 10)。如实施例 9 所述,通过菌落 PCR 筛选并鉴定具有 Pex16 缺失的细胞。

[0448] 在 93 个筛选的菌落中,88 个菌落具有在染色体随机位点整合的 Pex16 敲除片段,因此不是 $\Delta pex16$ 突变体(然而,由于存在 pYRH13,细胞能够在 ura⁻ 平板上生长)。这些随机整合的菌株中的两个,命名为 Y4305U-22 和 Y4305U-25,用作脂质生产实验中的对照物(下文)。

[0449] 筛选的剩余 5 个菌落(即,总计 93 个菌落中的剩余菌落)包含 Pex16 基因敲除。将在 Y4305U 菌株中的这五个 $\Delta pex16$ 突变体命名为 RHY20、RHY21、RHY22、RHY23 和 RHY24。如实施例 9 所述,通过如实施例 9 所述的定量实时 PCR 进行对 Y1Pex16 基因敲除的进一步确认。

[0450] 评估用于 EPA 生产的解脂耶氏酵母菌株 Y4305U($\Delta Pex10$) 和 Y4305U($\Delta Pex10$)($\Delta pex16$)

[0451] 为了评估多个 Pex 基因突变对总脂质级分中的 PUFA 百分比和细胞中的总脂质含量的效应,将 Y4305U($\Delta pex10$) 和 Y4305U($\Delta pex10$)($\Delta pex16$) 菌株在可比较的含油条件下生长。更具体地讲,菌株 Y4305U-22 和 Y4305U-25 具有在染色体随机位点整合的 Pex16 基因敲除片段,它们被认为是 Pex16 野生型, Pex10 基因敲除型(即, Y4305U($\Delta pex10$))。菌株 RHY22、RHY23 和 RHY24 是双敲除突变型菌株(即, Y4305U($\Delta pex10$)($\Delta pex16$))。在可比较的含油条件下平行培养每个菌株。

[0452] 具体地讲,在 25mL 合成右旋糖培养基(SD)中在起始 OD₆₀₀为~0.1 的情况下,在 125mL 烧瓶中培养 48 小时。在 50mL 圆锥管中,以 4300rpm 离心 5 分钟收获细胞。弃去上清液,将细胞重悬于 25mL HGM 中并转移到新的 125mL 烧瓶中。细胞在 30℃ 下透气培养另外 120 小时。

[0453] 为了测定干细胞重量(DCW),对来自 5mL HGM 生长培养物的细胞进行处理。培养的细胞在 4300rpm 下离心 5 分钟。使用 10mL 无菌水重悬沉淀物并在相同条件下二次离心。然后用 1mL 无菌 H₂O(第三次)重悬沉淀物,将其转移到预先称重的铝盘中。细胞悬浮液在 80℃ 真空炉中干燥过夜。测量细胞重量。

[0454] 为了测定总脂质含量,1mL 的 HGM 培养细胞通过在 13,000rpm 下离心 1 分钟进行收集,提取总脂质并通过酯交换制备 FAME,随后用 Hewlett-Packard 6890GC(一般方法)进行分析。

[0455] 表 15 显示了每个菌株的脂肪酸组成(即,16:0(棕榈酸))、16:1(棕榈油酸)、18:0、18:1(油酸)、18:2(LA)、18:3(ALA)、EDA、DGLA、ARA、ETrA、ETA 和 EPA)(以总脂肪酸(TFA)的重量百分比(wt. %)形式表示),以及 DCW(g/L)和总脂质含量(TFA% DCW)。菌株 Y4305U($\Delta pex10$) 和 Y4305U($\Delta pex10$)($\Delta pex16$) 的平均脂肪酸组成用灰色突出显示并用“Ave”指示。

[0456]

表 15
在解脂耶氏酵母菌株 Y4305U (Δ Pex10) 和 Y4305U (Δ Pex10) (Δ Pex16) 中的脂质组成

菌株	样本	DCW (g/L)	TFA %DCW	16:0	16:1	18:0	18:1	18:2	18:3	EDA	DGLA	ARA	ETra	ETA	EPA
Y4305U (Δ pex10)	Y4305U-22 #1	3.50	29	3.1	0.7	2.1	6.4	18.7	2.6	4.2	1.8	0.5	1.8	2.0	45.4
	Y4305U-22 #2	3.94	29	3.0	0.7	2.1	6.2	18.5	2.5	4.5	1.8	0.5	1.8	2.0	45.6
	Y4305U-25 #1	4.14	31	3.6	1.1	1.8	6.1	18.8	2.4	4.5	1.9	0.6	1.6	2.0	43.9
	Y4305U-25 #2	4.12	30	3.6	1.1	1.8	6.1	18.7	2.4	4.6	1.9	0.6	1.6	2.0	44.0
Y4305U (Δ pex10)	AVE	3.93	30	3.3	0.9	2.0	6.2	18.7	2.5	4.5	1.9	0.6	1.7	2.0	44.7
Y4305U (Δ pex10) (Δ pex16)	RHY22 #1	4.04	29	2.7	0.7	1.5	5.4	18.5	2.7	3.4	1.9	0.5	1.4	2.0	48.5
	RHY22 #2	3.82	32	2.7	0.6	1.5	5.5	18.4	3.0	3.0	2.0	0.5	1.5	2.0	48.8
	RHY23 #1	4.66	30	2.7	0.7	1.5	5.4	18.6	2.7	3.5	2.0	0.6	1.4	2.0	48.2
	RHY23 #2	4.18	30	2.7	0.7	1.5	5.4	18.4	2.6	3.5	1.9	0.6	1.4	2.0	48.5
	RHY24 #1	4.34	30	2.8	0.8	1.5	5.5	18.6	2.6	3.6	1.9	0.6	1.4	2.0	47.9
	RHY24 #2	4.58	30	2.7	0.7	1.5	5.6	18.8	2.6	3.6	2.0	0.6	1.4	2.0	47.8
Y4305U (Δ pex10) (Δ pex16)	AVE	4.27	30	2.7	0.7	1.5	5.5	18.6	2.7	3.4	2.0	0.6	1.4	2.0	48.3

[0457] 表 15 中的结果显示敲除 Y4305U (Δ pex10) (Δ pex16) 中的染色体 Pex16 基因与天然 Pex16p 未被敲除的菌株 Y4305U (Δ pex10) 中的 EPA% TFA 相比较, EPA% TFA 提高了大约 8%。此外, 与单突变菌株相比, EPA% DCW 也在双突变体中提高, 而 TFA% DCW 保持不变。

[0458] 因此,表 15 中的结果显示与对照 Y4305(Δ Pex10) 菌株相比, Y4305(Δ Pex10, Δ Pex16) 菌株平均具有更高的 EPA% TFA(48.3%对 44.7%) 和更高的 EPA% DCW(14.57%对 13.23%)。菌株 Y4305(Δ Pex10, Δ Pex16) 相对于菌株 Y4305(Δ Pex10) 的 EPA 相对于总 PUFA 的量(61%的 PUFA[% TFA] 对 58.3%的 PUFA[% TFA]) 仅提高了 1.05 倍,而 C20PUFA 相对于总 PUFA 的量的提高二者基本上相同(73%的 PUFA[% TFA] 对 72%的 PUFA[% TFA])。

[0459] 实施例 12

[0460] 在解脂耶氏酵母菌株 Y4036U 中的 Pex3 染色体缺失提高了积聚的 DGLA 百分比

[0461] 本实施例描述了使用构建体 pY157(图 12B;SEQ ID NO:82) 敲除在 Ura⁻、DGLA⁻ 生产耶氏酵母属菌株 Y4036U(实施例 1) 中的染色体 Pex3 基因(SEQ ID NO:3)。用 Pex3 敲除构建体转化解脂耶氏酵母菌株 Y4036U 导致生成菌株 Y4036(Dpex3)。测量 Pex3 敲除对 DGLA 含量的效应并与对照菌株 Y4036(Ura⁺ 菌株,它是菌株 Y4036U 的亲本) 比较。具体地讲,敲除 Pex3 提高 DGLA 占总脂肪酸的百分比并且与对照相比改善了大约 3 倍的 DGLA% DCW。

[0462] 构建体 pY157

[0463] 质粒 pY87(图 12A) 包含表达盒以敲除解脂耶氏酵母二酰基甘油酰基转移酶(DGAT2) 基因,如下表 16 所述:

[0464] 表 16

[0465] 质粒 pY87(SEQ ID NO:83) 的描述

[0466]

SEQ ID NO:83 中的 RE 位点和核苷酸	片段和嵌合基因组分的描述
SphI/PacI (1-721)	耶氏酵母属 DGAT2 基因的 5' 部分(SEQ ID NO:72 的碱基 1-720)(美国专利公开 7,267,976)
PacI/BglIII (721-2459)	LoxP::Ura3::LoxP,它包含: ● LoxP 序列(SEQ ID NO:84); ● 耶氏酵母属 Ura3 基因(GenBank 保藏号 AJ306421); ● LoxP 序列(SEQ ID NO:84)
BglIII/AscI (2459-3203)	耶氏酵母属 DGAT2 基因的 3' 部分(SEQ ID NO:72 的碱基 2468-3202)(美国专利公开 7,267,976)
AscI/SphI(3203-5910)	载体主链包括: ● ColE1 质粒复制起点 ● 用于在大肠杆菌中选择的氨苄青霉素抗性基因(Amp ^R)(4191-5051); ● 大肠杆菌 f1 复制起点

[0467] 质粒 pY157 来源于质粒 pY87。具体地讲,解脂耶氏酵母 Pex3 基因的 704bp 的 5' 启动子区域置换 pY87 的 SphI/PacI 片段,解脂耶氏酵母 Pex3 基因的 448bp 的 3' 终止子区域置换 pY87 的 BglIII/AscI 片段以生成 pYR157(SEQ ID NO:82;图 12B)。

[0468] 生成解脂耶氏酵母基因敲除菌株 Y4036(Δ Pex3)

[0469] 使用标准规程,用 Pex3 敲除构建体 pY157 的纯化 3648bp 的 AscI/SphI 片段来转化解脂耶氏酵母的菌株 Y4036U(实施例 1)(同上)。

[0470] 为了筛选具有 Pex3 缺失的细胞,使用 Taq 聚合酶 (Invitrogen ;Carlsbad, CA) 和 PCR 引物 UP 768(SEQ ID NO :85) 以及 LP 769(SEQ ID NO :86) 进行菌落 PCR。设计这组引物用于扩增完整 Pex3 基因的 2039bp 的野生型条带,以及当 Pex3 基因被定向敲除破坏时,3719bp 的敲除特异性条带。

[0471] 更具体地讲,菌落 PCR 使用 MasterAmp Taq 试剂盒 (EpicentreTechnologies, Madison, WI ;目录号 82250),并且按照制造商说明书进行,在 25 μ L 的反应中包含 :2.5 μ L 10X MasterAmp Taq 缓冲液、2.0 μ L 25mM MgCl₂、7.5 μ L 10X MasterAmp Enhancer、2.5 μ L 2.5mM dNTP(TaKaRa Bio Inc., Otsu Shiga, Japan)、1.0 μ L 10 μ M 上游引物、1.0 μ L 10 μ M 下游引物、0.25 μ L MasterAmp Taq DNA 聚合酶和 19.75 μ L 水。扩增如下进行 :95 $^{\circ}$ C 5 分钟进行初始变性,随后在 95 $^{\circ}$ C 30 秒进行 40 个循环的变性,在 56 $^{\circ}$ C 下退火 60 秒,在 72 $^{\circ}$ C 下延伸 4 分钟。最后在 72 $^{\circ}$ C 下进行 10 分钟的最终延伸循环,然后在 4 $^{\circ}$ C 下终止反应。

[0472] 筛选的 48 个菌落中,46 个菌落具有 2039bp 的预期条带,该条带来自野生型 (即,未破坏的)Pex3 基因,因此不是 Δ pex3 突变体。剩余的 2 个菌落只显示 2039bp 的弱条带,说明它们是在背景中存在一些污染的非转化细胞的 Δ pex3 突变体。通过划线接种 2 个推定敲除菌落到选择性平板上分离单菌落,这一点得到了证实。然后从每个推定基因敲除菌株的 3 个单菌落中来分离基因组 DNA 并用相同引物对来进行筛选。即,UP 768 和 LP 769(SEQ ID NO :85 和 86)。该方法被认为是比菌落 PCR 更灵敏的方法。来自主要转化体的所有三个单菌落都缺少 2039bp 的野生型条带,相反地具有 3719bp 的基因敲除特异性条带。将在 Y4036U 菌株中的 Δ pex3 突变体命名为 L134 和 L135。

[0473] 评估用于 DGLA 生产的解脂耶氏酵母菌株 Y4036 和 Y4036 (Δ Pex3)

[0474] 为了评估 Pex3 基因敲除对总脂质级分中的 PUFA 百分比和细胞中的总脂质含量的效应,Y4036 和 Y4036 (Δ pex3) 菌株 6 在可比较的含油条件下生长。将菌株 Y4036、L134 (即, Y4036 (Δ pex3)) 和 L135 (即, Y4036 (Δ pex3)) 接种到 25mL 的 CSM-Ura 并在振荡器中在 30 $^{\circ}$ C 下培养过夜。将预培养物等分到新的 25mL CSM-Ura 烧瓶中,最终 OD₆₀₀ 为 0.4。培养物在 30 $^{\circ}$ C 下在振荡器中生长。在 48 小时后,离心细胞 (它几乎不生长) 并重悬在新的 25mL CSM-Ura 中并继续生长 72 小时。将细胞离心、重悬于 25mL HGM 中,并继续如上文所述生长 72 小时。离心收获细胞,用蒸馏水洗涤一次并重悬于 25mL 水中,终体积为 20.5mL。使用等分试样 (1.5mL) 测量脂质含量,提取总脂质,通过酯交换制备 FAME 并通过 Hewlett-Packard 6890GC 进行分析 (一般方法)。如实施例 11 所述干燥剩余的等分试样以测量干细胞重量 (DCW)。

[0475] 表 17 显示了每个菌株的脂肪酸组成 (即,16:0 (棕榈酸)、16:1 (棕榈油酸)、18:0、18:1 (油酸)、18:2 (LA)、EDA 和 DGLA) (以占总脂肪酸 (TFA) 的重量百分比 (wt. %) 形式表示),以及总脂质含量 (TFA% DCW)。转化效率 (“CE”) 根据下面的公式测量 :([产物]/[底物+产物])*100,其中 ‘产物’ 包括中间产物和该途径中来源于它的所有产物。因此, Δ 12 去饱和酶转化效率 (Δ 12% CE) 计算如下 :([LA+EDA+DGLA]/[18:1+LA+EDA+DGLA])*100 ; Δ 9 延伸酶转化效率 (Δ 9e10% CE) 计算如下 :([EDA+DGLA]/[LA+EDA+DGLA])*100 ;并且 Δ 8 去饱和酶转化效率 (Δ 8% CE) 计算如下 :([DGLA]/[EDA+DGLA])*100。菌株 Y4036、L134 和

L135 的平均脂肪酸组成用灰色突出显示并用“Ave”表示,而“S. D. ”指示标准偏差。正如预期的那样, Δ pex3 菌株不在用油酸盐作唯一碳源的平板上生长。

[0476]

表 17

在解脂耶氏酵母菌株 Y4036 和 Y4036 (Δ Pex3) 中的脂质含量和组成

菌株	样本	TFA %DCW	16:0	16:1	18:0	18:1	18:2	EDA	DGLA	Δ 12 %CE	Δ 9 elo %CE	Δ 8 %CE
Y4036	Y4036-1	6.1	10	7	1	14	29	9	19	80	49	69
	Y4036-2	3.7	11	6	1	14	30	8	20	81	48	70
	Y4036-3	4.1	11	5	1	15	31	8	19	80	47	70
	Avg.	4.7	10	6	1	14	30	8	19	80	48	70
	S.D.	1.3	0.3	0.9	0.1	0.3	0.7	0.3	0.2	0.3	0.9	0.8
Y4036 (Δ pex3)	L134-1	6.2	7	5	1	8	12	10	45	89	83	81
	L134-2	5.4	7	5	1	8	11	10	47	90	83	82
	L134-3	6.7	6	5	1	8	12	11	46	90	83	82
	Avg.	6.1	7	5	1	8	12	10	46	90	83	82
	S.D.	0.6	0.5	0.1	0.1	0.3	0.1	0.1	1.0	0.5	0.4	0.3
Y4036 (Δ pex3)	L135-1	4.2	7	5	1	8	12	11	45	89	82	81
	L135-2	6.5	6	5	1	8	12	10	47	90	83	82
	L135-3	7.1	7	5	1	8	12	10	46	90	83	82
	Avg.	5.9	7	5	1	8	12	10	46	90	83	81
	S.D.	1.6	0.6	0.1	0.1	0.3	0.1	0.1	1.1	0.5	0.4	0.5

[0477] 表 17 中的结果显示敲除 Y4036 (Δ pex3) 中的染色体 Pex3 基因与天然 Pex3p 未被敲除的菌株 Y4036 中的 DGLA% TFA 相比较, DGLA% TFA 提高了大约 142%。具体地讲, Pex3 敲

除提高了 DGLA 含量,从在 Y4036 中的大约 19%提高到在 Y4036 (Δ pex3) 菌株 L134 和 L135 中的 46%。此外, Δ 9 延伸酶转化效率百分比从 Y4036 中的大约 48%提高到 Y4036 (Δ pex3) 菌株 L134 和 L135 中的 83%;并且菌株 L134 和 L135 中的 TFA% DCW 从 4.7%提高到 6%。LA% TFA 从 30%下降到 12%。Pex3 缺失实际上提高了脂肪酸的通量并从而提高了用于 Δ 9 延伸的底物可利用性。

[0478] 因此,表 17 中的结果显示与亲本菌株 Y4036 相比, Y4036 (Δ Pex3) 菌株具有平均更高的脂质含量 (TFA% DCW) (大约 6.0%对 4.7%),更高的 DGLA% TFA (46%对 19%),以及更高的 DGLA% DCW (大约 2.8%对 0.9%)。此外,菌株 Y4036 (Δ Pex3) 的 DGLA 相对于总 PUFA 的量提高了 2 倍 (67.7%的 PUFA[% TFA] 对 33.3%的 PUFA[% TFA]),而 C20PUFA 相对于总 PUFA 的量提高了 1.7 倍 (82%的 PUFA[% TFA] 对 47%的 PUFA[% TFA])。

[0479] 预测改善的 DGLA 产量也将导致在经基因工程化用于 EPA 生产的解脂耶氏酵母菌株中的 EPA 产量的改善 (例如实施例 10 所述并从中衍生出的解脂耶氏酵母菌株 Y4305U)。

序列表

<110>E. I. du Pont de Nemours & Co., Inc.

Zhu, Quinn

Xue, Zhixiong

<120> 破坏过氧化物酶体生物合成因子蛋白 (PEX) 以改变含油真核生物中多不饱和脂肪酸和总脂质含量

<130>CL3847

<160>89

<170>PatentIn 版本 3.4

<210>1

<211>1024

<212>PRT

<213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG82178)

<220>

<221>MISC_FEATURE

<222>(1).. (1024)

<223>Y1Pex1p

<400>1

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

Glu Asp Phe Asp Gly His Leu Arg Leu Glu Val Ser Leu Asp Val Ser		
165	170	175
Gln Leu Pro Pro Glu Phe His Gln Leu Lys Asn Val Ser Ile Lys Val		
180	185	190
Ile Thr Pro Pro Asn Leu Ala Ser Pro Gln Gln Ala Ala Ser Ile Ala		
195	200	205
Val Glu Glu Lys Ser Glu Glu Ser Leu Ser Gln Asn Lys Pro Pro Ser		
210	215	220
Ser Glu Pro Lys Val Glu Val Pro Pro Asp Ile Ile Asn Pro Ala Ser		
225	230	235
Glu Ile Val Ala Thr Leu Val Asn Asp Thr Thr Ser Pro Thr Gly His		
245	250	255
Ala Lys Leu Ser Tyr Ala Leu Ala Asp Ala Leu Gly Ile Pro Ser Ser		
260	265	270
Val Gly His Val Ile Arg Phe Glu Ser Ala Ser Lys Pro Leu Ser Gln		
275	280	285
Lys Pro Gly Ala Leu Val Ile His Arg Phe Ile Thr Lys Thr Val Gly		
290	295	300
Ala Ala Glu Gln Lys Ser Leu Arg Leu Lys Gly Glu Lys Asn Ala Asp		
305	310	315
Asp Gly Val Ser Ala Asp Asp Gln Phe Ser Leu Leu Glu Glu Leu Lys		
325	330	335
Lys Leu Gln Mer Leu Glu Gly Pro Ile Thr Asn Phe Gln Arg Leu Pro		
340	345	350
Pro Ile Pro Glu Leu Leu Pro Leu Gly Gly Val Ile Gly Leu Gln Asn		
355	360	365
Ser Glu Gly Trp Ile Gln Gly Gly Tyr Leu Gly Glu Glu Pro Ile Pro		
370	375	380
Phe Val Ser Gly Ser Glu Ile Leu Arg Ser Glu Ser Ser Leu Ser Pro		
385	390	395
Ser Asn Ile Glu Ser Glu Asp Lys Arg Val Val Gly Leu Asp Asn Met		
405	410	415
Leu Asn Lys Ile Asn Glu Val Leu Ser Arg Asp Ser Ile Gly Cys Leu		
420	425	430
Val Tyr Gly Ser Arg Gly Ser Gly Lys Ser Ala Val Leu Asn His Ile		
435	440	445
Lys Lys Glu Cys Lys Val Ser His Thr His Thr Val Ser Ile Ala Cys		
450	455	460
Gly Leu Ile Ala Gln Asp Arg Val Gln Ala Val Arg Glu Ile Leu Thr		

465	470	475	480
Lys Ala Phe Leu Glu Ala Ser Trp Phe Ser Pro Ser Val Leu Phe Leu			
	485	490	495
Asp Asp Ile Asp Ala Leu Met Pro Ala Glu Val Glu His Ala Asp Ser			
	500	505	510
Ser Arg Thr Arg Gln Leu Thr Gln Leu Phe Leu Glu Leu Ala Leu Pro			
	515	520	525
Ile Met Lys Ser Arg His Val Ser Val Val Ala Ser Ala Gln Ala Lys			
	530	535	540
Glu Ser Leu His Met Asn Leu Val Thr Gly His Val Phe Glu Glu Leu			
545	550	555	560
Phe His Leu Lys Ser Pro Asp Lys Glu Ala Arg Leu Ala Ile Leu Ser			
	565	570	575
Glu Ala Val Lys Leu Met Asp Gln Asn Val Ser Phe Ser Gln Asn Asp			
	580	585	590
Val Leu Glu Ile Ala Ser Gln Val Asp Gly Tyr Leu Pro Gly Asp Leu			
	595	600	605
Trp Thr Leu Ser Glu Arg Ala Gln His Glu Met Ala Leu Arg Gln Ile			
	610	615	620
Glu Ile Gly Leu Glu Asn Pro Ser Ile Gln Leu Ala Asp Phe Met Lys			
625	630	635	640
Ala Leu Glu Asp Phe Val Pro Ser Ser Leu Arg Gly Val Lys Leu Gln			
	645	650	655
Lys Ser Asn Val Lys Trp Asn Asp Ile Gly Gly Leu Lys Glu Thr Lys			
	660	665	670
Ala Val Leu Leu Glu Thr Leu Glu Trp Pro Thr Lys Tyr Ala Pro Ile			
	675	680	685
Phe Ala Ser Cys Pro Leu Arg Leu Arg Ser Gly Leu Leu Leu Tyr Gly			
	690	695	700
Tyr Pro Gly Cys Gly Lys Thr Tyr Leu Ala Ser Ala Val Ala Ala Gln			
705	710	715	720
Cys Gly Leu Asn Phe Ile Ser Ile Lys Gly Pro Glu Ile Leu Asn Lys			
	725	730	735
Tyr Ile Gly Ala Ser Glu Gln Ser Val Arg Glu Leu Phe Glu Arg Ala			
	740	745	750
Gln Ala Ala Lys Pro Cys Ile Leu Phe Phe Asp Glu Phe Asp Ser Ile			
	755	760	765
Ala Pro Lys Arg Gly His Asp Ser Thr Gly Val Thr Asp Arg Val Val			
	770	775	780

```

Asn Gln Met Leu Thr Gln Met Asp Gly Ala Glu Gly Leu Asp Gly Val
785                      790                      795                      800
Tyr Val Leu Ala Ala Thr Ser Arg Pro Asp Leu Ile Asp Pro Ala Leu
                        805                      810                      815
Leu Arg Pro Gly Arg Leu Asp Lys Met Leu Ile Cys Asp Leu Pro Ser
                        820                      825                      830
Tyr Glu Asp Arg Leu Asp Ile Leu Arg Ala Ile Val Asp Gly Lys Met
                        835                      840                      845
His Leu Asp Gly Glu Val Glu Leu Glu Tyr Val Ala Ser Arg Thr Asp
                        850                      855                      860
Gly Phe Ser Gly Ala Asp Leu Gln Ala Val Met Phe Asn Ala Tyr Leu
865                      870                      875                      880
Glu Ala Ile His Glu Val Val Asp Val Ala Asp Asp Thr Ala Ala Asp
                        885                      890                      895
Thr Pro Ala Leu Glu Asp Lys Arg Leu Glu Phe Phe Gln Thr Thr Leu
                        900                      905                      910
Gly Asp Ala Lys Lys Asp Pro Ala Ala Val Gln Asn Glu Val Met Asn
                        915                      920                      925
Ala Arg Ala Ala Val Ala Glu Lys Ala Arg Val Thr Ala Lys Leu Glu
                        930                      935                      940
Ala Leu Phe Lys Gly Met Ser Val Gly Val Asp Asn Asp Asp Asp Lys
945                      950                      955                      960
Pro Arg Lys Lys Ala Val Val Val Ile Lys Pro Gln His Met Asn Lys
                        965                      970                      975
Ser Leu Asp Glu Thr Ser Pro Ser Ile Ser Lys Lys Glu Leu Leu Lys
                        980                      985                      990
Leu Lys Gly Ile Tyr Ser Gln Phe Val Ser Gly Arg Ser Gly Asp Met
                        995                      1000                      1005
Pro Pro Gly Thr Ala Ser Thr Asp ValGly Gly Arg Ala Thr Leu
                        1010                      1015                      1020

```

Ala

<210>2

<211>381

<212>PRT

<213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG77647)

<220>

<221>MISC_FEATURE

<222>(1).. (381)

<223>Y1Pex2p

<400>2

```

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

```

Cys Leu Val Thr Lys Leu Ala Gln Gly Asp Gly Asp Gly Trp Asn Cys
 305 310 315 320
 Tyr Arg Cys Ala Lys Gln Val Gln Lys Met Lys Pro Trp Val Asp Val
 325 330 335
 Asp Glu Ala Ala Val Val Gly Ala Ala Glu Met His Glu Lys Val Asp
 340 345 350
 Val Ile Glu His Ala Glu Asp Asn Glu Gln Glu Glu Glu Glu Phe Asp
 355 360 365
 Asp Asp Asp Glu Asp Ser Asn Phe Gln Leu Met Lys Asp
 370 375 380

<210>3

<211>431

<212>PRT

<213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG78565)

<220>

<221>MISC_FEATURE

<222>(1)..(431)

<223>Y1Pex3p

<400>3

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

Gln Leu Asn Ile Leu Gly Arg Arg Asn Tyr Val Asn Ser Val Val Ala		
165	170	175
Leu Ala Gln Gln Gly Arg Glu Gly Asn Ala Glu Gly Arg Val Ala Pro		
180	185	190
Ser Phe Gly Asp Leu Ala Asp Met Gly Tyr Phe Gly Asp Leu Ser Gly		
195	200	205
Ser Ser Ser Phe Gly Glu Thr Ile Val Asp Pro Asp Leu Asp Glu Gln		
210	215	220
Tyr Leu Thr Phe Ser Trp Trp Leu Leu Asn Glu Gly Trp Val Ser Leu		
225	230	235
Ser Glu Arg Val Glu Glu Ala Val Arg Arg Val Trp Asp Pro Val Ser		
245	250	255
Pro Lys Ala Glu Leu Gly Phe Asp Glu Leu Ser Glu Leu Ile Gly Arg		
260	265	270
Thr Gln Met Leu Ile Asp Arg Pro Leu Asn Pro Ser Ser Pro Leu Asn		
275	280	285
Phe Leu Ser Gln Leu Leu Pro Pro Arg Glu Gln Glu Glu Tyr Val Leu		
290	295	300
Ala Gln Asn Pro Ser Asp Thr Ala Ala Pro Ile Val Gly Pro Thr Leu		
305	310	315
Arg Arg Leu Leu Asp Glu Thr Ala Asp Phe Ile Glu Ser Pro Asn Ala		
325	330	335
Ala Glu Val Ile Glu Arg Leu Val His Ser Gly Leu Ser Val Phe Met		
340	345	350
Asp Lys Leu Ala Val Thr Phe Gly Ala Thr Pro Ala Asp Ser Gly Ser		
355	360	365
Pro Tyr Pro Val Val Leu Pro Thr Ala Lys Val Lys Leu Pro Ser Ile		
70	375	380
Leu Ala Asn Met Ala Arg Gln Ala Gly Gly Met Ala Gln Gly Ser Pro		
385	390	395
Gly Val Glu Asn Glu Tyr Ile Asp Val Met Asn Gln Val Gln Glu Leu		
405	410	415
Thr Ser Phe Ser Ala Val Val Tyr Ser Ser Phe Asp Trp Ala Leu		
420	425	430

<210>4
 <211>395
 <212>PRT
 <213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG83356)
 <220>

<221>MISC_FEATURE

<222>(1).. (395)

<223>YlPex3Bp

<400>4

Met	Leu	Gln	Ser	Leu	Asn	Arg	Asn	Lys	Lys	Arg	Leu	Ala	Val	Ser	Thr
1				5				10						15	
Gly	Leu	Ile	Ala	Val	Ala	Tyr	Val	Val	Ile	Ser	Tyr	Thr	Thr	Lys	Arg
			20					25						30	
Leu	Ile	Glu	Lys	Gln	Glu	Gln	Lys	Leu	Glu	Glu	Glu	Arg	Ala	Lys	Glu
			35					40						45	
Arg	Leu	Lys	Gln	Leu	Phe	Ala	Gln	Thr	Gln	Asn	Glu	Ala	Ala	Phe	His
			50					55						60	
Thr	Ala	Ser	Val	Leu	Pro	Gln	Leu	Cys	Glu	Gln	Ile	Met	Glu	Phe	Val
65						70					75				80
Ala	Val	Glu	Lys	Ile	Ala	Glu	Gln	Leu	Gln	Asn	Met	Arg	Ala	Glu	Lys
						85					90				95
Arg	Lys	Lys	Gln	Asn	Met	Asp	Asp	Asp	Lys	His	Ser	Val	Leu	Ser	Leu
						100								110	
Gly	Thr	Glu	Thr	Thr	Ala	Ser	Met	Ala	Asp	Gly	Gln	Lys	Met	Ser	Lys
						115								125	
Ile	Gln	Leu	Trp	Asp	Glu	Leu	Lys	Ile	Glu	Ser	Leu	Thr	Arg	Ile	Val
						130								140	
Thr	Leu	Ile	Tyr	Cys	Val	Ser	Leu	Leu	Asn	Tyr	Leu	Ile	Arg	Leu	Gln
145						150								160	
Thr	Asn	Ile	Val	Gly	Arg	Lys	Arg	Tyr	Gln	Asn	Glu	Ala	Gly	Pro	Ala
						165								175	
Gly	Ala	Thr	Tyr	Asp	Met	Ser	Leu	Glu	Gln	Cys	Tyr	Thr	Trp	Leu	Leu
						180								190	
Thr	Arg	Gly	Trp	Lys	Ser	Val	Val	Asp	Asn	Val	Arg	Arg	Ser	Val	Gln
						195								205	
Gln	Val	Phe	Thr	Gly	Val	Asn	Pro	Arg	Gln	Asn	Leu	Ser	Leu	Asp	Glu
						210								220	
Phe	Ala	Thr	Leu	Leu	Lys	Arg	Val	Gln	Thr	Leu	Val	Asn	Ser	Pro	Pro
225						230								240	
Tyr	Ser	Thr	Thr	Pro	Asn	Thr	Phe	Leu	Thr	Ser	Leu	Leu	Pro	Pro	Arg
						245								255	
Glu	Leu	Glu	Gln	Leu	Arg	Leu	Glu	Lys	Glu	Lys	Gln	Ser	Leu	Ser	Pro
						260								270	
Asn	Tyr	Thr	Tyr	Gly	Ser	Pro	Leu	Lys	Asp	Leu	Val	Phe	Glu	Ser	Ala

275	280	285
Gln His Ile Gln Ser Pro	Gln Gly Met Ser Ser	Phe Arg Ala Ile Ile
290	295	300
Asp Gln Ser Phe Lys Val	Phe Leu Glu Lys Val	Asn Glu Ser Gln Tyr
305	310	315
Val Asn Pro Pro Ser Thr	Gly Gly Lys Arg Ile	Ala Val Gly Ala Leu
325	330	335
Gln Pro Pro Ile Ile Ser	Gly Gly Pro Lys Lys	Val Lys Leu Ala Ser
340	345	350
Leu Leu Ser Val Ala Thr	Arg Gln Ser Ser Val	Ile Ser His Ala Gln
355	360	365
Pro Asn Pro Tyr Val Asp	Ala Ile Asn Ser Val	Ala Glu Tyr Asn Gly
370	375	380
Leu Cys Ala Val Ile Tyr	Ser Ser Phe Glu Gln	
385	390	395
<210>5		
<211>153		
<212>PRT		
<213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG79130)		
<220>		
<221>MISC_FEATURE		
<222>(1).. (153)		
<223>Y1Pex4p		
<400>5		
Met Ala Ser Gln Lys Arg	Leu Ile Lys Glu Leu	Ala Ala Tyr Lys Lys
1	5	10
Asp Pro Asn Pro Cys Leu	Ala Ser Leu Thr	Ala Asp Gly Asp Ser Leu
20	25	30
Tyr Lys Trp Thr Ala Val	Met Arg Gly Thr	Glu Gly Thr Ala Tyr Glu
35	40	45
Asn Gly Leu Trp Gln Val	Glu Ile Asn Ile	Pro Glu Asn Tyr Pro Leu
50	55	60
Gln Pro Pro Thr Met Phe	Phe Arg Thr Lys	Ile Cys His Pro Asn Ile
65	70	75
His Phe Glu Thr Gly Glu	Val Cys Ile Asp	Val Leu Lys Thr Gln Trp
85	90	95
Ser Pro Ala Trp Thr Ile	Ser Ser Ala Cys	Thr Ala Val Ser Ala Met
100	105	110
Leu Ser Leu Pro Glu Pro	Asp Ser Pro Leu	Asn Ile Asp Ala Ala Asn

115	120	125
Leu Val Arg Cys Gly Asp Glu	Ser Ala Met Glu Gly Leu Val Arg Tyr	
130	135	140
Tyr Val Asn Lys Tyr Ala Ser Gly Asn		
145	150	
<210>6		
<211>598		
<212>PRT		
<213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG78803)		
<220>		
<221>MISC_FEATURE		
<222>(1).. (598)		
<223>Y1Pex5p		
<400>6		
Met Ser Phe Met Arg Gly Gly Ser Glu Cys Ser Thr Gly Arg Asn Pro		
1	5	10
Leu Ser Gln Phe Thr Lys His Thr Ala Glu Asp Arg Ser Leu Gln His		
20	25	30
Asp Arg Val Ala Gly Pro Ser Gly Gly Arg Val Gly Gly Met Arg Ser		
35	40	45
Asn Thr Gly Glu Met Ser Gln Gln Asp Arg Glu Met Met Ala Arg Phe		
50	55	60
Gly Ala Ala Gly Pro Glu Gln Ser Ser Phe Asn Tyr Glu Gln Met Arg		
65	70	75
His Glu Leu His Asn Met Gly Ala Gln Gly Gly Gln Ile Pro Gln Val		
85	90	95
Pro Ser Gln Gln Gly Ala Ala Asn Gly Gly Gln Trp Ala Arg Asp Phe		
100	105	110
Gly Gly Gln Gln Thr Ala Pro Gly Ala Ala Pro Gln Asp Ala Lys Asn		
115	120	125
Trp Asn Ala Glu Phe Gln Arg Gly Gly Ser Pro Ala Glu Ala Met Gln		
130	135	140
Gln Gln Gly Pro Gly Pro Met Gln Gly Gly Met Gly Met Gly Gly Met		
145	150	155
Pro Met Tyr Gly Met Ala Arg Pro Met Tyr Ser Gly Met Ser Ala Asn		
165	170	175
Met Ala Pro Gln Phe Gln Pro Gln Gln Ala Asn Ala Arg Val Val Glu		
180	185	190
Leu Asp Glu Gln Asn Trp Glu Glu Gln Phe Lys Gln Met Asp Ser Ala		

195	200	205
Val Gly Lys Gly Lys Glu	Val Glu Glu Gln Thr	Ala Glu Thr Ala Thr
210	215	220
Ala Thr Glu Thr Val Thr	Glu Thr Thr Thr	Glu Asp Lys Pro
225	230	235
Met Asp Ile Lys Asn Met	Asp Phe Glu Asn Ile	Trp Lys Asn Leu Gln
245	250	255
Val Asn Val Leu Asp Asn	Met Asp Glu Trp Leu	Glu Glu Thr Asn Ser
260	265	270
Pro Ala Trp Glu Arg Asp	Phe His Glu Tyr Thr	His Asn Arg Pro Glu
275	280	285
Phe Ala Asp Tyr Gln Phe	Glu Glu Asn Asn Gln	Phe Met Glu His Pro
290	295	300
Asp Pro Phe Lys Ile Gly	Val Glu Leu Met Glu	Thr Gly Gly Arg Leu
305	310	315
Ser Glu Ala Ala Leu Ala	Phe Glu Ala Ala Val	Gln Lys Asn Thr Glu
325	330	335
His Ala Glu Ala Trp Gly	Arg Leu Gly Ala Cys	Gln Ala Gln Asn Glu
340	345	350
Lys Glu Asp Pro Ala Ile	Arg Ala Leu Glu Arg	Cys Ile Lys Leu Glu
355	360	365
Pro Gly Asn Leu Ser Ala	Leu Met Asn Leu Ser	Val Ser Tyr Thr Asn
370	375	380
Glu Gly Tyr Glu Asn Ala	Ala Tyr Ala Thr Leu	Glu Arg Trp Leu Ala
385	390	395
Thr Lys Tyr Pro Glu Val	Val Asp Gln Ala Arg	Asn Gln Glu Pro Arg
405	410	415
Leu Gly Asn Glu Asp Lys	Phe Gln Leu His Ser	Arg Val Thr Glu Leu
420	425	430
Phe Ile Arg Ala Ala Gln	Leu Ser Pro Asp Gly	Ala Asn Ile Asp Ala
435	440	445
Asp Val Gln Val Gly Leu	Gly Val Leu Phe Tyr	Gly Asn Glu Glu Tyr
450	455	460
Asp Lys Ala Ile Asp Cys	Phe Asn Ala Ala Ile	Ala Val Arg Pro Asp
465	470	475
Asp Ala Leu Leu Trp Asn	Arg Leu Gly Ala Thr	Leu Ala Asn Ser His
485	490	495
Arg Ser Glu Glu Ala Ile	Asp Ala Tyr Tyr Lys	Ala Leu Glu Leu Arg
500	505	510

Pro Ser Phe Val Arg Ala Arg Tyr Asn Leu Gly Val Ser Cys Ile Asn
515 520 525
Ile Gly Cys Tyr Lys Glu Ala Ala Gln Tyr Leu Leu Gly Ala Leu Ser
530 535 540
Met His Lys Val Glu Gly Val Gln Asp Asp Val Leu Ala Asn Gln Ser
545v550 555 560
Thr Asn Leu Tyr Asp Thr Leu Lys Arg Val Phe Leu Gly Met Asp Arg
565 570 575
Arg Asp Leu Val Ala Lys Val Gly Asn Gly Met Asp Val Asn Gln Phe
580 585 590
Arg Asn Glu Phe Glu Phe
595

<210>7

<211>1024

<212>PRT

<213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG82306)

<220>

<221>MISC_FEATURE

<222>(1).. (1024)

<223>Y1Pex6p

<400>7

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

Cys	Glu	Pro	Val	Asn	Gln	Gly	Lys	Val	Val	Lys	Gly	Thr	Thr	Lys	Leu
145					150					155					160
Thr	Val	Ala	Lys	Glu	Gln	Glu	Thr	Ile	Gln	Pro	Ala	Asp	Glu	Ala	Ala
				165					170					175	
Asp	Val	Ala	Phe	Asp	Ile	Ala	Glu	Phe	Leu	Asp	Phe	Asp	Thr	Ser	Val
		180						185					190		
Ala	Lys	Thr	Arg	Glu	Ser	Thr	Asn	Leu	Gln	Val	Ala	Pro	Leu	Glu	Gly
	195						200					205			
Ala	Ile	Pro	Thr	Pro	Leu	Ser	Asp	Arg	Phe	Asp	Asp	Cys	Glu	Ser	Arg
	210						215				220				
Gly	Phe	Val	Lys	Ser	Glu	Thr	Met	Ser	Lys	Leu	Gly	Val	Phe	Ser	Gly
225				230					235					240	
Asp	Ile	Val	Ser	Ile	Lys	Thr	Lys	Asn	Gly	Ala	Glu	Arg	Val	Leu	Arg
				245					250					255	
Leu	Phe	Ala	Tyr	Pro	Glu	Pro	Asn	Thr	Val	Lys	Tyr	Asp	Val	Val	Tyr
		260						265					270		
Val	Ser	Pro	Ile	Leu	Tyr	His	Asn	Ile	Gly	Asp	Lys	Glu	Ile	Glu	Val
	275						280					285			
Thr	Pro	Asn	Gly	Glu	Thr	His	Lys	Ser	Val	Gly	Glu	Ala	Leu	Asp	Ser
	290						295				300				
Val	Leu	Glu	Ala	Ala	Glu	Glu	Val	Lys	Leu	Ala	Arg	Val	Leu	Gly	Pro
305				310					315					320	
Thr	Thr	Thr	Asp	Arg	Thr	Phe	Gln	Thr	Ala	Tyr	His	Ala	Gly	Leu	Gln
				325					330				335		
Ala	Tyr	Phe	Lys	Pro	Val	Lys	Arg	Ala	Val	Arg	Val	Gly	Asp	Leu	Ile
		340					345					350			
Pro	Ile	Pro	Phe	Asp	Ser	Ile	Leu	Ala	Arg	Thr	Ile	Gly	Glu	Asp	Pro
	355						360					365			
Glu	Met	Ser	His	Ile	Pro	Leu	Glu	Ala	Leu	Ala	Val	Lys	Pro	Asp	Ser
	370					375					380				
Val	Ala	Trp	Phe	Gln	Val	Thr	Ser	Leu	Asn	Gly	Ser	Glu	Asp	Pro	Ala
385				390					395					400	
Ser	Lys	Gln	Tyr	Leu	Val	Asp	Ser	Ser	Gln	Thr	Lys	Leu	Ile	Glu	Gly
				405					410					415	
Gly	Thr	Thr	Ser	Ser	Ala	Val	Ile	Pro	Thr	Ser	Val	Pro	Trp	Arg	Glu
		420						425				430			
Tyr	Leu	Gly	Leu	Asp	Thr	Leu	Pro	Lys	Phe	Gly	Ser	Glu	Phe	Ala	Tyr
	435						440					445			
Ala	Asp	Lys	Ile	Arg	Asn	Leu	Val	Gln	Ile	Ser	Thr	Ser	Ala	Leu	Ser

450	455	460
His Ala Lys Leu Asn Thr Ser Val Leu Leu His Ser Ala Lys Arg Gly		
465	470	475
Val Gly Lys Ser Thr Val Leu Arg Ser Val Ala Ala Gln Cys Gly Ile		480
	485	490
Ser Val Phe Glu Ile Ser Cys Phe Gly Leu Ile Gly Asp Asn Glu Ala		495
	500	505
Gln Thr Leu Gly Thr Leu Arg Ala Lys Leu Asp Arg Ala Tyr Gly Cys		510
	515	520
Ser Pro Cys Val Val Val Leu Gln His Leu Glu Ser Ile Ala Lys Lys		525
	530	535
Ser Asp Gln Asp Gly Lys Asp Glu Gly Ile Val Ser Lys Leu Val Asp		540
545	550	555
Val Leu Ala Asp Tyr Ser Gly His Gly Val Leu Leu Ala Ala Thr Ser		560
	565	570
Asn Asp Pro Asp Lys Ile Ser Glu Ala Ile Arg Ser Arg Phe Gln Phe		575
	580	585
Glu Ile Glu Ile Gly Val Pro Ser Glu Pro Gln Arg Arg Gln Ile Phe		590
	595	600
Ser His Leu Thr Lys Ser Gly Pro Gly Gly Asp Ser Ile Arg Asn Ala		605
	610	615
Pro Ile Ser Leu Arg Ser Asp Val Ser Val Glu Asn Leu Ala Leu Gln		620
625	630	635
Ser Ala Gly Leu Thr Pro Pro Asp Leu Thr Ala Ile Val Gln Thr Thr		640
	645	650
Arg Leu Arg Ala Ile Asp Arg Leu Asn Lys Leu Thr Lys Asp Ser Asp		655
	660	665
Thr Thr Leu Asp Asp Leu Leu Thr Leu Ser His Gly Thr Leu Gln Leu		670
	675	680
Thr Pro Ser Asp Phe Asp Asp Ala Ile Ala Asp Ala Arg Gln Lys Tyr		685
	690	695
Ser Asp Ser Ile Gly Ala Pro Arg Ile Pro Asn Val Gly Trp Asp Asp		700
705	710	715
Val Gly Gly Met Glu Gly Val Lys Lys Asp Ile Leu Asp Thr Ile Glu		720
	725	730
Thr Pro Leu Lys Tyr Pro His Trp Phe Ser Asp Gly Val Lys Lys Arg		735
	740	745
Ser Gly Ile Leu Phe Tyr Gly Pro Pro Gly Thr Gly Lys Thr Leu Leu		750
	755	760
		765

Ala Lys Ala Ile Ala Thr Thr Phe Ser Leu Asn Phe Phe Ser Val Lys
 770 775 780
 Gly Pro Glu Leu Leu Asn Met Tyr Ile Gly Glu Ser Glu Ala Asn Val
 785 790 795 800
 Arg Arg Val Phe Gln Lys Ala Arg Asp Ala Lys Pro Cys Val Val Phe
 805 810 815
 Phe Asp Glu Leu Asp Ser Val Ala Pro Gln Arg Gly Asn Gln Gly Asp
 820 825 830
 Ser Gly Gly Val Met Asp Arg Ile Val Ser Gln Leu Leu Ala Glu Leu
 835 840 845
 Asp Gly Met Ser Thr Ala Gly Gly Glu Gly Val Phe Val Val Gly Ala
 850 855 860
 Thr Asn Arg Pro Asp Leu Leu Asp Glu Ala Leu Leu Arg Pro Gly Arg
 865 870 875 880
 Phe Asp Lys Met Leu Tyr Leu Gly Ile Ser Asp Thr His Glu Lys Gln
 885 890 895
 Gln Thr Ile Met Glu Ala Leu Thr Arg Lys Phe Arg Leu Ala Ala Asp
 900 905 910
 Val Ser Leu Glu Ala Ile Ser Lys Arg Cys Pro Phe Thr Phe Thr Gly
 915 920 925
 Ala Asp Phe Tyr Ala Leu Cys Ser Asp Ala Met Leu Asn Ala Met Thr
 930 935 940
 Arg Thr Ala Asn Glu Val Asp Ala Lys Ile Lys Leu Leu Asn Lys Asn
 945 950 955 960
 Arg Glu Glu Ala Gly Glu Glu Pro Val Ser Ile Arg Trp Trp Phe Asp
 965 970 975
 His Glu Ala Thr Lys Ser Asp Ile Glu Val Glu Val Ala Gln Gln Asp
 980 985 990
 Phe Glu Lys Ala Lys Asp Glu Leu Ser Pro Ser Val Ser Ala Glu Glu
 995 1000 1005
 Leu Gln His Tyr Leu Lys Leu Arg Gln Gln Phe Glu Gly Gly Lys
 1010 1015 1020
 Lys
 <210>8
 <211>356
 <212>PRT
 <213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG78389)
 <220>
 <221>MISC_FEATURE

<222>(1).. (356)

<223>YlPex7p

<400>8

```

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

```

Ser Cys Ser Tyr Asp Met Thr Ala Arg Val Trp Arg Asp Ala Ser Thr
 290 295 300
 Gly Gly Asn Tyr Leu Ala Arg Gln Arg Gly Gly Thr Glu Val Lys Cys
 305 310 315 320
 Met Asp Arg His Thr Glu Phe Val Ile Gly Gly Asp Trp Ser Leu Trp
 325 330 335
 Gly Asp Pro Gly Trp Ile Thr Thr Val Gly Trp Asp Gln Met Val Tyr
 340 345 350
 Val Trp His Ala
 355

<210>9

<211>671

<212>PRT

<213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG80447)

<220>

<221>MISC_FEATURE

<222>(1).. (671)

<223>Y1Pex8p

<400>9

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

Leu Lys Arg Glu Phe Val Thr Thr Phe Asp Gln Cys Leu Ser Ile Asp		
165	170	175
Thr Arg Asn Arg Ser Asp Ala Thr Lys Trp Val Pro Val Leu Ala Cys		
180	185	190
Ile Ser Ile Ala Gln Val Tyr Ser Leu Leu Gly Asp Val Ala Ile Asn		
195	200	205
Tyr Arg Arg Phe Leu Gln Val Gly Leu Asp Leu Ile Phe Ser Asn Tyr		
210	215	220
Gly Leu Glu Met Gly Thr Ala Leu Ala Arg Leu His Ala Glu Ser Gly		
225	230	235
Gly Asp Ala Thr Thr Ala Gly Gly Leu Ile Gly Lys Lys Leu Lys Glu		
245	250	255
Pro Val Val Ala Leu Leu Asn Thr Phe Ala His Ile Ala Ser Ser Cys		
260	265	270
Ile Val His Val Asp Ile Asp Tyr Ile Asp Arg Ile Gln Asn Lys Ile		
275	280	285
Ile Leu Val Cys Glu Asn Gln Ala Glu Thr Trp Arg Ile Leu Thr Ile		
290	295	300
Glu Ser Pro Thr Val Met His His Gln Glu Ser Val Gln Tyr Leu Lys		
305	310	315
Trp Glu Leu Phe Thr Leu Cys Ile Ile Met Gln Gly Ile Ala Asn Met		
325	330	335
Leu Leu Thr Gln Lys Met Asn Gln Phe Met Tyr Leu Gln Leu Ala Tyr		
340	345	350
Lys Gln Leu Gln Ala Leu His Ser Ile Tyr Phe Ile Val Asp Gln Met		
355	360	365
Gly Ser Gln Phe Ala Ala Tyr Asp Tyr Val Phe Phe Ser Ala Ile Asp		
370	375	380
Val Leu Leu Ser Glu Tyr Ala Pro Tyr Ile Lys Asn Arg Gly Thr Ile		
385	390	395
Pro Pro Asn Lys Glu Phe Val Ala Glu Arg Leu Ala Ala Asn Leu Ala		
405	410	415
Gly Thr Ser Asn Val Gly Ser His Leu Pro Ile Asp Arg Ser Arg Val		
420	425	430
Leu Phe Ala Leu Asn Tyr Tyr Glu Gln Leu Val Thr Val Cys His Asp		
435	440	445
Ser Cys Val Glu Thr Ile Ile Tyr Pro Met Ala Arg Ser Phe Leu Tyr		
450	455	460
Pro Thr Ser Asp Ile Gln Gln Leu Lys Pro Leu Val Glu Ala Ala His		

465	470	475	480
Ser Val Ile Leu Ala Gly Leu Ala Val Pro Thr Asn Ala Val Val Asn			
	485	490	495
Ala Lys Leu Ile Pro Glu Tyr Met Gly Gly Val Leu Pro Leu Phe Pro			
	500	505	510
Gly Val Phe Ser Trp Asn Gln Phe Val Leu Ala Ile Gln Ser Ile Val			
	515	520	525
Asn Thr Val Ser Pro Pro Ser Glu ValPhe Lys Thr Asn Gln Lys Leu			
	530	535	540
Phe Arg Leu Val Leu Asp Ser Leu Met Lys Lys Cys Arg Asp Thr Pro			
545	550	555	560
Val Gly Ile Pro Val Pro His Ser Val Thr Val Ser Gln Glu Gln Glu			
	565	570	575
Asp Ile Pro Pro Thr Gln Arg Ala Val Val Met Leu Ala Leu Ile Asn			
	580	585	590
Ser Leu Pro Tyr Val Asp Ile Arg Ser Phe Glu Leu Trp Leu Gln Glu			
	595	600	605
Thr Trp Asn Met Ile Glu Ala Thr Pro Met Leu Ala Glu Asn Ala Pro			
	610	615	620
Asn Lys Glu Leu Ala His Ala Glu His Glu Phe Leu Val Leu Glu Met			
625	630	635	640
Trp Lys Met Ile Ser Gly Asn Ile Asp Gln Arg Leu Asn Asp Val Ala			
	645	650	655
Ile Arg Trp Trp Tyr Lys Lys Asn Ala Arg Val His Gly Thr Leu			
	660	665	670

<210>10
 <211>377
 <212>PRT
 <213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG81606)
 <220>
 <221>MISC_FEATURE
 <222>(1).. (377)
 <223>Y1Pex10p
 <400>10
 Met Trp Gly Ser Ser His Ala Phe Ala Gly Glu Ser Asp Leu Thr Leu
 1 5 10 15
 Gln Leu His Thr Arg Ser Asn Met Ser Asp Asn Thr Thr Ile Lys Lys
 20 25 30
 Pro Ile Arg Pro Lys Pro Ile Arg Thr Glu Arg Leu Pro Tyr Ala Gly

35						40						45							
Ala	Ala	Glu	Ile	Ile	Arg	Ala	Asn	Gln	Lys	Asp	His	Tyr	Phe	Glu	Ser				
50						55						60							
Val	Leu	Glu	Gln	His	Leu	Val	Thr	Phe	Leu	Gln	Lys	Trp	Lys	Gly	Val				
65						70						75						80	
Arg	Phe	Ile	His	Gln	Tyr	Lys	Glu	Glu	Leu	Glu	Thr	Ala	Ser	Lys	Phe				
85						90						95							
Ala	Tyr	Leu	Gly	Leu	Cys	Thr	Leu	Val	Gly	Ser	Lys	Thr	Leu	Gly	Glu				
100						105						110							
Glu	Tyr	Thr	Asn	Leu	Met	Tyr	Thr	Ile	Arg	Asp	Arg	Thr	Ala	Leu	Pro				
115						120						125							
Gly	Val	Val	Arg	Arg	Phe	Gly	Tyr	Val	Leu	Ser	Asn	Thr	Leu	Phe	Pro				
130						135						140							
Tyr	Leu	Phe	Val	Arg	Tyr	Met	Gly	Lys	Leu	Arg	Ala	Lys	Leu	Met	Arg				
145						150						155						160	
Glu	Tyr	Pro	His	Leu	Val	Glu	Tyr	Asp	Glu	Asp	Glu	Pro	Val	Pro	Ser				
165						170						175							
Pro	Glu	Thr	Trp	Lys	Glu	Arg	Val	Ile	Lys	Thr	Phe	Val	Asn	Lys	Phe				
180						185						190							
Asp	Lys	Phe	Thr	Ala	Leu	Glu	Gly	Phe	Thr	Ala	Ile	His	Leu	Ala	Ile				
195						200						205							
Phe	Tyr	Val	Tyr	Gly	Ser	Tyr	Tyr	Gln	Leu	Ser	Lys	Arg	Ile	Trp	Gly				
210						215						220							
Met	Arg	Tyr	Val	Phe	Gly	His	Arg	Leu	Asp	Lys	Asn	Glu	Pro	Arg	Ile				
225						230						235						240	
Gly	Tyr	Glu	Met	Leu	Gly	Leu	Leu	Ile	Phe	Ala	Arg	Phe	Ala	Thr	Ser				
245						250						255							
Phe	Val	Gln	Thr	Gly	Arg	Glu	Tyr	Leu	Gly	Ala	Leu	Leu	Glu	Lys	Ser				
260						265						270							
Val	Glu	Lys	Glu	Ala	Gly	Glu	Lys	Glu	Asp	Glu	Lys	Glu	Ala	Val	Val				
275						280						285							
Pro	Lys	Lys	Lys	Ser	Ser	Ile	Pro	Phe	Ile	Glu	Asp	Thr	Glu	Gly	Glu				
290						295						300							
Thr	Glu	Asp	Lys	Ile	Asp	Leu	Glu	Asp	Pro	Arg	Gln	Leu	Lys	Phe	Ile				
305						310						315						320	
Pro	Glu	Ala	Ser	Arg	Ala	Cys	Thr	Leu	Cys	Leu	Ser	Tyr	Ile	Ser	Ala				
325						330						335							
Pro	Ala	Cys	Thr	Pro	Cys	Gly	His	Phe	Phe	Cys	Trp	Asp	Cys	Ile	Ser				
340						345						350							

Glu Trp Val Arg Glu Lys Pro Glu Cys Pro Leu Cys Arg Gln Gly Val
 355 360 365
 Arg Glu Gln Asn Leu Leu Pro Ile Arg
 370 375
 <210>11
 <211>408
 <212>PRT
 <213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG81532)
 <220>
 <221>MISC_FEATURE
 <222>(1).. (408)
 <223>YIPex12p
 <400>11
 Met Asp Tyr Phe Ser Ser Leu Asn Ala Ser Gln Leu Asp Pro Asp Val
 1 5 10 15
 Pro Thr Leu Phe Glu Leu Leu Ser Ala Lys Gln Leu Glu Gly Leu Ile
 20 25 30
 Ala Pro Ser Val Arg Tyr Ile Leu Ala Phe Tyr Ala Gln Arg His Pro
 35 40 45
 Arg Tyr Leu Leu Arg Ile Val Asn Arg Tyr Asp Glu Leu Tyr Ala Leu
 50 55 60
 Phe Mer Gly Leu Val Glu Tyr Tyr Asn Leu Lys Thr Trp Asn Ala Ser
 65 70 75 80
 Phe Thr Glu Lys Phe Tyr Gly Leu Lys Arg Thr Gln Ile Leu Thr Asn
 85 90 95
 Pro Ala Leu Arg Thr Arg Gln Ala Val Pro Asp Leu Val Glu Ala Glu
 100 105 110
 Lys Arg Leu Ser Lys Lys Lys Ile Trp Gly Ser Leu Phe Phe Leu Ile
 115 120 125
 Val Val Pro Tyr Val Lys Glu Lys Leu Asp Ala Arg Tyr Glu Arg Leu
 130 135 140
 Lys Gly Arg Tyr Leu Ala Arg Asp Ile Asn Glu Glu Arg Ile Glu Ile
 145 150 155 160
 Lys Arg Thr Gly Thr Ala Gln Gln Ile Ala Val Phe Glu Phe Asp Tyr
 165 170 175
 Trp Leu Leu Lys Leu Tyr Pro Ile Val Thr Met Gly Cys Thr Thr Ala
 180 185 190
 Thr Leu Ala Phe His Met Leu Phe Leu Phe Ser Val Thr Arg Ala Tyr
 195 200 205

Ser Ile Asp Asp Phe Leu Leu Asn Ile Gln Phe Ser Arg Met Thr Arg
 210 215 220
 Tyr Asp Tyr Gln Met Glu Thr Gln Arg Asp Ser Arg Asn Ala Ala Asn
 225 230 235 240
 Val Ala His Thr Met Lys Ser Ile Ser Glu Tyr Pro Val Ala Glu Arg
 245 250 255
 Val Met Leu Leu Leu Thr Thr Lys Ala Gly Ala Asn Ala Met Arg Ser
 260 265 270
 Ala Ala Leu Ser Gly Leu Ser Tyr Val Leu Pro Thr Ser Ile Phe Ala
 275 280 285
 Leu Lys Phe Leu Glu Trp Trp Tyr Ala Ser Asp Phe Ala Arg Gln Leu
 290 295 300
 Asn Gln Lys Arg Arg Gly Asp Leu Glu Asp Asn Leu Pro Val Pro Asp
 305 310 315 320
 Lys Val Lys Gly Ala Asp Lys Leu Ala Glu Ser Val Ala Lys Trp Lys
 325 330 335
 Glu Asp Thr Ser Lys Cys Pro Leu Cys Ser Lys Glu Leu Val Asn Pro
 340 345 350
 Thr Val Ile Glu Ser Gly Tyr Val Phe Cys Tyr Thr Cys Ile Tyr Arg
 355 360 365
 His Leu Glu Asp Gly Asp Glu Glu Thr Gly Gly Arg Cys Pro Val Thr
 370 375 380
 Gly Gln Lys Leu Leu Gly Cys Arg Trp Gln Asp Asp Val Trp Gln Val
 385 390 395 400
 Thr Gly Leu Arg Arg Leu Met Val
 405

<210>12

<211>412

<212>PRT

<213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG81789)

<220>

<221>MISC_FEATURE

<222>(1).. (412)

<223>YlPex13p

<400>12

Met Ser Val Pro Arg Pro Lys Pro Trp Glu Gly Ala Ser Gly Ser Ser
 1 5 10 15
 Ala Ala Thr Ala Thr Pro Ala Ala Thr Ala Thr Pro Ala Ser Thr Asp
 20 25 30

Ala Val Ser Ser Ser Ala Gly Ser Ala Thr Gly Ala Pro Glu Leu Pro		
35	40	45
Ser Arg Pro Ser Ala Met Gly Ser Thr Ser Asn Ala Leu Ser Ser Pro		
50	55	60
Met Gly Ser Ser Met Asn Ser Gly Tyr Gly Gly Met Asn Ser Gly Tyr		
65	70	75
Gly Gly Met Gly Ser Ser Tyr Gly Ser Gly Tyr Gly Ser Ser Tyr Gly		
85	90	95
Met Gly Ser Ser Tyr Gly Ser Gly Tyr Gly Ser Gly Leu Gly Gly Tyr		
100	105	110
Gly Ser Tyr Gly Gly Met Gly Gly Met Gly Gly Met Tyr Gly Ser Arg		
115	120	125
Tyr Gly Gly Tyr Gly Ser Tyr Gly Gly Met Gly Gly Tyr Gly Gly Tyr		
130	135	140
Gly Gly Met Gly Gly Gly Pro Met Gly Gln Asn Gly Leu Ala Gly Gly		
145	150	155
Thr Gln Ala Thr Phe Gln Leu Ile Glu Ser Ile Val Gly Ala Val Gly		
165	170	175
Gly Phe Ala Gln Met Leu Glu Ser Thr Tyr Met Ala Thr Gln Ser Ser		
180	185	190
Phe Phe Ala Met Val Ser Val Ala Glu Gln Phe Gly Asn Leu Lys Asn		
195	200	205
Thr Leu Gly Ser Leu Leu Gly Ile Tyr Ala Ile Met Arg Trp Ala Arg		
210	215	220
Arg Leu Val Ala Lys Leu Ser Gly Gln Pro Val Thr Gly Ala Asn Gly		
225	230	235
Ile Thr Pro Ala Gly Phe Ala Lys Phe Glu Ala Thr Gly Gly Ala Ala		
245	250	255
Gly Pro Gly Arg Gly Pro Arg Pro Ser Tyr Lys Pro Leu Leu Phe Phe		
260	265	270
Leu Thr Ala Val Phe Gly Leu Pro Tyr Leu Leu Gly Arg Leu Ile Lys		
275	280	285
Ala Leu Ala Ala Lys Gln Glu Gly Met Tyr Asp Glu His Gly Asn Leu		
290	295	300
Leu Pro Gly Ala Gln Met Gly Met Gly Gly Pro Gly Met Glu Gly Gly		
305	310	315
Ala Glu Ile Asp Pro Ser Lys Leu Glu Phe Cys Arg Ala Asn Phe Asp		
325	330	335
Phe Val Pro Glu Asn Pro Gln Leu Glu Leu Glu Leu Arg Lys Gly Asp		

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

165																170						175											
Ala	Leu	Lys	Gly	Val	Glu	Glu	Thr	Ile	Asp	Gln	Leu	Lys	Ser	Gln	Thr																		
180																185						190											
Lys	Lys	Arg	Asp	Asp	Glu	Met	Lys	Leu	Val	Thr	Ala	Glu	Val	Glu	Arg																		
195																200						205											
Ile	Arg	Asp	Arg	Leu	Pro	Lys	Asn	Ile	Asp	Lys	Leu	Lys	Asp	Ser	Gln																		
210																215						220											
Glu	Gln	Gly	Leu	Ala	Asp	Ile	Gln	Ser	Glu	Leu	Lys	Ser	Leu	Lys	Gln																		
225																230						235						240					
Leu	Leu	Ser	Thr	Arg	Thr	Ala	Ala	Ser	Ser	Gly	Pro	Lys	Leu	Pro	Pro																		
245																250						255											
Ile	Pro	Pro	Pro	Ser	Ser	Tyr	Leu	Thr	Arg	Lys	Ala	Ser	Pro	Ala	Val																		
260																265						270											
Pro	Ala	Ala	Ala	Pro	Ala	Pro	Val	Thr	Pro	Gly	Ser	Pro	Val	His	Asn																		
275																280						285											
Val	Ser	Ser	Ser	Ser	Thr	Val	Pro	Ala	Asp	Arg	Asp	Asp	Phe	Ile	Pro																		
290																295						300											
Thr	Pro	Ala	Gly	Ala	Val	Pro	Met	Ile	Pro	Gln	Pro	Ala	Ser	Met	Ser																		
305																310						315						320					
Ser	Ser	Ser	Thr	Ser	Thr	Val	Pro	Asn	Ser	Ala	Ile	Ser	Ser	Ala	Pro																		
325																330						335											
Ser	Pro	Ile	Gln	Glu	Pro	Glu	Pro	Phe	Val	Pro	Glu	Pro	Gly	Asn	Ser																		
340																345						350											
Ala	Val	Lys	Lys	Pro	Ala	Pro	Lys	Ala	Ser	Ile	Pro	Ala	Trp	Gln	Leu																		
355																360						365											
Ala	Ala	Leu	Glu	Lys	Glu	Lys	Glu	Lys	Glu	Lys	Glu																						
370																375						380											

 $\langle 210 \rangle_{14}$

<211>391

 $\langle 212 \rangle$ PRT

<213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG79622)

 $\langle 220 \rangle$

<221>MISC FEATURE

 $\langle 222 \rangle (1) \dots (391)$ $\langle 223 \rangle$ Y1Pex16p $\langle 400 \rangle_{14}$

Met Thr Asp Lys Leu Val Lys Val Met Gln Lys Lys Lys Ser Ala Pro

1 5 10 15

Gln Thr Trp Leu Asp Ser Tyr Asp Lys Phe Leu Val Arg Asn Ala Ala

20	25	30
Ser Ile Gly Ser Ile Glu Ser Thr Leu Arg Thr Val Ser Tyr Val Leu		
35	40	45
Pro Gly Arg Phe Asn Asp Val Glu Ile Ala Thr Glu Thr Leu Tyr Ala		
50	55	60
Val Leu Asn Val Leu Gly Leu Tyr His Asp Thr Ile Ile Ala Arg Ala		
65	70	75
Val Ala Ala Ser Pro Asn Ala Ala Ala Val Tyr Arg Pro Ser Pro His		
85	90	95
Asn Arg Tyr Thr Asp Trp Phe Ile Lys Asn Arg Lys Gly Tyr Lys Tyr		
100	105	110
Ala Ser Arg Ala Val Thr Phe Val Lys Phe Gly Glu Leu Val Ala Glu		
115	120	125
Met Val Ala Lys Lys Asn Gly Gly Glu Met Ala Arg Trp Lys Cys Ile		
130	135	140
Ile Gly Ile Glu Gly Ile Lys Ala Gly Leu Arg Ile Tyr Met Leu Gly		
145	150	155
Ser Thr Leu Tyr Gln Pro Leu Cys Thr Thr Pro Tyr Pro Asp Arg Glu		
165	170	175
Val Thr Gly Glu Leu Leu Glu Thr Ile Cys Arg Asp Glu Gly Glu Leu		
180	185	190
Asp Ile Glu Lys Gly Leu Met Asp Pro Gln Trp Lys Met Pro Arg Thr		
195	200	205
Gly Arg Thr Ile Pro Glu Ile Ala Pro Thr Asn Val Glu Gly Tyr Leu		
210	215	220
Leu Thr Lys Val Leu Arg Ser Glu Asp Val Asp Arg Pro Tyr Asn Leu		
225	230	235
Leu Ser Arg Leu Asp Asn Trp Gly Val Val Ala Glu Leu Leu Ser Ile		
245	250	255
Leu Arg Pro Leu Ile Tyr Ala Cys Leu Leu Phe Arg Gln His Val Asn		
260	265	270
Lys Thr Val Pro Ala Ser Thr Lys Ser Lys Phe Pro Phe Leu Asn Ser		
275	280	285
Pro Trp Ala Pro Trp Ile Ile Gly Leu Val Ile Glu Ala Leu Ser Arg		
290	295	300
Lys Met Met Gly Ser Trp Leu Leu Arg Gln Arg Gln Ser Gly Lys Thr		
305	310	315
Pro Thr Ala Leu Asp Gln Met Glu Val Lys Gly Arg Thr Asn Leu Leu		
325	330	335

Gly Trp Trp Leu Phe Arg Gly Glu Phe Tyr Gln Ala Tyr Thr Arg Pro
 340 345 350
 Leu Leu Tyr Ser Ile Val Ala Arg Leu Glu Lys Ile Pro Gly Leu Gly
 355 360 365
 Leu Phe Gly Ala Leu Ile Ser Asp Tyr Leu Tyr Leu Phe Asp Arg Tyr
 370 375 380
 Tyr Phe Thr Ala Ser Thr Leu
 385 390
 <210>15
 <211>225
 <212>PRT
 <213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG84025)
 <220>
 <221>MISC_FEATURE
 <222>(1).. (225)
 <223>Y1Pex17p
 <400>15
 Met Ser Ala Phe Pro Glu Pro Ser Ser Phe Glu Ile Glu Phe Ala Lys
 1 5 10 15
 Gln Met Asn Arg Pro Arg Thr Val Gln Phe Lys Gln Leu Val Ala Val
 20 25 30
 Leu Tyr Ile Phe Gly Gly Thr Ser Ala Leu Ile Tyr Ile Ile Ser Lys
 35 40 45
 Thr Ile Leu Asn Pro Leu Phe Glu Glu Leu Thr Phe Ala Arg Ser Glu
 50 55 60
 Tyr Ala Ile His Ala Arg Arg Leu Met Glu Gln Leu Asn Ala Lys Leu
 65 70 75 80
 Ser Ser Met Ala Ser Tyr Ile Pro Pro Val Arg Ala Leu Gln Gly Gln
 85 90 95
 Arg Phe Val Asp Ala Gln Thr Gln Thr Glu Asp Glu Glu Gly Glu Asp
 100 105 110
 Ile Pro Asn Pro Ser Leu Gly Lys Ser Ser His Val Ser Phe Gly Glu
 115 120 125
 Ser Pro Met Gln Leu Lys Leu Ala Glu Lys Glu Lys Gln Gln Lys Leu
 130 135 140
 Ile Asp Asp Ser Val Asp Asn Leu Glu Arg Leu Ala Asp Ser Leu Lys
 145 150 155 160
 His Ala Gly Glu Val Ser Asp Leu Ser Ala Leu Ser Gly Phe Lys Tyr
 165 170 175

Gln Val Glu Glu Leu Thr Asn Tyr Ser Asp Gln Leu Ala Met Ser Gly
 180 185 190
 Tyr Ser Met Met Lys Ser Gly Leu Pro Gly His Glu Thr Ala Met Ser
 195 200 205
 Glu Thr Lys Lys Glu Ile Arg Ser Leu Lys Gly Ser Val Leu Ser Val
 210 215 220
 Arg
 225
 <210>16
 <211>324
 <212>PRT
 <213> 解脂耶氏酵母 (GenBank 保藏号 AAK84827)
 <220>
 <221>MISC_FEATURE
 <222>(1).. (324)
 <223>YlPex19p
 <400>16
 Met Ser His Glu Glu Asp Leu Asp Asp Leu Asp Asp Phe Leu Asp Glu
 1 5 10 15
 Phe Asp Glu Gln Val Leu Ser Lys Pro Pro Gly Ala Gln Lys Asp Ala
 20 25 30
 Thr Pro Thr Thr Ser Thr Ala Pro Thr Thr Ala Glu Ala Lys Pro Asp
 35 40 45
 Ala Thr Lys Lys Ser Thr Glu Thr Ser Gly Thr Asp Ser Lys Thr Glu
 50 55 60
 Gly Ala Asp Thr Ala Asp Lys Asn Ala Ala Thr Asp Ser Ala Glu Ala
 65 70 75 80
 Gly Ala Glu Lys Val Ser Leu Pro Asn Leu Glu Asp Gln Leu Ala Gly
 85 90 95
 Leu Lys Met Asp Asp Phe Leu Lys Asp Ile Glu Ala Asp Pro Glu Ser
 100 105 110
 Lys Ala Gln Phe Glu Ser Leu Leu Lys Glu Ile Asn Asn Val Thr Ser
 115 120 125
 Ala Thr Ala Ser Glu Lys Ala Gln Gln Pro Lys Ser Phe Lys Glu Thr
 130 135 140
 Ile Ser Ala Thr Ala Asp Arg Leu Asn Gln Ser Asn Gln Glu Met Gly
 145 150 155 160
 Asp Met Pro Leu Gly Asp Asp Met Leu Ala Gly Leu Met Glu Gln Leu
 165 170 175

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

<210>17

<211>417

<212>PRT

<213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG79226)

<220>

<221>MISC_FEATURE

<222>(1).. (417)

<223>Y1Pex20p

<400>17

Met Ala Ser Cys Gly Pro Ser Asn Ala Leu Gln Asn Leu Ser Lys His
 1 5 10 15
 Ala Ser Ala Asp Arg Ser Leu Gln His Asp Arg Met Ala Pro Gly Gly
 20 25 30
 Ala Pro Gly Ala Gln Arg Gln Gln Phe Arg Ser Gln Thr Gln Gly Gly
 35 40 45
 Gln Leu Asn Asn Glu Phe Gln Gln Phe Ala Gln Ala Gly Pro Ala His
 50 55 60
 Asn Ser Phe Glu Gln Ser Gln Met Gly Pro His Phe Gly Gln Gln His
 65 70 75 80
 Phe Gly Gln Pro His Gln Pro Gln Met Gly Gln His Ala Pro Met Ala

	85		90		95
His Gly Gln Gln Ser Asp Trp Ala Gln Ser Phe Ser Gln Leu Asn Leu					
100		105		110	
Gly Pro Gln Thr Gly Pro Gln His Thr Gln Gln Ser Asn Trp Gly Gln					
115		120		125	
Asp Phe Met Arg Gln Ser Pro Gln Ser His Gln Val Gln Pro Gln Met					
130		135		140	
Ala Asn Gly Val Met Gly Ser Met Ser Gly Met Ser Ser Phe Gly Pro					
145		150		155	160
Met Tyr Ser Asn Ser Gln Leu Met Asn Ser Thr Tyr Gly Leu Gln Thr					
165		170		175	
Glu His Gln Gln Thr His Lys Thr Glu Thr Lys Ser Ser Gln Asp Ala					
180		185		190	
Ala Phe Glu Ala Ala Phe Gly Ala Val Glu Glu Ser Ile Thr Lys Thr					
195		200		205	
Ser Asp Lys Gly Lys Glu Val Glu Lys Asp Pro Met Glu Gln Thr Tyr					
210		215		220	
Arg Tyr Asp Gln Ala Asp Ala Leu Asn Arg Gln Ala Glu His Ile Ser					
225		230		235	240
Asp Asn Ile Ser Arg Glu Glu Val Asp Ile Lys Thr Asp Glu Asn Gly					
245		250		255	
Glu Phe Ala Ser Ile Ala Arg Gln Ile Ala Ser Ser Leu Glu Glu Ala					
260		265		270	
Asp Lys Ser Lys Phe Glu Lys Ser Thr Phe Met Asn Leu Met Arg Arg					
275		280		285	
Ile Gly Asn His Glu Val Thr Leu Asp Gly Asp Lys Leu Val Asn Lys					
290		295		300	
Glu Gly Glu Asp Ile Arg Glu Glu Val Arg Asp Glu Leu Leu Arg Glu					
305		310		315	320
Gly Ala Ser Gln Glu Asn Gly Phe Gln Ser Glu Ala Gln Gln Thr Ala					
325		330		335	
Pro Leu Pro Val His His Glu Ala Pro Pro Pro Glu Gln Ile His Pro					
340		345		350	
His Thr Glu Thr Gly Asp Lys Gln Leu Glu Asp Pro Met Val Tyr Ile					
355		360		365	
Glu Gln Glu Ala Ala Arg Arg Ala Ala Glu Ser Gly Arg Thr Val Glu					
370		375		380	
Glu Glu Lys Leu Asn Phe Tyr Ser Pro Phe Glu Tyr Ala Gln Lys Leu					
385		390		395	400

Gly Pro Gln Gly Val Ala Lys Gln Ser Asn Trp Glu Glu Asp Tyr Asp
 405 410 415
 Phe
 <210>18
 <211>195
 <212>PRT
 <213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 CAG77876)
 <220>
 <221>MISC_FEATURE
 <222>(1).. (195)
 <223>YIPex22p
 <400>18
 Val Pro Arg Cys Thr Ser His Pro Cys Asn Leu Thr Leu His Leu Pro
 1 5 10 15
 Val Thr Thr Met Ala Pro Arg Lys Thr Arg Leu Pro Ala Val Ile Gly
 20 25 30
 Ala Ala Ala Ala Ala Ala Ala Val Ala Tyr Leu Val Tyr Ser Phe Val
 35 40 45
 Ala Lys Ser Asn Ser Asp Gln Asp Thr Phe Asp Ser Ser Val Gln Ser
 50 55 60
 Ser Ser Lys Ser Ser Thr Lys Ser Pro Lys Ser Thr Ala Thr Asn Ser
 65 70 75 80
 Lys Ile Thr Val Val Val Ser Gln Glu Leu Val Gln Ser Gln Leu Val
 85 90 95
 Asp Phe Lys His Leu Met Ser Val His Pro Asn Leu Val Val Ile Val
 100 105 110
 Pro Pro Met Val Ala Asn Lys Phe His Arg Ala Leu Lys Ser Ser Val
 115 120 125
 Gly His Asp His Gly Val Lys Val Ile Arg Cys Asp Thr Asp Val Gly
 130 135 140
 Val Ile His Val Ile Lys His Ile Arg Pro Asp Leu Ala Leu Ile Ala
 145 150 155 160
 Asp Gly Val Gly Asp Asn Ile Gln Gly Glu Ile Lys Arg Phe Val Gly
 165 170 175
 Ser Ser Glu Ala Leu Ser Gly Asp Val Asn Leu Ala Ala Glu Arg Leu
 180 185 190
 Thr Gly Leu
 195
 <210>19

<211>386

<212>PRT

<213> 解脂耶氏酵母 CLIB122 (GenBank 保藏号 NC_006072, 核苷酸 117230 至 118387 的反义翻译)

<220>

<221>MISC_FEATURE

<222>(1)..(386)

<223>Y1Pex26p

<400>19

```

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

```

Glu Met Thr His Lys Leu Asp Gln Ser Gln Lys His Ala Glu Gln Glu			
245	250	255	
His Lys Arg Leu Leu Glu Glu Ala Gln Glu Lys Glu Lys Ser Asp Ala			
260	265	270	
Lys Glu Lys Glu Arg Glu Glu Arg Val Ser Arg Asp Thr Gln Ser Arg			
275	280	285	
Glu Ile Lys Ser Pro Ile Val Asp Ser Ser Thr Ser Ser Arg Asp Val			
290	295	300	
Thr Arg Asp Thr Thr Arg Glu Leu Ser Lys Ser Ser Arg Gln Pro Arg			
305	310	315	320
Thr Leu Ser Gln Ile Ile Ser Thr Ser Leu Lys Ser Gln Phe Asp Gly			
325	330	335	
Asn Ala Ile Phe Arg Thr Leu Ala Leu Ile Val Ile Val Ser Leu Ser			
340	345	350	
Ala Ala Asn Pro Leu Ile Arg Lys Arg Val Val Asp Thr Leu Lys Met			
355	360	365	
Leu Trp Ile Lys Ile Leu Gln Thr Leu Ser Met Gly Phe Lys Val Ser			
370	375	380	

Tyr Leu

385

<210>20

<211>3387

<212>DNA

<213> 解脂耶氏酵母 (GenBank 保藏号 AB036770)

<400>20

ggtaccatca agggtaaaat caaggctatc atcaagggcc atatatcgca agtttggggg	60
aagataatat gttcatagtg aatcggggtg tggatttcct catctaacgg cattataact	120
agtcctggag ggtctttttt atggataacc tccatgtacg atgtatccaa gatctccacg	180
tactgtgttc tgtttcctaa gtaataccca acaacctctc caacaaacac ttgggaagat	240
gcacttgtgc tgagatgtca agatgttaga gagtagagac agtagcaagc gtaaaaggcg	300
gccgaggcca ccgagagAAC agcgtagcag ggcgcgtagt caccacaggg gacgcagaac	360
caaacaaatg acgaagaaga accacaagga gacgttttca aaggcaatgc aaacgaagag	420
ggcaatggaa ggattgagat tagagaactg gagactggag tggcgttttc ccgatgaacg	480
aacaaacacg cgaagctatg tggaccaaca tacaacacgg actgaaccag gtttttttat	540
gatttttttta ctggaaatag gtacgtgcc aagtggacca tgacactaaa cgtgtttaat	600
tagtaatatt cgtgtaagcg tacattcatt tcaaagggtta ttctttcacg gcaaagttat	660
aattaaatga atgtatatgc agaaaaaaa aaaaaaagta ctgtactgga tggagagaat	720
attaataaat aattgttacc caactacatc ttgtcgattg aaagagaccc ctaagacaga	780
taggatatct gcaacccgag gaatgaaccc cccagcaccg gcaccctttc tattaacaaa	840

atgccaaactg	aaatttgaaa	agttcaacta	aacttatttg	acccacaaaa	actcgtcaaa	900
agtggcggcg	aaagctggca	aatgatgaca	tccccttgga	accatgatata	cctctcggaa	960
tcttcgtccc	catttgccac	atctacttgc	aacgccacat	ctgcttacta	agcaacccaa	1020
atctgcctcg	gctcaaaatg	tggggaagtt	cacatgcatt	cgctgggtgaa	tctgatctga	1080
cactacaact	acacaccagg	tccaacatga	gcgacaatac	gacaatcaaa	aagccgatcc	1140
gacccaaacc	gatecggacg	gaacgcctgc	cttacgctgg	ggccgcagaa	atcatccgag	1200
ccaaccagaa	agaccactac	tttgagtccg	tgcttgaaca	gcatctcgtc	acgtttctgc	1260
agaaatggaa	gggagtacga	tttateccacc	agtacaagga	ggagctggag	acggcgtcca	1320
agtttgcata	tctcggtttg	tgtacgcttg	tgggctccaa	gactctcggga	gaagagtaca	1380
ccaatctcat	gtacactata	agagaccgaa	cagctctacc	gggggtgggtg	agacggtttg	1440
gctacgtgct	ttccaacact	ctgtttccat	acctgtttgt	gcgctacatg	ggcaagttgc	1500
gcgccaaact	gatgcgcgag	tatccccatc	tgggtggagta	cgacgaagat	gagcctgtgc	1560
ccagcccgga	aacatggaag	gagcgggtca	tcaagacgtt	tgtgaacaag	tttgacaagt	1620
tcacggcgct	ggaggggttt	accgcgatcc	acttggcgat	tttctacgtc	tacggctcgt	1680
actaccagct	cagtaagcgg	atctggggca	tgcgttatgt	atttggacac	cgactggaca	1740
agaatgagcc	tcgaatcggt	tacgagatgc	tcggtctgct	gattttcggc	cggtttgcca	1800
cgtcatttgt	gcagacggga	agagagtacc	tcggagcgct	gctggaaaag	agcgtggaga	1860
aagaggcagg	ggagaaggaa	gatgaaaagg	aagcggttgt	gccgaaaaag	aagtcgtcaa	1920
ttccgttcat	tgaggataca	gaagggggaga	cggaagacaa	gatcgatctg	gaggaccctc	1980
gacagctcaa	gttcattcct	gaggcgtcca	gagcgtgcac	tctgtgtctg	tcatacatta	2040
gtgcgccggc	atgtacgcca	tgtggacact	ttttctgttg	ggactgtatt	tccgaatggg	2100
tgagagagaa	gcccagatgt	cccttgtgtc	ggcaggggtgt	gagagagcag	aacttgttgc	2160
ctatcagata	atgacgaggt	ctggatggaa	ggactagtca	gcgagacaca	gagcatcagg	2220
gaccagacac	gaccaattca	atcgacaaca	ctgtgctgca	tagcagtga	cagaggtcct	2280
gggcatgaat	atatttttagc	attggagata	tgagtggtag	agcgtataca	gtattaattg	2340
tggaggtatc	tcgtcgcat	gatagagcaa	tacagttact	gctgaaggga	atgataccga	2400
gtatttcggc	ccgattcagt	tcttgatata	gtcattttgt	ctctattgtc	tacttttcag	2460
ataacctcaa	caaattcttc	acaaatctcc	cagtaaacag	tcagagatca	tatccgagat	2520
catatcagat	atgtcacgat	ccgagtacaa	taatggatat	taatctgctt	gattttgaat	2580
tctgtttcga	ttatgatttc	tttgatttcg	atatgaacac	atacggcgac	tcccagacct	2640
ttagaagctc	cagtttggat	tcttagcaat	ggttacactc	aactatatcc	caagtaatac	2700
ttggtaacaa	tatgccaaagt	tagtcattca	ttcgttatag	gagttagcaa	gtgtttgtca	2760
gctaaaaatg	gttagtcggt	cgattaccac	ttagatcttt	tcagcgtgga	acttgatgggt	2820
acgcttgaac	cgacacttgg	agtagtcggg	gctgttgatg	acgtagatga	cgtttcgtc	2880
agggtgagga	gtgcaatagt	agtactcctt	ggggccgtct	ctcagctcaa	aggttccatc	2940
ggcggcaatg	tcaaagaccg	agccctggag	cttgtagccg	tagtcgccgg	tccagaacaa	3000
agcctgcagc	tccagatagg	cgatgggcat	gtcgttaaca	gagaagggtgt	tgccctcgcc	3060
ctcggtgatg	gtgatgggtt	cgccgtcggt	ggaggcggtg	atcagggtcat	cttggttaggt	3120
gacgggcaga	gattcgaccg	attgggcgtc	tgatctggta	taggtcagct	tgtacttgtc	3180

tccgacagcc gccagagcgg tggtagcgac ggtgatgagg gagatgagtt tcatattggc 3240
 ggcaagttta gcaaaagatg gcagtgggat tgagggacaa gagtgtttat atagatatag 3300
 atacaacaca acgagtctga atgagacaac cgagacaacc actccccgaag cctcactaat 3360
 agttactaac ggcatatttc aggtacc 3387

<210>21

<211>1134

<212>DNA

<213> 解脂耶氏酵母 (GenBank 保藏号 AB036770, 核苷酸 1038-2171)

<220>

<221>CDS

<222>(1).. (1134)

<223>Pex10

<400>21

atg tgg gga agt tca cat gca ttc gct ggt gaa tct gat ctg aca cta 48
 Met Trp Gly Ser Ser His Ala Phe Ala Gly Glu Ser Asp Leu Thr Leu
 1 5 10 15

caa cta cac acc agg tcc aac atg agc gac aat acg aca atc aaa aag 96
 Gln Leu His Thr Arg Ser Asn Met Ser Asp Asn Thr Thr Ile Lys Lys
 20 25 30

ccg atc cga ccc aaa ccg atc cgg acg gaa cgc ctg cct tac gct ggg 144
 Pro Ile Arg Pro Lys Pro Ile Arg Thr Glu Arg Leu Pro Tyr Ala Gly
 35 40 45

gcc gca gaa atc atc cga gcc aac cag aaa gac cac tac ttt gag tcc 192
 Ala Ala Glu Ile Ile Arg Ala Asn Gln Lys Asp His Tyr Phe Glu Ser
 50 55 60

gtg ctt gaa cag cat ctc gtc acg ttt ctg cag aaa tgg aag gga gta 240
 Val Leu Glu Gln His Leu Val Thr Phe Leu Gln Lys Trp Lys Gly Val
 65 70 75 80

cga ttt atc cac cag tac aag gag gag ctg gag acg gcg tcc aag ttt 288
 Arg Phe Ile His Gln Tyr Lys Glu Glu Leu Glu Thr Ala Ser Lys Phe
 85 90 95

gca tat ctc ggt ttg tgt acg ctt gtg ggc tcc aag act ctc gga gaa 336
 Ala Tyr Leu Gly Leu Cys Thr Leu Val Gly Ser Lys Thr Leu Gly Glu
 100 105 110

gag tac acc aat ctc atg tac act atc aga gac cga aca gct cta ccg 384
 Glu Tyr Thr Asn Leu Met Tyr Thr Ile Arg Asp Arg Thr Ala Leu Pro
 115 120 125

ggg gtg gtg aga cgg ttt ggc tac gtg ctt tcc aac act ctg ttt cca 432
 Gly Val Val Arg Arg Phe Gly Tyr Val Leu Ser Asn Thr Leu Phe Pro

130	135	140	
tac ctg ttt gtg cgc	tac atg ggc aag ttg cgc gcc aaa ctg atg cgc		480
Tyr Leu Phe Val Arg	Tyr Met Gly Lys Leu Arg Ala Lys Leu Met Arg		
145	150	155	160
gag tat ccc cat ctg gtg gag tac gac gaa gat gag cct gtg ccc agc			528
Glu Tyr Pro His Leu Val	Glu Tyr Asp Glu Asp Glu Pro Val Pro Ser		
165	170	175	
ccg gaa aca tgg aag gag cgg gtc atc aag acg ttt gtg aac aag ttt			576
Pro Glu Thr Trp Lys Glu Arg Val Ile Lys Thr Phe Val Asn Lys Phe			
180	185	190	
gac aag ttc acg gcg ctg gag ggg ttt acc gcg atc cac ttg gcg att			624
Asp Lys Phe Thr Ala Leu Glu Gly Phe Thr Ala Ile His Leu Ala Ile			
195	200	205	
ttc tac gtc tac ggc tcg tac tac cag ctc agt aag cgg atc tgg ggc			672
Phe Tyr Val Tyr Gly Ser Tyr Tyr Gln Leu Ser Lys Arg Ile Trp Gly			
210	215	220	
atg cgt tat gta ttt gga cac cga ctg gac aag aat gag cct cga atc			720
Met Arg Tyr Val Phe Gly His Arg Leu Asp Lys Asn Glu Pro Arg Ile			
225	230	235	240
ggt tac gag atg ctc ggt ctg ctg att ttc gcc cgg ttt gcc acg tca			768
Gly Tyr Glu Met Leu Gly Leu Leu Ile Phe Ala Arg Phe Ala Thr Ser			
245	250	255	
ttt gtg cag acg gga aga gag tac ctc gga gcg ctg ctg gaa aag agc			816
Phe Val Gln Thr Gly Arg Glu Tyr Leu Gly Ala Leu Leu Glu Lys Ser			
260	265	270	
gtg gag aaa gag gca ggg gag aag gaa gat gaa aag gaa gcg gtt gtg			864
Val Glu Lys Glu Ala Gly Glu Lys Glu Asp Glu Lys Glu Ala Val Val			
275	280	285	
ccg aaa aag aag tcg tca att ccg ttc att gag gat aca gaa ggg gag			912
Pro Lys Lys Lys Ser Ser Ile Pro Phe Ile Glu Asp Thr Glu Gly Glu			
290	295	300	
acg gaa gac aag atc gat ctg gag gac cct cga cag ctc aag ttc att			960
Thr Glu Asp Lys Ile Asp Leu Glu Asp Pro Arg Gln Leu Lys Phe Ile			
305	310	315	320
cct gag gcg tcc aga gcg tgc act ctg tgt ctg tca tac att agt gcg			1008
Pro Glu Ala Ser Arg Ala Cys Thr Leu Cys Leu Ser Tyr Ile Ser Ala			
325	330	335	
ccg gca tgt acg cca tgt gga cac ttt ttc tgt tgg gac tgt att tcc			1056
Pro Ala Cys Thr Pro Cys Gly His Phe Phe Cys Trp Asp Cys Ile Ser			

340	345	350	
gaa tgg gtg aga gag aag ccc gag tgt ccc ttg tgt cgg cag ggt gtg			1104
Glu Trp Val Arg Glu Lys Pro Glu Cys Pro Leu Cys Arg Gln Gly Val			
355	360	365	
aga gag cag aac ttg ttg cct atc aga taa			1134
Arg Glu Gln Asn Leu Leu Pro Ile Arg			
370	375		
<210>22			
<211>377			
<212>PRT			
<213> 解脂耶氏酵母 (GenBank 保藏号 AB036770, 核苷酸 1038-2171)			
<400>22			
Met Trp Gly Ser Ser His Ala Phe Ala Gly Glu Ser Asp Leu Thr Leu			
1 5 10 15			
Gln Leu His Thr Arg Ser Asn Met Ser Asp Asn Thr Thr Ile Lys Lys			
20 25 30			
Pro Ile Arg Pro Lys Pro Ile Arg Thr Glu Arg Leu Pro Tyr Ala Gly			
35 40 45			
Ala Ala Glu Ile Ile Arg Ala Asn Gln Lys Asp His Tyr Phe Glu Ser			
50 55 60			
Val Leu Glu Gln His Leu Val Thr Phe Leu Gln Lys Trp Lys Gly Val			
65 70 75 80			
Arg Phe Ile His Gln Tyr Lys Glu Glu Leu Glu Thr Ala Ser Lys Phe			
85 90 95			
Ala Tyr Leu Gly Leu Cys Thr Leu Val Gly Ser Lys Thr Leu Gly Glu			
100 105 110			
Glu Tyr Thr Asn Leu Met Tyr Thr Ile Arg Asp Arg Thr Ala Leu Pro			
115 120 125			
Gly Val Val Arg Arg Phe Gly Tyr Val Leu Ser Asn Thr Leu Phe Pro			
130 135 140			
Tyr Leu Phe Val Arg Tyr Met Gly Lys Leu Arg Ala Lys Leu Met Arg			
145 150 155 160			
Glu Tyr Pro His Leu Val Glu Tyr Asp Glu Asp Glu Pro Val Pro Ser			
165 170 175			
Pro Glu Thr Trp Lys Glu Arg Val Ile Lys Thr Phe Val Asn Lys Phe			
180 185 190			
Asp Lys Phe Thr Ala Leu Glu Gly Phe Thr Ala Ile His Leu Ala Ile			
195 200 205			
Phe Tyr Val Tyr Gly Ser Tyr Tyr Gln Leu Ser Lys Arg Ile Trp Gly			

210	215	220	
Met Arg Tyr Val Phe Gly His Arg Leu Asp Lys Asn Glu Pro Arg Ile			
225	230	235	240
Gly Tyr Glu Met Leu Gly Leu Leu Ile Phe Ala Arg Phe Ala Thr Ser			
245	250	255	
Phe ValGln Thr Gly Arg Glu Tyr Leu Gly Ala Leu Leu Glu Lys Ser			
260	265	270	
ValGlu Lys Glu Ala Gly Glu Lys Glu Asp Glu Lys Glu Ala Val Val			
275	280	285	
Pro Lys Lys Lys Ser Ser Ile Pro Phe Ile Glu Asp Thr Glu Gly Glu			
290	295	300	
Thr Glu Asp Lys Ile Asp Leu Glu Asp Pro Arg Gln Leu Lys Phe Ile			
305	310	315	320
Pro Glu Ala Ser Arg Ala Cys Thr Leu Cys Leu Ser Tyr Ile Ser Ala			
325	330	335	
Pro Ala Cys Thr Pro Cys Gly His Phe Phe Cys Trp Asp Cys Ile Ser			
340	345	350	
Glu Trp Val Arg Glu Lys Pro Glu Cys Pro Leu Cys Arg Gln Gly Val			
355	360	365	
Arg Glu Gln Asn Leu Leu Pro Ile Arg			
370	375		
<210>23			
<211>1065			
<212>DNA			
<213> 解脂耶氏酵母 (GenBank 保藏号 AJ012084, 对应于 GenBank 保藏号 AB036770 的			
核			
苷酸 1107-2171)			
<220>			
<221>CDS			
<222>(1).. (1065)			
<223>Y1PEX10			
<400>23			
atg agc gac aat acg aca atc aaa aag ccg atc cga ccc aaa ccg atc			48
Met Ser Asp Asn Thr Thr Ile Lys Lys Pro Ile Arg Pro Lys Pro Ile			
1	5	10	15
cgg acg gaa cgc ctg cct tac gct ggg gcc gca gaa atc atc cga gcc			96
Arg Thr Glu Arg Leu Pro Tyr Ala Gly Ala Ala Glu Ile Ile Arg Ala			
20	25	30	
aac cag aaa gac cac tac ttt gag tcc gtg ctt gaa cag cat ctc gtc			144

Asn Gln Lys Asp His Tyr Phe Glu Ser Val Leu Glu Gln His Leu Val	
35 40 45	
acg ttt ctg cag aaa tgg aag gga gta cga ttt atc cac cag tac aag	192
Thr Phe Leu Gln Lys Trp Lys Gly Val Arg Phe Ile His Gln Tyr Lys	
50 55 60	
gag gag ctg gag acg gcg tcc aag ttt gca tat ctc ggt ttg tgt acg	240
Glu Glu Leu Glu Thr Ala Ser Lys Phe Ala Tyr Leu Gly Leu Cys Thr	
65 70 75 80	
ctt gtg ggc tcc aag act ctc gga gaa gag tac acc aat ctc atg tac	288
Leu Val Gly Ser Lys Thr Leu Gly Glu Glu Tyr Thr Asn Leu Met Tyr	
85 90 95	
act atc aga gac cga aca gct cta ccg ggg gtg gtg aga cgg ttt ggc	336
Thr Ile Arg Asp Arg Thr Ala Leu Pro Gly Val Val Arg Arg Phe Gly	
100 105 110	
tac gtg ctt tcc aac act ctg ttt cca tac ctg ttt gtg cgc tac atg	384
Tyr Val Leu Ser Asn Thr Leu Phe Pro Tyr Leu Phe Val Arg Tyr Met	
115 120 125	
ggc aag ttg cgc gcc aaa ctg atg cgc gag tat ccc cat ctg gtg gag	432
Gly Lys Leu Arg Ala Lys Leu Met Arg Glu Tyr Pro His Leu Val Glu	
130 135 140	
tac gac gaa gat gag cct gtg ccc agc ccg gaa aca tgg aag gag cgg	480
Tyr Asp Glu Asp Glu Pro Val Pro Ser Pro Glu Thr Trp Lys Glu Arg	
145 150 155 160	
gtc atc aag acg ttt gtg aac aag ttt gac aag ttc acg gcg ctg gag	528
Val Ile Lys Thr Phe Val Asn Lys Phe Asp Lys Phe Thr Ala Leu Glu	
165 170 175	
ggg ttt acc gcg atc cac ttg gcg att ttc tac gtc tac ggc tcg tac	576
Gly Phe Thr Ala Ile His Leu Ala Ile Phe Tyr Val Tyr Gly Ser Tyr	
180 185 190	
tac cag ctc agt aag cgg atc tgg ggc atg cgt tat gta ttt gga cac	624
Tyr Gln Leu Ser Lys Arg Ile Trp Gly Met Arg Tyr Val Phe Gly His	
195 200 205	
cga ctg gac aag aat gag cct cga atc ggt tac gag atg ctc ggt ctg	672
Arg Leu Asp Lys Asn Glu Pro Arg Ile Gly Tyr Glu Met Leu Gly Leu	
210 215 220	
ctg att ttc gcc cgg ttt gcc acg tca ttt gtg cag acg gga aga gag	720
Leu Ile Phe Ala Arg Phe Ala Thr Ser Phe Val Gln Thr Gly Arg Glu	
225 230 235 240	
tac ctc gga gcg ctg ctg gaa aag agc gtg gag aaa gag gca ggg gag	768

Tyr	Leu	Gly	Ala	Leu	Leu	Glu	Lys	Ser	Val	Glu	Lys	Glu	Ala	Gly	Glu		
				245				250						255			
aag	gaa	gat	gaa	aag	gaa	gcg	gtt	gtg	ccg	aaa	aag	aag	tcg	tca	att	816	
Lys	Glu	Asp	Glu	Lys	Glu	Ala	Val	Val	Pro	Lys	Lys	Lys	Ser	Ser	Ile		
				260				265						270			
ccg	ttc	att	gag	gat	aca	gaa	ggg	gag	acg	gaa	gac	aag	atc	gat	ctg	864	
Pro	Phe	Ile	Glu	Asp	Thr	Glu	Gly	Glu	Thr	Glu	Asp	Lys	Ile	Asp	Leu		
				275				280						285			
gag	gac	cct	cga	cag	ctc	aag	ttc	att	cct	gag	gcg	tcc	aga	gcg	tgc	912	
Glu	Asp	Pro	Arg	Gln	Leu	Lys	Phe	Ile	Pro	Glu	Ala	Ser	Arg	Ala	Cys		
				290				295						300			
act	ctg	tgt	ctg	tca	tac	att	agt	gcg	ccg	gca	tgt	acg	cca	tgt	gga	960	
Thr	Leu	Cys	Leu	Ser	Tyr	Ile	Ser	Ala	Pro	Ala	Cys	Thr	Pro	Cys	Gly		
305						310					315				320		
cac	ttt	ttc	tgt	tgg	gac	tgt	att	tcc	gaa	tgg	gtg	aga	gag	aag	ccc	1008	
His	Phe	Phe	Cys	Trp	Asp	Cys	Ile	Ser	Glu	Trp	Val	Arg	Glu	Lys	Pro		
						325					330				335		
gag	tgt	ccc	ttg	tgt	cgg	cag	ggt	gtg	aga	gag	cag	aac	ttg	ttg	cct	1056	
Glu	Cys	Pro	Leu	Cys	Arg	Gln	Gly	Val	Arg	Glu	Gln	Asn	Leu	Leu	Pro		
						340									350		
atc	aga	taa														1065	
Ile	Arg																
<210>	24																
<211>	354																
<212>	PRT																
<213>	解脂耶氏酵母 (GenBank 保藏号 AJ012084, 对应于 GenBank 保藏号 AB036770 的																
核																	
苷酸	1107-2171)																
<400>	24																
Met	Ser	Asp	Asn	Thr	Thr	Ile	Lys	Lys	Pro	Ile	Arg	Pro	Lys	Pro	Ile		
1				5					10					15			
Arg	Thr	Glu	Arg	Leu	Pro	Tyr	Ala	Gly	Ala	Ala	Glu	Ile	Ile	Arg	Ala		
				20					25					30			
Asn	Gln	Lys	Asp	His	Tyr	Phe	Glu	Ser	Val	Leu	Glu	Gln	His	Leu	Val		
				35					40					45			
Thr	Phe	Leu	Gln	Lys	Trp	Lys	Gly	Val	Arg	Phe	Ile	His	Gln	Tyr	Lys		
				50					55					60			
Glu	Glu	Leu	Glu	Thr	Ala	Ser	Lys	Phe	Ala	Tyr	Leu	Gly	Leu	Cys	Thr		
65						70									80		

Leu Val Gly Ser Lys Thr Leu Gly Glu Glu Tyr Thr Asn Leu Met Tyr		
85	90	95
Thr Ile Arg Asp Arg Thr Ala Leu Pro Gly Val Val Arg Arg Phe Gly		
100	105	110
Tyr Val Leu Ser Asn Thr Leu Phe Pro Tyr Leu Phe Val Arg Tyr Met		
115	120	125
Gly Lys Leu Arg Ala Lys Leu Met Arg Glu Tyr Pro His Leu Val Glu		
130	135	140
Tyr Asp Glu Asp Glu Pro Val Pro Ser Pro Glu Thr Trp Lys Glu Arg		
145	150	155
Val Ile Lys Thr Phe Val Asn Lys Phe Asp Lys Phe Thr Ala Leu Glu		
165	170	175
Gly Phe Thr Ala Ile His Leu Ala Ile Phe Tyr Val Tyr Gly Ser Tyr		
180	185	190
Tyr Gln Leu Ser Lys Arg Ile Trp Gly Met Arg Tyr Val Phe Gly His		
195	200	205
Arg Leu Asp Lys Asn Glu Pro Arg Ile Gly Tyr Glu Met Leu Gly Leu		
210	215	220
Leu Ile Phe Ala Arg Phe Ala Thr Ser Phe Val Gln Thr Gly Arg Glu		
225	230	235
Tyr Leu Gly Ala Leu Leu Glu Lys Ser Val Glu Lys Glu Ala Gly Glu		
245	250	255
Lys Glu Asp Glu Lys Glu Ala Val Val Pro Lys Lys Lys Ser Ser Ile		
260	265	270
Pro Phe Ile Glu Asp Thr Glu Gly Glu Thr Glu Asp Lys Ile Asp Leu		
275	280	285
Glu Asp Pro Arg Gln Leu Lys Phe Ile Pro Glu Ala Ser Arg Ala Cys		
290	295	300
Thr Leu Cys Leu Ser Tyr Ile Ser Ala Pro Ala Cys Thr Pro Cys Gly		
305	310	315
His Phe Phe Cys Trp Asp Cys Ile Ser Glu Trp Val Arg Glu Lys Pro		
325	330	335
Glu Cys Pro Leu Cys Arg Gln Gly Val Arg Glu Gln Asn Leu Leu Pro		
340	345	350
Ile Arg		
<210>25		
<211>38		
<212>PRT		
<213> 解脂耶氏酵母		

<220>

<221>misc_feature

<222>(2)..(3)

<223>Xaa 可以是任何天然存在的氨基酸

<220>

<221>misc_feature

<222>(5)..(15)

<223>Xaa 可以是任何天然存在的氨基酸

<220>

<221>misc_feature

<222>(17)..(17)

<223>Xaa 可以是任何天然存在的氨基酸

<220>

<221>misc_feature

<222>(19)..(20)

<223>Xaa 可以是任何天然存在的氨基酸

<220>

<221>misc_feature

<222>(22)..(23)

<223>Xaa 可以是任何天然存在的氨基酸

<220>

<221>misc_feature

<222>(25)..(34)

<223>Xaa 可以是任何天然存在的氨基酸

<220>

<221>misc_feature

<222>(36)..(37)

<223>Xaa 可以是任何天然存在的氨基酸

<400>25

Cys Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys

1

5

10

15

Xaa His Xaa Xaa Cys Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

20

25

30

Xaa Xaa Cys Xaa Xaa Cys

35

<210>26

<211>345

<212>PRT

<213>解脂耶氏酵母

<400>26

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

Thr Glu Asp Lys Ile Asp Leu Glu Asp Pro Arg Gln Leu Lys Phe Ile
 305 310 315 320
 Pro Glu Ala Ser Arg Ala Cys Thr Leu Cys Leu Ser Tyr Ile Ser Ala
 325 330 335
 Pro Ala Cys Thr Pro Cys Gly His Phe
 340 345

<210>27

<211>2987

<212>DNA

<213> 解脂耶氏酵母

<220>

<221>misc_feature

<223> 具有 W497L 突变的突变型乙酰羟酸合酶 (AHAS)

<300>

<302> 解脂耶氏酵母的高二十碳五烯酸生产菌株

<310>US 2006-0115881-A1

<311>2005-11-02

<312>2006-06-01

<313>(1).. (2987)

<300>

<302> 解脂耶氏酵母的高二十碳五烯酸生产菌株

<310>W0 2006/052870

<311>2005-11-03

<312>2006-05-18

<313>(1).. (2987)

<400>27

ttccctagtc ccagtgtaca cccgccgata tcgcttaccc tgcagccgga ttaaggttgg	60
caatttttca cgtccttgtc tccgcaatta ctcaccgggt ggttttataag attgcaagcg	120
tcttgatttg tctctgtata ctaacatgca atcgcgactc gcccgcacggg ccactaacct	180
ggccagaatc tccagatcca agtattctct tgggtctgcga tatgtttcca acacaaaagc	240
ccctgctgcc cagccggcaa ctgctgagtg agtatctcct gccataaacg acccagaacc	300
actgtatagt gtttggaagc actagtcaga agaccagcga aaacaggtgg aaaaaactga	360
gacgaaaagc aacgaccaga aatgtaatgt gtggaaaagc gacacacaca gagcagataa	420
agaggtgaca aataacgaca aatgaaatat cagtatcttc ccacaatcac tacctctcag	480
ctgtctgaag gtgcggctga tatatccatc ccacgtctaa cgtatggagt gtgatagaat	540
atgacgacac aagcatgaga actcgctctc tatccaacca ccgaaacact gtcactacag	600
ccgttcttgt tgcctcattc gcttttgtga ttccatgcct tctctgggtga ctgacaacat	660
tccttctttt tctccagccc tgttgttate tgctcatgac ctacggccac tctctatcgc	720
ataactaacat agacgatecc agcccgtcc ccacttccag ggcaccgttg gcaagcctcc	780

tatcctcaag aaggctgagg ctgccaacgc tgacatggac gagtccttca tcggaatgtc	840
tggaggagag atcttccacg agatgatgct gcgacacaac gtcgacactg tcttcgggta	900
ccccggtgga gccattctcc ccgtctttga cgccattcac aactctgagt acttcaactt	960
tgtgtctccct cgacacgagc aggggtgccg ccacatggcc gagggctacg ctcgagcctc	1020
tggtaagccc ggtgtcgttc tcgtcacctc tggccccggt gccaccaacg tcataccccc	1080
catgcaggac gctctttccg atgggtacccc catggttgtc ttcaccgggc aggtcctgac	1140
ctccgttata ggcactgacg ctttccagga ggccgatgtt gtcggcatct cccgatcttg	1200
caccaagtgg aacgtcatgg tcaagaacgt tgcctgagctc ccccgacgaa tcaacgaggc	1260
ctttgagatt gctacttccg gccgacccgg tcccgttctc gtcgatctgc ccaaggatgt	1320
tactgtctgcc atcctgcgag agcccatccc caccaagtcc accattccct cgcattctct	1380
gaccaacctc acctctgccg ccgccaccga gttccagaag caggctatcc agcgagccgc	1440
caacctcatc aaccagtcca agaagccgt cctttacgtc ggacagggta tccttggctc	1500
cgaggagggt cctaagctgc ttaaggagct ggctgagaag gccgagattc ccgtcaccac	1560
tactctgcag ggtcttgggtg cctttgacga gcgagacccc aagtctctgc acatgctcgg	1620
tatgcacggt tccggctacg ccaacatggc catgcagaac gctgactgta tcattgctct	1680
cggcgccga tttgatgacc gagttaccgg ctccatcccc aagtttgccc ccgaggctcg	1740
agccgctgcc cttgagggtc gaggtgggtat tgttacttt gagatccagg ccaagaacat	1800
caacaaggtt gttcaggcca ccgaagccgt tgaggggagac gttaccgagt ctgtccgaca	1860
gctcatcccc ctcatcaaca aggtctctgc cgctgagcga gctccctgga ctgagactat	1920
ccagtcctgg aagcagcagt tccccttccct ctccgaggct gaaggtgagg atggtgttat	1980
caagccccag tccgtcattg ctctgtctct tgacctgaca gagaacaaca aggacaagac	2040
catcatcacc accggtgttg gtcagcatca gatgtggact gccagcatt tccgatggcg	2100
acaccctcga accatgatca cttctgggtg tcttggaaact atgggttacg gcctgcccgc	2160
cgctatcggc gccaaaggttg cccgacctga ctgcgacgtc attgacatcg atggtgacgc	2220
ttctttcaac atgactctga ccgagctgtc caccgccgtt cagttcaaca ttggcgtcaa	2280
ggctattgtc ctcaacaacg aggaacagggt tatgggtacc cagctgcagt ctctcttcta	2340
cgagaaccga tactgccaca ctcatcagaa gaaccccgac ttcatagaagc tggccgagtc	2400
catgggcatg aagggtatcc gaatactca cattgaccag ctggaggccg gtctcaagga	2460
gatgctcgca tacaagggcc ctgtgctcgt tgaggttgtt gtcgacaaga agatccccgt	2520
tcttcccatg gttcccgtg gtaaggcttt gcatgagttc cttgtctacg acgctgacgc	2580
cgaggctgct tctcgaccg atcgactgaa gaatgcccc gccctcacg tccaccagac	2640
cacctttgag aactaagtgg aaaggaacac aagcaatccg aaccaaaaaat aattggggtc	2700
ccgtgcccac agagtctagt gcagacctaa aatgaccaca gtaaattata gctgttatta	2760
aacatgagat tttgaccaac aagagcgtag gaatgttatt agctactact tgtacataca	2820
cagcatttgt tttaaataat gttgcctcca ggggcagtga gatcaggacc cagatccgtg	2880
gccagctctc tgacttcaga ccgcttgtag ttaagcagct cgcaacactg ttgtcgagga	2940
ttgaacttgc catattcgat tttgtgggtca tgaatccagc acacctc	2987

<210>28

<211>14688

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pZKLeuN-29E3

<400>28

cgattgttgt	ctactaacta	tcgtacgata	acttcgtata	gcatacatta	tacgaagtta	60
tcgcgtcgac	gagtatctgt	ctgactcgtc	attgccgcct	ttggagtacg	actccaacta	120
tgagtgtgct	tggatcactt	tgacgataca	ttcttcgttg	gaggctgtgg	gtctgacagc	180
tgcgttttcg	gcgcggttgg	ccgacaacaa	tatcagctgc	aacgtcattg	ctggctttca	240
tcatgatcac	atTTTTgtcg	gcaaaggcga	cgcccagaga	gccattgacg	ttcttttctaa	300
tttggaccga	tagccgtata	gtccagtcta	tctataagtt	caactaactc	gtaactatta	360
ccataacata	tacttccactg	ccccagataa	ggttccgata	aaaagttctg	cagactaaat	420
ttatttccagt	ctcctcttca	ccaccaaaat	gccctcctac	gaagctcgag	ctaacgtcca	480
caagtcgcgc	tttgccgctc	gagtgtctca	gtctgtggca	gccaaagaaa	ccaacctgtg	540
tgcttctctg	gatgttacca	ccaccaagga	gtcattgag	cttgccgata	aggctcggacc	600
ttatgtgtgc	atgatcaaaa	cccatatcga	catcattgac	gacttcacct	acgccggcac	660
tgtgtctccc	ctcaaggaac	ttgtctctta	gcacggtttc	ttcctgttcg	aggacagaaa	720
gttcgcagat	attggcaaca	ctgtcaagca	ccagtaccgg	tgtcaccgaa	tcgccgagtg	780
gtccgatata	accaacgccc	acgggtgtacc	cggaaccgga	atcattgctg	gcctgcgagc	840
tggtgccgag	gaaactgtct	ctgaacagaa	gaaggaggac	gtctctgact	acgagaactc	900
ccagtacaag	gagttcctag	tccccctctcc	caacgagaag	ctggccagag	gtctgtctcat	960
gctggccgag	ctgtctttgca	agggctctct	ggccactggc	gagtactcca	agcagacccat	1020
tgagcttgcc	cgatccgacc	ccgagtttgt	ggttggcttc	attgcccaga	accgacctaa	1080
gggcgactct	gaggactggc	ttattctgac	ccccgggggtg	ggtcttgacg	acaagggaga	1140
cgctctcgga	cagcagtacc	gaactgttga	ggatgtcatg	tctaccggaa	cggatatcat	1200
aattgtcggc	cgaggtctgt	acggccagaa	ccgagatcct	attgaggagg	ccaagcgata	1260
ccagaaggct	ggctgggagg	cttaccagaa	gattaactgt	tagaggttag	actatggata	1320
tgtaatttaa	ctgtgtatat	agagagcgtg	caagtatgga	gcgcttgttc	agcttgtatg	1380
atggtcagac	gacctgtctg	atcgagtatg	tatgatactg	cacaacctgt	gtatccgcct	1440
gatctgtcca	atggggcatg	ttgttgtgtt	tctcgatacg	gagatgctgg	gtacagtgtc	1500
aatacgttga	actacttata	cttatatgag	gtcgaagaa	agctgacttg	tgtatgactt	1560
attctcaact	acatccccag	tcacaatacc	accactgcac	taccactaca	ccaaaaccat	1620
gatcaaacca	cccatggact	tcctggaggc	agaagaactt	gttatggaaa	agctcaagag	1680
agagatcata	acttcgtata	gcatacatta	tacgaagtta	tcctgcagg	aaaggaattc	1740
tggagtttct	gagagaaaaa	ggcaagatac	gtatgtaaca	aagcgacgca	tggtaacaata	1800
ataccggagg	catgtatcat	agagagttag	tggttcgatg	atggcactgg	tgcctgggat	1860
gactttatac	ggctgactac	atatttgtcc	tcagacatac	aattacagtc	aagcacttac	1920
ccttggacat	ctgtaggtac	ccccgggcca	agacgatctc	agcgtgtcgt	atgtcggatt	1980
ggcgtagctc	cctcgctcgt	caattggctc	ccatctactt	tcttctgctt	ggctacaccc	2040

agcatgtctg	ctatggctcg	ttttcgtgcc	ttatctatcc	tcccagtatt	accaactcta	2100
aatgacatga	tgtgattggg	tctacacitt	catatcagag	ataaggagta	gcacagttgc	2160
ataaaaagcc	caactctaata	cagcttcttc	ctttcttgta	attagtacaa	aggtgattag	2220
cgaatcttg	aagcttagtt	ggccctaaaa	aatcaaaaa	aagcaaaaaa	cgaaaaacga	2280
aaaaccacag	ttttgagaac	aggagggtaa	cgaaggatcg	tatatatata	tatatatata	2340
tatacccacg	gateccgaga	ccggcctttg	attcttccct	acaaccaacc	attctcacca	2400
ccctaattca	caacatgga	gtctggacce	atgcctgctg	gcattccctt	ccctgagtac	2460
tatgacttct	ttatggactg	gaagactccc	ctggccatcg	ctgccaccta	cactgctgcc	2520
gtcggctctt	tcaaccccaa	ggttggcaag	gtctcccag	tggttgcaa	gtcggctaac	2580
gcaaacgctg	ccgagcgaac	ccagtcggga	gctgccatga	ctgccttcgt	ctttgtgcac	2640
aacctcattc	tgtgtgtcta	ctctggcate	accttctact	acatgtttcc	tgctatggtc	2700
aagaacttcc	gaaccacac	actgcacgaa	gcctactgcg	acacggatca	gtccctctgg	2760
aacaacgcac	ttggctactg	gggttacctc	ttctacctgt	ccaagttcta	cgaggtcatt	2820
gacaccatca	tcatcatcct	gaaggagcga	cggctcctgc	tgtctcagac	ctaccacat	2880
gctggagcca	tgattacat	gtggtctggc	atcaactacc	aagccactcc	catttggatc	2940
tttgtggtct	tcaactcctt	cattcacacc	atcatgtact	gttactatgc	cttcacctct	3000
atcgatttcc	atcctcctgg	caaaaagtac	ctgacttcga	tgcagattac	tcagtttctg	3060
gtcggtatca	ccattgccgt	gtcctacctc	ttcgttccctg	gctgcatccg	aacacccggt	3120
gctcagatgg	ctgtctggat	caacgtcggc	tacctgtttc	ccttgacctc	tctgttcgtg	3180
gactttgcca	agcgaacctc	ctccaagcga	tctgccattg	ccgctcagaa	aaaggctcag	3240
taagcggccg	cattgatgat	tggaaacaca	cacatgggtt	atatctaggt	gagagttagt	3300
tggacagtta	tatatataat	cagctatgcc	aacggtaact	tcattcatgt	caacgaggaa	3360
ccagtactg	caagtaatat	agaatttgac	caccttgcca	ttctcttgca	ctcctttact	3420
atatctcatt	tatttcttat	atacaaatca	cttcttcttc	ccagcatcga	gctcggaaac	3480
ctcatgagca	ataacatcgt	ggatctcgtc	aatagagggc	tttttggaact	ccttgctggt	3540
ggccaccttg	tccttgctgt	ctggctcatt	ctgtttcaac	gccttttaac	taacggagta	3600
ggtctcgggtg	tcggaagcga	cgcagatcc	gtcatcctcc	tttcgctctc	caaagtagat	3660
acctccgacg	agctctcgga	caatgatgaa	gtcgggtgcc	tcaacgtttc	ggatggggga	3720
gagatcggcg	agcttgggcg	acagcagctg	gcagggtcgc	aggttggcgt	acaggttcag	3780
gtcctttcgc	agcttgagga	gacctgtctc	gggtcgcacg	tcggttcgtc	cgtcgggagt	3840
ggtccatacg	gtgttggcag	cgcctccgac	agcaccgagc	ataatagagt	cagcctttcg	3900
gcagatgtcg	agagtagcgt	cggtagatgg	ctcgccctcc	ttctcaatgg	cagctcctcc	3960
aatgagtcgg	tcctcaaaac	caaactcggt	gccggaggcc	tcagcaacag	acttgagcac	4020
cttgacggcc	tcggcaatca	cctcggggcc	acagaagtcg	ccgccgagaa	gaacaatctt	4080
cttggagtca	gtcttgggtt	tcttagtttc	gggttccatt	gtggatgtgt	gtggttgtat	4140
gtgtgatgtg	gtgtgtggag	tgaatatctg	tggctggcaa	acgctcttgt	atatatacgc	4200
acttttgccc	gtgctatgtg	gaagactaaa	cctccgaaga	ttgtgactca	ggtagtgccg	4260
tatcggttag	ggacccaaac	cttgctgatg	ccgatagcat	gcgacgtcgg	gcccatttcg	4320
ccctatagtg	agtcgtatta	caattcactg	gccgtcgttt	tacaacgtcg	tgactgggaa	4380

aaccctggcg	ttaccecaact	taategcett	gcagcacatc	cccctttcgc	cagctggcgt	4440
aatagegaag	aggccccgac	cgatcgccct	tcccaacagt	tgcgcagcct	gaatggcgaa	4500
tggacgcgcc	ctgtageggc	gcattaagcg	cggcgggtgt	ggtggttacg	cgcagcgtga	4560
ccgtacact	tgccagegcc	ctagegcccc	ctcctttcgc	tttcttccct	tcctttctcg	4620
ccacgttcgc	cggttttccc	cgteaagctc	taaatcgggg	gctcccttta	gggttccgat	4680
ttagtgcttt	acggcacctc	gacccccaaa	aacttgatta	gggtgatggg	tcacgtagtg	4740
ggccatcgcc	ctgatagacg	gtttttcgcc	ctttgacggt	ggagtccacg	ttctttaata	4800
gtggactctt	gttccaaact	ggaacaacac	tcaaccctat	ctcgggtctat	tcttttgatt	4860
tataagggat	tttgccgatt	tcggcctatt	ggttaaaaaa	tgagctgatt	taacaaaaat	4920
ttaacgcgaa	ttttaacaaa	atattaacgc	ttacaatttc	ctgatgcggg	attttctcct	4980
tacgcatctg	tgcggtatth	cacaccgcat	cagggtggcac	ttttcgggga	aatgtgcgcg	5040
gaacccttat	ttgtttatth	ttctaaatac	attcaaatat	gtatccgctc	atgagacaat	5100
aaccctgata	aatgcttcaa	taatatlgaa	aaaggaagag	tatgagtatt	caacatttcc	5160
gtgtcgccct	tattccctth	tttgccgcat	tttgcccttc	tgthttttgct	caccagaaaa	5220
cgctgggtgaa	agtaaaagat	gctgaagatc	agttgggtgc	acgagtgggt	tacatcgaac	5280
tggatctcaa	cagcggtaag	atccttgaga	gttttcgccc	cgaagaacgt	tttccaatga	5340
tgagcactth	taaagtthctg	ctatgtggcg	cgggtattatc	ccgtattgac	gccgggcaag	5400
agcaactcgg	tcgccgcata	cactattctc	agaatgactt	ggttgagtac	tcaccagtca	5460
cagaaaagca	tcttacggat	ggcatgacag	taagagaatt	atgcagtgtc	gccataacca	5520
tgagtataaa	cactgcggcc	aacttacttc	tgacaacgat	cggaggaccg	aaggagctaa	5580
ccgctthttt	gcacaacatg	ggggatcatg	taactcgcct	tgatcgthtg	gaaccggagc	5640
tgaatgaagc	cataccaaac	gacgagcgtg	acaccacgat	gcctgtagca	atggcaacaa	5700
cgttgcgcaa	actattaact	ggcgaactac	ttactctagc	ttcccggcaa	caattaatag	5760
actggatgga	ggcggataaa	gttgcaggac	cacttctgcg	ctcggccctt	ccggctggct	5820
ggtthattgc	tgataaatct	ggagccgggtg	agcgtgggtc	tcgcggtatc	attgcagcac	5880
tggggccaga	tggtaagccc	tcccgtatcg	tagttatcta	cacgacgggg	agtcaggcaa	5940
ctatggatga	acgaaataga	cagatcgctg	agatagggtc	ctcactgatt	aagcattggt	6000
aactgtcaga	ccaagthttac	tcatatatac	thtagattga	thtaaaactt	cattthttaat	6060
thtaaaaggat	ctaggtgaag	atcctthttg	ataatctcat	gaccaaatac	ccttaacgtg	6120
agthtttcgt	ccactgagcg	tcagaccccc	tagaaaagat	caaaggatct	tcttgagatc	6180
ctthttthct	gcgcgtaatc	tgctgcttgc	aaacaaaaaa	accaccgcta	ccagcggthg	6240
thtgthtgcc	ggatcaagag	ctaccaactc	thtttccgaa	ggtaactggc	thcagcagag	6300
cgcagatacc	aaatactgtt	cttctagthg	agccgtagth	aggccaccac	thcaagaact	6360
ctgtagcacc	gcctacatac	ctcgtcttgc	taatcctgtt	accagtggct	gctgccagth	6420
gcgataagtc	gtgtcttacc	gggttggact	caagacgata	gttaccggat	aaggcgcagc	6480
ggtcgggctg	aacggggggg	tcgtgcacac	agcccagctt	ggagcgaacg	acctacaccg	6540
aactgagata	cctacagcgt	gagctatgag	aaagcgccac	gcttcccga	gggagaaaagg	6600
cggacaggta	tccggtaagc	ggcagggtcg	gaacaggaga	gcgcacgagg	gagcttccag	6660
ggggaaacgc	ctggtatctt	tatagtcctg	tcgggtthtcg	ccacctctga	cttgagcgtc	6720

gatttttgtg atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct	6780
ttttacgggtt cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc	6840
ctgattctgt ggataaccgt attaccgcct ttgagtgage tgataccgct cgccgcagcc	6900
gaacgaccga ggcgcagcag tcagtgagcg aggaagcgga agagcgccca atacgcaaac	6960
cgcctctccc cgcgcgttgg ccgattcatt aatgcagctg gcgcgcccac tgagctcgtc	7020
taacggactt gatatacaac caattaaaac aaatgaaaag aaatacagtt ctttgtatca	7080
tttgtaacaa ttaccctgta caaactaagg tattgaaatc ccacaatatt cccaaagtcc	7140
acccttttcc aaattgtcat gcctacaact catataccaa gcactaacct accaaacacc	7200
actaaaaccc cacaaaatat atcttaccga atatacagta acaagctacc accacactcg	7260
ttgggtgcag tcgccagctt aaagatatct atccacatca gccacaactc ccttccttta	7320
ataaaccgac tacacccttg gctattgagg ttatgagtga atatactgta gacaagacac	7380
tttcaagaag actgtttcca aaacgtacca ctgtcctcca ctacaaacac acccaatctg	7440
cttcttctag tcaaggttgc tacaccggta aattataaat catcatttca ttagcagggc	7500
agggcccttt ttatagagtc ttatacacta gcggaccctg ccggtagacc aaccgcgagg	7560
cgcgtcagtt tgctccttcc atcaatgcgt cgtagaaacg acttactcct tcttgagcag	7620
ctccttgacc ttgtttggcaa caagtctccg acctcggagg tggaggaaga gcctccgata	7680
tcggcggtag tgataccagc ctgcacggac tccttgacgg cagcctcaac agcgtcaccg	7740
gcgggcttca tgttaagaga gaacttgagc atcatggcgg cagacagaat ggtggcgtag	7800
gcaactaaca tgaatgaata cgatatacat caaagactat gatacgcagt attgcacact	7860
gtacgagtaa gagcactagc cactgcactc aagtgaacc gttgcccggg tacgagtatg	7920
agtatgtaca gtatgtttag tattgtactt ggacagtgtc tgtatcgtac attctcaagt	7980
gtcaaacata aatatccgtt gctatatcct cgcaccacca cgtagctcgc tataatccctg	8040
tgttgaatcc atccatcttg gattgccaat tgtgcacaca gaaccgggca ctcaattccc	8100
catccacact tgcggccgct taagcaacgg gcttgataac agcggggggg gtgcccacgt	8160
tgttgcggtt gcggaagaac agaacaccct taccagcacc ctgcgcacca gcgctgggct	8220
caaccactg gcacatacgc gcaactgcgt acatggcgcg gatgaagcca cgaggaccat	8280
cctggacatc agcccggtag tgcttgccca tgatgggctt aatggcctcg gtggcctcgt	8340
ccgcgttgta gaaggggatg ctgctgacgt agtgggtggag gacatgagtc tcgatgatgc	8400
cgtggagaag gtggcggccg atgaagccca tctcacggtc aatggtagca gcggcaccac	8460
ggacgaagtt ccactcgtcg ttgggtgtagt ggggaagggt agggctcggg tgctggagga	8520
aggtgatggc aacgagccag tggttaaccc agaggtaggg aacaaagtac cagatggcca	8580
tgttgtagaa accgaacttc tgaacgagga agtacagagc agtggccatc agaccgatac	8640
caatategct gaggacgatg agcttagcgt cactgttctc gtacagaggg ctgcggggat	8700
cgaagtgggt aacaccaccg ccgaggccgt tatgcttgcc cttgccgcga ccctcacgct	8760
ggcgctcgtg gtagttgttg ccggtaacat tgggtgatgag gtagttgggc cagccaacga	8820
gctgctgaag gacgagcatg agaagagtga aagcgggggt ctccctcagta agatgagcga	8880
gctcgtgggt catctttccg agacgagtag cctgctgctc gcgggttcgg ggaacgaaga	8940
ccatgtcacg ctccatgttg ccagtggcct tgtgggtgctt tcgggtgggag atttgccagc	9000
tgaagtaggg gacaaggagg gaagagtga gaaccagcc agtaatgtcg ttgatgatgc	9060

gagaatcgga	gaaagcaccg	tgaccgcact	catgggcaat	aacccagaga	ccagtaccga	9120
aaagaccctg	aagaacggtg	tacacggccc	acagaccagc	gcgggcgggg	gtggagggga	9180
tatatteggg	ggtcacaaag	ttgtaccaga	tgctgaaagt	ggtagtcagg	aggacaatgt	9240
cgcggaggat	ataaccgtat	cccttgagag	cggagcgctt	gaagcagtgc	ttagggatgg	9300
cattgtagat	gtccttgatg	gtaaagtcgg	gaacctcgaa	ctggttgccg	taggtgtcga	9360
gcatgacacc	atactcggac	ttgggcttgg	cgatatcaac	ctcggacatg	gacgagagcg	9420
atgtggaaga	ggccgagtg	cggggagagt	ctgaaggaga	gacggcggca	gactcagaat	9480
ccgtcacagt	agttgaggtg	acgggtgcgtc	taagcgcagg	gttctgcttg	ggcagagccg	9540
aagtggacgc	catggttgat	gtgtgtttta	ttcaagaatg	aatatagaga	agagaagaag	9600
aaaaaagatt	caattgagcc	ggcgatgcag	acccttatat	aaatgttgcc	ttggacagac	9660
ggagcaagcc	cgcccaaacc	tacgttcggt	ataatatggt	aagcttttta	acacaaaggt	9720
ttggcttggg	gtaacctgat	gtgggtgcaa	agaccgggcg	ttggcgagcc	attgcgcggg	9780
cgaatggggc	cgtgactcgt	ctcaaattcg	agggcgtgcc	tcaattcgtg	cccccggtggc	9840
tttttccgc	cgtttccgcc	ccgtttgcac	cactgcagcc	gcttcttttg	ttcggacacc	9900
ttgctgcgag	ctaggtgcct	tgtgctactt	aaaaagtggc	ctcccaacac	caacatgaca	9960
tgagtgcgtg	ggccaagaca	cgttggcggg	gtcgcagtcg	gctcaatggc	ccggaaaaaa	10020
cgctgctgga	gctggttcgg	acgcagtcgc	ccgcggcgta	tggatatccg	caaggttcca	10080
tagcgccatt	gccctccgtc	ggcgtctatc	ccgcaacctc	taaatagagc	gggaatataa	10140
cccaagcttc	ttttttttcc	tttaacacgc	acacccccaa	ctatcatggt	gctgctgctg	10200
tttgactcta	ctctgtggag	gggtgctccc	acccaacca	acctacaggt	ggatccggcg	10260
ctgtgattgg	ctgataagtc	tcctatccgg	actaattctg	accaatggga	catgcgcgca	10320
ggacccaaat	gcegcgaatta	cgtaacccca	acgaaatgcc	taccctctct	tggagcccag	10380
cggccccaaa	tcccccaag	cagcccgggt	ctaccggctt	ccatctccaa	gcacaagcag	10440
cccggttcta	ccggcttcca	tctccaagca	cccccttctc	cacacccac	aaaaagaccc	10500
gtgcaggaca	tcctactgcg	tcgacatcat	ttaaattcct	tcacttcaag	ttcattcttc	10560
atctgcttct	gttttacttt	gacaggcaaa	tgaagacatg	gtacgacttg	atggaggcca	10620
agaacgccat	ttcaccccg	gacaccgaag	tgccgtgaaat	cctggctgcc	cccattgata	10680
acatcgga	ctacggtatt	ccggaaagtg	tatatagaac	ctttccccag	cttgtgtctg	10740
tggatatgga	tgggtgtaac	ccctttgagt	actcgtcttg	gcttctctcc	gagcagtatg	10800
aggctctcta	atctagcgca	tttaatatct	caatgtatct	atatatttat	cttctcatgc	10860
ggccgctcac	tgaatctttt	tggtctccct	gtgcttccctg	acgatatacg	tttgacata	10920
gaaattcaag	aacaaacaca	agactgtgcc	aacataaaaag	taattgaaga	accagccaaa	10980
catctctate	ccatcttggc	gataacaggg	aatgttccctg	tacttccaga	caatgtagaa	11040
accaacattg	aattgaatga	tctgcattga	tgtaatcagg	gattttggca	tggggaactt	11100
cagcttgatc	aatctgggtc	aataataacc	gtacatgac	cagtggatga	aaccattcaa	11160
cagcaca	atccaaacag	cttcatttctg	gtaattatag	aacagccaca	tatccatcgg	11220
tgccccaaa	tgatggaaga	attgcaacca	ggtcagaggc	ttgcccata	gtggcaata	11280
gaaggagtca	atatactcca	ggaacttgct	caaatagaac	aactgcgtgg	tgatcctgaa	11340
gacgttggtg	tcaaaagcct	tctcgcagtt	gtcagacata	acaccgatgg	tgtacatggc	11400

atatgccatt	gagaggaatg	atcccaacga	ataaatggac	atgagaaggt	tgtaattggt	11460
gaaaacaaac	ttcatacgag	actgaccttt	tggaccaagg	gggccaagag	tgaacttcaa	11520
gatgacaaat	gcgatggaca	agtaaagcac	ctcacagtga	ctggcatcac	tccagagttg	11580
ggcataatca	actggttggg	taaaacttcc	tgcccaattg	agactatttc	attcaccacc	11640
tccatggcca	ttgctgtaga	tatgtcttgt	gtgtaagggg	gttgggggtg	ttgttttgtt	11700
tcttgacttt	tgtgttagca	agggaagacg	ggcaaaaaag	tgagtgtggt	tgggagggag	11760
agacgagcct	tatatataat	gcttgtttgt	gtttgtgcaa	gtggacgccg	aaacgggcag	11820
gagccaaact	aaacaaggca	gacaatgcga	gcttaattgg	attgcctgat	gggcaggggt	11880
tagggctcga	tcaatggggg	tgcgaagtga	caaaattggg	aattagggtc	gcaagcaagg	11940
ctgacaagac	tttggcccaa	acatttgtac	gcggtggaca	acaggagcca	cccatcgtct	12000
gtcacgggct	agccggtcgt	gcgtcctgtc	aggctccacc	taggctccat	gccactccat	12060
acaatcccac	tagtgtaccg	ctaggccgct	tttagctccc	atctaagacc	cccccaaac	12120
ctccactgta	cagtgcactg	tactgtgtgg	cgatcaaggg	caagggaana	aaggcgcaaa	12180
catgcacgca	tggaatgacg	taggtaaggc	gttactagac	tgaaaagtgg	cacatttcgg	12240
cgtgccaaag	ggtcctaggt	gcgtttcgcg	agctgggcgc	caggccaagc	cgctccaaaa	12300
cgcctctccg	actccctcca	gcggcctcca	tatccccatc	cctctccaca	gcaatgttgt	12360
taagccttgc	aaacgaaaaa	atagaaaggc	taataagctt	ccaatattgt	ggtgtacgct	12420
gcataacgca	acaatgagcg	caaacaaca	cacacacaca	gcacacagca	gcattaacca	12480
cgatgaacag	catgacatta	caggtgggtg	tgtaatcagg	gccctgattg	ctgggtgggtg	12540
gagcccccac	catgggcaga	tctgcgtaca	ctgttttaac	agtgtacgca	gatctactat	12600
agaggaacat	ttaaattgcc	ccggagaaga	cggccaggcc	gcctagatga	caaattcaac	12660
aactcacagc	tgactttctg	ccattgccac	tagggggggg	cctttttata	tggccaagcc	12720
aagctctcca	cgtcggttgg	gctgcaccca	acaataaatg	ggtaggggtg	caccaacaaa	12780
gggatgggat	ggggggtaga	agatacgagg	ataacggggc	tcaatggcac	aaataagaac	12840
gaatactgcc	attaagactc	gtgatccagc	gactgacacc	attgcatcat	ctaagggcct	12900
caaaactacc	tcggaactgc	tgcgctgata	tggacaccac	agaggttccg	agcactttag	12960
gttgcaccaa	atgtcccacc	aggtgcaggc	agaaaacgct	ggaacagcgt	gtacagtttg	13020
tcttaacaaa	aagttagggc	gctgaggtcg	agcagggttg	tgtgacttgt	tatagccttt	13080
agagctgcga	aagcgcgtat	ggatttggct	catcaggcca	gattgagggt	ctgtggacac	13140
atgtcatgtt	agtgtacttc	aatcgccccc	tggatatagc	cccgacaata	ggccgtggcc	13200
tcattttttt	gccttccgca	catttccatt	gctcgatacc	cacaccttgc	ttctcctgca	13260
cttgccaacc	ttaatactgg	tttacattga	ccaacatctt	acaagcgggg	ggcttgtcta	13320
gggtatatat	aaacagtggc	tctcccaatc	ggttgccagt	ctcttttttc	ctttctttcc	13380
ccacagattc	gaaatctaaa	ctacacatca	cagaattccg	agccgtgagt	atccacgaca	13440
agatcagtgt	cgagacgacg	cgtttttgtt	aatgacacaa	tccgaaagtc	gctagcaaca	13500
cacactctct	acacaaacta	accagctctt	ggtaccatgg	aggctcgtga	cgaaatcgct	13560
tccattggcc	aggaggttct	tccaaggctc	gactatgctc	agctctggct	tgatgcctcg	13620
cactgcgagg	tgtgtacctt	ctccatcgcc	ttcgtcatcc	tgaagttcac	ccttggctcct	13680
ctcggaccca	agggtcagtc	tcgaatgaag	tttgtgttca	ccaactacaa	cctgctcatg	13740

```

tccatctact cgctgggctc cttectctct atggcctacg ccatgtacac cattggtgtc 13800
atgtccgaca actgcgagaa ggctttcgac aacaatgtct tccgaatcac cactcagctg 13860
ttctacctca gcaagttcct cgagtacatt gactccttct atctgccccct catgggcaag 13920
cctctgacct gggtgcagtt ctttcacat ctcggagctc ctatggacat gtggctgttc 13980
tacaactacc gaaacgaagc cgtttgatc tttgtgctgc tcaacggctt cattcactgg 14040
atcatgtacg gctactattg gacccgactg atcaagctca agttccctat gccaagtcc 14100
ctgattactt ctatgcagat cattcagttc aacgttggtt tctacatcgt ctggaagtac 14160
cggaacattc cctgctaccg acaagatgga atgagaatgt ttggctgggt tttcaactac 14220
ttctacgttg gtactgtcct gtgtctgttc ctcaacttct acgtgcagac ctacatcgtc 14280
cgaaagcaca agggagccaa aaagattcag tgagcggccg catgtacata caagattatt 14340
tatagaaatg aatcgcgatc gaacaaagag tacgagtgtg cgagtagggg atgatgataa 14400
aagtgggaaga agttccgcat ctttgatttt atcaacgtgt aggacgatac ttcctgtaaa 14460
aatgcaatgt ctttaccata ggttctgctg tagatgttat taactaccat taacatgtct 14520
acttgtacag ttgcagacca gttggagtat agaatggtag acttaccaaa aagtgttgat 14580
ggttgtaact acgatatata aaactgttga cgggatcccc gctgatatgc ctaaggaaca 14640
atcaaagagg aagatattaa ttcagaatgc tagtatacag ttagggat 14688

```

<210>29

<211>1434

<212>DNA

<213> 串珠镰孢菌

<220>

<221>CDS

<222>(1).. (1434)

<223> Δ -12 去饱和酶

<300>

<302> 适于改变含油酵母中多不饱和脂肪酸含量的 Δ -12 去饱和酶

<310>WO 2005/047485

<311>2004-11-12

<312>2005-05-26

<313>(1).. (1434)

<300>

<302> 适于改变含油酵母中多不饱和脂肪酸含量的 Δ -12 去饱和酶

<310>US 2005-0216975-A1

<311>2004-11-10

<312>2005-09-29

<313>(1).. (1434)

<400>29

```

atg gcg tcc act tcg gct ctg ccc aag cag aac cct gcg ctt aga cgc      48
Met Ala Ser Thr Ser Ala Leu Pro Lys Gln Asn Pro Ala Leu Arg Arg

```

1	5	10	15	
acc gtc acc tca act act gtg acg gat tct gag tct gcc gcc gtc tct				96
Thr Val Thr Ser Thr Thr Val Thr Asp Ser Glu Ser Ala Ala Val Ser				
20	25	30		
cct tca gac tct ccc cgc cac tgc gcc tct tcc aca tgc ctc tgc tcc				144
Pro Ser Asp Ser Pro Arg His Ser Ala Ser Ser Thr Ser Leu Ser Ser				
35	40	45		
atg tcc gag gtt gat atc gcc aag ccc aag tcc gag tat ggt gtc atg				192
Met Ser Glu Val Asp Ile Ala Lys Pro Lys Ser Glu Tyr Gly Val Met				
50	55	60		
ctc gac acc tac ggc aac cag ttc gag gtt ccc gac ttt acc atc aag				240
Leu Asp Thr Tyr Gly Asn Gln Phe Glu Val Pro Asp Phe Thr Ile Lys				
65	70	75	80	
gac atc tac aat gcc atc cct aag cac tgc ttc aag cgc tcc gct ctc				288
Asp Ile Tyr Asn Ala Ile Pro Lys His Cys Phe Lys Arg Ser Ala Leu				
85	90	95		
aag gga tac ggt tat atc ctc cgc gac att gtc ctc ctg act acc act				336
Lys Gly Tyr Gly Tyr Ile Leu Arg Asp Ile Val Leu Leu Thr Thr Thr				
100	105	110		
ttc agc atc tgg tac aac ttt gtg acc ccc gaa tat atc ccc tcc acc				384
Phe Ser Ile Trp Tyr Asn Phe Val Thr Pro Glu Tyr Ile Pro Ser Thr				
115	120	125		
ccc gcc cgc gct ggt ctg tgg gcc gtg tac acc gtt ctt cag ggt ctt				432
Pro Ala Arg Ala Gly Leu Trp Ala Val Tyr Thr Val Leu Gln Gly Leu				
130	135	140		
ttc ggt act ggt ctc tgg gtt att gcc cat gag tgc ggt cac ggt gct				480
Phe Gly Thr Gly Leu Trp Val Ile Ala His Glu Cys Gly His Gly Ala				
145	150	155	160	
ttc tcc gat tct cgc atc atc aac gac att act ggc tgg gtt ctt cac				528
Phe Ser Asp Ser Arg Ile Ile Asn Asp Ile Thr Gly Trp Val Leu His				
165	170	175		
tct tcc ctc ctt gtc ccc tac ttc agc tgg caa atc tcc cac cga aag				576
Ser Ser Leu Leu Val Pro Tyr Phe Ser Trp Gln Ile Ser His Arg Lys				
180	185	190		
cac cac aag gcc act ggc aac atg gag cgt gac atg gtc ttc gtt ccc				624
His His Lys Ala Thr Gly Asn Met Glu Arg Asp Met Val Phe Val Pro				
195	200	205		
cga acc cgc gag cag cag gct act cgt ctc gga aag atg acc cac gag				672
Arg Thr Arg Glu Gln Gln Ala Thr Arg Leu Gly Lys Met Thr His Glu				

210	215	220	
ctc gct cat ctt act gag gag acc ccc gct ttc act ctt ctc atg ctc			720
Leu Ala His Leu Thr Glu Glu Thr Pro Ala Phe Thr Leu Leu Met Leu			
225	230	235	240
gtc ctt cag cag ctc gtt ggc tgg ccc aac tac ctc atc acc aat gtt			768
Val Leu Gln Gln Leu Val Gly Trp Pro Asn Tyr Leu Ile Thr Asn Val			
245	250	255	
acc ggc cac aac tac cac gag cgc cag cgt gag ggt cgc ggc aag ggc			816
Thr Gly His Asn Tyr His Glu Arg Gln Arg Glu Gly Arg Gly Lys Gly			
260	265	270	
aag cat aac ggc ctc ggc ggt ggt gtt aac cac ttc gat ccc cgc agc			864
Lys His Asn Gly Leu Gly Gly Gly Val Asn His Phe Asp Pro Arg Ser			
275	280	285	
cct ctg tac gag aac agt gac gct aag ctc atc gtc ctc agc gat att			912
Pro Leu Tyr Glu Asn Ser Asp Ala Lys Leu Ile Val Leu Ser Asp Ile			
290	295	300	
ggt atc ggt ctg atg gcc act gct ctg tacttc ctc gtt cag aag ttc			960
Gly Ile Gly Leu Met Ala Thr Ala Leu Tyr Phe Leu ValGln Lys Phe			
305	310	315	320
ggt ttc tac aac atg gcc atc tgg tac ttt gtt ccc tac ctc tgg gtt			1008
Gly Phe Tyr Asn Met Ala Ile Trp Tyr Phe Val Pro Tyr Leu Trp Val			
325	330	335	
aac cac tgg ctc gtt gcc atc acc ttc ctc cag cac acc gac cct acc			1056
Asn His Trp Leu Val Ala Ile Thr Phe Leu Gln His Thr Asp Pro Thr			
340	345	350	
ctt ccc cac tac acc aac gac gag tgg aac ttc gtc cgt ggt gcc gct			1104
Leu Pro His Tyr Thr Asn Asp Glu Trp Asn Phe Val Arg Gly Ala Ala			
355	360	365	
gct acc att gac cgt gag atg ggc ttc atc ggc cgc cac ctt ctc cac			1152
Ala Thr Ile Asp Arg Glu Met Gly Phe Ile Gly Arg His Leu Leu His			
370	375	380	
ggc atc atc gag act cat gtc ctc cac cac tac gtc agc agc atc ccc			1200
Gly Ile Ile Glu Thr His Val Leu His His Tyr Val Ser Ser Ile Pro			
385	390	395	400
ttc tac aac gcg gac gag gcc acc gag gcc att aag ccc atc atg ggc			1248
Phe Tyr Asn Ala Asp Glu Ala Thr Glu Ala Ile Lys Pro Ile Met Gly			
405	410	415	
aag cac tac cgg gct gat gtc cag gat ggt cct cgt ggc ttc atc cgc			1296
Lys His Tyr Arg Ala Asp Val Gln Asp Gly Pro Arg Gly Phe Ile Arg			

420	425	430	
gcc atg tac cgc agt gcg cgt atg tgc cag tgg gtt gag ccc agc gct			1344
Ala Met Tyr Arg Ser Ala Arg Met Cys Gln Trp Val Glu Pro Ser Ala			
435	440	445	
ggt gcc gag ggt gct ggt aag ggt gtt ctg ttc ttc cgc aac cgc aac			1392
Gly Ala Glu Gly Ala Gly Lys Gly Val Leu Phe Phe Arg Asn Arg Asn			
450	455	460	
aac gtg ggc acc ccc ccc gct gtt atc aag ccc gtt gct taa			1434
Asn Val Gly Thr Pro Pro Ala Val Ile Lys Pro Val Ala			
465	470	475	

<210>30

<211>477

<212>PRT

<213> 串珠镰孢菌

<400>30

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

His His Lys Ala Thr Gly Asn Met Glu Arg Asp Met Val Phe Val Pro		
195	200	205
Arg Thr Arg Glu Gln Gln Ala Thr Arg Leu Gly Lys Met Thr His Glu		
210	215	220
Leu Ala His Leu Thr Glu Glu Thr Pro Ala Phe Thr Leu Leu Met Leu		
225	230	235
Val Leu Gln Gln Leu Val Gly Trp Pro Asn Tyr Leu Ile Thr Asn Val		
245	250	255
Thr Gly His Asn Tyr His Glu Arg Gln Arg Glu Gly Arg Gly Lys Gly		
260	265	270
Lys His Asn Gly Leu Gly Gly Gly Val Asn His Phe Asp Pro Arg Ser		
275	280	285
Pro Leu Tyr Glu Asn Ser Asp Ala Lys Leu Ile Val Leu Ser Asp Ile		
290	295	300
Gly Ile Gly Leu Met Ala Thr Ala Leu Tyr Phe Leu Val Gln Lys Phe		
305	310	315
Gly Phe Tyr Asn Met Ala Ile Trp Tyr Phe Val Pro Tyr Leu Trp Val		
325	330	335
Asn His Trp Leu Val Ala Ile Thr Phe Leu Gln His Thr Asp Pro Thr		
340	345	350
Leu Pro His Tyr Thr Asn Asp Glu Trp Asn Phe Val Arg Gly Ala Ala		
355	360	365
Ala Thr Ile Asp Arg Glu Met Gly Phe Ile Gly Arg His Leu Leu His		
370	375	380
Gly Ile Ile Glu Thr His Val Leu His His Tyr Val Ser Ser Ile Pro		
385	390	395
Phe Tyr Asn Ala Asp Glu Ala Thr Glu Ala Ile Lys Pro Ile Met Gly		
405	410	415
Lys His Tyr Arg Ala Asp Val Gln Asp Gly Pro Arg Gly Phe Ile Arg		
420	425	430
Ala Met Tyr Arg Ser Ala Arg Met Cys Gln Trp Val Glu Pro Ser Ala		
435	440	445
Gly Ala Glu Gly Ala Gly Lys Gly Val Leu Phe Phe Arg Asn Arg Asn		
450	455	460
Asn Val Gly Thr Pro Pro Ala Val Ile Lys Pro Val Ala		
465	470	475

<210>31

<211>777

<212>DNA

<213> 小眼虫

<220>

<221>CDS

<222>(1).. (777)

<223> 合成 Δ -9 延伸酶 (经密码子优化用于解脂耶氏酵母)

<300>

<302> Δ -9 延伸酶及其在制备多不饱和脂肪酸中的应用

<310>WO 2007/061742

<311>2006-11-16

<312>2007-05-31

<313>(1).. (777)

<300>

<302> Δ -9 延伸酶及其在制备多不饱和脂肪酸中的应用

<310>US-2007-0117190-A1

<311>2006-11-16

<312>2007-05-24

<313>(1).. (777)

<400>31

atg gag gtc gtg aac gaa atc gtc tcc att ggc cag gag gtt ctt ccc 48

Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1 5 10 15

aag gtc gac tat gct cag ctc tgg tct gat gcc tcg cac tgc gag gtg 96

Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val

20 25 30

ctg tac ctc tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct 144

Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro

35 40 45

ctc gga ccc aag ggt cag tct cga atg aag ttt gtg ttc acc aac tac 192

Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr

50 55 60

aac ctg ctc atg tcc atc tac tcg ctg ggc tcc ttc ctc tct atg gcc 240

Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala

65 70 75 80

tac gcc atg tac acc att ggt gtc atg tcc gac aac tgc gag aag gct 288

Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala

85 90 95

ttc gac aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc agc 336

Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser

100 105 110

aag ttc ctc gag tac att gac tcc ttc tat ctg ccc ctc atg ggc aag	384
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys	
115 120 125	
cct ctg acc tgg ttg cag ttc ttt cac cat ctc gga gct cct atg gac	432
Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp	
130 135 140	
atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtg	480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val	
145 150 155 160	
ctg ctc aac ggc ttc att cac tgg atc atg tac ggc tac tat tgg acc	528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr	
165 170 175	
cga ctg atc aag ctc aag ttc cct atg ccc aag tcc ctg att act tct	576
Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser	
180 185 190	
atg cag atc att cag ttc aac gtt ggc ttc tac atc gtc tgg aag tac	624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr	
195 200 205	
cgg aac att ccc tgc tac cga caa gat gga atg aga atg ttt ggc tgg	672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp	
210 215 220	
ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac	720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn	
225 230 235 240	
ttc tac gtg cag acc tac atc gtc cga aag cac aag gga gcc aaa aag	768
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys	
245 250 255	
att cag tga	777
Ile Gln	
<210>32	
<211>258	
<212>PRT	
<213> 小眼虫	
<400>32	
Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro	
1 5 10 15	
Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val	
20 25 30	
Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro	

35	40	45
Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr		
50	55	60
Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala		
65	70	75
Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala		
85	90	95
Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser		
100	105	110
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys		
115	120	125
Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp		
130	135	140
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val		
145	150	155
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr		
165	170	175
Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser		
180	185	190
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr		
195	200	205
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp		
210	215	220
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn		
225	230	235
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys		
245	250	255

Ile Gln

<210>33

<211>34

<212>DNA

<213> 大肠杆菌

<400>33

ataacttcgt ataatgtatg ctatacgaag ttat

<210>34

<211>828

<212>DNA

<213> 高山被孢霉

<220>

34

<221>CDS

<222>(1).. (828)

<223> 合成 C16/18 延伸酶（经密码子优化用于解脂耶氏酵母）

<300>

<302> 高山被孢霉 C16/18 脂肪酸延伸酶

<310>US 2007-0087420-A1

<311>2005-10-19

<312>2007-04-19

<313>(1).. (828)

<300>

<302> 高山被孢霉 C16/18 脂肪酸延伸酶

<310>W0 2007/046817

<311>2005-11-04

<312>2007-04-26

<313>(1).. (828)

<400>34

atg gag tct gga ccc atg cct gct ggc att ccc ttc cct gag tac tat 48

Met Glu Ser Gly Pro Met Pro Ala Gly Ile Pro Phe Pro Glu Tyr Tyr

1 5 10 15

gac ttc ttt atg gac tgg aag act ccc ctg gcc atc gct gcc acc tac 96

Asp Phe Phe Met Asp Trp Lys Thr Pro Leu Ala Ile Ala Ala Thr Tyr

20 25 30

act gct gcc gtc ggt ctc ttc aac ccc aag gtt ggc aag gtc tcc cga 144

Thr Ala Ala Val Gly Leu Phe Asn Pro Lys Val Gly Lys Val Ser Arg

35 40 45

gtg gtt gcc aag tcg gct aac gca aag cct gcc gag cga acc cag tcc 192

Val Val Ala Lys Ser Ala Asn Ala Lys Pro Ala Glu Arg Thr Gln Ser

50 55 60

gga gct gcc atg act gcc ttc gtc ttt gtg cac aac ctc att ctg tgt 240

Gly Ala Ala Met Thr Ala Phe Val Phe Val His Asn Leu Ile Leu Cys

65 70 75 80

gtc tac tct ggc atc acc ttc tac tac atg ttt cct gct atg gtc aag 288

Val Tyr Ser Gly Ile Thr Phe Tyr Tyr Met Phe Pro Ala Met Val Lys

85 90 95

aac ttc cga acc cac aca ctg cac gaa gcc tac tgc gac acg gat cag 336

Asn Phe Arg Thr His Thr Leu His Glu Ala Tyr Cys Asp Thr Asp Gln

100 105 110

tcc ctc tgg aac aac gca ctt ggc tac tgg ggt tac ctc ttc tac ctg 384

Ser Leu Trp Asn Asn Ala Leu Gly Tyr Trp Gly Tyr Leu Phe Tyr Leu

115	120	125	
tcc aag ttc tac gag gtc att gac acc atc atc atc atc ctg aag gga			432
Ser Lys Phe Tyr Glu Val Ile Asp Thr Ile Ile Ile Ile Leu Lys Gly			
130	135	140	
cga cgg tcc tcg ctg ctt cag acc tac cac cat gct gga gcc atg att			480
Arg Arg Ser Ser Leu Leu Gln Thr Tyr His His Ala Gly Ala Met Ile			
145	150	155	160
acc atg tgg tct ggc atc aac tac caa gcc act ccc att tgg atc ttt			528
Thr Met Trp Ser Gly Ile Asn Tyr Gln Ala Thr Pro Ile Trp Ile Phe			
165	170	175	
gtg gtc ttc aac tcc ttc att cac acc atc atg tac tgt tac tat gcc			576
Val Val Phe Asn Ser Phe Ile His Thr Ile Met Tyr Cys Tyr Tyr Ala			
180	185	190	
ttc acc tct atc gga ttc cat cct cct ggc aaa aag tac ctg act tcg			624
Phe Thr Ser Ile Gly Phe His Pro Pro Gly Lys Lys Tyr Leu Thr Ser			
195	200	205	
atg cag att act cag ttt ctg gtc ggt atc acc att gcc gtg tcc tac			672
Met Gln Ile Thr Gln Phe Leu Val Gly Ile Thr Ile Ala Val Ser Tyr			
210	215	220	
ctc ttc gtt cct ggc tgc atc cga aca ccc ggt gct cag atg gct gtc			720
Leu Phe Val Pro Gly Cys Ile Arg Thr Pro Gly Ala Gln Met Ala Val			
225	230	235	240
tgg atc aac gtc ggc tac ctg ttt ccc ttg acc tat ctg ttc gtg gac			768
Trp Ile Asn Val Gly Tyr Leu Phe Pro Leu Thr Tyr Leu Phe Val Asp			
245	250	255	
ttt gcc aag cga acc tac tcc aag cga tct gcc att gcc gct cag aaa			816
Phe Ala Lys Arg Thr Tyr Ser Lys Arg Ser Ala Ile Ala Ala Gln Lys			
260	265	270	
aag gct cag taa			828
Lys Ala Gln			
275			
<210>35			
<211>275			
<212>PRT			
<213> 高山被孢霉			
<400>35			
Met Glu Ser Gly Pro Met Pro Ala Gly Ile Pro Phe Pro Glu Tyr Tyr			
1	5	10	15
Asp Phe Phe Met Asp Trp Lys Thr Pro Leu Ala Ile Ala Ala Thr Tyr			

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

<210>36

<211>8739

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pY116

<400>36

ggccgccacc	gcgccccgag	attccggcct	cttcggccgc	caagcgaccc	gggtggacgt	60
ctagaggtag	ctagcaatta	acagatagtt	tgccggtgat	aattctctta	acctcccaca	120
ctcctttgac	ataacgattt	atgtaacgaa	actgaaattt	gaccagatat	tgtgtccgcg	180
gtggagctcc	agctttttgt	cccttttagtg	agggttttaa	cgagcttggc	gtaatcatgg	240
tcatagctgt	tccctgtgtg	aaattgttat	ccgctcacia	ttccacacia	cgtacgagcc	300
ggaagcataa	agtgtaaagc	ctgggggtgcc	taatgagtga	gctaactcac	attaattgcg	360
ttgcgctcac	tgcccgcctt	ccagtcggga	aacctgtcgt	gccagctgca	ttaatgaatc	420
ggccaacgcg	cggggagagg	cggtttgctg	attgggcgct	cttccgcttc	ctcgctcact	480
gactcgctgc	gctcggtcgt	tcggctgcgg	cgagcgggat	cagctcactc	aaaggcggta	540
atacggttat	ccacagaatc	aggggataac	gcaggaaaga	acatgtgagc	aaaaggccag	600
caaaaggcca	ggaaccgtaa	aaaggccgcg	ttgctggcgt	ttttccatag	gctccgcccc	660
cctgacgagc	atcacaaaaa	tcgacgctca	agtcagaggt	ggcgaaaccc	gacaggacta	720
taaagatacc	aggcgcttcc	ccctggaagc	tccctcgtgc	gctctcctgt	tccgacctgt	780
ccgcttaccg	gatacctgtc	cgcctttctc	ccttcgggaa	gcgtggcgct	ttctcatagc	840
tcacgctgta	ggtatctcag	ttcggtgtag	gtcgttcgct	ccaagctggg	ctgtgtgcac	900
gaaccccccg	ttcagcccga	ccgctgcgcc	ttatccggta	actatcgtct	tgagtccaac	960
ccggtaaagc	acgacttata	gccactggca	gcagccactg	gtaacaggat	tagcagagcg	1020
aggtatgtag	gcggtgctac	agagtctctg	aagtgggtgc	ctaactacgg	ctacactaga	1080
aggacagtat	ttggtatctg	cgctctgctg	aagccagtta	ccttcggaaa	aagagttggt	1140
agctcttgat	ccggcaaaaca	aaccaccgct	ggtagcggtg	gtttttttgt	ttgcaagcag	1200
cagattacgc	gcagaaaaaa	aggatctcaa	gaagatcctt	tgatcttttc	tacggggctc	1260
gacgctcagt	ggaacgaaaa	ctcacgttaa	gggatttttg	tcatgagatt	atcaaaaagg	1320
atcttcacct	agatcctttt	aaattaaaaa	tgaagtttta	aatcaatcta	aagtatatat	1380
gagtaaaact	ggtctgacag	ttaccaatgc	ttaatcagtg	aggcacctat	ctcagcgatc	1440
tgtctatttc	gttcatccat	agttgcctga	ctccccgctg	tgtagataac	tacgatacgg	1500
gagggccttac	catctggccc	cagtgtctga	atgataccgc	gagaccacac	ctcaccggct	1560
ccagatttat	cagcaataaa	ccagccagcc	ggaagggccg	agcgcagaag	tggctcctga	1620
actttatccg	cctccatcca	gtctattaat	tgttgccggg	aagctagagt	aagtagttcg	1680
ccagttaata	gtttgcgcaa	cgttgtttgc	attgctacag	gcatcgtggt	gtcacgctcg	1740
tcgttttggt	tggcttcatt	cagctccggt	tcccaacgat	caaggcgagt	tacatgatcc	1800
cccatgttgt	gcaaaaaagc	ggttagctcc	ttcggtcctc	cgatcgttgt	cagaagtaag	1860
ttggccgcag	tgttatcact	catggttatg	gcagcactgc	ataattctct	tactgtcatg	1920
ccatccgtaa	gatgcttttc	tgtgactggt	gagtactcaa	ccaagtcatt	ctgagaatag	1980
tgtatgcggc	gaccgagttg	ctcttgcccc	gcgtcaatac	gggataatac	cgcgccacat	2040
agcagaactt	taaaagtgtc	catcattgga	aaacgttctt	cggggcgaaa	actctcaagg	2100
atcttaccgc	tgttgagatc	cagttcgatg	taaccacttc	gtgcacccaa	ctgatcttca	2160
gcatctttta	ctttaccagc	cgtttctggg	tgagcaaaaa	caggaaggca	aaatgccgca	2220
aaaaagggaa	taagggcgac	acggaaatgt	tgaatactca	tactcttcct	ttttcaatat	2280

tattgaagca	tttatcaggg	ttattgtctc	atgagcggat	acatatttga	atgtattttag	2340
aaaaataaac	aaataggggt	tccgcgcaca	tttccccgaa	aagtgccacc	tgacgcgccc	2400
tgtagecgcg	cattaagegc	ggcgggtgtg	gtggttacgc	gcagcgtgac	cgctacactt	2460
gccagegccc	tagegcegc	tccttteget	ttcttccctt	cctttctcgc	cacgttcgcc	2520
ggctttcccc	gtcaagctct	aaateggggg	ctcccttttag	ggttccgatt	tagtgcttta	2580
cggcacctcg	acccccaaaa	acttgattag	ggtgatgggt	cacgtagtgg	gccatcgccc	2640
tgatagacgg	tttttgcgcc	tttgacgttg	gagtcacagt	tctttaatag	tggactcttg	2700
ttccaaactg	gaacaacact	caaccctatc	tcggtctatt	cttttgattt	ataagggtt	2760
ttgccgattt	cggcctattg	gttaaaaaat	gagctgattt	aacaaaaatt	taacgcgaat	2820
tttaacaaaa	tattaacgct	tacaatttcc	attcgccatt	caggctgcgc	aactgttggg	2880
aagggcgatc	ggtgcgggcc	tcttcgctat	tacgccagct	ggcgaaaggg	ggatgtgctg	2940
caaggcgatt	aagttgggta	acgccagggt	tttcccagtc	acgacgttgt	aaaacgacgg	3000
ccagtgaatt	gtaatacgac	tcactatagg	gcgaattggg	taccgggccc	cccctcgagg	3060
tcgatggtgt	cgataagctt	gatatcgaat	tcatgtcaca	caaaccgatc	ttcgccctcaa	3120
ggaaacctaa	ttctacatcc	gagagactgc	cgagatccag	tctacactga	ttaattttcg	3180
ggccaataat	ttaaaaaaat	cgtgtttatat	aatattatat	gtattatata	tatacatcat	3240
gatgatactg	acagtcatgt	cccattgcta	aatagacaga	ctccatctgc	cgccctccaac	3300
tgatgttctc	aatattttaag	gggtcatctc	gcattgttta	ataataaaca	gactccatct	3360
accgcctcca	aatgatgttc	tcaaaatata	ttgtatgaac	ttatttttat	tacttagtat	3420
tattagacaa	cttacttgct	ttatgaaaaa	cacttcctat	ttaggaaaca	atttataatg	3480
gcagttcggt	catttaacaa	tttatgtaga	ataaatgtta	taaatgcgta	tgggaaatct	3540
taaatatgga	tagcataaat	gatatctgca	ttgcctaatt	cgaaatcaac	agcaacgaaa	3600
aaaatccctt	gtacaacata	aatagtcatc	gagaaatata	aactatcaaa	gaacagctat	3660
tcacacgtta	ctattgagat	tattattgga	cgagaatcac	acactcaact	gtctttctct	3720
cttctagaaa	tacaggtaca	agtatgtact	attctcattg	ttcatacttc	tagtcatttc	3780
atcccacata	ttccttggat	ttctctccaa	tgaatgacat	tctatcttgc	aaattcaaca	3840
attataataa	gatataccaa	agtagcggta	tagtggcaat	caaaaagctt	ctctgggtgtg	3900
cttctcgat	ttatttttat	tctaataatc	cattaaaggt	atatattttat	ttcttgttat	3960
ataatccctt	tgtttattac	atgggctgga	tacataaagg	tatttttgatt	taattttttg	4020
cttaaattca	atccccctc	gttcagtgtc	aactgtaatg	gtaggaaatt	accatacttt	4080
tgaagaagca	aaaaaaatga	aagaaaaaaa	aaatcgatatt	tccagggttag	acgttccgca	4140
gaatctagaa	tgcggtatgc	ggtacattgt	tcttcgaacg	taaaagttgc	gtccctgag	4200
atattgtaca	tttttgcttt	tacaagtaca	agtacatcgt	acaactatgt	actactgttg	4260
atgcatccac	aacagtttgt	tttgtttttt	tttgtttttt	ttttttctaa	tgattcatta	4320
ccgctatgta	tacctacttg	tacttgtagt	aagccgggtt	attggcggtc	aattaatcat	4380
agacttatga	atctgcacgg	tgtgcgctgc	gagttacttt	tagcttatgc	atgctacttg	4440
ggtgtaatat	tgggatctgt	tcggaaatca	acggatgctc	aaccgatttc	gacagtaatt	4500
aattaatttg	aatcgaatcg	gagcctaaaa	tgaacccgag	tatatctcat	aaaattctcg	4560
gtgagaggtc	tgtgactgtc	agtacaagggt	gccttcatta	tgccctcaac	cttaccatac	4620

ctcactgaat	gtagtgtacc	tctaaaaatg	aaatacagtg	ccaaaagcca	aggcactgag	4680
ctcgtctaac	ggacttgata	tacaaccaat	taaaacaaat	gaaaagaaat	acagttcttt	4740
gtatcatttg	taacaattac	cctgtacaaa	ctaaggtatt	gaaatcccac	aatattccca	4800
aagtcacccc	ctttccaaat	tgteatgcct	acaactcata	taccaagcac	taacctacca	4860
aacaccacta	aaaccccaca	aaatatatct	taccgaatat	acagtaacaa	gctaccacca	4920
cactcgttgg	gtgcagtcgc	cagcttaaag	atatctatcc	acatcagcca	caactccctt	4980
cctttaataa	accgactaca	cccttggcta	ttgaggttat	gagtgaatat	actgtagaca	5040
agacactttc	aagaagactg	tttccaaaac	gtaccactgt	cctccactac	aaacacaccc	5100
aatctgcttc	ttctagtcaa	ggttgctaca	ccggtaaatt	ataaatcatc	atttcattag	5160
cagggcaggg	ccctttttat	agagtcttat	acactagcgg	accctgccgg	tagaccaacc	5220
cgcaggcgcg	tcagtttgct	ccttccatca	atgcgtcgta	gaaacgactt	actccttctt	5280
gagcagctcc	ttgaccttgt	tggcaacaag	tctccgacct	cggagggtgga	ggaagagcct	5340
ccgatatcgg	cggtagtgat	accagcctcg	acggactcct	tgacggcagc	ctcaacagcg	5400
tcaccggcgg	gcttcatgtt	aagagagAAC	ttgagcatca	tggcggcaga	cagaatgggtg	5460
gcaatggggg	tgaccttctg	cttgccgaga	tcgggggcag	atccgtgaca	gggctcgtac	5520
agaccgaacg	cctcgtttgt	gtcgggcaga	gaagccagag	aggcggagggg	cagcagaccc	5580
agagaaccgg	ggatgacgga	ggcctcgtcg	gagatgatat	cgccaaacat	gttgggtgggtg	5640
atgatgatac	cattcatctt	ggagggtcgc	ttgatgagga	tcatggcggc	cgagtcgac	5700
agctgggtgg	tgagctcgag	ctgggggaat	tcgtccttga	ggactcgagt	gacagtcttt	5760
cgccaaagtc	gagaggaggc	cagcacgttg	gccttgtcaa	gagaccacac	gggaagaggg	5820
gggttgtgct	gaagggccag	gaaggcggcc	attcgggcaa	ttcgctcaac	ctcaggaacg	5880
gagtaggtct	cgggtgctgga	agcgacgcca	gatccgtcat	cctcctttcg	ctctccaaag	5940
tagatacctc	cgacgagctc	tcggacaatg	atgaagtcgg	tgccctcaac	gtttcggatg	6000
ggggagagat	cggcgagctt	gggcgacagc	agctggcagg	gtcgcaggtt	ggcgtagagg	6060
ttcaggtcct	ttcgagctt	gaggagaccc	tgctcgggtc	gcacgtcggg	tcgtccgtcg	6120
ggagtgggtc	atacgttgtt	ggcagcgcct	ccgacagcac	cgagcataat	agagtcagcc	6180
tttcggcaga	tgctgagagt	agcgtcgggtg	atgggctcgc	cctccttctc	aatggcagct	6240
cctccaatga	gtcggctctc	aaacacaaac	tcggtgccgg	aggcctcagc	aacagacttg	6300
agcaccttga	cggcctcggc	aatcacctcg	gggccacaga	agtcgccgcc	gagaagaaca	6360
atcttcttgg	agtcagtctt	ggtcttctta	gtttcgggtt	ccattgtgga	tgtgtgtggt	6420
tgtatgtgtg	atgtggtgtg	tggagtgaag	atctgtggct	ggcaaacgct	cttgtatata	6480
tacgcacttt	tgcccgtgct	atgtggaaga	ctaaacctcc	gaagattgtg	actcaggtag	6540
tgcggtatcg	gctagggacc	caaaccttgt	cgatgccgat	agcgctatcg	aacgtacccc	6600
agccggccgg	gagtatgtcg	gaggggacat	acgagatcgt	caagggtttg	tggccaactg	6660
gtattttaaat	gtagctaacg	gtagcaggcg	aactactggg	acataacctcc	cccgggaatat	6720
gtacaggcat	aatgcgtatc	tgtgggacat	gtggtcgttg	cgccattatg	taagcagcgt	6780
gtactctctc	gactgtccat	atggtttgct	ccatctcacc	ctcatcgttt	tcattgttca	6840
cagcgggcca	caaaaaaact	gtcttctctc	cttctctctt	cgccttagtc	tactcggacc	6900
agtttttagtt	tagcttggcg	ccactggata	aatgagacct	caggccttgt	gatgaggagg	6960

tcacttatga agcatgttag gaggtgcttg tatggataga gaagcaccca aaataataag	7020
aataataata aaacaggggg cgttgctcatt tcatatcgtg ttttcacat caatacacct	7080
ccaacaatg cccttcattg ggccagcccc aatattgtcc tgtagttcaa ctctatgcag	7140
ctcgtatctt attgagcaag taaaactctg tcagccgata ttgcccgacc cgcgacaagg	7200
gtcaacaagg tgggtgtaagg ccttcgcaga agtcaaaact gtgccaaaca aacatctaga	7260
gtctcttttg tgtttctcgc atatatitwa tggctgtct tacgtatttg cgcctcggt	7320
ccggactaat ttcggatcat cccaatacgt ctttttcttc gcagctgtca acagtgtcca	7380
tgatctatcc acctaaatgg gtcatatgag gcgtataatt tcgtgggtgt gataataatt	7440
cccatatatt tgacacaaaa ctccccccc tagacataca tctcacaatc tcacttcttg	7500
tgcttctgtc acacatctcc tccagctgac ttcaactcac acctctgccc cagttgggtct	7560
acagcggtat aaggtttctc cgcatagagg tgcaccactc ctcccgatac ttgtttgtgt	7620
gacttgtggg tcacgacata tatactctaca cacattgcgc cacccttttg ttcttccagc	7680
acaacaaaaa cacgacacgc taacctggc caatttactg accgtacacc aaaatttgcc	7740
tgcattaccg gtcgatgcaa cgagtgatga ggttcgcaag aacctgatgg acatgttcag	7800
ggatcgccag gcgttttctg agcatacctg gaaaatgctt ctgtccgttt gccggctgtg	7860
ggcggcatgg tgcaagttga ataaccggaa atggtttccc gcagaacctg aagatgttcg	7920
cgattatctt ctatatcttc aggcgcgcgg tctggcagta aaaactatcc agcaacattt	7980
gggccagcta aacatgcttc atcgtcggtc cgggctgcca cgaccaagt acagcaatgc	8040
tgtttactg gttatgcggc ggatccgaaa agaaaacgtt gatgccggtg aacgtgcaaa	8100
acaggctcta gcgttcgaac gcactgattt cgaccagggt cgttactca tggaaaatag	8160
cgatcgctgc caggatatac gtaatctggc atttctgggg attgcttata acaccctgtt	8220
acgtatagcc gaaattgcca ggatcagggt taaagatata tcacgtactg acggtgggag	8280
aatgttaate catattggca gaacgaaaac gctggttagc accgcagggt tagagaaggc	8340
acttagcctg ggggtaacta aactggctga gcgatggatt tccgtctctg gtgtagctga	8400
tgatccgaat aactacctgt tttgccgggt cagaaaaaat ggtgttgccg cgccatctgc	8460
caccagccag ctatcaactc gcgccctgga agggattttt gaagcaactc atcgattgat	8520
ttacggcgct aaggatgact ctggctcagag atacctggcc tggctctggac acagtggccg	8580
tgctcgagcc gcgcgagata tggcccgcgc tggagtttca ataccggaga tcatgcaagc	8640
tgggtggctgg accaatgtaa atattgtcat gaactatata cgtaacctgg atagtgaac	8700
aggggcaatg gtgcgcctgc tggaagatgg cgattaagc	8739

<210>37

<211>15337

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pK02UF8289

<400>37

cgatcgagga agaggacaag cggctgcttc ttaagtttgt gacatcagta tccaaggcac	60
cattgcaagg attcaaggct ttgaaccgt catttgccat tcgtaacgct ggtagacagg	120

ttgategggtt	ccctacggcc	tecacctgtg	tcaatcttct	caagctgcct	gactatcagg	180
acattgatca	acttcggaag	aaacttttgt	atgccattcg	atcacatgct	ggtttcgatt	240
tgtcttagag	gaacgcata	acagtaatca	tagagaataa	acgatattca	tttattaaag	300
tagatagttg	aggtagaagt	tgtaaagagt	gataaatagc	ggccgctcac	tgaatctttt	360
tggtccctt	gtgttctctg	acgatatacg	tttgacata	gaaattcaag	aacaaacaca	420
agactgtgcc	aacataaaaag	taattgaaga	accagccaaa	catcctcatc	ccatcttggc	480
gataacaggg	aatgttctctg	tacttccaga	caatgtagaa	accaacattg	aattgaatga	540
tctgcattga	tgtaatcagg	gatttttgga	tggggaactt	cagcttgatc	aatctgggtc	600
aataataacc	gtacatgac	cagtggatga	aaccattcaa	cagcacaaaa	atccaaacag	660
cttcatttctg	gtaattatag	aacagccaca	tatccatcgg	tgcccccaaa	tgatggaaga	720
attgcaacca	ggtcagaggc	ttgcccatca	gtggcaaata	gaaggagtca	atatactcca	780
ggaacttgct	caaatagaac	aactgcgtgg	tgatcctgaa	gacgttggtg	tcaaaagcct	840
tctgcagtt	gtcagacata	acaccgatgg	tgtacatggc	atatgccatt	gagaggaatg	900
atcccaacga	ataaatggac	atgagaaggt	tgtaattggt	gaaaacaaac	ttcatacgag	960
actgaccttt	tggaccaagg	gggccaagag	tgaacttcaa	gatgacaaat	gcgatggaca	1020
agtaaagcac	ctcacagtga	ctggcatcac	tccagagttg	ggcataatca	actggttggg	1080
taaaacttcc	tgcccaattg	agactatttc	attcaccacc	tccatgggtta	gcgtgtcgtg	1140
tttttgttgt	gctggaagaa	caaagggtg	gcgcaatgtg	tgtagatata	tatgtcgtga	1200
cccacaagtc	acacaaacaa	gtatcggggag	gagtgggtgca	cctctatgcg	gagaaacctt	1260
ataccgctgt	agaccaactg	gggcagaggt	gtgagttgaa	gtcagctgga	ggagatgtgt	1320
gacagaagca	caagaagtga	gattgtgaga	tgtatgtcta	gggggggaag	ttttgtgtca	1380
aatatatggg	aattattatc	agcaccacga	aattatacgc	ctcatatgac	ccatttaggt	1440
ggatagatca	tggacactgt	tgacagctgc	gaagaaaaag	cgtattgggg	atgatccgaa	1500
attagtccgg	taccgagcg	caaatacgta	agacagccga	twaatatat	gcgagaaaca	1560
caaagagac	tctagatgtt	tgtttggcac	agttttgact	tctgcgaagg	ccttacacca	1620
ccttgttgac	ccttgctcgcg	ggtcgggcaa	tatcggtctga	cagagtttta	cttgctcaat	1680
aagatacgag	ctgcatagag	ttgaactaca	ggacaatatt	ggggctggcc	acatgaaggg	1740
cattgttttg	aggtgtattg	atggtgaaaa	cacgatatga	aatgacaacg	ccccctgttt	1800
tattattatt	cttattatct	tgggtgtctc	tctatccata	caagcacctc	ctaactgct	1860
tcataagtga	cctcctcatc	acaaggcctg	aggtctcatt	tatccagtgg	cgccaagcta	1920
aactaaaact	ggtccgagta	gactaaggcg	aagagagaag	gagagaagac	agtttttttg	1980
tggccgcctg	tgaacaatga	aaacgatgag	ggtgagatgg	agcaaaccat	atggtttaaa	2040
cagtcagagg	agtacacgct	gcttacataa	tggcgcaacg	accacatgtc	ccacagatac	2100
gcatcgattc	gattcaaat	aattaaaagg	cgttgaaaca	gaatgagcca	gacagcaagg	2160
acaaggtggc	caacagcaag	gagtccaaaa	agccctctat	tgacgagatc	cacgatgtta	2220
ttgctcatga	ggtttccgag	ctcgatgctg	ggaagaagaa	gtgatttgta	tataagaaat	2280
aaatgagata	tagtaaagga	gtgcaagaga	atggcaaggt	ggtcaaattc	tatattactt	2340
gcagtcactg	gttctctgtt	gacatgaatg	aagttaccgt	tggcatagct	gatttaatat	2400
ataactgtcc	aactaactct	cacctagata	taacccatgt	gtgtgtttcc	aatcatcaat	2460

gcggccgctt	actgagcctt	ggcaccgggc	tgtttctcgg	ccattcgagc	gaactgggac	2520
aggtatcgga	gcaggatgac	gagaccctca	tggggcagag	ggtttcggta	ggggaggttg	2580
tgtttctggc	acagctgttc	cacctggtag	gaaacggcag	tgaggttgtg	tcgaggcagg	2640
gtgggccaga	gatggtgctc	gatctggtag	ttcaggcctc	caaagaacca	gtcagtaaatg	2700
atgcctcgtc	gaatgttcat	ggtctcatgg	atctgaccca	cagagaagcc	atgtccgtcc	2760
cagacggaat	caccgatctt	ctccagaggg	tagtggttca	tgaagaccac	gatggcaatt	2820
ccgaagccac	cgacgagctc	ggaaacaaag	aacaccagca	tcgaggtcag	gatggagggc	2880
ataaagaaga	ggtggaacag	ggtcttgaga	gtccagtgc	gagcgagtc	aatggcctct	2940
ttcttgtact	gagatecgga	gaactgggtg	tctcggtcct	tgagggatcg	aacggtcagc	3000
acagactgga	aacaccagat	gaatcgagc	agaatacaga	tgaccaggaa	atagtactgt	3060
tggaactgaa	tgagctttcg	ggagatggga	gaagctcgag	tgacatcgtc	ctcggaccag	3120
gcgagcagag	gcaggttata	aatgtcggga	tcgtgaccct	gaacgttggg	agcagaatga	3180
tgggcgttgt	gtctgtcctt	ccaccaggtc	acggagaagc	cctggagtc	gttgccaaag	3240
accagaccca	ggacgttatt	ccagtttcgg	ttcttgaagg	tctggtgggtg	gcagatgtca	3300
tgagacagcc	atcccatgtg	ctggtagtgc	ataccgagca	cgagagcacc	aatgaagtac	3360
aggtggtact	ggaccagcat	gaagaaggca	agcacgccaa	gaccaggggt	ggtcaagatc	3420
ttgtacgagt	accagagggg	agaggcgtca	aacatgccag	tggcgatcag	ctcttctcgg	3480
agctttcgga	aatcctcctg	agcttcgttg	acggcagcct	ggggaggcag	ctcgggaagcc	3540
tggttgatct	tgggcattcg	cttgagcttg	tcgaaggctt	cctgagagtg	cataaccatg	3600
aaggcgctag	tagcatctcg	tccttggtag	ttctcaatga	tttcagctcc	accaggggtg	3660
aagttcaccc	aagcggagac	gtcgtacacc	tttccgtcga	tgacgagggg	cagagcctgt	3720
cgagaagcct	tcaccatggc	cattgtctga	gatatgtctt	gtgtgtaagg	gggttgggggt	3780
ggttgtttgt	gttcttgact	tttgtgttag	caagggaaga	cgggcaaaaa	agtgagtgtg	3840
gttgggaggg	agagacgagc	cttatatata	atgcttgttt	gtgtttgtgc	aagtggacgc	3900
cgaacggggc	aggagccaaa	ctaaacaagg	cagacaatgc	gagcttaatt	ggattgcctg	3960
atgggcaggg	gttagggctc	gatcaatggg	ggtgcgaagt	gacaaaattg	ggaattaggt	4020
tcgcaagcaa	ggctgacaag	actttggccc	aaacatttgt	acgcggtgga	caacaggagc	4080
caccatcggt	ctgtcacggg	ctagccggtc	gtgcgtcctg	tcaggctcca	cctaggctcc	4140
atgccactcc	atacaatccc	actagtgtac	cgctaggccg	cttttagctc	ccatctaaga	4200
ccccccaaa	acctccactg	tacagtgcac	tgtactgtgt	ggcgatcaag	ggcaagggaa	4260
aaaaggcgca	aacatgcacg	catggaatga	cgtaggtaag	gcgttactag	actgaaaagt	4320
ggcacatttc	ggcgtgccaa	agggtcctag	gtgcgttttcg	cgagctgggc	gccaggccaa	4380
gccgtccaa	aacgcctctc	cgactccctc	cagcggcctc	catatcccca	tcctctccca	4440
cagcaatgtt	gttaagcctt	gcaaacgaaa	aaatagaaaag	gctaataagc	ttccaatatt	4500
gtggtgtacg	ctgcataacg	caacaatgag	cgccaaacaa	cacacacaca	cagcacacag	4560
cagcattaac	cacgatgttt	aaacagtgtg	cgcagatccc	gtcaacagtt	ttatatatcg	4620
tagttacaac	catcaacact	ttttggtaag	tgtaccattc	tatactccaa	ctggtctgca	4680
actgtacaag	tagacatgtt	aatggtagtt	aataacatct	acagcagaac	ctatggtaaa	4740
gacattgcat	ttttacagga	agtatcgctc	tacacgttga	taaatccaaa	gatgcggaac	4800

ttcttccact	tttatcatca	tcccctactc	gtacactcgt	actctttgtt	cgatcgcgat	4860
tcattttctat	aaataatctt	gtatgtacat	gcgcccgctt	aagcaacggg	cttgataaca	4920
gcgggggggg	tgcccacgtt	gttgcggttg	cggaagaaca	gaacaccctt	accagcaccc	4980
tcggcaccag	cgttgggctc	aaccctactg	cacatacgcg	cactgcggta	catggcgcg	5040
atgaagccac	gaggaccatc	ctggacatca	gcccggtagt	gcttgcccat	gatgggctta	5100
atggcctcgg	tggcctcgtc	cgcgtttagt	aaggggatgc	tgctgacgta	gtggtggagg	5160
acatgagtct	cgatgatgcc	gtggagaagg	tggcggccga	tgaagcccat	ctcacgggtca	5220
atggtagcag	cggcaccacg	gacgaagttc	cactcgtcgt	tggtgtagt	gggaagggtta	5280
gggtcgggtg	gctggaggaa	ggtgatggca	acgagccagt	ggttaaccca	gaggtaggga	5340
acaaagtacc	agatggccat	gttgtagaaa	ccgaacttct	gaacgaggaa	gtacagagca	5400
gtggccatca	gaccgatacc	aatatcgctg	aggacgatga	gcttagcgtc	actgttctcg	5460
tacagagggc	tgcggggatc	gaagtgggta	acaccaccgc	cgaggccggt	atgcttgccc	5520
ttgccgcgac	cctcacgctg	gcgctcgtgg	tagttgtggc	cggtaacatt	ggtgatgagg	5580
tagttgggcc	agccaacgag	ctgctgaagg	acgagcatga	gaagagtga	agcgggggtc	5640
tcctcagtaa	gatgagcgag	ctcgtgggtc	atctttccga	gacgagtagc	ctgctgctcg	5700
cgggttcggg	gaacgaagac	catgtcacgc	tccatgttgc	cagtggcctt	gtggtgcttt	5760
cgggtgggaga	tttgccagct	gaagtagggg	acaaggaggg	aagagtgaag	aaccagcca	5820
gtaatgtcgt	tgatgatgcg	agaatcggag	aaagcaccgt	gaccgcactc	atgggcaata	5880
accagagac	cagtaccgaa	aagaccctga	agaacgggtg	acacggccca	cagaccagcg	5940
cgggcggggg	tggaggggat	atattcgggg	gtcacaaggt	tgtaccagat	gctgaaagt	6000
gtagtcagga	ggacaatgtc	gcggaggata	taaccgtatc	ccttgagagc	ggagcgcttg	6060
aagcagtgt	tagggatggc	attgttagatg	tccttgatgg	taaagtcggg	aacctcgaac	6120
tggttgccgt	aggtgtcgag	catgacacca	tactcggact	tgggcttggc	gatatcaacc	6180
tcggacatgg	acgagagcga	tgtggaagag	gccgagtggc	ggggagagtc	tgaaggagag	6240
acggcggcag	actcagaatc	cgtcacagta	gttgagggtga	cgggtcgtct	aagcgcaggg	6300
ttctgcttgg	gcagagccga	agtggacgcc	atggttgtga	attaggggtg	tgagaatggt	6360
tggttgtagg	gaagaatcaa	aggccggtct	cgggatccgt	gggtatatat	atatatatat	6420
atatatacga	tccttcgtta	cctccctgtt	ctcaaaactg	tggtttttcg	tttttcgttt	6480
tttgcTTTTT	ttgatttttt	tagggccaac	taagcttcca	gatttcgcta	atcacctttg	6540
tactaattac	aagaaaggaa	gaagctgatt	agagttgggc	tttttatgca	actgtgctac	6600
tccttatctc	tgatatgaaa	gtgtagaccc	aatcacatca	tgtcatttag	agttggtaat	6660
actgggagga	tagataaggc	acgaaaacga	gccatagcag	acatgctggg	tgtagccaag	6720
cagaagaaag	tagatgggag	ccaattgacg	agcgagggag	ctacgccaat	ccgacatacg	6780
acacgctgag	atcgtcttgg	ccggggggta	cctacagatg	tccaagggtta	agtgcctgac	6840
tgtaattgta	tgtctgagga	caaatatgta	gtcagccgta	taaagtcata	ccaggcacca	6900
gtgccatcat	cgaaccacta	actctctatg	atacatgcct	ccggtattat	tgtaccatgc	6960
gtcgttttgt	tacatacgta	tcttgccctt	ttctctcaga	aactccagac	tttggtattt	7020
ggtcgagata	agcccggacc	atagttagtc	tttcacactc	tacatttctc	ccttgctcca	7080
actatttaaa	ttgccccgga	gaagacggcc	aggccgccta	gatgacaaat	tcaacaactc	7140

acagctgact	ttctgccatt	gccactaggg	gggggccttt	ttatatggcc	aagccaagct	7200
ctccacgtcg	gttgggctgc	accaacaat	aatgggtag	ggttgacca	acaaagggat	7260
gggatggggg	gtagaagata	cgaggataac	ggggctcaat	ggcacaaata	agaacgaata	7320
ctgccattaa	gactcgtgat	ccagecactg	acaccattgc	atcatctaag	ggcctcaaaa	7380
ctacctcgga	actgctgcgc	tgatctggac	accacagagg	ttccgagcac	tttaggttgc	7440
accaaagtgc	ccaccaggtg	caggcagaaa	acgctggaac	agcgtgtaca	gtttgtctta	7500
acaaaaagtg	agggcgctga	ggtcgagcag	ggtgggtgtga	cttggttatag	cctttagagc	7560
tgcgaaagcg	cgtatggatt	tggtcatca	ggccagattg	agggtctgtg	gacacatgtc	7620
atgttagtgt	acttcaatcg	ccccctggat	atagccccga	caataggccg	tggcctcatt	7680
tttttgcctt	ccgcacattt	ccattgctcg	gtaccacac	cttgcttctc	ctgcacttgc	7740
caaccttaat	actggtttac	attgaccaac	atcttacaag	cggggggcctt	gtctagggta	7800
tatataaaca	gtggctctcc	caatcggttg	ccagtctctt	ttttcctttc	tttccccaca	7860
gattcgaaat	ctaaactaca	catcacagaa	ttccgagccg	tgagtatcca	cgacaagatc	7920
agtgtcgaga	cgacgcgttt	tgtgtaatga	cacaatccga	aagtcgctag	caacacacac	7980
tctctacaca	aactaaccca	gctctggtac	catggtgaag	gcttctcgac	aggctctgcc	8040
cctcgtcate	gacggaaagg	tgtacgacgt	ctccgcttgg	gtgaacttcc	accctgggtg	8100
agctgaaatc	attgagaact	accagggacg	agatgctact	gacgccttca	tggttatgca	8160
ctctcaggaa	gccttcgaca	agctcaagcg	aatgcccagg	atcaaccagg	cttccgagct	8220
gcctccccag	gctgccgtca	acgaagctca	ggaggatttc	cgaaagctcc	gagaagagct	8280
gatcgccact	ggcatgtttg	acgcctctcc	cctctggtac	tcgtacaaga	tcttgaccac	8340
cctgggtctt	ggcgtgcttg	ccttcttcat	gctgggtccag	taccacctgt	acttcattgg	8400
tgctctcgtg	ctcggtatgc	actaccagca	aatgggatgg	ctgtctcatg	acatctgcc	8460
ccaccagacc	ttcaagaacc	gaaactggaa	taacgtcctg	ggtctggtct	ttggcaacgg	8520
actccagggc	ttctccgtga	cctgggtggaa	ggacagacac	aacgcccac	attctgctac	8580
caacgttcag	ggtcacgac	ccgacattga	taacctgcct	ctgctcgcct	ggtccgagga	8640
cgatgtcact	cgagcttctc	ccatctccc	aaagctcatt	cagttccaac	agtactat	8700
cctgggtcate	tgtattctcc	tgcgattcat	ctgggtgttc	cagtctgtgc	tgaccgttcg	8760
atccctcaag	gaccgagaca	accagttcta	ccgatctcag	tacaagaaag	aggccattgg	8820
actcgctctg	cactggactc	tcaagacct	gttccacctc	ttctttatgc	cctccatcct	8880
gacctcgatg	ctgggtgtct	ttgtttccga	gctcgctcgg	ggcttcggaa	ttgccatcgt	8940
ggtcttcatg	aaccactacc	ctctggagaa	gatcggtgat	tccgtctggg	acggacatgg	9000
cttctctgtg	ggtcagatcc	atgagaccat	gaacattcga	cgaggcatca	ttactgactg	9060
gttcttttga	ggcctgaact	accagatcga	gcaccatctc	tggccccacc	tgccctcgaca	9120
caacctcact	gccgtttcct	accaggtgga	acagctgtgc	cagaagcaca	acctccccta	9180
ccgaaacctt	ctgccccatg	aaggctctcg	catcctgctc	cgataacctgt	cccagttcgc	9240
tcgaatggcc	gagaagcagc	ccggtgccaa	ggctcagtaa	gcggccgcaa	gtgtggatgg	9300
ggaagtgagt	gcccggttct	gtgtgcacaa	ttggcaatcc	aagatggatg	gattcaacac	9360
agggatatag	cgagctacgt	ggtgggtgcga	ggatatagca	acggatat	atgtttgaca	9420
cttgagaatg	tacgatacaa	gcactgtcca	agtacaatac	taaacatact	gtacatactc	9480

atactcgtac	ccgggcaacg	gtttcacttg	agtgacgtgg	ctagtgtctt	tactcgtaca	9540
gtgtgcaata	ctgcgtatca	tagtctttga	tgtatatcgt	attcattcat	gttagttgcg	9600
tacgggtgaa	gcttccactg	gtcggcgtgg	tagtggggca	gagtggggtc	ggtgtgctgc	9660
aggtaggtga	tggccacgag	ccagtgggtg	acccacaggt	aggggatcag	gtagtagagg	9720
gtgacggaag	ccaggeccca	tcggttgatg	gagtatgcga	tgacggacat	ggtgatacca	9780
ataccgacgt	tagagatcca	gatgttgaac	cagtccttct	tctcaaacag	cggggcgttg	9840
gggttgaagt	ggttgacagc	ccatttgttg	agcttggggg	acttctgtcc	ggtaacgtaa	9900
gacagcagat	acagaggcca	tccaaacacc	tgctgggtga	tgaggccgta	gagggtcatg	9960
aggggagcgt	cctcagcaag	ctcagaccag	tcatgggcgc	ctcggttctc	cataaactcc	10020
tttcggtcct	tgggcacaaa	caccatatca	cgggtgaggt	gaccagtgga	cttgtgggtc	10080
atggagtggg	tcagcttcca	ggcgtagtaa	gggaccagca	tggaggagtg	cagaacccat	10140
ccggtgacgt	tgttgacggt	gtagagtcg	gagaaagcag	agtggccaca	ctcgtgggca	10200
agaaccacaca	gaccggtgcc	aaacagaccc	tggacaatgg	agtacatggc	ccaggccaca	10260
gctcggccgg	aagccgaggg	aataagaggg	aggtacgcgt	aggccatgta	ggcaaaaacg	10320
gcgataaaga	agcaggcgcg	ccagctgcat	taatgaatcg	gccaacgcgc	ggggagaggc	10380
ggtttgcgta	ttgggcgctc	ttccgcttcc	tcgctcactg	actcgtctcg	ctcggctcgtt	10440
cggctgcggc	gagcgggtatc	agctcactca	aaggcggtaa	tacggttatc	cacagaatca	10500
ggggataacg	caggaaagaa	catgtgagca	aaaggccagc	aaaaggccag	gaaccgtaaa	10560
aaggccgcgt	tgctggcggtt	tttccatagg	ctccgcccc	ctgacgagca	tcacaaaaat	10620
cgacgctcaa	gtcagaggtg	gcgaaaccgc	acaggactat	aaagatacca	ggcgtttccc	10680
cctggaagct	ccctcgtgcg	ctctcctgtt	ccgaccctgc	cgcttaccgg	atacctgtcc	10740
gcctttctcc	cttcgggaag	cgtggcgctt	tctcatagct	cacgctgtag	gtatctcagt	10800
tcgggtgtagg	tcgttcgctc	caagctgggc	tgtgtgcacg	aacccccctg	tcagcccgcg	10860
cgtcgcgctt	tatccggtaa	ctatcgtctt	gagtccaacc	cggtaagaca	cgacttatcg	10920
ccactggcag	cagccactgg	taacaggatt	agcagagcga	ggtatgtagg	cgggtgctaca	10980
gagttcttga	agtgggtggc	taactacggc	tacactagaa	gaacagtatt	tggatatctgc	11040
gctctgctga	agccagttac	cttcggaaaa	agagttggta	gctcttgatc	cggcaaaacaa	11100
accaccgctg	gtagcgggtg	tttttttgtt	tgcaagcagc	agattacgcg	cagaaaaaaaa	11160
ggatctcaag	aagatccttt	gatcttttct	acggggctctg	acgctcagtg	gaacgaaaaac	11220
tcacgttaag	ggatttttgt	catgagatta	tcaaaaagga	tcttcaccta	gatcctttta	11280
aattaaaaat	gaagttttaa	atcaatctaa	agtatatatg	agtaaacttg	gtctgacagt	11340
taccaatgct	taatcagtga	ggcacctatc	tcagcgatct	gtctatttctg	ttcatccata	11400
gttgcctgac	tccccgtcgt	gtagataact	acgatacggg	agggcttacc	atctggcccc	11460
agtgtctgcaa	tgataccgcg	agaccacgc	tcaccggctc	cagattttatc	agcaataaac	11520
cagccagccg	gaagggccga	gcgcagaagt	ggtcctgcaa	ctttatccgc	ctccatccag	11580
tctattaatt	gttgcgggga	agctagagta	agtagttcgc	cagttaatag	tttgcgcaac	11640
gttgttgcca	ttgctacagg	catcgtgggtg	tcacgctcgt	cgtttgggtat	ggcttcattc	11700
agctccgggtt	cccaacgata	aaggcgaggtt	acatgatccc	ccatgttgtg	caaaaaagcg	11760
gttagctcct	tcggctctcc	gatcgttgctc	agaagtaagt	tggccgcagt	gttatcactc	11820

atggttatgg	cagcactgca	taattctctt	actgtcatgc	catccgtaag	atgcttttct	11880
gtgactggtg	agtactcaac	caagtcattc	tgagaatagt	gtatgcggcg	accgagttgc	11940
tcttgcccgg	cgtcaatacg	ggataatacc	gcgccacata	gcagaacttt	aaaagtgctc	12000
atcattggaa	aacgtttctt	ggggcgaaaa	ctctcaagga	tcttaccgct	gttgagatcc	12060
agttcgatgt	aaccctactg	tgcacccaac	tgatcttcag	catcttttac	tttcaccagc	12120
gtttctgggt	gagcaaaaac	aggaaggcaa	aatgccgcaa	aaaagggaat	aagggcgaca	12180
cggaaatgtt	gaatactcat	actcttcctt	tttcaatatt	attgaagcat	ttatcagggt	12240
tattgtctca	tgagecgata	catatttgaa	tgtatttaga	aaaataaaca	aatagggggt	12300
ccgcgcacat	ttccccgaaa	agtgccacct	gatgcgggtg	gaaataccgc	acagatgcgt	12360
aaggagaaaa	taccgcatca	ggaaattgta	agcgtaata	ttttgttaaa	attcgcgtta	12420
aatttttgtt	aaatcagctc	attttttaac	caataggccg	aaatcggcaa	aatcccttat	12480
aaatcaaaag	aatagaccga	gatagggttg	agtgttggtc	cagtttgga	caagagtcca	12540
ctattaaaga	acgtggactc	caacgtcaaa	gggcgaaaaa	ccgtctatca	gggcgatggc	12600
ccactacgtg	aaccatcacc	ctaatcaagt	tttttggggt	cgagggtccg	taaagcacta	12660
aatcggaacc	ctaaaggag	ccccgattt	agagcttgac	ggggaaagcc	ggcgaacgtg	12720
gcgagaaagg	aagggaagaa	agcgaaagga	gcgggcgcta	gggcgctggc	aagtgtagcg	12780
gtcacgctgc	gcgtaaccac	cacaccgcgc	gcgcttaatg	cgccgctaca	gggcgcgtcc	12840
attcgccatt	caggctgcgc	aactgttggg	aagggcgata	ggtgcgggcc	tcttcgctat	12900
tacgccagct	ggcgaaaggg	ggatgtgctg	caaggcgatt	aagttgggta	acgccagggt	12960
tttcccagtc	acgacgttgt	aaaacgacgg	ccagtgaatt	gtaatacgac	tcactatagg	13020
gcgaattggg	cccgaacgtc	catgcttgaa	tctacaagta	ggagggttgg	agtgattaag	13080
tgaactttct	ttaacggctc	tatgccagtt	ctattgatat	ccgaaacatc	agtatgaagg	13140
tctgataagg	gtgacttctt	cccacagatt	cgtatcagta	cgagtacgag	accggtactt	13200
gtaacagtat	tgatactaaa	gggaaactac	aacggttgtc	agcgtaatgt	gacttcgccc	13260
atgaacgcag	acacgcagtg	ccgagtgcgg	tgatatcgcc	tactcgttac	gtccatggac	13320
tacacaaccc	ctcggtctcg	cttggcttag	cctcgggctc	ggtgctgttc	agttaaaaca	13380
caatcaaata	acatttctac	tttttagaag	gcaggccgtc	aggagcaact	ccgactccat	13440
tgacgtttct	aaacatctga	atgccttctt	taccttcaac	aaactggcag	gttcgggcga	13500
cagtgtaaag	agacttgatg	aagttgggtg	cgtcgtgtcg	gtagtgttg	cccatgacct	13560
tcttgatctt	ctcagtggcg	attcgggcgt	tgtagaaggg	aattccttta	cctgcaggat	13620
aacttcgtat	aatgtatgct	atacgaagtt	atgatctctc	tcttgagctt	ttccataaca	13680
agttcttctg	cctccaggaa	gtccatgggt	ggtttgatca	tggttttgg	gtagtggtag	13740
tgcagtgggtg	gtattgtgac	tggggatgta	gttgagaata	agtcatacac	aagtcagctt	13800
tcttcgagcc	tcatataagt	ataagtagtt	caacgtatta	gcactgtacc	cagcatctcc	13860
gtatcgagaa	acacaacaac	atgccccatt	ggacagatca	tgcggataca	caggttgtgc	13920
agtatcatac	atactcgatc	agacaggctg	tctgaccatc	atacaagctg	aacaagcgt	13980
ccatacttgc	acgtctctta	tatacacagt	taaattacat	atccatagtc	taacctctaa	14040
cagttaatct	tcttgtaagc	ctcccagcca	gccttctgg	atcgcttggc	ctcctcaata	14100
ggatctcggt	tcttgccgta	cagacctcgg	ccgacaatta	tgatatccgt	tccggtagac	14160

atgacatcct caacagtteg gtactgctgt ccgagagcgt ctcccttgte gtcaagaccc 14220
 accccggggg tcagaataag ccagtcctca gagtcgccct taggtcgggt ctgggcaatg 14280
 aagccaacca caaactcggg gtcggatcgg gcaagctcaa tggctctgctt ggagtactcg 14340
 ccagtggcca gagagccctt gcaagacage tcggccagca tgagcagacc tctggccagc 14400
 ttctcgttgg gagaggggac taggaactcc ttgtactggg agttctcgta gtcagagacg 14460
 tcctccttct tctgttcaga gacagtttcc tcggcaccag ctgcagggcc agcaatgatt 14520
 ccggttccgg gtacaccgtg ggcgttggtg atatcggacc actcggcgat tcggtgacac 14580
 cggtagtggg gcttgacagt gttgccaata tctgcgaact ttctgtcctc gaacaggaag 14640
 aaaccgtgct taagagcaag ttccttgagg gggagcacag tgccggcgta ggtgaagtcg 14700
 tcaatgatgt cgatatgggt ttgatcatg cacacataag gtccgacctt atcggcaagc 14760
 tcaatgagct ccttggtggg ggtaacatcc agagaagcac acaggttggg tttcttggct 14820
 gccacgagct tgagcactcg agcggcaaag gcgacttgt ggacgttagc tcgagcttcg 14880
 taggagggca ttttggtggg gaagaggaga ctgaaataaa tttagtctgc agaacttttt 14940
 atcggaacct tatctggggc agtgaagtat atgttatggg aatagttacg agttagttga 15000
 acttatagat agactggact atacggctat cgggtccaaat tagaaagaac gtcaatggct 15060
 ctctgggcgt cgcctttgcc gacaaaaatg tgatcatgat gaaagccagc aatgacgttg 15120
 cagctgatat tgttgteggc caaccgcgcc gaaaacgcag ctgtcagacc cacagcctcc 15180
 aacgaagaat gtatcgtaa agtgatccaa gcacactcat agttggagtc gtactccaaa 15240
 ggcggcaatg acgagtcaga cagatactcg tcgacgcgat aacttcgtat aatgtatgct 15300
 atacgaagtt atcgtagat agttagtaga caacaat 15337

<210>38

<211>1272

<212>DNA

<213> 人工序列

<220>

<223> 突变型 EgD8M Δ -8 去饱和酶 (也称为“EgD8S-23”)

<220>

<221>CDS

<222>(2).. (1270)

<400>38

c atg gtg aag gct tct cga cag gct ctg ccc ctc gtc atc gac gga aag 49
 Met Val Lys Ala Ser Arg Gln Ala Leu Pro Leu Val Ile Asp Gly Lys
 1 5 10 15
 gtg tac gac gtc tcc gct tgg gtg aac ttc cac cct ggt gga gct gaa 97
 Val Tyr Asp Val Ser Ala Trp Val Asn Phe His Pro Gly Gly Ala Glu
 20 25 30
 atc att gag aac tac cag gga cga gat gct act gac gcc ttc atg gtt 145
 Ile Ile Glu Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val
 35 40 45

atg cac tct cag gaa gcc ttc gac aag ctc aag cga atg ccc aag atc	193
Met His Ser Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile	
50 55 60	
aac cag gct tcc gag ctg cct ccc cag gct gcc gtc aac gaa gct cag	241
Asn Gln Ala Ser Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln	
65 70 75 80	
gag gat ttc cga aag ctc cga gaa gag ctg atc gcc act ggc atg ttt	289
Glu Asp Phe Arg Lys Leu Arg Glu Glu Leu Ile Ala Thr Gly Met Phe	
85 90 95	
gac gcc tct ccc ctc tgg tac tcg tac aag atc ttg acc acc ctg ggt	337
Asp Ala Ser Pro Leu Trp Tyr Ser Tyr Lys Ile Leu Thr Thr Leu Gly	
100 105 110	
ctt ggc gtg ctt gcc ttc ttc atg ctg gtc cag tac cac ctg tac ttc	385
Leu Gly Val Leu Ala Phe Phe Met Leu Val Gln Tyr His Leu Tyr Phe	
115 120 125	
att ggt gct ctc gtg ctc ggt atg cac tac cag caa atg gga tgg ctg	433
Ile Gly Ala Leu Val Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu	
130 135 140	
tct cat gac atc tgc cac cac cag acc ttc aag aac cga aac tgg aat	481
Ser His Asp Ile Cys His His Gln Thr Phe Lys Asn Arg Asn Trp Asn	
145 150 155 160	
aac gtc ctg ggt ctg gtc ttt ggc aac gga ctc cag ggc ttc tcc gtg	529
Asn Val Leu Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val	
165 170 175	
acc tgg tgg aag gac aga cac aac gcc cat cat tct gct acc aac gtt	577
Thr Trp Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val	
180 185 190	
cag ggt cac gat ccc gac att gat aac ctg cct ctg ctc gcc tgg tcc	625
Gln Gly His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser	
195 200 205	
gag gac gat gtc act cga gct tct ccc atc tcc cga aag ctc att cag	673
Glu Asp Asp Val Thr Arg Ala Ser Pro Ile Ser Arg Lys Leu Ile Gln	
210 215 220	
ttc caa cag tac tat ttc ctg gtc atc tgt att ctc ctg cga ttc atc	721
Phe Gln Gln Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile	
225 230 235 240	
tgg tgt ttc cag tct gtg ctg acc gtt cga tcc ctc aag gac cga gac	769
Trp Cys Phe Gln Ser Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp	
245 250 255	

aac cag ttc tac cga tct cag tac aag aaa gag gcc att gga ctc gct	817
Asn Gln Phe Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala	
260 265 270	
ctg cac tgg act ctc aag acc ctg ttc cac ctc ttc ttt atg ccc tcc	865
Leu His Trp Thr Leu Lys Thr Leu Phe His Leu Phe Phe Met Pro Ser	
275 280 285	
atc ctg acc tcg atg ctg gtg ttc ttt gtt tcc gag ctc gtc ggt ggc	913
Ile Leu Thr Ser Met Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly	
290 295 300	
ttc gga att gcc atc gtg gtc ttc atg aac cac tac cct ctg gag aag	961
Phe Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys	
305 310 315 320	
atc ggt gat tcc gtc tgg gac gga cat ggc ttc tct gtg ggt cag atc	1009
Ile Gly Asp Ser Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile	
325 330 335	
cat gag acc atg aac att cga cga ggc atc att act gac tgg ttc ttt	1057
His Glu Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe	
340 345 350	
gga ggc ctg aac tac cag atc gag cac cat ctc tgg ccc acc ctg cct	1105
Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro	
355 360 365	
cga cac aac ctc act gcc gtt tcc tac cag gtg gaa cag ctg tgc cag	1153
Arg His Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln	
370 375 380	
aag cac aac ctc ccc tac cga aac cct ctg ccc cat gaa ggt ctc gtc	1201
Lys His Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val	
385 390 395 400	
atc ctg ctc cga tac ctg tcc cag ttc gct cga atg gcc gag aag cag	1249
Ile Leu Leu Arg Tyr Leu Ser Gln Phe Ala Arg Met Ala Glu Lys Gln	
405 410 415	
ccc ggt gcc aag gct cag taa gc	1272
Pro Gly Ala Lys Ala Gln	
420	

<210>39

<211>422

<212>PRT

<213> 人工序列

<220>

<223> 合成构建体

<400>39

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

atg gag gtg gtg aat gaa ata gtc tca att ggg cag gaa gtt tta ccc
Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro

1	5	10	15	
aaa gtt gat tat gcc caa ctc tgg agt gat gcc agt cac tgt gag gtg				96
Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val				
20	25	30		
ctt tac ttg tcc atc gca ttt gtc atc ttg aag ttc act ctt ggc ccc				144
Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro				
35	40	45		
ctt ggt cca aaa ggt cag tct cgt atg aag ttt gtt ttc acc aat tac				192
Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr				
50	55	60		
aac ctt ctc atg tcc att tat tcg ttg gga tca ttc ctc tca atg gca				240
Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala				
65	70	75	80	
tat gcc atg tac acc atc ggt gtt atg tct gac aac tgc gag aag gct				288
Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala				
85	90	95		
ttt gac aac aac gtc ttc agg atc acc acg cag ttg ttc tat ttg agc				336
Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser				
100	105	110		
aag ttc ctg gag tat att gac tcc ttc tat ttg cca ctg atg ggc aag				384
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys				
115	120	125		
cct ctg acc tgg ttg caa ttc ttc cat cat ttg ggg gca ccg atg gat				432
Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp				
130	135	140		
atg tgg ctg ttc tat aat tac cga aat gaa gct gtt tgg att ttt gtg				480
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val				
145	150	155	160	
ctg ttg aat ggt ttc atc cac tgg atc atg tac ggt tat tat tgg acc				528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr				
165	170	175		
aga ttg atc aag ctg aag ttc ccc atg cca aaa tcc ctg att aca tca				576
Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser				
180	185	190		
atg cag atc att caa ttc aat gtt ggt ttc tac att gtc tgg aag tac				624
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr				
195	200	205		
agg aac att ccc tgt tat cgc caa gat ggg atg agg atg ttt ggc tgg				672
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp				

210	215	220	
ttc ttc aat tac ttt tat gtt ggc aca gtc ttg tgt ttg ttc ttg aat			720
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn			
225	230	235	240
ttc tat gtg caa acg tat atc gtc agg aag cac aag gga gcc aaa aag			768
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys			
	245	250	255
att cag tga			777
Ile Gln			
<210>41			
<211>258			
<212>PRT			
<213>小眼虫			
<400>41			
Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro			
1	5	10	15
Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val			
	20	25	30
Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro			
	35	40	45
Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr			
50	55	60	
Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala			
65	70	75	80
Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala			
	85	90	95
Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser			
	100	105	110
Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys			
	115	120	125
Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp			
	130	135	140
Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val			
145	150	155	160
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr			
	165	170	175
Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser			
	180	185	190
Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr			

195	200	205
Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp		
210	215	220
Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn		
225	230	235
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys		
245	250	255

Ile Gln

<210>42

<211>13707

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pZKSL-555R

<400>42

```

aaacagtgtgta cgcagatctg cccatgatgg gggctccac caccagcaat cagggccctg      60
attacacacc cacctgtaat gtcatgctgt tcatcgtggg taatgctgct gtgtgctgtg      120
tgtgtgtgtgt gtttggecgt cattgttgcg ttatgcagcg tacaccacaa tattggaagc      180
ttattagcct ttctatTTTT tcgtttgcaa ggcttaacaa cattgctgtg gagagggatg      240
gggatatgga ggccgctgga gggagtcgga gaggcgtttt ggagcggcctt ggcctggcgc      300
ccagctcgcg aaacgcacct aggacccttt ggcacgccga aatgtgccac ttttcagtct      360
agtaacgcct tacctacgtc attccatgcg tgcattgttg cgcctTTTTT cccttgccct      420
tgatgccac acagtacagt gcactgtaca gtggagggtt tgggggggtc ttagatggga      480
gctaaaagcg gcctagcggg acactagtgg gattgtatgg agtggcatgg agcctagggt      540
gagcctgaca ggacgcacga ccggctagcc cgtgacagac gatgggtggc tcctgttgtc      600
caccgcgtac aaatgttttg gccaaagtct tgtcagcctt gcttgcaaac ctaattccca      660
atTTTgtcac ttgcacccc cattgatega gccctaacc cTgcccata ggcaatccaa      720
ttaagctcgc attgtctgcc ttgttttagt ttgctcctgc ccgtttcggc gtccacttgc      780
acaacacaaa acaagcatta tatataaggc tcgtctctcc ctcccaacca cactcacttt      840
tttgcccgtc ttcccttgct aacacaaaag tcaagaacac aaacaaccac cccaaccccc      900
ttacacacaa gacatatcta cagcaatggc catggctctc tcccttacta ccgagcagct      960
gctcgagega cccgacctgg ttgccatcga cggcattctc tacgatctgg aaggtcttgc      1020
caaggtccat cccggaggcg acttgatcct cgcttctggg gcctccgatg cttctcctct      1080
gttctactcc atgcaccctt acgtcaagcc cgagaactcg aagctgcttc aacagtctgt      1140
gcgaggcaag cacgaccgaa cctccaagga cattgtctac acctacgact ctccctttgc      1200
acaggacgtc aagcgaacta tgcgagaggt catgaaaggt cggaactggg atgccacacc      1260
tggattctgg ctgcgaaccg ttggcatcat tgctgtcacc gcctttttgcg agtggcactg      1320
ggctactacc ggaatgggtg tgtgggggtc cttgactgga ttcatgcaca tgcagatcgg      1380
cctgtccatt cagcacgatg cctctcatgg tgccatcagc aaaaagccct ggggtcaacgc      1440

```

tctctttgcc	tacggcatcg	acgtcattgg	atcgccaga	tggatctggc	tgcagtctca	1500
catcatcgca	catcacacct	acaccaatca	gcatggtctc	gacctggatg	ccgagtccgc	1560
agaaccattc	cttgtgttcc	acaactacce	tgtctccaac	actgctcgaa	agtggtttca	1620
ccgattccag	gcctggtaca	tgtacctcgt	gcttggagcc	tacggcgttt	cgctggtgta	1680
caaccctctc	tacatcttcc	gaatgcagca	caacgacacc	attccccgagt	ctgtcacagc	1740
catgcgagag	aacggcttcc	tgcgacggta	ccgaaccttt	gcattcgtta	tgcgagcttt	1800
cttcatcttt	cgaaccgcct	tcttgccctg	gtatctcact	ggaacctccc	tgtctcatcac	1860
cattcctctg	gtgcccactg	ctaccgggtg	cttcctcacc	ttctttttca	tcttgtctca	1920
caacttcgat	ggctcggagc	gaatccccga	caagaactgc	aagggtcaaga	gctccgagaa	1980
ggacgttgaa	gccgatcaga	tcgactggta	cagagctcag	gtggagacct	cttccaccta	2040
cgggtggacc	attgccatgt	tctttactgg	cgggtctcaac	ttccagatcg	agcatcacct	2100
ctttcctcga	atgtcgtctt	ggcactatcc	cttcgtgcag	caagctgtcc	gagagtgttg	2160
cgaacgacac	ggagttcggg	acgtcttcta	ccctaccatt	gtgggcaaca	tcatttccac	2220
cctcaagtac	atgcacaaag	tcggtgtggg	tcactgtgtc	aaggacgctc	aggattccta	2280
agcggccgca	agtgtggatg	gggaagtgag	tgcccggttc	tgtgtgcaca	attggcaatc	2340
caagatggat	ggattcaaca	cagggatata	gcgagctacg	tgggtggtgcg	aggatatagc	2400
aacggatatt	tatgtttgac	acttgagaat	gtacgataca	agcactgtcc	aagtacaata	2460
ctaaacatac	tgtacatact	catactcgta	cccgggcaac	ggtttcaact	gagtgcagtg	2520
gctagtgtct	ttactcgtac	agtgtgcaat	actgcgtatc	atagtctttg	atgtatatcg	2580
tattcattca	tgttagttgc	gtacgctgtg	ttgttgtatg	tgggtgaagct	tgacaatgga	2640
tgggtgtgtc	tatcaggctg	gggaacaatt	gtgcttaagt	atgctgcagt	tgagtaagag	2700
tcatcgctcc	accaaataa	agtttgccat	tagggttgga	gagagagatg	gtggctggaa	2760
gaattaaatg	acatcaagct	gaggattgtg	ggtgtgcaat	aacacatggt	aggggtgacc	2820
tgtggctcga	aatctgataa	ttattttgta	actttatgat	tattcttaga	ttttttaata	2880
ttcctctata	taacacataa	gtagctgtcg	tctagttgtt	catagcctga	ctcctgcaat	2940
agattagtgc	agagtgattt	tgtgcaattg	agagccacgg	ttgagtcaag	tgactttgtg	3000
tgtgaagtca	tcttacgttt	caagtctcac	aggttactca	attggttggg	tgtctgccct	3060
ttacagatat	ttacagtacc	tgagegtaaa	gtcgttcac	cacggaatga	ctgttcctgt	3120
cacgcagtca	tgatcatgga	tgtggctggg	caggaacct	tttgatagg	agacttaggg	3180
attggactat	tattgaaaaa	actgagccga	atatgatata	gttctatttg	aatgcagaac	3240
ttctgatggg	caattcactt	atttcaggca	tatcggtcat	ggtggcagct	gccacgatgt	3300
tatctcgttg	gaaacctcgg	cgcgccagct	gcattaatga	atcggccaac	gcgcggggag	3360
aggcggtttg	cgtattgggc	gctcttccgc	ttcctcgtc	actgactcgc	tgcgctcggt	3420
cgttcggctg	cggcgagcgg	tatcagctca	ctcaaaggcg	gtaatacggg	tatccacaga	3480
atcaggggat	aacgcaggaa	agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaccg	3540
taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	ccccctgacg	agcatcacaa	3600
aaatcgacgc	tcaagtcaga	ggtggcgaaa	cccgcagga	ctataaagat	accaggcggt	3660
tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	ctgccgctta	ccggataacct	3720
gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcat	agctcacgct	gtaggtatct	3780

cagttcgggtg	taggtcgttc	gctccaagct	gggctgtgtg	cacgaacccc	ccgttcagcc	3840
cgaccgctgc	gccttatecg	gtaactatcg	tcttgagtcc	aacccggtaa	gacacgactt	3900
atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	gcgaggatatg	taggcgggtgc	3960
tacagagttc	ttgaagtgg	ggcctaacta	cggctacact	agaagaacag	tatttgggtat	4020
ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	ggtagctctt	gatccggcaa	4080
acaaccacc	gctggtagcg	gtggtttttt	tgtttgcaag	cagcagatta	cgcgcagaaa	4140
aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	tctgacgctc	agtggaacga	4200
aaactcacgt	taagggattt	tggtcatgag	attatcaaaa	aggatcttca	cctagatcct	4260
tttaaattaa	aatgaagtt	ttaaataaat	ctaaagtata	tatgagtaaa	cttgggtctga	4320
cagttacca	tgcttaatca	gtgaggcacc	tatctcagcg	atctgtctat	ttcgttcac	4380
catagttgcc	tgactccccg	tcgtgtagat	aactacgata	cgggaggggct	taccatctgg	4440
ccccagtgt	gcaatgatac	cgcgagaccc	acgtcaccg	gctccagatt	tatcagcaat	4500
aaaccagcca	gccggaaggg	ccgagcgcag	aagtggctct	gcaactttat	ccgcctccat	4560
ccagtctatt	aattgtttgc	gggaagctag	agtaagtagt	tcgccagtta	atagtttgcg	4620
caacgttggt	gccattgcta	caggcatcgt	ggtgtcacgc	tcgtcgtttg	gtatggcttc	4680
attcagctcc	ggttcccaac	gatcaaggcg	agttacatga	tcccccatgt	tgtgcaaaaa	4740
agcggttagc	tccttcggtc	ctccgatcgt	tgtcagaagt	aagttggccg	cagtgttatc	4800
actcatgggt	atggcagcac	tgcataatc	tcttactgtc	atgccatccg	taagatgctt	4860
ttctgtgact	ggtgagtact	caaccaagtc	attctgagaa	tagtgtatgc	ggcgaccgag	4920
ttgctcttgc	ccggcgctca	tacgggataa	taccgcgcca	catagcagaa	ctttaaaagt	4980
gctcatcatt	ggaaaacgtt	cttcggggcg	aaaactctca	aggatcttac	cgtgtttgag	5040
atccagttcg	atgtaacca	ctcgtgcacc	caactgatct	tcagcatctt	ttactttcac	5100
cagcgtttct	gggtgagcaa	aaacaggaag	gcaaaatgcc	gcaaaaaagg	gaataagggc	5160
gacacggaaa	tgttgaatac	tcatactctt	cttttttcaa	tattattgaa	gcatttatca	5220
gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	aacaaatagg	5280
ggttccgcgc	acatttcccc	gaaaagtgcc	acctgatgcg	gtgtgaaata	ccgcacagat	5340
gcgtaaggag	aaaataccgc	atcaggaaat	tgtaaagcgt	aatattttgt	taaaattcgc	5400
gttaaatttt	tgttaaatac	gctcatTTTT	taaccaatag	gccgaaatcg	gcaaaatccc	5460
ttataaatca	aaagaataga	ccgagatagg	gttgagtgtt	gttccagttt	ggaacaagag	5520
tccactatta	aagaacgtgg	actccaacgt	caaaggcgca	aaaaccgtct	atcaggggcg	5580
tggeccacta	cgtgaacat	cacctaatc	aagttttttg	gggtcgaggt	gccgtaaagc	5640
actaaatcgg	aaccctaaag	ggagcccccg	atttagagct	tgacggggaa	agccggcgaa	5700
cgtggcgaga	aaggaaggga	agaaagcgaa	aggagcgggc	gctagggcg	tggaagtgt	5760
agcggtcacg	ctgcgcgtaa	ccaccacacc	cgcgcgcgtt	aatgcgccgc	tacagggcgc	5820
gtccattcgc	cattcaggct	gcgcaactgt	tgggaagggc	gatcgggtgcg	ggcctcttcg	5880
ctattacgcc	agctggcgaa	agggggatgt	gctgcaaggc	gattaagttg	ggtaacgcca	5940
gggttttccc	agtcacgacg	ttgtaaaacg	acggccagtg	aattgtaata	cgactcacta	6000
tagggcgaa	tgggccccgac	gtcgcagtga	ttccatagcc	acacctttgc	ctatggcttc	6060
acaaccgaag	gcaattcgag	aggtcgcgct	tatggaatcg	actcgtataa	agctgaaggg	6120

aaaggagac	gttccgagcg	ctcagatgca	atagtcgtcc	agctaattgtg	gattcaaaaa	6180
caaccccaac	agtaattcttg	aaaatttgaa	cggatcaatc	tgaacactct	tgctccaggt	6240
cattcttcta	acgcacatcc	ccagagtcta	gagggagttg	tgttgtgaac	atcctaataa	6300
acaatgcaat	ggattcggga	tattctctgt	ctcgccccct	actcgatgtc	gagtaaaccg	6360
atcaccaact	aacaatactc	ctccgcgttc	tgccattgac	tctcaaacag	acatcgctat	6420
caacggaaca	gcataatttta	gcttcttagg	acaataaata	ttgataatgc	cggctctccc	6480
tcggtatatt	aagcaatcca	ttcatacact	cattcatcag	gttaatttta	tatatataat	6540
ttgtctatct	aaacaccgta	aattactggg	accatcatct	cctccttttc	aaatacacgt	6600
ctatttgcat	taatgaaatt	actcgccaat	tcgcagaacg	tgtttgtcga	acagagcctt	6660
agctcgggtc	cagacaggag	cagtgtctcg	ctgaggaagc	tgccaggagag	ttaatatact	6720
cacctgcagg	attgagacta	tgaatggatt	cccgtgcccc	tattactcta	ctaatttgat	6780
cttggaaacg	gaaaatacgt	ttctaggact	ccaaagaatc	tcaactcttg	tccttactaa	6840
atatactacc	catagttgat	ggtttacttg	aacagagagg	acatgttcac	ttgacccaaa	6900
gtttctcgca	tctcttggat	atttgaacaa	cggcgctccac	tgaccgtcag	ttatccagtc	6960
acaaaacccc	cacattcata	cattcccatg	tacgtttaca	aagttctcaa	ttccatcgtg	7020
caaatcaaaa	tcacatctat	tcattcatca	tatataaacc	catcatgtct	actaacactc	7080
acaactccat	agaaaacatc	gactcagaac	acacgctcca	tgccggccgct	taggaatcct	7140
gtgcgtcctt	cacgcagtgg	acgacacca	ccttatgcat	gtacttcagg	gtggagatga	7200
tgttgccgac	gatggtaggg	tagaaaacat	atcgactccc	atgtcgttcg	caacactccc	7260
ggaccgcctg	ctggacgaag	gggtagtggc	aagacgacat	ccggggaaag	aggtggtgct	7320
cgatctggaa	attgagaccg	ccagtgaaga	acatggcgat	ggggccaccg	tatgtggagg	7380
acgtctccac	ctgcgcccga	taccagtcaa	tttggtcagc	ctcaacgtcc	ttctcagatc	7440
gcttaacctt	gcagttcttg	tcggggatcc	gttcggagcc	atcaaaattg	tgggacaaaa	7500
tgaagaagaa	cgtcaagaag	gcaccagtgg	cgggtgggcac	cagaggaatg	gtgatcagca	7560
atgaggtccc	agtgaggtac	cagggcaaga	atgcggtccg	gaagatgaag	aaagctcgca	7620
tcacgaatgc	aagtgtgcgg	tagcgcgcga	gaaagccatt	ttcccgcatt	gccgtgacag	7680
actctgggat	ggtgtcattg	tgctgcatcc	ggaaaatgta	gagcgggttg	tacaccagcg	7740
ataccccgta	tgcccccagc	acaaggtaca	tgtaccaagc	ctggaagcgg	tggaaccact	7800
ttcgggcggt	gtttgcggcg	gggtagtgtg	ggaacaccag	gaacggctct	gccgactccg	7860
catccaggtc	gaggccgtgc	tggttggtgt	aggtgtggtg	ccgcatgatg	tgccgactgca	7920
gccaaatcca	ccgggacgat	ccgatgacgt	caatgccgta	ggcgaagagg	gcgttgaccc	7980
aaggtttctt	gctgatggcc	ccgtgggacg	catcatgctg	gatggataag	ccgatctgca	8040
tgtgcatgaa	tccagtcaac	aggccccaca	gcaccatccc	cgtggtagcc	cagtgccact	8100
cgcaaaaggc	cgtcacggcg	atgatcccaa	cgggtgcgcag	ccagaagcca	gggggttcgt	8160
accagttcct	ccctttcacc	acctcgcgca	ttgtccgctt	aacgtcttgt	gcgaaggagg	8220
aatcatacgt	gtagacaatg	tccttcgagg	tgccggtcatg	cttccctcgg	acgaactgtt	8280
gaagcaattt	ggagtctctc	ggtttgacgt	atggatgcat	tgaataaaaag	agaggggagg	8340
catcagaggc	accagaagcg	agaatcaaat	ctcctcctgg	atgaactttg	gcaagccctt	8400
caaggctcgt	gaggatgcca	tcaatcgcaa	ccaaatcagg	gcgtttctaac	agctgtttctg	8460

tggttaagact	gagagccatg	gagagctggg	ttagttttgtg	tagagagtgt	gtgttgctag	8520
cgacttttcgg	attgtgtcat	tacacaaaac	gcgtcgtctc	gacactgate	ttgtcgtgga	8580
tactcacggc	tcggacatcg	tcgccgacga	tgacaccgga	ctttcgctta	aggacgtcag	8640
taacaggcat	tgtgtgatgt	gtagttttaga	tttcgaatct	gtggggaaag	aaaggaaaaa	8700
agagactggc	aaccgattgg	gagagccact	gtttatatat	accctagaca	agccccccgc	8760
ttgtaagatg	ttggtcaatg	taaaccagta	ttaaggttgg	caagtgcagg	agaagcaagg	8820
tgtgggtacc	gagcaatgga	aatgtgcgga	aggcaaaaaa	atgaggccac	ggcctattgt	8880
cggggctata	tccagggggc	gattgaagta	cactaacatg	acatgtgtcc	acagaccctc	8940
aatctggcct	gatgagccaa	atccatacgc	gctttcgcag	ctctaaaggc	tataacaagt	9000
cacaccaccc	tgctcgacct	cagcgccctc	actttttgtt	aagacaaact	gtacacgctg	9060
ttccagcggt	ttctgcctgc	acctgggtggg	acatttggtg	caacctaaag	tgctcggaac	9120
ctctgtgggtg	tccagatcag	cgcagcagtt	ccgaggtagt	tttgaggccc	ttagatgatg	9180
caatgggtgtc	agtcgctgga	tcacgagtct	taatggcagt	attcgttctt	atttgtgcca	9240
ttgagccccg	ttatcctcgt	atctttctacc	ccccatccca	tccctttgtt	ggtgcaaccc	9300
taccatttta	ttgtttgggtg	cagcccaacc	gacgtggaga	gcttggttg	gccatataaa	9360
aaggcccccc	cctagtggca	atggcagaaa	gtcagctgtg	agttgttgaa	tttgtcatct	9420
aggcggcctg	gccgtcttct	ccggggcaat	tggggctgtt	ttttgggaca	caaatacgcc	9480
gccaacccgg	tctctcctga	attccgtcgt	cgcctgagtc	gacatcattt	atttaccagt	9540
tggccacaaa	cccttgacga	tctcgtatgt	ccccccgac	atactcccgg	ccggctgggg	9600
tacgttcgat	agcgtatcgc	gcatcgacaa	ggtttgggtc	cctagccgat	accgcactac	9660
ctgagtcaca	atcttcggag	gtttagtctt	ccacatagca	cgggcaaaag	tgcttatata	9720
tacaagagcg	tttgccagcc	acagattttc	actccacaca	ccacatcaca	catacaacca	9780
cacacatcca	caatggaacc	cgaactaag	aagaccaaga	ctgactccaa	gaagattgtt	9840
cttctcggcg	gcgacttctg	tggccccgag	gtgattgccg	aggccgtcaa	ggtgctcaag	9900
tctgttgctg	aggcctccgg	caccgagttt	gtgtttgagg	accgactcat	tggaggagct	9960
gccattgaga	aggagggcga	gcccatacc	gacgtacttc	tcgacatctg	ccgaaaggct	10020
gactctatta	tgctcgggtc	tgteggaggc	gctgccaaaca	ccgtatggac	cactcccgc	10080
ggacgaaccg	acgtgcgacc	cgagcagggt	ctcctcaagc	tgcgaaagga	cctgaacctg	10140
tacgccaacc	tgcgaccctg	ccagctgctg	tcgcccgaagc	tcgccgatct	ctcccccatc	10200
cgaaacgttg	agggcaccga	cttcatcatt	gtccgagagc	tcgtcggagg	tatctacttt	10260
ggagagcgaa	aggaggatga	cggatctggc	gtcgtttccg	acaccgagac	ctactccgtt	10320
cctgaggttg	agcgaattgc	ccgaatggcc	gccttccttg	cccttcagca	caacccccct	10380
cttcccgtgt	ggtctcttga	caaggccaac	gtgctggcct	cctctcgact	ttggcgaaaag	10440
actgtcactc	gagtcctcaa	ggacgaactc	ccccagctcg	agctcaacca	ccagctgate	10500
gactcggccg	ccatgatect	catcaagcag	ccctccaaga	tgaatgggtat	catcatcacc	10560
accaacatgt	ttggcgatat	catctccgac	gaggcctccg	tcatccccgg	ttctctgggt	10620
ctgctgccct	ccgcctctct	ggcttctctg	cccgacacca	acgaggcggt	cggctctgtac	10680
gagccctgtc	acggatctgc	ccccgatctc	ggcaagcaga	aggtcaaccc	cattgccacc	10740
attctgtctg	ccgcatgat	gtcgaagttc	tctcttaaca	tgaagcccg	cgggtgacgt	10800

gttgaggctg	ccgtcaagga	gtccgtcgag	gctgggtatca	ctaccgccga	tatcggaggc	10860
tcttctcca	cctccgaggt	cggagacttg	ttgccaacaa	ggtcaaggag	ctgctcaaga	10920
aggagtaagt	cgtttctacg	acgcattgat	ggaaggagca	aactgacgcg	cctgcggggt	10980
ggtctaccgg	cagggtccgc	tagtgtataa	gactctataa	aaagggccct	gccctgctaa	11040
tgaaatgatg	atttataatt	taccgggtga	gcaaccttga	ctagaagaag	cagattgggt	11100
gtgtttgtag	tggaggacag	tggtacgttt	tggaaacagt	cttcttgaaa	gtgtcttgct	11160
tacagtatat	tcactcataa	cctcaatagc	caagggtgta	gtcggtttat	taaaggaagg	11220
gagttgtggc	tgatgtggat	atcgatagtt	ggagcaaggg	agaaatgtag	agtgtgaaag	11280
actcactatg	gtccgggctt	atctcgacca	atagccaaag	tctggagttt	ctgagagaaa	11340
aaggcaagat	acgtatgtaa	caaagcgacg	catgggtacaa	taataccgga	ggcatgtatc	11400
atagagagtt	agtggttcga	tgatggcact	ggtgcctggg	atgactttat	acggctgact	11460
acataattgt	cctcagacat	acaattacag	tcaagcactt	acccttggac	atctgtaggt	11520
acccccggc	caagacgate	tcagcgtgtc	gtatgtcgga	ttggcgtagc	tccctcgctc	11580
gtcaattggc	tccatctac	tttcttctgc	ttggctacac	ccagcatgtc	tgctatggct	11640
cgttttcgtg	ccttatctat	cctcccagta	ttaccaactc	taaatgacat	gatgtgattg	11700
ggtctacact	ttcatatcag	agataaggag	tagcacagtt	gcataaaaag	ccaactcta	11760
atcagcttct	tcctttcttg	taattagtag	aaagggtgatt	agcgaaatct	ggaagcttag	11820
ttggccctaa	aaaaatcaaa	aaaagcaaaa	aacgaaaaac	gaaaaaccac	agttttgaga	11880
acagggaggt	aacgaaggat	cgtatatata	tatatatata	tatataccca	cggatcccga	11940
gaccggcctt	tgattcttcc	ctacaaccaa	ccattctcac	caccctaatt	cacaacctatg	12000
gctcccgacg	ccgacaagct	gcgacagcga	aaggctcagt	ccatccagga	cactgccgat	12060
tctcaggcta	ccgagctcaa	gattggcacc	ctgaaggggc	tccaaggcac	cgagatcgct	12120
attgatggcg	acatctacga	catcaaagac	ttcgatcacc	ctggaggcga	atccatcatg	12180
acctttgggtg	gcaacgacgt	tactgccacc	tacaagatga	ttcatcccta	ccactcgaag	12240
catcacctgg	agaagatgaa	aaaggctcgg	cgagtgcctg	actacacctc	cgagtacaag	12300
ttcgatactc	ccttcgaacg	agagatcaaa	caggaggtct	tcaagattgt	gcgaagaggt	12360
cgagagtttg	gaacacctgg	ctacttcttt	cgagccttct	gctacatcgg	tctcttcttt	12420
tacctgcagt	atctctgggt	taccactcct	accactttcg	cccttgctat	cttctacggt	12480
gtgtctcagg	ccttcattgg	cctgaacgtc	cagcacgacg	ccaaccacgg	agctgcctcc	12540
aaaaagccct	ggatcaacaa	tttgctcggc	ctgggtgccg	actttatcgg	aggctccaag	12600
tggtcttgga	tgaaccagca	ctggacctat	cacacttaca	ccaaccatca	cgagaaggat	12660
cccgcgccc	tgggtgcaga	gcctatgctg	ctcttcaacg	actatccctt	gggtcacccc	12720
aagcgaacce	tcattcatca	cttccaagcc	ttctactatc	tgtttgtcct	tgctggctac	12780
tgggtgtctt	cgggtgttcaa	ccctcagatc	ctggacctcc	agcaccgagg	tgcccaggct	12840
gtcggcatga	agatggagaa	cgactacatt	gccaagtctc	gaaagtacgc	tatcttctctg	12900
cgactcctgt	acatctacac	caacattgtg	gtccccatcc	agaaccaagg	cttttcgctc	12960
accgtcgttg	ctcacattct	tactatgggt	gtcgcctcca	gcctgaccct	cgctactctg	13020
ttcgccctct	cccacaactt	cgagaacgca	gatcgggatc	ccacctacga	ggctcgaaaag	13080
ggaggcgagc	ctgtctgttg	gttcaagtcg	cagggtgaaa	cctcctctac	ttacgggtggc	13140

```

ttcatttccg gttgccttac aggcggactc aactttcagg tcgagcatca cctgtttcct 13200
cgaatgtcct ctgcctggta cccctacatc gctcctaccg ttcgagaggt ctgcaaaaag 13260
cacggcgtea agtacgccta ctatccctgg gtgtggcaga acctcatctc gaccgtcaag 13320
tacctgcate agtccggaac tggtcgaac tggaagaacg gtgccaatcc ctactctggc 13380
aagctgtaag cgcccgcatg tacatacaag attatttatagaaatgaatc gcgatcgaac 13440
aaagagtacg agtgtacgag taggggatga tgataaaagt ggaagaagtt ccgcatcttt 13500
ggatttatca acgtgtagga cgatacttcc tgtaaaaatg caatgtcttt accatagggtt 13560
ctgctgtaga tgttattaac taccattaac atgtctactt gtacagttgc agaccagttg 13620
gagtatagaa tggtagactt accaaaaagt gttgatgggt gtaactacga tatataaac 13680
tgttgacggg atctgcgtac actgttt 13707

```

<210>43

<211>1350

<212>DNA

<213> 小眼虫

<220>

<221>CDS

<222>(1).. (1350)

<223> 合成 Δ -5 去饱和酶 (经密码子优化用于解脂耶氏酵母)

<400>43

```

atg gct ctc tcc ctt act acc gag cag ctg ctc gag cga ccc gac ctg 48
Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu
1          5          10          15
gtt gcc atc gac ggc att ctc tac gat ctg gaa ggt ctt gcc aag gtc 96
Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val
20          25          30
cat ccc gga ggc gac ttg atc ctc gct tct ggt gcc tcc gat gct tct 144
His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser
35          40          45
cct ctg ttc tac tcc atg cac cct tac gtc aag ccc gag aac tcg aag 192
Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys
50          55          60
ctg ctt caa cag ttc gtg cga ggc aag cac gac cga acc tcc aag gac 240
Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp
65          70          75          80
att gtc tac acc tac gac tct ccc ttt gca cag gac gtc aag cga act 288
Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr
85          90          95
atg cga gag gtc atg aaa ggt cgg aac tgg tat gcc aca cct gga ttc 336
Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe

```

100	105	110	
tgg ctg cga acc gtt ggc atc att gct gtc acc gcc ttt tgc gag tgg			384
Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp			
115	120	125	
cac tgg gct act acc gga atg gtg ctg tgg ggt ctc ttg act gga ttc			432
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe			
130	135	140	
atg cac atg cag atc ggc ctg tcc att cag cac gat gcc tct cat ggt			480
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ser His Gly			
145	150	155	160
gcc atc agc aaa aag ccc tgg gtc aac gct ctc ttt gcc tac ggc atc			528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile			
165	170	175	
gac gtc att gga tgc tcc aga tgg atc tgg ctg cag tct cac atc atg			576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met			
180	185	190	
cga cat cac acc tac acc aat cag cat ggt ctc gac ctg gat gcc gag			624
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu			
195	200	205	
tcc gca gaa cca ttc ctt gtg ttc cac aac tac cct gct gcc aac act			672
Sar Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr			
210	215	220	
gct cga aag tgg ttt cac cga ttc cag gcc tgg tac atg tac ctc gtg			720
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val			
225	230	235	240
ctt gga gcc tac ggc gtt tgc ctg gtg tac aac cct ctc tac atc ttc			768
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe			
245	250	255	
cga atg cag cac aac gac acc att ccc gag tct gtc aca gcc atg cga			816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg			
260	265	270	
gag aac ggc ttt ctg cga cgg tac cga acc ctt gca ttc gtt atg cga			864
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg			
275	280	285	
gct ttc ttc atc ttt cga acc gcc ttc ttg ccc tgg tat ctc act gga			912
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly			
290	295	300	
acc tcc ctg ctc atc acc att cct ctg gtg ccc act gct acc ggt gcc			960
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala			

305	310	315	320	
ttc ctc acc ttc ttt ttc atc ttg tct cac aac ttc gat ggc tcg gag				1008
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu				
	325	330	335	
cga atc ccc gac aag aac tgc aag gtc aag agc tcc gag aag gac gtt				1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val				
	340	345	350	
gaa gcc gat cag atc gac tgg tac aga gct cag gtg gag acc tct tcc				1104
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser				
	355	360	365	
acc tac ggt gga ccc att gcc atg ttc ttt act ggc ggt ctc aac ttc				1152
Thr Tyr Gly Gly Pro Ile Ala Met Phe Phe Thr Gly Gly Leu Asn Phe				
	370	375	380	
cag atc gag cat cac ctc ttt cct cga atg tcg tct tgg cac tat ccc				1200
Gln Ile Glu His His Leu Phe Pro Arg Met Ser Ser Trp His Tyr Pro				
385	390	395	400	
ttc gtg cag caa gct gtc cga gag tgt tgc gaa cga cac gga gtt cgg				1248
Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg				
	405	410	415	
tac gtc ttc tac cct acc att gtg ggc aac atc att tcc acc ctc aag				1296
Tyr ValPhe Tyr Pro Thr Ile Val Gly Asn Ile Ile Ser Thr Leu Lys				
	420	425	430	
tac atg cac aaa gtc ggt gtg gtt cac tgt gtc aag gac gct cag gat				1344
Tyr Met His Lys Val Gly Val Val His Cys Val Lys Asp Ala Gln Asp				
	435	440	445	
tcc taa				1350
Ser				
<210>44				
<211>449				
<212>PRT				
<213> 小眼虫				
<400>44				
Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu				
1	5	10	15	
Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val				
	20	25	30	
His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser				
	35	40	45	
Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys				

50	55	60
Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp		
65	70	75
Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr		80
	85	90
Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe		95
	100	105
Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp		110
	115	120
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe		125
	130	135
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ser His Gly		140
145	150	155
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile		160
	165	170
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met		175
	180	185
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu		190
	195	200
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr		205
	210	215
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val		220
225	230	235
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe		240
	245	250
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg		255
	260	265
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg		270
	275	280
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly		285
	290	295
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala		300
305	310	315
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu		320
	325	330
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Ser Ser Glu Lys Asp Val		335
	340	345
Glu Ala Asp Gln Ile Asp Trp Tyr Arg Ala Gln Val Glu Thr Ser Ser		350
	355	360
		365

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

Ser

<210>45

<211>1392

<212>DNA

<213> 多甲藻属 CCMP626

<220>

<221>CDS

<222>(1).. (1392)

<223> 合成 Δ -5 去饱和酶 (经密码子优化用于解脂耶氏酵母)

<400>45

atg gct ccc gac gcc gac aag ctg cga cag cga aag gct cag tcc atc	48
Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Lys Ala Gln Ser Ile	
1 5 10 15	
cag gac act gcc gat tct cag gct acc gag ctc aag att ggc acc ctg	96
Gln Asp Thr Ala Asp Ser Gln Ala Thr Glu Leu Lys Ile Gly Thr Leu	
20 25 30	
aag ggt ctc caa ggc acc gag atc gtc att gat ggc gac atc tac gac	144
Lys Gly Leu Gln Gly Thr Glu Ile Val Ile Asp Gly Asp Ile Tyr Asp	
35 40 45	
atc aaa gac ttc gat cac cct gga ggc gaa tcc atc atg acc ttt ggt	192
Ile Lys Asp Phe Asp His Pro Gly Gly Glu Ser Ile Met Thr Phe Gly	
50 55 60	
ggc aac gac gtt act gcc acc tac aag atg att cat ccc tac cac tcg	240
Gly Asn Asp Val Thr Ala Thr Tyr Lys Met Ile His Pro Tyr His Ser	
65 70 75 80	
aag cat cac ctg gag aag atg aaa aag gtc ggt cga gtg ccc gac tac	288
Lys His His Leu Glu Lys Met Lys Lys Val Gly Arg Val Pro Asp Tyr	
85 90 95	
acc tcc gag tac aag ttc gat act ccc ttc gaa cga gag atc aaa cag	336

Thr	Ser	Glu	Tyr	Lys	Phe	Asp	Thr	Pro	Phe	Glu	Arg	Glu	Ile	Lys	Gln	
100				105				110								
gag	gtc	ttc	aag	att	gtg	cga	aga	ggt	cga	gag	ttt	gga	aca	cct	ggc	384
Glu	Val	Phe	Lys	Ile	Val	Arg	Arg	Gly	Arg	Glu	Phe	Gly	Thr	Pro	Gly	
115				120				125								
tac	ttc	ttt	cga	gcc	ttc	tgc	tac	atc	ggt	ctc	ttc	ttt	tac	ctg	cag	432
Tyr	Phe	Phe	Arg	Ala	Phe	Cys	Tyr	Ile	Gly	Leu	Phe	Phe	Tyr	Leu	Gln	
130				135				140								
tat	ctc	tgg	gtt	acc	act	cct	acc	act	ttc	gcc	ctt	gct	atc	ttc	tac	480
Tyr	Leu	Trp	Val	Thr	Thr	Pro	Thr	Thr	Phe	Ala	Leu	Ala	Ile	Phe	Tyr	
145				150				155				160				
ggt	gtg	tct	cag	gcc	ttc	att	ggc	ctg	aac	gtc	cag	cac	gac	gcc	aac	528
Gly	Val	Ser	Gln	Ala	Phe	Ile	Gly	Leu	Asn	Val	Gln	His	Asp	Ala	Asn	
165				170				175								
cac	gga	gct	gcc	tcc	aaa	aag	ccc	tgg	atc	aac	aat	ttg	ctc	ggc	ctg	576
His	Gly	Ala	Ala	Ser	Lys	Lys	Pro	Trp	Ile	Asn	Asn	Leu	Leu	Gly	Leu	
180				185				190								
ggt	gcc	gac	ttt	atc	gga	ggc	tcc	aag	tgg	ctc	tgg	atg	aac	cag	cac	624
Gly	Ala	Asp	Phe	Ile	Gly	Gly	Ser	Lys	Trp	Leu	Trp	Met	Asn	Gln	His	
195				200				205								
tgg	acc	cat	cac	act	tac	acc	aac	cat	cac	gag	aag	gat	ccc	gac	gcc	672
Trp	Thr	His	His	Thr	Tyr	Thr	Asn	His	His	Glu	Lys	Asp	Pro	Asp	Ala	
210				215				220								
ctg	ggt	gca	gag	cct	atg	ctg	ctc	ttc	aac	gac	tat	ccc	ttg	ggt	cac	720
Leu	Gly	Ala	Glu	Pro	Met	Leu	Leu	Phe	Asn	Asp	Tyr	Pro	Leu	Gly	His	
225				230				235				240				
ccc	aag	cga	acc	ctc	att	cat	cac	ttc	caa	gcc	ttc	tac	tat	ctg	ttt	768
Pro	Lys	Arg	Thr	Leu	Ile	His	His	Phe	Gln	Ala	Phe	Tyr	Tyr	Leu	Phe	
245				250				255								
gtc	ctt	gct	ggc	tac	tgg	gtg	tct	tcg	gtg	ttc	aac	cct	cag	atc	ctg	816
Val	Leu	Ala	Gly	Tyr	Trp	Val	Ser	Ser	Val	Phe	Asn	Pro	Gln	Ile	Leu	
260				265				270								
gac	ctc	cag	cac	cga	ggt	gcc	cag	gct	gtc	ggc	atg	aag	atg	gag	aac	864
Asp	Leu	Gln	His	Arg	Gly	Ala	Gln	Ala	Val	Gly	Met	Lys	Met	Glu	Asn	
275				280				285								
gac	tac	att	gcc	aag	tct	cga	aag	tac	gct	atc	ttc	ctg	cga	ctc	ctg	912
Asp	Tyr	Ile	Ala	Lys	Ser	Arg	Lys	Tyr	Ala	Ile	Phe	Leu	Arg	Leu	Leu	
290				295				300								
tac	atc	tac	acc	aac	att	gtg	gct	ccc	atc	cag	aac	caa	ggc	ttt	tcg	960

Tyr Ile Tyr Thr Asn Ile Val Ala Pro Ile Gln Asn Gln Gly Phe Ser	
305	310 315 320
ctc acc gtc gtt gct cac att ctt act atg ggt gtc gcc tcc agc ctg	1008
Leu Thr Val Val Ala His Ile Leu Thr Met Gly Val Ala Ser Ser Leu	
325	330 335
acc ctc gct act ctg ttc gcc ctc tcc cac aac ttc gag aac gca gat	1056
Thr Leu Ala Thr Leu Phe Ala Leu Ser His Asn Phe Glu Asn Ala Asp	
340	345 350
cgg gat ccc acc tac gag gct cga aag gga ggc gag cct gtc tgt tgg	1104
Arg Asp Pro Thr Tyr Glu Ala Arg Lys Gly Gly Glu Pro Val Cys Trp	
355	360 365
ttc aag tcg cag gtg gaa acc tcc tct act tac ggt ggc ttc att tcc	1152
Phe Lys Ser Gln Val Glu Thr Ser Ser Thr Tyr Gly Gly Phe Ile Ser	
370	375 380
ggt tgc ctt aca ggc gga ctc aac ttt cag gtc gag cat cac ctg ttt	1200
Gly Cys Leu Thr Gly Gly Leu Asn Phe Gln Val Glu His His Leu Phe	
385	390 395 400
cct cga atg tcc tct gcc tgg tac ccc tac atc gct cct acc gtt cga	1248
Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala Pro Thr Val Arg	
405	410 415
gag gtc tgc aaa aag cac ggc gtc aag tac gcc tac tat ccc tgg gtg	1296
Glu Val Cys Lys Lys His Gly Val Lys Tyr Ala Tyr Tyr Pro Trp Val	
420	425 430
tgg cag aac ctc atc tcg acc gtc aag tac ctg cat cag tcc gga act	1344
Trp Gln Asn Leu Ile Ser Thr Val Lys Tyr Leu His Gln Ser Gly Thr	
435	440 445
ggc tcg aac tgg aag aac ggt gcc aat ccc tac tct ggc aag ctg taa	1392
Gly Ser Asn Trp Lys Asn Gly Ala Asn Pro Tyr Ser Gly Lys Leu	
450	455 460

<210>46

<211>463

<212>PRT

<213> 多甲藻属 CCMP626

<400>46

Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Lys Ala Gln Ser Ile
1 5 10 15
Gln Asp Thr Ala Asp Ser Gln Ala Thr Glu Leu Lys Ile Gly Thr Leu
20 25 30
Lys Gly Leu Gln Gly Thr Glu Ile Val Ile Asp Gly Asp Ile Tyr Asp

35					40					45						
Ile	Lys	Asp	Phe	Asp	His	Pro	Gly	Gly	Glu	Ser	Ile	Met	Thr	Phe	Gly	
50					55					60						
Gly	Asn	Asp	Val	Thr	Ala	Thr	Tyr	Lys	Met	Ile	His	Pro	Tyr	His	Ser	
65					70					75					80	
Lys	His	His	Leu	Glu	Lys	Met	Lys	Lys	Val	Gly	Arg	Val	Pro	Asp	Tyr	
85					90					95						
Thr	Ser	Glu	Tyr	Lys	Phe	Asp	Thr	Pro	Phe	Glu	Arg	Glu	Ile	Lys	Gln	
100					105					110						
Glu	Val	Phe	Lys	Ile	Val	Arg	Arg	Gly	Arg	Glu	Phe	Gly	Thr	Pro	Gly	
115					120					125						
Tyr	Phe	Phe	Arg	Ala	Phe	Cys	Tyr	Ile	Gly	Leu	Phe	Phe	Tyr	Leu	Gln	
130					135					140						
Tyr	Leu	Trp	Val	Thr	Thr	Pro	Thr	Thr	Phe	Ala	Leu	Ala	Ile	Phe	Tyr	
145					150					155					160	
Gly	Val	Ser	Gln	Ala	Phe	Ile	Gly	Leu	Asn	Val	Gln	His	Asp	Ala	Asn	
165					170					175						
His	Gly	Ala	Ala	Ser	Lys	Lys	Pro	Trp	Ile	Asn	Asn	Leu	Leu	Gly	Leu	
180					185					190						
Gly	Ala	Asp	Phe	Ile	Gly	Gly	Ser	Lys	Trp	Leu	Trp	Met	Asn	Gln	His	
195					200					205						
Trp	Thr	His	His	Thr	Tyr	Thr	Asn	His	His	Glu	Lys	Asp	Pro	Asp	Ala	
210					215					220						
Leu	Gly	Ala	Glu	Pro	Met	Leu	Leu	Phe	Asn	Asp	Tyr	Pro	Leu	Gly	His	
225					230					235					240	
Pro	Lys	Arg	Thr	Leu	Ile	His	His	Phe	Gln	Ala	Phe	Tyr	Tyr	Leu	Phe	
245					250					255						
Val	Leu	Ala	Gly	Tyr	Trp	Val	Ser	Ser	Val	Phe	Asn	Pro	Gln	Ile	Leu	
260					265					270						
Asp	Leu	Gln	His	Arg	Gly	Ala	Gln	Ala	Val	Gly	Met	Lys	Met	Glu	Asn	
275					280					285						
Asp	Tyr	Ile	Ala	Lys	Ser	Arg	Lys	Tyr	Ala	Ile	Phe	Leu	Arg	Leu	Leu	
290					295					300						
Tyr	Ile	Tyr	Thr	Asn	Ile	Val	Ala	Pro	Ile	Gln	Asn	Gln	Gly	Phe	Ser	
305					310					315					320	
Leu	Thr	Val	Val	Ala	His	Ile	Leu	Thr	Met	Gly	Val	Ala	Ser	Ser	Leu	
325					330					335						
Thr	Leu	Ala	Thr	Leu	Phe	Ala	Leu	Ser	His	Asn	Phe	Glu	Asn	Ala	Asp	
340					345					350						

Arg Asp Pro Thr Tyr Glu Ala Arg Lys Gly Gly Glu Pro ValCys Trp
 355 360 365
 Phe Lys Ser Gln Val Glu Thr Ser Ser Thr Tyr Gly Gly Phe Ile Ser
 370 375 380
 Gly Cys Leu Thr Gly Gly Leu Asn Phe Gln ValGlu His His Leu Phe
 385 390 395 400
 Pro Arg Met Ser Ser Ala Trp Tyr Pro Tyr Ile Ala Pro Thr Val Arg
 405 410 415
 Glu ValCys Lys Lys His Gly ValLys Tyr Ala Tyr Tyr Pro Trp Val
 420 425 430
 Trp Gln Asn Leu Ile Ser Thr Val Lys Tyr Leu His Gln Ser Gly Thr
 435 440 445
 Gly Ser Asn Trp Lys Asn Gly Ala Asn Pro Tyr Ser Gly Lys Leu
 450 455 460

<210>47

<211>1350

<212>DNA

<213> 小眼虫

<220>

<221>CDS

<222>(1).. (1350)

<223> Δ -5 去饱和酶

<400>47

atg gct ctc agt ctt acc aca gaa cag ctg tta gaa cgc cct gat ttg	48
Met Ala Leu Ser Leu Thr Thr Glu Gln Leu Leu Glu Arg Pro Asp Leu	
1 5 10 15	
gtt gcg att gat ggc atc ctc tac gac ctt gaa ggg ctt gcc aaa gtt	96
Val Ala Ile Asp Gly Ile Leu Tyr Asp Leu Glu Gly Leu Ala Lys Val	
20 25 30	
cat cca gga gga gat ttg att ctc gct tct ggt gcc tct gat gcc tcc	144
His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser	
35 40 45	
cct ctc ttt tat tca atg cat cca tac gtc aaa ccg gag aat tcc aaa	192
Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys	
50 55 60	
ttg ctt caa cag ttc gtc cga ggg aag cat gac cgc acc tcg aag gac	240
Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp	
65 70 75 80	
att gtc tac acg tat gat tct ccc ttc gca caa gac gtt aag cgg aca	288

Ile	Val	Tyr	Thr	Tyr	Asp	Ser	Pro	Phe	Ala	Gln	Asp	Val	Lys	Arg	Thr	
85				90				95								
atg	cgc	gag	gtg	atg	aaa	ggg	agg	aac	tgg	tac	gca	acc	cct	ggc	ttc	336
Met	Arg	Glu	Val	Met	Lys	Gly	Arg	Asn	Trp	Tyr	Ala	Thr	Pro	Gly	Phe	
100				105				110								
tgg	ctg	cgc	acc	gtt	ggg	atc	atc	gcc	gtg	acg	gcc	ttt	tgc	gag	tgg	384
Trp	Leu	Arg	Thr	Val	Gly	Ile	Ile	Ala	Val	Thr	Ala	Phe	Cys	Glu	Trp	
115				120				125								
cac	tgg	gct	acc	acg	ggg	atg	gtg	ctg	tgg	ggc	ctg	ttg	act	gga	ttc	432
His	Trp	Ala	Thr	Thr	Gly	Met	Val	Leu	Trp	Gly	Leu	Leu	Thr	Gly	Phe	
130				135				140								
atg	cac	atg	cag	atc	ggc	tta	tcc	atc	cag	cat	gat	gcg	tcc	cac	ggg	480
Met	His	Met	Gln	Ile	Gly	Leu	Ser	Ile	Gln	His	Asp	Ala	Ser	His	Gly	
145				150				155				160				
gcc	atc	agc	aag	aag	cct	tgg	gtc	aac	gcc	ctc	ttc	gcc	tac	ggc	att	528
Ala	Ile	Ser	Lys	Lys	Pro	Trp	Val	Asn	Ala	Leu	Phe	Ala	Tyr	Gly	Ile	
165				170				175								
gac	gtc	atc	gga	tcg	tcc	cgg	tgg	att	tgg	ctg	cag	tcg	cac	atc	atg	576
Asp	Val	Ile	Gly	Ser	Ser	Arg	Trp	Ile	Trp	Leu	Gln	Ser	His	Ile	Met	
180				185				190								
cgg	cac	cac	acc	tac	acc	aac	cag	cac	ggc	ctc	gac	ctg	gat	gcg	gag	624
Arg	His	His	Thr	Tyr	Thr	Asn	Gln	His	Gly	Leu	Asp	Leu	Asp	Ala	Glu	
195				200				205								
tcg	gca	gag	ccg	ttc	ctg	gtg	ttc	cac	aac	tac	ccc	gcc	gca	aac	acc	672
Ser	Ala	Glu	Pro	Phe	Leu	Val	Phe	His	Asn	Tyr	Pro	Ala	Ala	Asn	Thr	
210				215				220								
gcc	cga	aag	tgg	ttc	cac	cgc	ttc	caa	gct	tgg	tac	atg	tac	ctt	gtg	720
Ala	Arg	Lys	Trp	Phe	His	Arg	Phe	Gln	Ala	Trp	Tyr	Met	Tyr	Leu	Val	
225				230				235				240				
ctg	ggg	gca	tac	ggg	gta	tcg	ctg	gtg	tac	aac	ccg	ctc	tac	att	ttc	768
Leu	Gly	Ala	Tyr	Gly	Val	Ser	Leu	Val	Tyr	Asn	Pro	Leu	Tyr	Ile	Phe	
245				250				255								
cgg	atg	cag	cac	aat	gac	acc	atc	cca	gag	tct	gtc	acg	gcc	atg	cgg	816
Arg	Met	Gln	His	Asn	Asp	Thr	Ile	Pro	Glu	Ser	Val	Thr	Ala	Met	Arg	
260				265				270								
gag	aat	ggc	ttt	ctg	cgg	cgc	tac	cgc	aca	ctt	gca	ttc	gtg	atg	cga	864
Glu	Asn	Gly	Phe	Leu	Arg	Arg	Tyr	Arg	Thr	Leu	Ala	Phe	Val	Met	Arg	
275				280				285								
gct	ttc	ttc	atc	ttc	cgg	acc	gca	ttc	ttg	ccc	tgg	tac	ctc	act	ggg	912

Ala	Phe	Phe	Ile	Phe	Arg	Thr	Ala	Phe	Leu	Pro	Trp	Tyr	Leu	Thr	Gly	
290				295				300								
acc	tca	ttg	ctg	atc	acc	att	cct	ctg	gtg	ccc	act	gca	act	ggt	gcc	960
Thr	Ser	Leu	Leu	Ile	Thr	Ile	Pro	Leu	Val	Pro	Thr	Ala	Thr	Gly	Ala	
305				310				315				320				
ttc	ttg	acg	ttc	ttc	ttc	att	ttg	tcc	cac	aat	ttt	gat	ggc	tcc	gaa	1008
Phe	Leu	Thr	Phe	Phe	Phe	Ile	Leu	Ser	His	Asn	Phe	Asp	Gly	Ser	Glu	
325				330				335								
cgg	atc	ccc	gac	aag	aac	tgc	aag	gtt	aag	agc	tct	gag	aag	gac	gtt	1056
Arg	Ile	Pro	Asp	Lys	Asn	Cys	Lys	Val	Lys	Ser	Ser	Glu	Lys	Asp	Val	
340				345				350								
gag	gct	gac	caa	att	gac	tgg	tat	cgg	gcg	cag	gtg	gag	acg	tcc	tcc	1104
Glu	Ala	Asp	Gln	Ile	Asp	Trp	Tyr	Arg	Ala	Gln	Val	Glu	Thr	Ser	Ser	
355				360				365								
aca	tac	ggt	ggc	ccc	atc	gcc	atg	ttc	ttc	act	ggc	ggt	ctc	aat	ttc	1152
Thr	Tyr	Gly	Gly	Pro	Ile	Ala	Met	Phe	Phe	Thr	Gly	Gly	Leu	Asn	Phe	
370				375				380								
cag	atc	gag	cac	cac	ctc	ttt	ccc	cgg	atg	tcg	tct	tgg	cac	tac	ccc	1200
Gln	Ile	Glu	His	His	Leu	Phe	Pro	Arg	Met	Ser	Ser	Trp	His	Tyr	Pro	
385				390				395				400				
ttc	gtc	cag	cag	gcg	gtc	cgg	gag	tgt	tgc	gaa	cgc	cat	gga	gtg	cga	1248
Phe	Val	Gln	Gln	Ala	Val	Arg	Glu	Cys	Cys	Glu	Arg	His	Gly	Val	Arg	
405				410				415								
tat	gtt	ttc	tac	cct	acc	atc	gtc	ggc	aac	atc	atc	tcc	acc	ctg	aag	1296
Tyr	Val	Phe	Tyr	Pro	Thr	Ile	Val	Gly	Asn	Ile	Ile	Ser	Thr	Leu	Lys	
420				425				430								
tac	atg	cat	aag	gtg	ggt	gtc	gtc	cac	tgc	gtg	aag	gac	gca	cag	gat	1344
Tyr	Met	His	Lys	Val	Gly	Val	Val	His	Cys	Val	Lys	Asp	Ala	Gln	Asp	
435				440				445								
tcc	tga															1350
Ser																
<210>48																
<211>449																
<212>PRT																
<213> 小眼虫																
<400>48																
Met	Ala	Leu	Ser	Leu	Thr	Thr	Glu	Gln	Leu	Leu	Glu	Arg	Pro	Asp	Leu	
1				5					10					15		
Val	Ala	Ile	Asp	Gly	Ile	Leu	Tyr	Asp	Leu	Glu	Gly	Leu	Ala	Lys	Val	

20	25	30
His Pro Gly Gly Asp Leu Ile Leu Ala Ser Gly Ala Ser Asp Ala Ser		
35	40	45
Pro Leu Phe Tyr Ser Met His Pro Tyr Val Lys Pro Glu Asn Ser Lys		
50	55	60
Leu Leu Gln Gln Phe Val Arg Gly Lys His Asp Arg Thr Ser Lys Asp		
65	70	75
Ile Val Tyr Thr Tyr Asp Ser Pro Phe Ala Gln Asp Val Lys Arg Thr		
85	90	95
Met Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly Phe		
100	105	110
Trp Leu Arg Thr Val Gly Ile Ile Ala Val Thr Ala Phe Cys Glu Trp		
115	120	125
His Trp Ala Thr Thr Gly Met Val Leu Trp Gly Leu Leu Thr Gly Phe		
130	135	140
Met His Met Gln Ile Gly Leu Ser Ile Gln His Asp Ala Ser His Gly		
145	150	155
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile		
165	170	175
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met		
180	185	190
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu		
195	200	205
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr		
210	215	220
Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val		
225	230	235
Leu Gly Ala Tyr Gly Val Ser Leu Val Tyr Asn Pro Leu Tyr Ile Phe		
245	250	255
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg		
260	265	270
Glu Asn Gly Phe Leu Arg Arg Tyr Arg Thr Leu Ala Phe Val Met Arg		
275	280	285
Ala Phe Phe Ile Phe Arg Thr Ala Phe Leu Pro Trp Tyr Leu Thr Gly		
290	295	300
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala		
305	310	315
Phe Leu Thr Phe Phe Phe Ile Leu Ser His Asn Phe Asp Gly Ser Glu		
325	330	335

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

Ser

<210>49

<211>13066

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pZP3-Pa777U

<400>49

tctcgggtcta ttcttttgat ttataaggga ttttgccgat ttcggcctat tggttaaaaa	60
atgagctgat ttaacaaaaa tttaacgcga attttaacaa aatattaacg cttacaattt	120
cctgatgcgg tattttctcc ttacgcattc gtgcgggtatt tcacaccgca tcagggtggca	180
cttttcgggg aaatgtgcgc ggaaccctta tttgtttatt tttctaaata cattcaaata	240
tgtatccgct catgagacaa taaccctgat aaatgcttca ataatttga aaaaggaaga	300
gtatgagtat tcaacatttc cgtgtgcgcc ttattccctt ttttgccgca ttttgccttc	360
ctgtttttgc tcaccagaa acgctggtga aagtaaaaga tgctgaagat cagttgggtg	420
cacgagtggg ttacatcgaa ctggatctca acagcggtaa gatccttgag agttttcgcc	480
ccgaagaacg ttttccaatg atgagcactt ttaaagttct gctatgtggc gcggtattat	540
cccgtattga cgccgggcaa gagcaactcg gtcgccgcat acactattct cagaatgact	600
tggttgagta ctaccagtc acagaaaagc atcttacgga tggcatgaca gtaagagaat	660
tatgcagtgc tgccataacc atgagtgata acactgcggc caacttactt ctgacaacga	720
tcggaggacc gaaggagcta accgcttttt tgcacaacat gggggatcat gtaactcgcc	780
ttgatcggtg ggaaccggag ctgaatgaag ccataccaaa cgacgagcgt gacaccacga	840
tgccctgtage aatggcaaca acgttgcgca aactattaac tggcgaacta cttactctag	900
cttcccggca acaattaata gactggatgg aggcggataa agttgcagga ccacttctgc	960
gctcggccct tccggctggc tggttttatt ctgataaatc tggagccggt gagcgtgggt	1020

ctcgcggtat	cattgcagca	ctggggccag	atggtaagcc	ctcccgtatc	gtagttatct	1080
acacgacggg	gagtcaggca	actatggatg	aacgaaatag	acagatcgct	gagatagggtg	1140
cctcactgat	taagcattgg	taactgtcag	accaagttta	ctcatatata	cttttagattg	1200
atttaaaact	tcatttttaa	tttaaaagga	tctaggtgaa	gacccctttt	gataatctca	1260
tgacccaaat	cccttaacgt	gagttttcgt	tccactgagc	gtcagacccc	gtagaaaaga	1320
tcaaaggatc	ttcttgagat	ccttttttct	tgcgcgtaat	ctgctgcttg	caaacaaaaa	1380
aaccaccgct	accagcgggtg	gtttgtttgc	cggatcaaga	gctaccaact	ctttttccga	1440
aggtaactgg	cttcagcaga	gcgcagatac	caaatactgt	tcttctagt	tagccgtagt	1500
taggccacca	cttcaagaac	tctgtagcac	cgcctacata	cctcgcctctg	ctaatacctgt	1560
taccagtggc	tgctgccagt	ggcgataagt	cgtgtcttac	cgggttggac	tcaagacgat	1620
agttaccgga	taaggcgcag	cggtcgggct	gaacgggggg	ttcgtgcaca	cagcccagct	1680
tggagcgaac	gacctacacc	gaactgagat	acctacagcg	tgagctatga	gaaagcgcca	1740
cgtttccga	aggagaaaag	gcggacaggt	atccggtaag	cggcagggtc	ggaacaggag	1800
agcgcacgag	ggagcttcca	gggggaaacg	cctggatatct	ttatagtcct	gtcgggtttc	1860
gccacctctg	acttgagcgt	cgatttttgt	gatgctcgtc	aggggggcg	agcctatgga	1920
aaaacgccag	caacgcggcc	tttttacggt	tcctggcctt	ttgctggcct	tttgctcaca	1980
tgttctttcc	tgcgttatcc	cctgattctg	tggataaccg	tattaccgcc	tttgagttag	2040
ctgataccgc	tgcgcgcagc	cgaacgaccg	agcgcagcga	gtcagtgagc	gaggaagcgg	2100
aagagcgccc	aatacgcaaa	ccgcctctcc	ccgcgcgttg	gccgattcat	taatgcagct	2160
ggcgcgccac	caatcacaat	tctgaaaagc	acatcttgat	ctcctcattg	cggggagtc	2220
aacggtggtc	ttattccccc	gaatttcccg	ctcaatctcg	ttccagaccg	acccggacac	2280
agtgtttaac	gccgttccga	aactctaccg	cagatatgct	ccaacggact	gggctgcata	2340
gatgtgatcc	tggccttggg	gaaatggata	aaagccggcc	aaaaaaaaag	cggaaaaaag	2400
cggaaaaaaa	gagaaaaaaa	atcgcaaaat	ttgaaaaata	gggggaaaag	acgcaaaaac	2460
gcaaggaggg	gggagtatat	gacactgata	agcaagctca	caacgggttc	tcttattttt	2520
ttcctcatct	tctgcctagg	ttcccaaaat	cccagatgct	tctctccagt	gccaaaagta	2580
agtacccac	aggttttcgg	ccgaaaattc	cacgtgcagc	aacgtcgtgt	ggggtgttaa	2640
aatgtggggg	gggggaacca	ggacaagagg	ctcttgtggg	agccgaatga	gagcacaag	2700
cgggcgggtg	tgataagggc	atttttgccc	attttccctt	ctcctgtctc	tccgacgggtg	2760
atggcgttgt	gcgtcctcta	tttcttttta	tttctttttg	ttttattttt	ctgactaccg	2820
atttggtttg	atttctctca	ccccacacaa	ataagctcgg	gccgaggaat	atatatatac	2880
acggacacag	tgcacctgtg	gacaacacgt	cactacctct	acgatacaca	ccgtacgttg	2940
tgtggaagct	tgtgagcgga	taacaatttc	acacaggaaa	cagctatgac	catgattacg	3000
ccaagctcga	aattaaccct	cactaaaggg	aacaaaagct	ggagctccac	cgcgacacac	3060
atatctggtc	aaatttcagt	ttcgttacat	ttaaattcct	tcacttcaag	ttcattcttc	3120
atctgcttct	gttttacttt	gacaggcaaa	tgaagacatg	gtacgacttg	atggaggcca	3180
agaacgccat	ttcaccccg	gacaccgaag	tgcctgaaat	cctggctgcc	cccattgata	3240
acatcggaag	ctacgggtatt	ccggaaagtg	tatatagaac	ctttccccag	cttgtgtctg	3300
tggatatgga	tgggtgtaac	ccctttgagt	actcgtcttg	gcttctctcc	gagcagtatg	3360

aggctctcta atctagecga tttaatatct caatgtatct atatatctat cttctcatgc	3420
ggccgcttag ttggctttgg tcttggcage cttggcctcc ttgagggtaa acatcttggc	3480
atccttgtcg accacgccgt acttggcgta cataagacca attcggatga aggtgggaat	3540
gatgggagaa gccgacttcc gcaccagttc gggaaaggcc tgagcgaagg cagcagtggc	3600
ctcgttgage ttgtagttag gaatgatggg aaacagatgg tggatctgat gtgtaccaat	3660
gttgtgggac aggtttgtcga tgagggtccc gtagcttcgg tccacagagg acaagttgcc	3720
cttgacatag gtccactccg aatcggcgta ccaggagatt tcctcgtcgt tgtgatggag	3780
gaaggtagtg acaaccagca tgggtggcgaa tccaaagaga ggtgcgaagt aatacagagc	3840
catggtcttg aggccgtaga cgtaggtaag gtaggcgtac agaccagcaa aggccacgag	3900
agagccgagg gaaatgatga cggcagacat tcttcgcagg tagagaggct cccagggatt	3960
gaagtgggtg acctttcggg gaggaaatcc agcaacgagg taggcaaacc aagccgaacc	4020
aaggagatg accatgtgtc gggacagggg atgagagtcg gcttctcgtc gagggtagaa	4080
gatctcatcc ttgtcgatgt tgccggtgtt cttgtgatgg tgtcgatggc tgatcttcca	4140
cgactcgtag ggagtcagaa tgatggagtg aatgagtgtg ccaacagaga agttgagcag	4200
gtgggatcgc gagaaggcac catgtccaca gtcgtgaccg atggtaaaga atccccagaa	4260
cacgataccc tggagcagaa ttagccaggt gcaaaggacg gcacgcagca gtgcaaactc	4320
ctgcacgata gcaagggtc gagcatagta cagtccgaga gcaagggaac cggcaatgcc	4380
cagagctcgc acggtatagt agagggacca gggaaacagag gcttcgaagc agtgggcagg	4440
cagggatcgc ttgatctcgg tgagagtagg gaactcgtag ggagcggcaa cggtagagga	4500
agccatggtt gtgaattagg gtgggtagaa tggttgggtg tagggaagaa tcaaaggccg	4560
gtctcgggat ccgtgggtat atatatatat atatatatat acgatccttc gttacctccc	4620
tgttctcaaa actgtggttt ttctgttttc gttttttgct ttttttgatt tttttagggc	4680
caactaagct tccagatttc gctaatacacc tttgtactaa ttacaagaaa ggaagaagct	4740
gattagagtt gggcttttta tgcaactgtg ctactcctta tctctgatat gaaagtgtag	4800
accaatacac atcatgtcat ttagagttgg taatactggg aggatagata aggcacgaaa	4860
acgagccata gcagacatgc tgggtgttag caagcagaag aaagtagatg ggagccaatt	4920
gacgagcgag ggagctacgc caatccgaca tacgacacgc tgagatcgtc ttggccgggg	4980
ggtacctaca gatgtccaag ggtaagtgtc tgactgtaat tgtatgtctg aggacaaata	5040
tgtagttagc cgtataaagt cataccaggc accagtgcc aatcgaacc actaactctc	5100
tatgatacat gcctccggtg ttattgtacc atgcgtcgtc ttgttacata cgtatcttgc	5160
ctttttctct cagaaactcc agactttggc tattgggtcga gataagcccc gaccatagtg	5220
agtctttcac actctacatt tctcccttgc tccaactatc gattgttgtc tactaactat	5280
cgtacgataa cttcgtatag catacattat acgaagtat cgcgtcgacg agtatctgtc	5340
tgactcgtca ttgccgcctt tggagtacga ctccaactat gagtgtgctt ggatcacttt	5400
gacgatacat tcttcgttgg aggcgtgtgg tctgacagct gcgttttcgg cgcggttggc	5460
cgacaacaat atcagctgca acgtcattgc tggctttcat catgatcaca tttttgtcgg	5520
caaaggcgac gccagagag ccattgacgt tctttctaatt ttggaccgat agccgtatag	5580
tccagtctat ctataagttc aactaactcg taactattac cataacatat acttcaactgc	5640
cccagataag gtcccgataa aaagtcttgc agactaaatt tatttcagtc tcctcttcac	5700

caccaaaatg	ccctcctacg	aagctcgagc	taacgtccac	aagtcgcct	ttgccgctcg	5760
agtgtcgaag	ctcgtggcag	ccaagaaaac	caacctgtgt	gcttctcttg	atgttaccac	5820
caccaaggag	ctcattgagc	ttgccgataa	ggtcggacct	tatgtgtgca	tgatcaaaac	5880
ccatategac	atcattgacg	acttcaccta	cgccggcact	gtgctcccc	tcaaggaact	5940
tgctcttaag	cacggtttct	tctgttctga	ggacagaaag	ttcgagata	ttggcaacac	6000
tgtcaagcac	cagtaccggt	gtcaccgaat	cgccgagtgg	tccgatatca	ccaacgcccc	6060
cgggtgtacc	ggaaccggaa	tcattgtctg	cctgcgagct	gggtccgagg	aaactgtctc	6120
tgaacagaag	aaggaggacg	tctctgacta	cgagaactcc	cagtacaagg	agttcctagt	6180
cccctctccc	aacgagaagc	tgccagagg	tctgtctatg	ctggccgagc	tgtcttgcaa	6240
gggctctctg	gccactggcg	agtactccaa	gcagaccatt	gagcttgccc	gatccgaccc	6300
cgagtttgtg	gttggcttca	ttgccagaa	ccgacctaac	ggcgactctg	aggactggct	6360
tattctgacc	cccggggtgg	gtcttgacga	caagggagac	gctctcggac	agcagtaccg	6420
aactgttgag	gatgtcatgt	ctaccggaac	ggatatcata	attgtcggcc	gaggtctgta	6480
cggccagAAC	cgagatccta	ttgaggaggc	caagcgatac	cagaaggctg	gctgggaggc	6540
ttaccagaag	attaactgtt	agagggtaga	ctatggatat	gtaatttaac	tgtgtatata	6600
gagagcgtgc	aagtattggag	cgcttgttca	gcttgtatga	tggtcagacg	acctgtctga	6660
tcgagtatgt	atgatactgc	acaacctgtg	tatccgcatg	atctgtccaa	tggggcatgt	6720
tgttgtgttt	ctcgatacgg	agatgctggg	tacagtgtca	atacgttgaa	ctacttatac	6780
ttatatgagg	ctcgaagaaa	gctgacttgt	gtatgactta	ttctcaacta	catccccagt	6840
cacaatacca	ccactgcact	accactacac	caaaaccatg	atcaaaccac	ccatggactt	6900
cctggaggca	gaagaacttg	ttatggaaaa	gctcaagaga	gagatcataa	cttcgtatag	6960
catacattat	acgaagttaa	cctgcaggta	aaggaattca	tgctgttcat	cgtgggtaat	7020
gctgctgtgt	gctgtgtgtg	tgtgttgttt	ggcgctcatt	gttgcgttat	gcagcgtaca	7080
ccacaatatt	ggaagcttat	tagcctttct	attttttcgt	ttgcaaggct	taacaacatt	7140
gctgttgaga	gggatgggga	tatggaggcc	gctggaggga	gtcggagagg	cgttttggag	7200
cggcttggcc	tggcgcccag	ctcgcgaaac	gcacctagga	ccctttggca	cgccgaaatg	7260
tgccactttt	cagtctagta	acgccttacc	tacgtcattc	catgcgtgca	tgtttgcgcc	7320
ttttttccct	tgcccttgat	cgccacacag	tacagtgcac	tgtacagtgg	aggttttggg	7380
ggggtcttag	atgggagcta	aaagcggcct	agcgggtacac	tagtgggatt	gtatggagtg	7440
gcatggagcc	taggtggagc	ctgacaggac	gcacgaccgg	ctagcccgtg	acagacgatg	7500
ggtggctcct	gttgtccacc	gcgtacaaat	gtttgggcca	aagtcttgtc	agccttgctt	7560
gcgaacctaa	ttcccaattt	tgtcacttct	cacccccatt	gatcgagccc	taacccttgc	7620
ccatcaggca	atccaattaa	gtcgcatttg	tctgccttgt	ttagtttggc	tcctgcccgt	7680
ttcggcgctc	acttgcacaa	acacaaacaa	gcattatata	taaggctcgt	ctctccctcc	7740
caaccacact	cacttttttg	cccgtcttcc	cttgctaaca	caaaagtcaa	gaacacaaac	7800
aaccacccca	acccctttac	acacaagaca	tatctacagc	aatggccatg	gcttcttcca	7860
ctgttctgtc	gccgtacgag	ttcccagcgc	tgacggagat	caagcgctcg	ctgccagcgc	7920
actgctttga	ggcctcggtc	ccgtggctgc	tctactacac	cgtgcgcgcg	ctgggcacgc	7980
ccggctcgtc	cgcgctcggc	ctctactacg	cgcgcgcgct	cgcgatcgtg	caggagtttg	8040

ccctgctgga	tgcggtgctc	tgcacgggggt	acattctgct	gcagggcatc	gtattctggg	8100
ggtttcttcac	catcggccat	gactgcggcc	acggcgcggtt	ctcgcgttcg	cacctgctca	8160
acttcagegt	cggcacgctc	attcactega	tcatectcac	gccgtacgag	tcatggaaga	8220
tctcgcaaccg	ccaccaccac	aagaacacgg	gcaacatcga	caaggacgag	attttctacc	8280
cgcagegcga	ggccgactcg	cacccactgt	cccgaacat	ggtgatctcg	ctcggctcgg	8340
cctggttcgc	gtacctcggt	gegggtcttc	ctcctcgcaa	ggtgaaccac	ttcaaccctt	8400
gggaaccggt	gtacctgcgc	cgcatgtctg	ccgtcatcat	ctcactcggc	tcgctcgtgg	8460
cgttcgcggg	cttgtatgcg	tatctcacct	acgtctatgg	ccttaagacc	atggcgctgt	8520
actacttcgc	ccctctcttt	gggttcgcca	cgatgctcgt	ggtcactacc	tttttgcacc	8580
acaatgacga	ggaaacgcca	tggtacgccg	actcggagtg	gacgtacgtc	aagggaacc	8640
tctcgtccgt	ggaccgctcg	tacggcgcg	tcacgacaa	cctgagccac	aacatcggca	8700
cgcaccagat	ccaccacctg	tttcgcatca	tcccgcacta	caagctgaac	gaggcgacgg	8760
cagcgttcgc	gcaggcggtc	ccggagctcg	tgcgcaagag	cgcgtcgccg	atcatcccga	8820
cgttcacccg	catcgggctc	atgtacgcca	agtacggcgt	cgtggacaag	gacgccaaga	8880
tgtttacgct	caaggaggcc	aaggccgcca	agaccaaggc	caactaggcg	gccgcattga	8940
tgatttgaaa	cacacacatg	ggttatatct	aggtgagagt	tagttggaca	gttatatatt	9000
aaatcagcta	tgccaacggt	aacttcattc	atgtcaacga	ggaaccagtg	actgcaagta	9060
atatagaatt	tgaccacctt	gccattctct	tgcactcctt	tactatatct	cattttatttc	9120
ttatatacaa	atcatttctt	cttcccagca	tcgagctcgg	aaacctcatg	agcaataaca	9180
tcgttgatct	cgtcaataga	gggttttttg	gactccttgc	tgttggccac	cttgtccttg	9240
ctgttttaaac	agtgtacgca	gatctactat	agaggaacat	ttaaattgcc	ccggagaaga	9300
cggccaggcc	gcctagatga	caaattcaac	aactcacagc	tgactttctg	ccattgccac	9360
tagggggggg	cctttttata	tggccaagcc	aagctctcca	cgtcggttgg	gctgcaccca	9420
acaataaatg	ggtagggttg	caccaacaaa	gggatgggat	ggggggtaga	agatacgagg	9480
ataacggggc	tcaatggcac	aaataagaac	gaatactgcc	attaagactc	gtgatccagc	9540
gactgacacc	attgcatcat	ctaagggcct	caaaactacc	tcggaactgc	tgcgctgac	9600
tggacaccac	agaggttccg	agcactttag	gttgcaccaa	atgtcccacc	aggtgcaggc	9660
agaaaacgct	ggaacagcgt	gtacagtttg	tcttaacaaa	aagtgagggc	gctgaggctg	9720
agcagggtgg	tgtgacttgt	tatagccttt	agagctgcga	aagcgcgtat	ggatttggct	9780
catcaggcca	gattgagggt	ctgtggacac	atgtcatgtt	agtgtacttc	aatcgccccc	9840
tggatatagc	cccgacaata	ggccgtggcc	tcattttttt	gccttccgca	catttccatt	9900
gctcgggtacc	cacaccttgc	ttctcctgca	cttgccaacc	ttaatactgg	tttacattga	9960
ccaacatctt	acaagcgggg	ggcttgtcta	gggtatatat	aaacagtggc	tctcccaatc	10020
ggttgccagt	ctcttttttc	ctttcttttc	ccacagattc	gaaatctaaa	ctacacatca	10080
cagaattccg	agccgtgagt	atccacgaca	agatcagtgt	cgagacgacg	cgttttgtgt	10140
aatgacacaa	tccgaaagtc	gctagcaaca	cacactctct	acacaaacta	acccagctct	10200
ggtaccatgg	cttctttccac	tgttgctgcg	ccgtacgagt	tcccgcgcgt	gacggagatc	10260
aagegctcgc	tgccagegca	ctgctttgag	gcctcgggtc	cgtggctcgt	ctactacacc	10320
gtgcgcgcgc	tgggcacgcg	cggctcgcgc	gcgctcggcc	tctactacgc	gcgcgcgcgc	10380

gcgatcgtgc	aggagtttgc	cctgctggat	gcggtgctct	gcacggggta	cattctgctg	10440
cagggcatcg	tattctgggg	gttcttcacc	atcgcccatg	actgcggcca	cggcgcgttc	10500
tcgcgttcgc	acctgctcaa	cttcagcgtc	ggcacgctca	ttcactcgat	catcctcacg	10560
ccgtacgagt	catggaagat	ctcgcaccgc	caccaccaca	agaacacggg	caacatcgac	10620
aaggacgaga	ttttctaccc	gcagcgcgag	gccgactcgc	acccactgtc	ccgacacatg	10680
gtgatctcgc	tcggctcggc	ctggttcgcg	tacctcgttg	cgggcttccc	tcctcgcaag	10740
gtgaaccact	tcaacccttg	ggaaccgttg	tacctgcgcc	gcatgtctgc	cgtcacatc	10800
tcactcggtc	cgctcgtggc	gttcgcgggc	ttgtatgcgt	atctcaccta	cgtctatggc	10860
cttaagacca	tggcgctgta	ctacttcgcc	cctctctttg	ggttcgccac	gatgctcgtg	10920
gtcactacct	ttttgcacca	caatgacgag	gaaacgccat	ggtacgccga	ctcggagtgg	10980
acgtacgtca	agggcaacct	ctcgtccgtg	gaccgctcgt	acggcgcgct	catcgacaac	11040
ctgagccaca	acatcggcac	gcaccagatc	caccacctgt	ttccgatcat	cccgcactac	11100
aagctgaacg	aggcgacggc	agcgttcgcg	caggcgttcc	cggagctcgt	gcgcaagagc	11160
gcgtcgccga	tcateccgac	gttcatccgc	atcgggctca	tgtacgccaa	gtacggcgtc	11220
gtggacaagg	acgccaagat	gtttacgctc	aaggaggcca	aggccgccaa	gaccaaggcc	11280
aactaggcgg	ccgcatggag	cgtgtgttct	gagtcgatgt	tttctatgga	gttgtgagtg	11340
ttagtagaca	tgatgggttt	atatatgatg	aatgaataga	tgtgattttg	atttgcacga	11400
tggaattgag	aactttgtaa	acgtacatgg	gaatgtatga	atgtgggggt	tttgtgactg	11460
gataactgac	ggtcagtgga	cgccgttggt	caaatatcca	agagatgcga	gaaactttgg	11520
gtcaagtga	catgtcctct	ctgttcaagt	aaaccatcaa	ctatgggtag	tatatttagt	11580
aaggacaaga	gttgagattc	tttggagtcc	tagaaacgta	ttttcgcgtt	ccaagatcaa	11640
attagtagag	taatacgggc	acgggaatcc	attcatagtc	tcaatcctgc	aggtgagtta	11700
attaagatga	cgacatttgc	gagctggacg	aggaatagat	ggagcgtgtg	ttctgagtcg	11760
atgttttcta	tggagtgttg	agtgttagta	gacatgatgg	gtttatatat	gatgaatgaa	11820
tagatgtgat	tttgatttgc	acgatggaat	tgagaacttt	gtaaacgtac	atgggaatgt	11880
atgaatgtgg	gggttttgtg	actggataac	tgacggtcag	tggacgccgt	tgttcaaata	11940
tccaagagat	gcgagaaact	ttgggtcaag	tgaacatgtc	ctctctgttc	aagtaaacca	12000
tcaactatgg	gtagtatatt	tagtaaggac	aagagttgag	attcttttga	gtcctagaaa	12060
cgtattttcg	cgttccaaga	tcaaatagat	agagtaatac	gggcacggga	atccattcat	12120
agtcctcaatt	ttcccatagg	tgtgctacaa	ggtgttgaga	tgtggtacag	taccaccatg	12180
attcgaggta	aagagcccag	aagtcattga	tgagggtcaag	aaatacacag	atctacagct	12240
caatacaatg	aatatcttct	ttcatattct	tcagggtgaca	ccaaggggtg	ctattttccc	12300
cagaaatgcg	tgaaaaggcg	cgtgtgttagc	gtggagtatg	ggttcggttg	gcgtatcctt	12360
catatatcga	cgaaatagta	gggcaagaga	tgacaaaaag	tatctatatg	tagacagcgt	12420
agaatatgga	tttgatttgg	ataaattcat	ttattgcgtg	tctcaciaat	actctcgata	12480
agttgggggt	aaactggaga	tggaacaatg	tcgatatctc	gacgcatgcg	acgtcggggc	12540
caattcgccc	tatagttagt	cgtattacaa	ttcactggcc	gtcgtttttac	aacgtcgtga	12600
ctgggaaaac	cctggcggtta	cccaacttaa	tcgccttgca	gcacatcccc	ctttcgccag	12660
ctggcgtaat	agcgaagagg	cccgccacca	tcgccccttc	caacagttgc	gcagcctgaa	12720

tggcgaatgg acgcgccctg tagcggcgca ttaagcgcg cgggtgtggt ggttacgcgc 12780
 agcgtgaccg ctacacttgc cagegcccta gcgcccgtc ctttcgcttt cttcccttcc 12840
 tttctcgcca cgttcgccgg ctttccccgt caagctctaa atcgggggct ccctttaggg 12900
 ttccgattta gtgctttacg gcacctcgac cccaaaaaac ttgattaggg tgatggttca 12960
 cgtagtgggc categccctg atagacgggt tttcgccctt tgacgttgga gtccacgttc 13020
 tttaatagtg gactcttggt ccaactgga acaacactca acccta 13066

<210>50

<211>1080

<212>DNA

<213> 瓜果腐霉菌

<220>

<221>CDS

<222>(1).. (1080)

<223> 合成 Δ -17 去饱和酶 (经密码子优化用于解脂耶氏酵母)

<400>50

atg gct tcc tct acc gtt gcc gct ccc tac gag ttc cct act ctc acc 48
 Met Ala Ser Ser Thr Val Ala Ala Pro Tyr Glu Phe Pro Thr Leu Thr
 1 5 10 15

gag arc aag cga tcc ctg cct gcc cac tgc ttc gaa gcc tct gtt ccc 96
 Glu Ile Lys Arg Ser Leu Pro Ala His Cys Phe Glu Ala Ser Val Pro
 20 25 30

tgg tcc ctc tac tat acc gtg cga gct ctg ggc att gcc ggt tcc ctt 144
 Trp Ser Leu Tyr Tyr Thr Val Arg Ala Leu Gly Ile Ala Gly Ser Leu
 35 40 45

gct ctc gga ctg tac tat gct cga gcc ctt gct atc gtg cag gag ttt 192
 Ala Leu Gly Leu Tyr Tyr Ala Arg Ala Leu Ala Ile Val Gln Glu Phe
 50 55 60

gca ctg ctc gat gcc gtc ctt tgc act ggc tac att ctg ctc cag ggt 240
 Ala Leu Leu Asp Ala Val Leu Cys Thr Gly Tyr Ile Leu Leu Gln Gly
 65 70 75 80

atc gtg ttc tgg gga ttc ttt acc atc ggt cac gac tgt gga cat ggt 288
 Ile Val Phe Trp Gly Phe Phe Thr Ile Gly His Asp Cys Gly His Gly
 85 90 95

gcc ttc tcg cga tcc cac ctg ctc aac ttc tct gtt ggc aca ctc att 336
 Ala Phe Ser Arg Ser His Leu Leu Asn Phe Ser Val Gly Thr Leu Ile
 100 105 110

cac tcc atc att ctg act ccc tac gag tcg tgg aag atc agc cat cga 384
 His Ser Ile Ile Leu Thr Pro Tyr Glu Ser Trp Lys Ile Ser His Arg
 115 120 125

cac cat cac aag aac acc ggc aac atc gac aag gat gag atc ttc tac	432
His His His Lys Asn Thr Gly Asn Ile Asp Lys Asp Glu Ile Phe Tyr	
130 135 140	
cct cag cga gaa gcc gac tct cat ccc ctg tcc cga cac atg gtc atc	480
Pro Gln Arg Glu Ala Asp Ser His Pro Leu Ser Arg His Met Val Ile	
145 150 155 160	
tcc ctt ggt tgc gct tgg ttt gcc tac ctc gtt gct gga ttt cct ccc	528
Ser Leu Gly Ser Ala Trp Phe Ala Tyr Leu Val Ala Gly Phe Pro Pro	
165 170 175	
cga aag gtc aac cac ttc aat ccc tgg gag cct ctc tac ctg cga aga	576
Arg Lys Val Asn His Phe Asn Pro Trp Glu Pro Leu Tyr Leu Arg Arg	
180 185 190	
atg tct gcc gtc atc att tcc ctc ggc tct ctc gtg gcc ttt gct ggt	624
Met Ser Ala Val Ile Ile Ser Leu Gly Ser Leu Val Ala Phe Ala Gly	
195 200 205	
ctg tac gcc tac ctt acc tac gtc tac ggc ctc aag acc atg gct ctg	672
Leu Tyr Ala Tyr Leu Thr Tyr Val Tyr Gly Leu Lys Thr Met Ala Leu	
210 215 220	
tat tac ttc gca cct ctc ttt gga ttc gcc acc atg ctg gtt gtc act	720
Tyr Tyr Phe Ala Pro Leu Phe Gly Phe Ala Thr Met Leu Val Val Thr	
225 230 235 240	
acc ttc ctc cat cac aac gac gag gaa act ccc tgg tac gcc gat tcg	768
Thr Phe Leu His His Asn Asp Glu Glu Thr Pro Trp Tyr Ala Asp Ser	
245 250 255	
gag tgg acc tat gtc aag ggc aac ttg tcc tct gtg gac cga agc tac	816
Glu Trp Thr Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr	
260 265 270	
gga gcc ctc atc gac aac ctg tcc cac aac att ggt aca cat cag atc	864
Gly Ala Leu Ile Asp Asn Leu Ser His Asn Ile Gly Thr His Gln Ile	
275 280 285	
cac cat ctg ttt ccc atc att cct cac tac aag ctc aac gag gcc act	912
His His Leu Phe Pro Ile Ile Pro His Tyr Lys Leu Asn Glu Ala Thr	
290 295 300	
gct gcc ttc gct cag gcc ttt ccc gaa ctg gtg cga aag tcg gct tct	960
Ala Ala Phe Ala Gln Ala Phe Pro Glu Leu Val Arg Lys Ser Ala Ser	
305 310 315 320	
ccc atc att ccc acc ttc atc cga att ggt ctt atg tac gcc aag tac	1008
Pro Ile Ile Pro Thr Phe Ile Arg Ile Gly Leu Met Tyr Ala Lys Tyr	
325 330 335	

```

ggc gtg gtc gac aag gat gcc aag atg ttt acc ctc aag gaggcc aag      1056
Gly ValVal Asp Lys Asp Ala Lys Met Phe Thr Leu Lys Glu Ala Lys
          340          345          350

gct gcc aag acc aaa gcc aac taa      1080
Ala Ala Lys Thr Lys Ala Asn
          355

<210>51
<211>359
<212>PRT
<213> 瓜果腐霉菌
<400>51
Met Ala Ser Ser Thr Val Ala Ala Pro Tyr Glu Phe Pro Thr Leu Thr
1          5          10          15
Glu Ile Lys Arg Ser Leu Pro Ala His Cys Phe Glu Ala Ser Val Pro
          20          25          30
Trp Ser Leu Tyr Tyr Thr Val Arg Ala Leu Gly Ile Ala Gly Ser Leu
          35          40          45
Ala Leu Gly Leu Tyr Tyr Ala Arg Ala Leu Ala Ile Val Gln Glu Phe
          50          55          60
Ala Leu Leu Asp Ala Val Leu Cys Thr Gly Tyr Ile Leu Leu Gln Gly
65          70          75          80
Ile Val Phe Trp Gly Phe Phe Thr Ile Gly His Asp Cys Gly His Gly
          85          90          95
Ala Phe Ser Arg Ser His Leu Leu Asn Phe Ser Val Gly Thr Leu Ile
          100          105          110
His Ser Ile Ile Leu Thr Pro Tyr Glu Ser Trp Lys Ile Ser His Arg
          115          120          125
His His His Lys Asn Thr Gly Asn Ile Asp Lys Asp Glu Ile Phe Tyr
          130          135          140
Pro Gln Arg Glu Ala Asp Ser His Pro Leu Ser Arg His Met Val Ile
145          150          155          160
Ser Leu Gly Ser Ala Trp Phe Ala Tyr Leu Val Ala Gly Phe Pro Pro
          165          170          175
Arg Lys Val Asn His Phe Asn Pro Trp Glu Pro Leu Tyr Leu Arg Arg
          180          185          190
Met Ser Ala Val Ile Ile Ser Leu Gly Ser Leu Val Ala Phe Ala Gly
          195          200          205
Leu Tyr Ala Tyr Leu Thr Tyr Val Tyr Gly Leu Lys Thr Met Ala Leu
          210          215          220

```


Tyr Tyr Phe Ala Pro Leu Phe Gly Phe Ala Thr Met Leu Val Val Thr
 225 230 235 240
 Thr Phe Leu His His Asn Asp Glu Glu Thr Pro Trp Tyr Ala Asp Ser
 245 250 255
 Glu Trp Thr Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr
 260 265 270
 Gly Ala Leu Ile Asp Asn Leu Ser His Asn Ile Gly Thr His Gln Ile
 275 280 285
 His His Leu Phe Pro Ile Ile Pro His Tyr Lys Leu Asn Glu Ala Thr
 290 295 300
 Ala Ala Phe Ala Gln Ala Phe Pro Glu Leu Val Arg Lys Ser Ala Ser
 305 310 315 320
 Pro Ile Ile Pro Thr Phe Ile Arg Ile Gly Leu Met Tyr Ala Lys Tyr
 325 330 335
 Gly Val Val Asp Lys Asp Ala Lys Met Phe Thr Leu Lys Glu Ala Lys
 340 345 350
 Ala Ala Lys Thr Lys Ala Asn
 355

<210>52

<211>1080

<212>DNA

<213> 瓜果腐霉菌

<220>

<221>CDS

<222>(1).. (1080)

<223> Δ -17 去饱和酶

<400>52

atg gct tct tcc act gtt gct gcg ccg tac gag ttc ccg acg ctg acg 48
 Met Ala Ser Ser Thr Val Ala Ala Pro Tyr Glu Phe Pro Thr Leu Thr
 1 5 10 15
 gag atc aag cgc tcg ctg cca gcg cac tgc ttt gag gcc tcg gtc ccg 96
 Glu Ile Lys Arg Ser Leu Pro Ala His Cys Phe Glu Ala Ser Val Pro
 20 25 30
 tgg tcg ctc tac tac acc gtg cgc gcg ctg ggc atc gcc ggc tcg ctc 144
 Trp Ser Leu Tyr Tyr Thr Val Arg Ala Leu Gly Ile Ala Gly Ser Leu
 35 40 45
 gcg ctc ggc ctc tac tac gcg cgc gcg ctc gcg atc gtg cag gag ttt 192
 Ala Leu Gly Leu Tyr Tyr Ala Arg Ala Leu Ala Ile Val Gln Glu Phe
 50 55 60

gcc ctg ctg gat gcg gtg ctc tgc acg ggg tac att ctg ctg cag ggc	240
Ala Leu Leu Asp Ala Val Leu Cys Thr Gly Tyr Ile Leu Leu Gln Gly	
65 70 75 80	
atc gta ttc tgg ggg ttc ttc acc atc ggc cat gac tgc ggc cac ggc	288
Ile Val Phe Trp Gly Phe Phe Thr Ile Gly His Asp Cys Gly His Gly	
85 90 95	
gcg ttc tcg cgt tcg cac ctg ctc aac ttc agc gtc ggc acg ctc att	336
Ala Phe Ser Arg Ser His Leu Leu Asn Phe Ser ValGly Thr Leu Ile	
100 105 110	
cac tcg atc atc ctc acg ccg tac gag tca tgg aag atc tcg cac cgc	384
His Ser Ile Ile Leu Thr Pro Tyr Glu Ser Trp Lys Ile Ser His Arg	
115 120 125	
cac cac cac aag aac acg ggc aac atc gac aag gac gag att ttc tac	432
His His His Lys Asn Thr Gly Asn Ile Asp Lys Asp Glu Ile Phe Tyr	
130 135 140	
ccg cag cgc gag gcc gac tcg cac cca ctg tcc cga cac atg gtg atc	480
Pro Gln Arg Glu Ala Asp Ser His Pro Leu Ser Arg His Met Val Ile	
145 150 155 160	
tcg ctc ggc tcg gcc tgg ttc gcg tac ctc gtt gcg ggc ttc cct cct	528
Ser Leu Gly Ser Ala Trp Phe Ala Tyr Leu Val Ala Gly Phe Pro Pro	
165 170 175	
cgc aag gtg aac cacttc aac cct tgg gaa ccg ttg tac ctg cgc cgc	576
Arg Lys Val Asn His Phe Asn Pro Trp Glu Pro Leu Tyr Leu Arg Arg	
180 185 190	
atg tct gcc gtc a tc atctca ctc ggc tcg ctc gtg gcg ttc gcg ggc	624
Met Ser Ala Val Ile Ile Ser Leu Gly Ser Leu Val Ala Phe Ala Gly	
195 200 205	
ttg tat gcg tat ctc acc tac gtc tat ggc ctt aag acc atg gcg ctg	672
Leu Tyr Ala Tyr Leu Thr Tyr Val Tyr Gly Leu Lys Thr Met Ala Leu	
210 215 220	
tac tac ttc gcc cct ctc ttt ggg ttc gcc acg atg ctc gtg gtc act	720
Tyr Tyr Phe Ala Pro Leu Phe Gly Phe Ala Thr Met Leu Val Val Thr	
225 230 235 240	
acc ttt ttg cac cac aat gac gag gaa acg cca tgg tac gcc gac tcg	768
Thr Phe Leu His His Asn Asp Glu Glu Thr Pro Trp Tyr Ala Asp Ser	
245 250 255	
gag tgg acg tac gtc aag ggc aac ctc tcg tcc gtg gac cgc tcg tac	816
Glu Trp Thr Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr	
260 265 270	

ggc gcg ctc atc gac aac ctg agc cac aac atc ggc acg cac cag atc	864
Gly Ala Leu Ile Asp Asn Leu Ser His Asn Ile Gly Thr His Gln Ile	
275 280 285	
cac cac ctg ttt ccg atc atc ccg cac tac aag ctg aac gag gcg acg	912
His His Leu Phe Pro Ile Ile Pro His Tyr Lys Leu Asn Glu Ala Thr	
290 295 300	
gca gcg ttc gcg cag gcg ttc ccg gag ctc gtg cgc aag agc gcg tcg	960
Ala Ala Phe Ala Gln Ala Phe Pro Glu Leu Val Arg Lys Ser Ala Ser	
305 310 315 320	
ccg atc atc ccg acg ttc atc cgc atc ggg ctc atg tac gcc aag tac	1008
Pro Ile Ile Pro Thr Phe Ile Arg Ile Gly Leu Met Tyr Ala Lys Tyr	
325 330 335	
ggc gtc gtg gac aag gac gcc aag atg ttt acg ctc aag gag gcc aag	1056
Gly Val Val Asp Lys Asp Ala Lys Met Phe Thr Leu Lys Glu Ala Lys	
340 345 350	
gcc gcc aag acc aag gcc aac tag	1080
Ala Ala Lys Thr Lys Ala Asn	
355	

<210>53

<211>359

<212>PRT

<213> 瓜果腐霉菌

<400>53

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

His His His Lys Asn Thr Gly Asn Ile Asp Lys Asp Gln Ile Phe Tyr	
130	135 140
Pro Gln Arg Glu Ala Asp Ser His Pro Leu Ser Arg His Met Val Ile	
145	150 155 160
Ser Leu Gly Ser Ala Trp Phe Ala Tyr Leu Val Ala Gly Phe Pro Pro	
165	170 175
Arg Lys Val Asn His Phe Asn Pro Trp Glu Pro Leu Tyr Leu Arg Arg	
180	185 190
Met Ser Ala Val Ile Ile Ser Leu Gly Ser Leu Val Ala Phe Ala Gly	
195	200 205
Leu Tyr Ala Tyr Leu Thr Tyr Val Tyr Gly Leu Lys Thr Met Ala Leu	
210	215 220
Tyr Tyr Phe Ala Pro Leu Phe Gly Phe Ala Thr Met Leu Val Val Thr	
225	230 235 240
Thr Phe Leu His His Asn Asp Glu Glu Thr Pro Trp Tyr Ala Asp Ser	
245	250 255
Glu Trp Thr Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr	
260	265 270
Gly Ala Leu Ile Asp Asn Leu Ser His Asn Ile Gly Thr His Gln Ile	
275	280 285
His His Leu Phe Pro Ile Ile Pro His Tyr Lys Leu Asn Glu Ala Thr	
290	295 300
Ala Ala Phe Ala Gln Ala Phe Pro Glu Leu Val Arg Lys Ser Ala Ser	
305	310 315 320
Pro Ile Ile Pro Thr Phe Ile Arg Ile Gly Leu Met Tyr Ala Lys Tyr	
325	330 335
Gly Val Val Asp Lys Asp Ala Lys Met Phe Thr Leu Lys Glu Ala Lys	
340	345 350
Ala Ala Lys Thr Lys Ala Asn	
355	

<210>54

<211>9570

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pY117

<400>54

ggccgccacc gcgccccgag attccggcct ctccggccgc caagcgaccc ggggtggacgt 60

ctagaggtac ctagcaatta acagatagtt tgccggtgat aattctctta acctcccaca 120

ctcctttgac	ataacgattt	atgtaacgaa	actgaaattt	gaccagatat	tgtgtccgcg	180
gtggagctcc	agcttttgtt	cccttttagtg	agggttttaa	cgagcttggc	gtaatcatgg	240
tcatagctgt	tacctgtgtg	aaattgttat	ccgctcacia	ttccacacia	cgtacgagcc	300
ggaagcataa	agtgtaaagc	ctgggggtgcc	taatgagtga	gctaactcac	attaattgcg	360
ttgcgctcac	tgeccgcttt	ccagtcggga	aacctgtcgt	gccagctgca	ttaatgaatc	420
ggccaacgcg	cggggagagg	cggtttgcgt	attgggcgct	cttccgcttc	ctcgctcact	480
gactcgctgc	gctcggtcgt	tcggctgcgg	cgagcgggat	cagctcactc	aaaggcggta	540
atacggttat	ccacagaatc	aggggataac	gcaggaaaga	acatgtgagc	aaaaggccag	600
caaaaggcca	ggaaccgtaa	aaaggccgcg	ttgctggcgt	ttttccatag	gctccgcccc	660
cctgacgagc	atcacaaaaa	tcgacgtca	agtcagaggt	ggcgaaaccc	gacaggacta	720
taaagatacc	aggcggttcc	ccctggaagc	tccctcgtgc	gctctcctgt	tccgaccctg	780
ccgcttaccg	gatacctgtc	cgcctttctc	ccttcgggaa	gcgtggcgct	ttctcatagc	840
tcacgctgta	ggtatctcag	ttcggtgtag	gtcgttcgct	ccaagctggg	ctgtgtgcac	900
gaaccccccg	ttcagcccca	ccgctgcgcc	ttatccggta	actatcgctc	tgagtccaac	960
ccggttaagac	acgacttata	gccactggca	gcagccactg	gtaacaggat	tagcagagcg	1020
aggtatgtag	gcggtgctac	agagttcttg	aagtgggtggc	ctaactacgg	ctacactaga	1080
aggacagtat	ttggtatctg	cgctctgctg	aagccagtta	ccttcggaaa	aagagttggg	1140
agctcttgat	ccggcaaaaa	aaccaccgct	ggtagcgggtg	gtttttttgt	ttgcaagcag	1200
cagattacgc	gcagaaaaaa	aggatctcaa	gaagatcctt	tgatcttttc	tacgggggtct	1260
gacgctcagt	ggaacgaaaa	ctcacgttaa	gggatttttg	tcatgagatt	atcaaaaagg	1320
atcttcacct	agatcctttt	aaattaaaaa	tgaagtttta	aatcaatcta	aagtatatat	1380
gagtaaacct	ggtctgacag	ttaccaatgc	ttaatcagtg	aggcacctat	ctcagcgatc	1440
tgtctatttc	gttcatccat	agttgcctga	ctccccgtcg	tgtagataac	tacgatacgg	1500
gagggtttac	catctggccc	cagtgtctga	atgataccgc	gagaccacac	ctcaccggct	1560
ccagatttat	cagcaataaa	ccagccagcc	ggaaggggccg	agcgcagaag	tggctcctgca	1620
actttatccg	cctccatcca	gtctattaat	tgttgccggg	aagctagagt	aagtagttcg	1680
ccagtttaata	gtttgcgcaa	cgttgttgcc	attgctacag	gcacgttggt	gtcacgctcg	1740
tcgtttggta	tggcttcatt	cagctccggg	tcccaacgat	caaggcgagt	tacatgatcc	1800
cccatgttgt	gcaaaaaagc	ggttagctcc	ttcggtcctc	cgatcgttgt	cagaagtaag	1860
ttggccgcag	tgttatcact	catggttatg	gcagcactgc	ataattctct	tactgtcatg	1920
ccatccgtaa	gatgcttttc	tgtgactggg	gagtactcaa	ccaagtcatt	ctgagaatag	1980
tgtatgcggc	gaccgagttg	ctcttgcccc	gcgtcaatac	gggataatac	cgcgccacat	2040
agcagaactt	taaaagtgtc	catcattgga	aaacgttctt	cggggcgaaa	actctcaagg	2100
atcttaccgc	tgttgagatc	cagttcgatg	taaccacactc	gtgcacccaa	ctgatcttca	2160
gcacctttta	ctttcaccag	cgtttctggg	tgagcaaaaa	caggaaggca	aaatgccgca	2220
aaaaagggaa	taagggcgac	acggaaatgt	tgaatactca	tactcttcct	ttttcaatat	2280
tattgaagca	tttatcaggg	ttattgtctc	atgagcggat	acatatattga	atgtatttag	2340
aaaaataaac	aaataggggt	tccgcgcaca	tttccccgaa	aagtgccacc	tgacgcgccc	2400
tgtagcggcg	cattaagcgc	ggcgggtgtg	gtggttacgc	gcagcgtgac	cgctacactt	2460

gccagegccc	tagegcegc	tcctttegct	ttcttccctt	cctttctcgc	cacgttcgcc	2520
ggctttcccc	gtcaagctct	aaatcggggg	ctcccttttag	ggttccgatt	tagtgcttta	2580
cggcacctcg	acccccaaaa	acttgattag	ggtgatgggt	cacgtagtgg	gccatcgccc	2640
tgatagacgg	tttttgeccc	tttgacgttg	gagtcacagt	tctttaatag	tggactcttg	2700
ttccaaactg	gaacaacact	caaccctatc	tcggtctatt	cttttgattt	ataagggatt	2760
ttgccgattt	cggcctattg	gttaaaaaat	gagctgattt	aacaaaaatt	taacgcgaat	2820
tttaacaaaa	tattaacgct	tacaatttcc	attcgccatt	caggctgcgc	aactgttggg	2880
aagggcgatc	ggtgcgggce	tcttcgctat	tacgccagct	ggcgaaaggg	ggatgtgctg	2940
caaggcgatt	aagttgggta	acgccagggt	tttcccagtc	acgacgttgt	aaaacgacgg	3000
ccagtgaatt	gtaatacgac	tcactatagg	gcgaattggg	taccgggccc	cccctcgagg	3060
tcgatggtgt	cgataagctt	gatatcgaat	tcatgtcaca	caaaccgatc	ttcgccctcaa	3120
ggaaacctaa	ttctacatcc	gagagactgc	cgagatccag	tctacactga	ttaatttttcg	3180
ggccaataat	ttaaaaaaat	cgtgttatat	aatattatat	gtattatata	tatacatcat	3240
gatgatactg	acagtcatgt	cccattgcta	aatagacaga	ctccatctgc	cgcctccaac	3300
tgatgttctc	aatattttaag	gggtcatctc	gcattgttta	ataataaaca	gactccatct	3360
accgcctcca	aatgatgttc	tcaaaatata	ttgtatgaac	ttatttttat	tacttagtat	3420
tattagacaa	cttacttgct	ttatgaaaaa	cacttcctat	ttaggaaaca	atttataatg	3480
gcagttcggt	catttaacaa	tttatgtaga	ataaatgtta	taaatgcgta	tgggaaatct	3540
taaatatgga	tagcataaat	gatatctgca	ttgcctaatt	cgaaatcaac	agcaacgaaa	3600
aaaatccctt	gtacaacata	aatagtcac	gagaaatata	aactatcaaa	gaacagctat	3660
tcacacgtta	ctattgagat	tattattgga	cgagaatcac	acactcaact	gtctttctct	3720
cttctagaaa	tacaggtaca	agtatgtact	attctcattg	ttcatacttc	tagtcatttc	3780
atcccacata	ttccttggat	ttctctccaa	tgaatgacat	tctatcttgc	aaattcaaca	3840
attataataa	gatataccaa	agtagcggtg	tagtggcaat	caaaaagctt	ctctggtgtg	3900
cttctcgat	ttatttttat	tctaattgat	cattaaaggt	atatatttat	ttcttgttat	3960
ataatccttt	tgtttattac	atgggctgga	tacataaagg	tattttgatt	taattttttg	4020
cttaaattca	atccccctc	gttcagtgtc	aactgtaatt	gtaggaaatt	accatacttt	4080
tgaagaagca	aaaaaaatga	aagaaaaaaa	aaatcgattt	tccagggttag	acgttccgca	4140
gaatctagaa	tgcggtatgc	ggtacattgt	tcttcgaacg	taaaagttgc	gctccctgag	4200
atattgtaca	tttttgcttt	tacaagtaca	agtacatcgt	acaactatgt	actactgttg	4260
atgcateccac	aacagtttgt	tttgtttttt	tttgtttttt	ttttttctaa	tgattcatta	4320
ccgctatgta	tacctacttg	tacttgtagt	aagccggggt	attggcggtc	aattaatcat	4380
agacttatga	atctgcacgg	tgtgcgctgc	gagttacttt	tagcttatgc	atgctacttg	4440
ggtgtaatat	tgggatctgt	tcggaaatca	acggatgctc	aaccgatttc	gacagtaatt	4500
aattaattcc	ctagtcccag	tgtacacccg	ccgatatcgc	ttaccctgca	gccggattaa	4560
ggttggcaat	ttttcacgtc	cttgtctccg	caattactca	ccgggtgggt	tataagattg	4620
caagcgtctt	gatttgtctc	tgtatactaa	catgcaatcg	cgactcgccc	gacgggccac	4680
taacctggcc	agaatctcca	gatccaagta	ttctcttggt	ctgcgatatg	tttccaacac	4740
aaaagccctt	gctgcccagc	cggcaactgc	tgagtgagta	ttccttgcca	taaacgaccc	4800

agaaccactg	tatagtgttt	ggaagcacta	gtcagaagac	cagcgaaaac	aggtggaaaa	4860
aactgagacg	aaaagcaacg	accagaaatg	taatgtgtgg	aaaagcgaca	cacacagagc	4920
agataaagag	gtgacaaata	acgacaaatg	aaatatcagt	atcttcccac	aatcactacc	4980
tctcagctgt	ctgaaggtgc	ggctgatata	tccatcccac	gtctaacgta	tggagtgtga	5040
tagaatatga	cgacacaagc	atgagaacte	gctctctatc	caaccaccga	aacactgtca	5100
ctacagccgt	tcttgttget	ccattegctt	ttgtgattcc	atgccttctc	tggtgactga	5160
caacattcct	tccttttctc	cagccctggt	gttatctgct	catgacctac	ggccactctc	5220
tatcgcatac	taacatagac	gateccagcc	cgctccccac	ttccagggca	ccgttggcaa	5280
gcctcctatc	ctcaagaagg	ctgaggtgc	caacgctgac	atggacgagt	ccttcacggt	5340
aatgtctgga	ggagagatct	tccacgagat	gatgctgcga	cacaacgtcg	acactgtctt	5400
cggttacccc	ggtggagcca	ttctccccgt	ctttgacgcc	attcacaact	ctgagtactt	5460
caactttgtg	ctccctcgac	acgagcaggg	tgccggccac	atggccgagg	gctacgctcg	5520
agcctctggt	aagcccgtg	tcgttctcgt	cacctctggc	cccggtgcca	ccaacgtcat	5580
cacccccatg	caggacgctc	ttccgatgg	tacccccatg	gttgtcttca	ccggtcaggt	5640
cctgacctcc	gttatcgcca	ctgacgcctt	ccaggaggcc	gatgttgctg	gcctctcccg	5700
atcttgcacc	aagtggaaacg	tcatgggtcaa	gaacgttgct	gagctcccc	gacgaatcaa	5760
cgaggccctt	gagattgcta	cttccggccg	acccggctcc	gttctcgtcg	atctgccccaa	5820
ggatgttact	gctgccatcc	tgcgagagcc	catccccacc	aagtccacca	ttccctcgca	5880
ttctctgacc	aacctcacct	ctgccgccgc	caccgagttc	cagaagcagg	ctatccagcg	5940
agccgccaac	ctcatcaacc	agtccaagaa	gcccgtcctt	tacgtcggac	agggtatcct	6000
tggctccgag	gagggctcta	agctgcttaa	ggagctggct	gagaaggccg	agattcccgt	6060
caccactact	ctgcagggtc	ttgggtgcctt	tgacgagcga	gaccccaagt	ctctgcacat	6120
gctcggtatg	cacggttccg	gctacgcca	catggccatg	cagaacgctg	actgtatcat	6180
tgtctctggc	gcccgatattg	atgaccgagt	taccggctcc	atccccaagt	ttgccccga	6240
ggctcgagcc	gctgcccttg	agggtcgagg	tggatttggt	cactttgaga	tccaggccaa	6300
gaacatcaac	aaggttggtc	aggccaccga	agccgttgag	ggagacgtta	ccgagtctgt	6360
ccgacagctc	atccccctca	tcaacaaggt	ctctgccgct	gagcgagctc	cctggactga	6420
gactatccag	tcctggaagc	agcagttccc	cttccctctc	gaggctgaag	gtgaggatgg	6480
tgttatcaag	ccccagtcg	tcattgtctt	gctctctgac	ctgacagaga	acaacaagga	6540
caagaccatc	atcaccaccg	gtgttggtca	gcacagatg	tggactgccc	agcatttccg	6600
atggcgacac	cctcgaacca	tgatcacttc	tgggtggtctt	ggaactatgg	gttacggcct	6660
gcccgcgct	atcggcgcca	aggttgcccc	acctgactgc	gacgtcattg	acatcgatgg	6720
tgacgtttct	ttcaacatga	ctctgaccga	gctgtccacc	gccgttcagt	tcaacattgg	6780
cgtaaggct	attgtcctca	acaacgagga	acagggtatg	gtcaccacgc	tgcagtctct	6840
cttctacgag	aaccgatact	gccacactca	tcagaagaac	cccgaattca	tgaagctggc	6900
cgagtccatg	ggcatgaagg	gtatccgaat	cactcacatt	gaccagctgg	aggccggctt	6960
caaggagatg	ctcgcataca	agggccctgt	gctcggttgg	gttggtgtcg	acaagaagat	7020
ccccgttctt	cccatgggtc	ccgttggtaa	ggctttgcat	gagttccttg	tctacgacgc	7080
tgacgcccag	gctgcttctc	gacccgatcg	actgaagaat	gcccccgccc	ctcacgtcca	7140

ccagaccacc	tttgagaact	aagtggaaag	gaacacaagc	aatccgaacc	aaaaataatt	7200
gggggtcccg	gcccacagag	tctagtgcag	acctaaaatg	accacagtaa	attatagctg	7260
ttattaaaca	tgagattttg	accaacaaga	gcgtaggaat	gttattagct	actacttgta	7320
catacacage	atttgtttta	aataatgttg	cctccagggg	cagtgcagtc	aggaccacga	7380
tccgtggcca	gctctctgac	ttcagaccgc	ttgtacttaa	gcagctcgca	acactgttgt	7440
cgaggattga	acttgccata	ttcgattttg	tggctcatgaa	tccagcacac	ctcattttaa	7500
tgtagctaac	ggtagcaggc	gaactactgg	tacatacctc	ccccggaata	tgtacaggca	7560
taatgcgtat	ctgtgggaca	tgtggctcgt	gcgccattat	gtaagcagcg	tgtactcctc	7620
tgactgtcca	tatggtttgc	tccatctcac	cctcatcggt	ttcattgttc	acaggcggcc	7680
acaaaaaac	tgtcttctct	ccttctctct	tcgccttagt	ctactcggac	cagttttagt	7740
ttagcttggc	gccactggat	aaatgagacc	tcaggccttg	tgatgaggag	gtcacttatg	7800
aagcatgtta	ggaggtgctt	gtatggatag	agaagcaccc	aaaataataa	gaataataat	7860
aaaacagggg	gcgttgtcat	ttcatatcgt	gttttcacca	tcaatacacc	tccaaacaat	7920
gcccttcatg	tggccagccc	caatattgtc	ctgtagttca	actctatgca	gctcgtatct	7980
tattgagcaa	gtaaaactct	gtcagccgat	attgcccgac	ccgcgacaag	ggtaacaag	8040
gtgggtgtaag	gccttcgcag	aagtcaaaac	tgtgccaaac	aaacatctag	agtctctttg	8100
gtgtttctcg	catatatttw	atcggtgtgc	ttacgtatct	gcgcctcggt	accggactaa	8160
tttcggatca	tccccaatac	gctttttctt	cgcagctgtc	aacagtgtcc	atgatctatc	8220
cacctaaatg	ggtcatatga	ggcgtataat	ttcgtgggtg	tgataataat	tcccatatat	8280
ttgacacaaa	acttcccccc	ctagacatac	atctcacaat	ctcacttctt	gtgcttctgt	8340
cacacatctc	ctccagctga	cttcaactca	cacctctgcc	ccagttgggtc	tacagcggta	8400
taaggtttct	ccgcatagag	gtgcaccact	cctcccgata	cttgttttgt	tgacttgtgg	8460
gtcacgacat	atatactctac	acacattgcg	ccaccctttg	gttcttccag	cacaacaaaa	8520
acacgacacg	ctaaccatgg	ccaatttact	gaccgtacac	caaaatttgc	ctgcattacc	8580
ggtcgatgca	acgagtgatg	aggttcgcaa	gaacctgatg	gacatgttca	gggatcgcca	8640
ggcgttttct	gagcatacct	ggaaaatgct	tctgtccggt	tgccgggtcgt	gggcggcatg	8700
gtgcaagtgt	aataaccgga	aatggtttcc	cgcagaacct	gaagatgttc	gcgattatct	8760
tctatatctt	caggcgcgcg	gtctggcagt	aaaaactatc	cagcaacatt	tgggccagct	8820
aaacatgctt	catcgtcggg	ccgggctgcc	acgaccaagt	gacagcaatg	ctgttttact	8880
ggttatgcgg	cggatccgaa	aagaaaacgt	tgatgccggg	gaacgtgcaa	aacaggctct	8940
agcggttcgaa	cgcactgatt	tcgaccaggt	tcgttacttc	atggaaaata	gcgatcgctg	9000
ccaggatata	cgtaatctgg	catttctggg	gattgcttat	aacaccctgt	tacgtatagc	9060
cgaaattgcc	aggatcaggg	ttaaagatat	ctcacgtact	gacggtggga	gaatgttaat	9120
ccatattggc	agaacgaaaa	cgctgggttag	caccgcaggt	gtagagaagg	cacttagcct	9180
gggggtaact	aaactggctg	agcgatggat	ttccgtctct	ggtgtagctg	atgatccgaa	9240
taactacctg	ttttgccggg	tcagaaaaaa	tgggtgttgc	gcgccatctg	ccaccagcca	9300
gctatcaact	cgcgccctgg	aagggttttt	tgaagcaact	catcgattga	tttacggcgc	9360
taaggatgac	tctggctcaga	gatacctggc	ctggctctgga	cacagtgtccc	gtgtcggagc	9420
cgcgcgagat	atggcccgcg	ctggagtttc	aataccggag	atcatgcaag	ctgggtggctg	9480

gaccaatgta aatattgtca tgaactatat ccgtaacctg gatagtgaaa caggggcaat	9540
ggtgcgcctg ctggaagatg gcgattaagc	9570
<210>55	
<211>15743	
<212>DNA	
<213> 人工序列	
<220>	
<223> 质粒 pZP2-2988	
<400>55	
ggccgcatgt acatacaaga ttatttatag aatgaatcg cgatcgaaca aagagtacga	60
gtgtacgagt aggggatgat gataaaagtg gaagaagttc cgcatctttg gatttatcaa	120
cgtgtaggac gatacttcct gtaaaaatgc aatgtcttta ccataggttc tgctgtagat	180
gttattaact accattaaca tgtctacttg tacagttgca gaccagttgg agtatagaat	240
ggtacactta ccaaaaagtg ttgatggttg taactacgat atataaaaact gttgacggga	300
tctgtatatt cggtaagata tattttgtgg ggtttttagtg gtgttttaac agtgtacgca	360
gtactataga ggaacaattg ccccgagaa gacggccagg ccgcctagat gacaaattca	420
acaactcaca gctgactttc tgccattgcc actagggggg gcccttttta tatggccaag	480
ccaagctctc cacgtcggtt gggctgcacc caacaataaa tgggtagggt tgcaccaaca	540
aagggatggg atggggggta gaagatacga ggataacggg gctcaatggc acaaataaga	600
acgaatactg ccattaagac tcgtgatcca gcgactgaca ccattgcac atctaagggc	660
ctcaaaacta cctcggaact gctgcgtga tctggacacc acagaggttc cgagcacttt	720
aggttgcacc aaatgtccca ccagggtgcag gcagaaaacg ctggaacagc gtgtacagtt	780
tgtcttaaca aaaagtgagg gcgctgaggt cgagcagggt ggtgtgactt gttatagcct	840
ttagagctgc gaaagcgcgt atggatttgg ctcatcaggc cagattgagg gtctgtggac	900
acatgtcatg ttagtgtact tcaatcgccc cctggatata gccccgacaa taggccgtgg	960
cctcattttt ttgccttcg cacatttcca ttgctcggtg cccacacctt gcttctcctg	1020
cacttgccaa ccttaatact ggtttacatt gaccaacatc ttacaagcgg ggggcttgtc	1080
tagggatatat ataaacagtg gctctcccaa tcggttgcca gtctcttttt tcctttcttt	1140
ccccacagat tcgaaatcta aactacacat cacaccatgg aggtcgtgaa cgaaatcgtc	1200
tccattggcc aggaggttct tcccaaggtc gactatgctc agctctggtc tgatgcctcg	1260
cactgcgagg tgctgtacct ctccatcgcc ttctcatcc tgaagttcac ccttggtcct	1320
ctcggaccca agggtcagtc tcgaatgaag tttgtgttca ccaactacaa cctgctcatg	1380
tccatctact cgctgggctc ctctctctct atggcctacg ccatgtacac cattgggtgc	1440
atgtccgaca actgcgagaa ggctttcgac aacaatgtct tccgaatcac cactcagctg	1500
ttctacctca gcaagttcct cgagtacatt gactccttct atctgccccct catgggcaag	1560
cctctgacct ggttgcagtt ctttcacat ctcggagctc ctatggacat gtggctgttc	1620
tacaactacc gaaacgaagc cgtttggatc tttgtgctgc tcaacggctt cattcactgg	1680
atcatgtacg gctactattg gacccgactg atcaagctca agttccctat gcccgaagtc	1740
ctgattactt ctatgcagat cattcagttc aacgttggct tctacatcgt ctggaagtac	1800

cggaacattc	cctgctaccg	acaagatgga	atgagaatgt	ttggctgggt	tttcaactac	1860
ttctacgttg	gtactgtcct	gtgtctgttc	ctcaacttct	acgtgcagac	ctacatcgtc	1920
cgaaagcaca	agggagccaa	aaagattcag	tgagcggccg	caagtgtgga	tggggaagtg	1980
agtgcccggt	tctgtgtgca	caattggcaa	tccaagatgg	atggattcaa	cacagggata	2040
tagcgagcta	cgtgggtggg	cgaggatata	gcaacggata	tttatgtttg	acacttgaga	2100
atgtacgata	caagcactgt	ccaagtacaa	tactaaacat	actgtacata	ctcatactcg	2160
tacccgggca	acggtttcac	ttgagtgcag	tggttagtgc	tcttactcgt	acagtgtgca	2220
atactgcgta	tcatagtctt	tgatgtatat	cgtattcatt	catgttagtt	gcgtacgggc	2280
gtcgttgctt	gtgtgatttt	tgaggacceca	tccctttggg	atataagtat	actctggggg	2340
taaggttgcc	cgtgtagtct	aggttatagt	tttcatgtga	aataaccgaga	gccgaggggag	2400
aataaacggg	ggtattttgga	cttgtttttt	tgcgggaaaa	gcgtcgaatc	aaccctgcgg	2460
gccttgcacc	atgtccacga	cgtgtttctc	gccccaatc	gccccttgca	cgtcaaaatt	2520
aggcctccat	ctagaccctt	ccataacatg	tgactgtggg	gaaaagtata	agggaaacca	2580
tgcaaccata	gacgacgtga	aagacgggga	ggaaccaatg	gaggccaaag	aaatggggta	2640
gcaacagtcc	aggagacaga	caaggagaca	aggagagggc	gcccgaaga	tcggaaaaac	2700
aaacatgtcc	aattggggca	gtgacggaaa	cgacacggac	acttcagtac	aatggaccga	2760
ccatctccaa	gccagggtta	ttccgggtatc	accttggccg	taacctcccg	ctggtacctg	2820
atattgtaca	cgttcacatt	caatatactt	tcagctacaa	taagagaggc	tgtttgtcgg	2880
gcatgtgtgt	ccgtcgtatg	gggtgatgtc	cgagggcgaa	attcgctaca	agcttaactc	2940
tggcgcttgt	ccagtatgaa	tagacaagtc	aagaccagtg	gtgccatgat	tgacaggggag	3000
gtacaagact	tcgatactcg	agcattactc	ggacttgtgg	cgattgaaca	gacgggcgat	3060
cgcttctccc	ccgtattgcc	ggcgcgccag	ctgcattaat	gaatcggcca	acgcgcgggg	3120
agaggcggtt	tgcgtattgg	gcgctcttcc	gcttctctgc	tactgactc	gctgcgctcg	3180
gtcgttcggc	tgcggcgagc	ggtatcagct	cactcaaagg	cggttaatacg	gttatccaca	3240
gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaaag	gccagcaaaa	ggccaggaac	3300
cgtaaaaagg	ccgcgttgct	ggcggttttc	cataggctcc	gccccctga	cgagcatcac	3360
aaaaatcgac	gctcaagtca	gaggtggcga	aacccgacag	gactataaag	ataccaggcg	3420
tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	taccggatac	3480
ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	atagctcacg	ctgtaggtat	3540
ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	tgcacgaacc	ccccgttcag	3600
cccgaccgct	gcgccttata	cggtaactat	cgtcttgagt	ccaacccggg	aagacacgac	3660
ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	tgtaggcggt	3720
gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaagaac	agtatttggt	3780
atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	ttgatccggc	3840
aaacaaacca	ccgttggtag	cggtgggttt	tttgtttgca	agcagcagat	tacgcgcaga	3900
aaaaaaggat	ctcaagaaga	tcctttgatc	ttttctacgg	ggtctgacgc	tcagtggaac	3960
gaaaactcac	gttaagggat	tttggtcattg	agattatcaa	aaaggatctt	cacctagatc	4020
cttttaaatt	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	aacttgggtct	4080
gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	atttcgttca	4140

tccatagttg	cctgactccc	cgtcgtgtag	ataactacga	tacgggaggg	cttaccatct	4200
ggccccagt	ctgcaatgat	accgcgagac	ccacgctcac	cggtccaga	tttatcagca	4260
ataaaccage	cagccggaag	ggccgagcgc	agaagtggtc	ctgcaacttt	atccgcctcc	4320
atccagtcta	ttaattgttg	ccgggaagct	agagtaagta	gttcgccagt	taatagtttg	4380
cgcaacgttg	ttgccattgc	tacaggcate	gtgggtgtcac	gctcgtcggt	tggatatggct	4440
tcattcagct	ccggttccca	acgatcaagg	cgagttacat	gatcccccat	gttgtgcaaa	4500
aaagcggtta	gctccttcgg	tectccgate	gttgtcagaa	gtaagttggc	cgcagtgtta	4560
tcactcatgg	ttatggcagc	actgcataat	tctcttactg	tcatgccatc	cgtaagatgc	4620
ttttctgtga	ctgggtgagta	ctcaaccaag	tcattctgag	aatagtgtat	gcggcgaccg	4680
agttgctctt	gcccggcgtc	aatacgggat	aataccgcgc	cacatagcag	aactttaaaa	4740
gtgctcatca	ttggaaaacg	ttcttcgggg	cgaaaactct	caaggatctt	accgctgttg	4800
agatccagtt	cgatgtaacc	cactcgtgca	cccaactgat	cttcagcatc	ttttactttc	4860
accagcgttt	ctgggtgagc	aaaaacagga	aggcaaaatg	ccgcaaaaaa	gggaataagg	4920
gcgacacgga	aatgttgaat	actcatactc	ttcctttttc	aatattattg	aagcatttat	4980
cagggttatt	gtctcatgag	cggatacata	tttgaatgta	tttagaaaaa	taaacaataa	5040
ggggttccgc	gcacatttcc	ccgaaaagtg	ccacctgatg	cggtgtgaaa	taccgcacag	5100
atgcgtaagg	agaaaatacc	gcatacaggaa	attgtaagcg	ttaatatattt	gttaaaattc	5160
gcgttaaatt	tttgttaaat	cagctcattt	tttaaccaat	aggccgaaat	cggcaaaatc	5220
ccttataaat	caaaagaata	gaccgagata	gggttgagtg	ttgttccagt	ttggaacaag	5280
agtcactat	taaagaacgt	ggactccaac	gtcaaagggc	gaaaaaccgt	ctatcagggc	5340
gatggcccac	tacgtgaacc	atcacccata	tcaagttttt	tggggtcgag	gtgccgtaaa	5400
gcactaaatc	ggaaccctaa	aggagacccc	cgatttagag	cttgacgggg	aaagccggcg	5460
aacgtggcga	gaaaggaagg	gaagaaagcg	aaaggagcgg	gcgctagggc	gctggcaagt	5520
gtagcggta	cgctgcgcgt	aaccaccaca	cccgcgcgc	ttaatgcgcc	gctacagggc	5580
gcgtccattc	gccattcagg	ctgcgcaact	gttgggaagg	gcgatcggtg	cgggcctctt	5640
cgctattacg	ccagctggcg	aaagggggat	gtgctgcaag	gcgattaagt	tgggtaacgc	5700
cagggttttc	ccagtcacga	cgttgtaaaa	cgacggccag	tgaattgtaa	tacgactcac	5760
tatagggcga	attggggccg	acgtcgcatg	cgctgatgac	actttggtct	gaaagagatg	5820
cattttgaat	cccaaacttg	cagtgcceaa	gtgacataca	tctccgcgtt	ttggaaaatg	5880
ttcagaaaca	gttgattgtg	ttggaatggg	gaatggggaa	tggaaaaatg	actcaagtat	5940
caattccaaa	aacttctctg	gctggcagta	cctactgtcc	atactactgc	attttctcca	6000
gtcaggccac	tctatactcg	acgacacagt	agtaaaaccc	agataatttc	gacataaaca	6060
agaaaacaga	cccaataata	tttatatata	gtcagccggt	tgtccagttc	agactgtaat	6120
agccgaaaaa	aaatccaaag	tttctattct	aggaaaatat	attccaatat	ttttaattct	6180
taatctcatt	tattttattc	tagcgaaata	catttcagct	acttgagaca	tgtgataccc	6240
acaaatcgga	ttcggactcg	gttggtcaga	agagcatatg	gcattcgtgc	tcgcttggtc	6300
acgtattctt	cctgttccat	ctcttgcccg	acaatcacac	aaaaatgggg	tttttttttt	6360
aattctaatt	attcattaca	gcaaaattga	gatatagcag	accacgtatt	ccataatcac	6420
caaggaagtt	cttgggcgct	ttaattaact	cacctgcagg	attgagacta	tgaatggatt	6480

cccgtgcccc	tattactcta	ctaatttgat	cttggaaacgc	gaaaatacgt	ttctaggact	6540
ccaaagaate	tcaactcttg	tccttactaa	atatactacc	catagttgat	ggtttacttg	6600
aacagagagg	acatgttcac	ttgacccaaa	gtttctcgc	tctcttggat	atttgaacaa	6660
cggcgtccac	tgaccgtcag	ttatccagtc	acaaaacccc	cacattcata	cattcccatg	6720
tacgtttaca	aagttctcaa	ttccatcggt	caaatcaaaa	tcacatctat	tcattcatca	6780
tatataaacc	catcatgtct	actaacactc	acaactccat	agaaaacatc	gactcagaac	6840
acacgctcca	tgcggccgct	tactgagcct	tggcaccggg	ctgcttctcg	gccattcgag	6900
cgaactggga	caggtatcgg	agcaggatga	cgagaccttc	atggggcaga	gggtttcgg	6960
aggggaggtt	gtgcttcttg	cacagctgtt	ccacctggta	ggaaacggca	gtgaggttgt	7020
gtcagagcag	ggtgggccag	agatgggtgt	cgatctggta	gttcaggcct	ccaaagaacc	7080
agtcagtaat	gatgcctcgt	cgaatgttca	tgggtctcatg	gatctgaccc	acagagaagc	7140
catgtccgtc	ccagacggaa	tcaccgatct	tctccagagg	gtagtggttc	atgaagacca	7200
cgatggcaat	tccgaagcca	ccgacgagct	cggaaacaaa	gaacaccagc	atcgagggtca	7260
ggatggaggg	cataaagaag	aggtggaaca	gggtcttgag	agtccagtgc	agagcgagtc	7320
caatggcctc	tttcttgtac	tgagatcggt	agaactgggt	gtctcgggtcc	ttgagggatc	7380
gaacggtcag	cacagactgg	aaacaccaga	tgaatcgcag	gagaatacag	atgaccagga	7440
aatagtactg	ttggaactga	atgagctttc	gggagatggg	agaagctcga	gtgacatcgt	7500
cctcggaaca	ggcgagcaga	ggcaggttat	caatgtcggg	atcgtgaccc	tgaacgttgg	7560
tagcagaatg	atgggcgttg	tgtctgtcct	tccaccaggt	cacggagaag	ccctggagtc	7620
cgttgccaaa	gaccagaccc	aggacgttat	tccagtttcg	gttcttgaag	gtctggtggt	7680
ggcagatgtc	atgagacagc	catcccatit	gctggtagtg	cataccgagc	acgagagcac	7740
caatgaagta	caggttggtac	tggaccagca	tgaagaaggc	aagcacgcca	agaccaggg	7800
tgggtcaagat	cttgtacgag	taccagaggg	gagaggcgtc	aaacatgcca	gtggcgatca	7860
gctcttctcg	gagctttcgg	aaatcctcct	gagcttcgtt	gacggcagcc	tggggaggca	7920
gctcggaagc	ctggttgatc	ttgggcattc	gcttgagctt	gtcgaaggct	tcctgagagt	7980
gcataacat	gaaggcgtca	gtagcatctc	gtccctggta	gttctcaatg	atttcagctc	8040
caccagggtg	gaagttcacc	caagcggaga	cgtcgtacac	ctttccgtcg	atgacgaggg	8100
gcagagcctg	tcgagaagcc	ttcaccatgg	ttgtgaatta	gggtgggtgag	aatggttggt	8160
tgtagggaag	aatcaaaggc	cggtctcggg	atccgtgggt	atatatatat	atatatatat	8220
atacgatcct	tcgttacctc	cctgttctca	aaactgtggt	ttttcgtttt	tcgttttttg	8280
ctttttttga	tttttttagg	gccaaactaag	cttccagatt	tcgctaataca	cctttgtact	8340
aattacaaga	aaggaagaag	ctgattagag	ttgggctttt	tatgcaactg	tgctactcct	8400
tatctctgat	atgaaagtgt	agaccaatc	acatcatgtc	atttagagtt	ggtaatactg	8460
ggaggataga	taaggcacga	aaacgagcca	tagcagacat	gctgggtgta	gccaagcaga	8520
agaaagtaga	tgggagccaa	ttgacgagcg	agggagctac	gccaatccga	catacgacac	8580
gctgagatcg	tcttggccgg	ggggtacctc	cagatgtcca	agggttaagt	cttgactgta	8640
attgtatgtc	tgaggacaaa	tatgtagtca	gccgtataaa	gtcataaccag	gcaccagtgc	8700
catcatcgaa	ccactaactc	tctatgatac	atgcctccgg	tattattgta	ccatgcgtcg	8760
ctttgttaca	tacgtatctt	gcctttttct	ctcagaaaact	ccagactttg	gctattgggtc	8820

gagataagcc	cggaccatag	tgagtctttc	acactctaca	tttctccctt	gctccaacta	8880
tttaaattcc	ttcacttcaa	gttcattctt	catctgcttc	tgttttactt	tgacaggcaa	8940
atgaagacat	ggtacgactt	gatggaggcc	aagaacgcca	tttcacccccg	agacaccgaa	9000
gtgcctgaaa	tcctggctgc	ccccattgat	aacatcgga	actacggtat	tccggaagt	9060
gtatatagaa	cctttcccca	gcttgtgtct	gtggatatgg	atgggtgta	cccccttgag	9120
tactcgtctt	ggcttctctc	cgagcagtat	gaggctctct	aatctagcgc	atttaatatc	9180
tcaatgtatt	tatatattta	tcttctcatg	cggccgctta	ctgagccttg	gcaccgggct	9240
gcttctcggc	cattcgagcg	aactgggaca	ggtatcgga	caggatgacg	agaccttcat	9300
ggggcagagg	gtttcggtag	gggaggttgt	gcttctggca	cagctgttcc	acctggtagg	9360
aaacggcagt	gaggttgtgt	cgaggcaggg	tgggccagag	atgggtgctcg	atctggtagt	9420
tcaggcctcc	aaagaaccag	tcagtaatga	tgcctcgtcg	aatgttcatg	gtctcatgga	9480
tctgaccac	agagaagcca	tgctcgctcc	agacggaatc	accgatcttc	tccagagggt	9540
agtggttcat	gaagaccacg	atggcaattc	cgaagccacc	gacgagctcg	gaaacaaaga	9600
acaccagcat	cgaggtcagg	atggagggca	taaagaagag	gtggaacagg	gtcttgagag	9660
tccagtgcag	agcgagtcca	atggcctctt	tcttgtactg	agatcggtag	aactggttgt	9720
ctcggtcctt	gagggatcga	acggtcagca	cagactggaa	acaccagatg	aatcgcagga	9780
gaatacagat	gaccaggaaa	tagtactgtt	ggaactgaat	gagctttcgg	gagatgggag	9840
aagctcgagt	gacatcgctc	tcggaccagg	cgagcagagg	caggttatca	atgtcgggat	9900
cgtgaccctg	aacgttggta	gcagaatgat	gggcgttgtg	tctgtccttc	caccagggtca	9960
cggagaagcc	ctggagtccg	ttgccaaaga	ccagaccag	gacgttattc	cagtttcggt	10020
tcttgaaggt	ctggtgggtg	cagatgtcat	gagacagcca	tcccatttgc	tggtagtgc	10080
taccgagcac	gagagcacca	atgaagtaca	ggtgggtactg	gaccagcatg	aagaaggcaa	10140
gcacgccaa	accagggtg	gtcaagatct	tgtacgagta	ccagagggga	gaggcgtcaa	10200
acatgccagt	ggcgatcagc	tcttctcgga	gctttcgga	atcctcctga	gcttcgttga	10260
cggcagcctg	gggaggcagc	tcggaagcct	ggttgatctt	gggcattcgc	ttgagcttgt	10320
cgaaggcttc	ctgagagtgc	ataacatga	aggcgtcagt	agcatctcgt	ccctggtagt	10380
tctcaatgat	ttcagctcca	ccagggtgga	agttcaccca	agcggagacg	tcgtacacct	10440
ttccgtcgat	gacgaggggc	agagcctgtc	gagaagcctt	caccatgggc	aggacctgtg	10500
ttagtacatt	gtcggggagt	catcaattgg	ttcgacaggt	tgtcgactgt	tagtatgagc	10560
tcaattgggc	tctggtgggt	cgatgacact	tgtcatctgt	ttctgttggg	tcatgtttcc	10620
atcaccttct	atggtactca	caattcgtcc	gattcgcccc	aatccgttaa	taccgacttt	10680
gatggccatg	ttgatgtgtg	tttaattcaa	gaatgaatat	agagaagaga	agaagaaaaa	10740
agattcaatt	gagccggcga	tgcagaccct	tatataaatg	ttgccttggga	cagacggagc	10800
aagcccgcgc	aaacctacgt	tcggtataat	atgttaagct	ttttaacaca	aaggtttggc	10860
ttggggtaac	ctgatgtggt	gcaaaagacc	gggcgttggc	gagccattgc	gcgggcgaat	10920
ggggccgtga	ctcgtctcaa	attcgagggc	gtgcctcaat	tcgtgcccc	gtggcttttt	10980
cccgcgttt	ccgccccgtt	tgcaccactg	cagccgcttc	tttggttcgg	acaccttget	11040
gcgagctagg	tgccttgtgc	tacttaaaaa	gtggcctccc	aacaccaaca	tgacatgagt	11100
gcgtgggcca	agacacgttg	gcggggctgc	agtcggctca	atggccccga	aaaaacgctg	11160

ctggagctgg	ttcggacgca	gtccgccgcg	gcgtatggat	atccgcaagg	ttccatagcg	11220
ccattgccct	ccgtcgcggt	ctatcccgcg	acctctaaat	agagcgggaa	tataacccaa	11280
gcttcttttt	tttcttttaa	cacgcacacc	cccaactatc	atgttgctgc	tgctgtttga	11340
ctctactctg	tggaggggtg	ctcccaccca	acccaacctt	cagggtggatc	cggcgctgtg	11400
attggctgat	aagtctccta	tccggactaa	ttctgaccaa	tgggacatgc	gcgcaggacc	11460
caaatgccgc	aattacgtaa	ccccaacgaa	atgcctaccc	ctctttggag	cccagcggcc	11520
ccaaatcccc	ccaagcagcc	cggttctacc	ggcttccatc	tccaagcaca	agcagcccgg	11580
aattccttta	cctgcaggat	aacttcgtat	aatgtatgct	atacgaagtt	atgatctctc	11640
tcttgagctt	ttccataaca	agttcttctg	cctccaggaa	gtccatgggt	ggtttgatca	11700
tggttttggg	gtagtggtag	tgcagtgggtg	gtattgtgac	tggggatgta	gttgagaata	11760
agtcatacac	aagtcagctt	tcttcgagcc	tcatataagt	ataagtagtt	caacgtatta	11820
gcactgtacc	cagcatctcc	gtatcgagaa	acacaacaac	atgccccatt	ggacagatca	11880
tgcggataca	caggttgtgc	agtatcatac	atactcgatc	agacaggctcg	tctgaccatc	11940
atacaagctg	aacaagcgct	ccatacttgc	acgtctctta	tatacacagt	taaattacat	12000
atccatagtc	taacctctaa	cagttaatct	tctggtaagc	ctcccagcca	gccttctggg	12060
atcgcttggc	ctcctcaata	ggatctcggt	tctggccgta	cagacctcgg	ccgacaatta	12120
tgatatccgt	tccggtagac	atgacatcct	caacagttcg	gtactgctgt	ccgagagcgt	12180
ctcccttgtc	gtcaagaccc	accccggggg	tcagaataag	ccagtcctca	gagtcgccct	12240
taggtcgggt	ctgggcaatg	aagccaacca	caaactcggg	gtcggatcgg	gcaagctcaa	12300
tggctctgct	ggagtactcg	ccagtggcca	gagagccctt	gcaagacagc	tcggccagca	12360
tgagcagacc	tctggccagc	ttctcgttgg	gagagggggac	taggaactcc	ttgtactggg	12420
agttctcgta	gtcagagacg	tctccttctt	tctgttcaga	gacagtttcc	tcggcaccag	12480
ctcgcaggcc	agcaatgatt	ccggttccgg	gtacaccgtg	ggcgttgggtg	atatcggacc	12540
actcggcgat	tcggtgacac	cggtagtggg	gcttgacagt	gttgccaata	tctgcgaact	12600
ttctgtcctc	gaacaggaag	aaaccgtgct	taagagcaag	ttccttgagg	gggagcacag	12660
tgccggcgta	ggtgaagtcg	tcaatgatgt	cgatatgggt	tttgatcatg	cacacataag	12720
gtccgacctt	atcggaagc	tcaatgagct	ccttgggtgg	ggtaacatcc	agagaagcac	12780
acaggttggg	tttcttggct	gccacgagct	tgagcactcg	agcggcaaaag	gcggaactgt	12840
ggacgttagc	tcgagcttcg	taggagggca	ttttgggtgg	gaagaggaga	ctgaaaataaa	12900
tttagtctgc	agaacttttt	atcggaacct	tatctggggc	agtgaagtat	atgttatggg	12960
aatagttacg	agttagttag	acttatagat	agactggact	atacggctat	cgggtccaaat	13020
tagaaagaac	gtcaatggct	ctctgggcgt	cgcctttgcc	gacaaaaatg	tgatcatgat	13080
gaaagccagc	aatgacgttg	cagctgatat	tgttgtcggc	caaccgcgcc	gaaaacgcag	13140
ctgtcagacc	cacagcctcc	aacgaagaat	gtatcgtcaa	agtgatccaa	gcacactcat	13200
agttggagtc	gtactccaaa	ggcggcaatg	acgagtcaga	cagatactcg	tcgacgcgat	13260
aacttcgtat	aatgtatgct	atacgaagtt	atcgtacgat	agttagtaga	caacaatcga	13320
taacgtctcg	taccaaccac	agattacgac	ccattcgcag	tcacagttca	ctagggtttg	13380
ggttgcatcc	gttgagagcg	gtttgttttt	aaccttctcc	atgtgctcac	tcagggtttg	13440
ggttcagatc	aaatcaaggc	gtgaaccact	ttgtttgagg	acaaatgtga	cacaaccaac	13500

cagtgtcagg	ggcaagtc	tgacaaagg	gaagatacaa	tgcaattact	gacagttaca	13560
gactgcctcg	atgcccta	cttgcccca	aataagacaa	ctgtcctcgt	ttaagcgcaa	13620
ccctattcag	cgtcacgt	taatagcg	tggatagcac	tagtctatga	ggagcgtttt	13680
atgttgccgt	gagggcgatt	ggtgctcata	tgggttcaat	tgagggtggcg	gaacgagc	13740
agtcttcaat	tgagggtg	gcgacaca	tgggtgtcac	gtggccta	tgacctcggg	13800
tcgtggagtc	cccagttata	cagcaacc	gagggtgc	ggtaggag	gtcaccag	13860
aatagggttt	tttttggact	ggagagggt	gggcaaaag	gctcaacggg	ctgtttgggg	13920
agctgtgggg	gaggaattg	cgatattt	gagggtta	gctccgatt	gcgtgtttt	13980
tcgtcctgc	atctcccat	accatata	tccctccca	cctctttcca	cgataatttt	14040
acggatcagc	aataagg	cttctcct	tttccacgt	catatatat	tatgctgc	14100
cgctcttttc	gtgacatc	caaacacat	acaacaat	ctgttactga	cgctcctta	14160
cgaagtc	gtgtcatc	cggcgacg	gtccgagcc	tgagtatcca	cgacaagat	14220
agtgtcgaga	cgacgcgt	tgtgtaat	cacaatcc	aagtcgctag	caacacac	14280
tctctacaca	aactaacca	gctctccat	gcctccac	cggtctctg	caagcaga	14340
cctgccctcc	gacgaacc	cacttccac	actgtgacc	actcggagt	tgctgccgt	14400
tctccctccg	attctccc	acactcg	tcctctaca	cgctgtctt	catgtccg	14460
gtggacattg	ccaagccca	gtccgagta	ggtgtcat	tggataccta	cggcaacc	14520
ttcgaagt	ccgacttc	catcaagg	atctacaac	ctattccca	gcactgctt	14580
aagcgatctg	ctctcaagg	atacggtac	attcttcg	acattgtct	cctgactacc	14640
actttcagca	tcttggtaca	ctttgtgac	cccaggtaca	ttccctcc	tcctgtcga	14700
gccggtctgt	gggtgtgt	caccgttct	cagggact	tcgggtact	actgtgggt	14760
attgccacg	agtgtggaca	tgggtctt	tcgattccc	gaatcatcaa	cgacattact	14820
ggctgggtgc	ttactctt	cctgcttgt	ccctactt	gctggcaaat	ctcccacc	14880
aagcatcaca	aggccact	aaacatgg	cgagacat	ttctcg	tcgaaccc	14940
gagcagcaag	ctactcgact	cggcaagat	accacgaac	tcgccc	taccgagg	15000
actcctgctt	tcaccctg	catgcttgt	cttcagca	tggtcggt	gcccactat	15060
ctcattacca	acgttact	acacaact	catgagc	agcgagagg	tcgaggca	15120
ggaaagcaca	acggtctt	cggtggagt	aaccatttc	atccccgat	tcctctgt	15180
gagaacagcg	acgccaag	catcgtg	tcgacatt	gcattggt	tatggccacc	15240
gctctgtact	ttctcg	gaagtccga	ttctaca	tgccatct	gtacttcgt	15300
ccctacttgt	gggttaacca	ctggctcgt	gccattac	ttctgcag	cacagatc	15360
actcttcccc	actacacca	cgacgagt	aactttgt	gagggtccg	tgcaaccat	15420
gaccgagaga	tgggcttcat	tggacgtcat	ctgctccac	gcattatcga	gactcacgt	15480
ctgcatcact	acgtctctt	cattccctt	tacaatgc	acgaagctac	cgaggccat	15540
aaacctatca	tgggcaag	ctatcgag	gatgtccag	acggtcctc	aggattcatt	15600
cgagccatgt	accgatctg	acgaatgtg	cagtgggtt	aacctccgc	tgggtccgag	15660
ggagctggca	agggtgtc	gttctttcga	aaccgaa	atgtgggc	tcctcccgt	15720
gtcatcaagc	ccgttgccta	agc				15743

<210>56

<211>1434		
<212>DNA		
<213> 串珠镰刀菌		
<220>		
<221>CDS		
<222>(1).. (1434)		
<223> 合成 Δ -12 去饱和酶 (经密码子优化用于解脂耶氏酵母)		
<300>		
<302> 适于改变含油酵母中多不饱和脂肪酸含量的 Δ -12 去饱和酶		
<310>WO 2005/047485		
<311>2004-11-12		
<312>2005-05-26		
<313>(1).. (1434)		
<300>		
<302> 适于改变含油酵母中多不饱和脂肪酸含量的 Δ -12 去饱和酶		
<310>US 2005-0216975-A1		
<311>2004-11-10		
<312>2005-09-29		
<313>(1).. (1434)		
<400>56		
atg gcc tcc acc tcg gct ctg ccc aag cag aac cct gcc ctc cga cga	48	
Met Ala Ser Thr Ser Ala Leu Pro Lys Gln Asn Pro Ala Leu Arg Arg		
1 5 10 15		
acc gtc act tcc acc act gtg acc gac tcg gag tct gct gcc gtc tct	96	
Thr Val Thr Ser Thr Thr Val Thr Asp Ser Glu Ser Ala Ala Val Ser		
20 25 30		
ccc tcc gat tct ccc aga cac tcg gcc tcc tct aca tcg ctg tct tcc	144	
Pro Ser Asp Ser Pro Arg His Ser Ala Ser Ser Thr Ser Leu Ser Ser		
35 40 45		
atg tcc gag gtg gac att gcc aag ccc aag tcc gag tac ggt gtc atg	192	
Met Ser Glu Val Asp Ile Ala Lys Pro Lys Ser Glu Tyr Gly Val Met		
50 55 60		
ctg gat acc tac ggc aac cag ttc gaa gtt ccc gac ttc acc atc aag	240	
Leu Asp Thr Tyr Gly Asn Gln Phe Glu Val Pro Asp Phe Thr Ile Lys		
65 70 75 80		
gac atc tac aac gct att ccc aag cac tgc ttc aag cga tct gct ctc	288	
Asp Ile Tyr Asn Ala Ile Pro Lys His Cys Phe Lys Arg Ser Ala Leu		
85 90 95		
aag gga tac ggc tac att ctt cga gac att gtc ctc ctg act acc act	336	

Lys	Gly	Tyr	Gly	Tyr	Ile	Leu	Arg	Asp	Ile	Val	Leu	Leu	Thr	Thr	Thr	
100				105				110								
ttc	agc	atc	tgg	tac	aac	ttt	gtg	aca	ccc	gag	tac	att	ccc	tcc	act	384
Phe	Ser	Ile	Trp	Tyr	Asn	Phe	Val	Thr	Pro	Glu	Tyr	Ile	Pro	Ser	Thr	
115				120				125								
cct	gct	cga	gcc	ggt	ctg	tgg	gct	gtg	tac	acc	gtt	ctt	cag	gga	ctc	432
Pro	Ala	Arg	Ala	Gly	Leu	Trp	Ala	Val	Tyr	Thr	Val	Leu	Gln	Gly	Leu	
130				135				140								
ttc	ggt	act	gga	ctg	tgg	gtc	att	gcc	cac	gag	tgt	gga	cat	ggt	gct	480
Phe	Gly	Thr	Gly	Leu	Trp	Val	Ile	Ala	His	Glu	Cys	Gly	His	Gly	Ala	
145				150				155				160				
ttc	tcc	gat	tcc	cga	atc	atc	aac	gac	att	act	ggc	tgg	gtg	ctt	cac	528
Phe	Ser	Asp	Ser	Arg	Ile	Ile	Asn	Asp	Ile	Thr	Gly	Trp	Val	Leu	His	
165				170				175								
tct	tcc	ctg	ctt	gtt	ccc	tac	ttc	agc	tgg	caa	atc	tcc	cac	cgg	aag	576
Ser	Ser	Leu	Leu	Val	Pro	Tyr	Phe	Ser	Trp	Gln	Ile	Ser	His	Arg	Lys	
180				185				190								
cat	cac	aag	gcc	act	gga	aac	atg	gag	cga	gac	atg	gtc	ttc	gtt	cct	624
His	His	Lys	Ala	Thr	Gly	Asn	Met	Glu	Arg	Asp	Met	Val	Phe	Val	Pro	
195				200				205								
cga	acc	cga	gag	cag	caa	gct	act	cga	ctc	ggc	aag	atg	acc	cac	gaa	672
Arg	Thr	Arg	Glu	Gln	Gln	Ala	Thr	Arg	Leu	Gly	Lys	Met	Thr	His	Glu	
210				215				220								
ctc	gcc	cat	ctt	acc	gag	gaa	act	cct	gct	ttc	acc	ctg	ctc	atg	ctt	720
Leu	Ala	His	Leu	Thr	Glu	Glu	Thr	Pro	Ala	Phe	Thr	Leu	Leu	Met	Leu	
225				230				235				240				
gtg	ctt	cag	caa	ctg	gtc	ggt	tgg	ccc	aac	tat	ctc	att	acc	aac	gtt	768
Val	Leu	Gln	Gln	Leu	Val	Gly	Trp	Pro	Asn	Tyr	Leu	Ile	Thr	Asn	Val	
245				250				255								
act	gga	cac	aac	tac	cat	gag	cgg	cag	cga	gag	ggt	cga	ggc	aag	gga	816
Thr	Gly	His	Asn	Tyr	His	Glu	Arg	Gln	Arg	Glu	Gly	Arg	Gly	Lys	Gly	
260				265				270								
aag	cac	aac	ggt	ctt	ggc	ggt	gga	gtt	aac	cat	ttc	gat	ccc	cga	tct	864
Lys	His	Asn	Gly	Leu	Gly	Gly	Gly	Val	Asn	His	Phe	Asp	Pro	Arg	Ser	
275				280				285								
cct	ctg	tac	gag	aac	agc	gac	gcc	aag	ctc	atc	gtg	ctc	tcc	gac	att	912
Pro	Leu	Tyr	Glu	Asn	Ser	Asp	Ala	Lys	Leu	Ile	Val	Leu	Ser	Asp	Ile	
290				295				300								
ggc	att	ggt	ctt	atg	gcc	acc	gct	ctg	tac	ttt	ctc	gtt	cag	aag	ttc	960

Gly	Ile	Gly	Leu	Met	Ala	Thr	Ala	Leu	Tyr	Phe	Leu	Val	Gln	Lys	Phe	
305					310					315					320	
gga	ttc	tac	aac	atg	gcc	atc	tgg	tac	ttc	gtt	ccc	tac	ttg	tgg	gtt	1008
Gly	Phe	Tyr	Asn	Met	Ala	Ile	Trp	Tyr	Phe	Val	Pro	Tyr	Leu	Trp	Val	
				325					330					335		
aac	cac	tgg	ctc	gtc	gcc	att	acc	ttt	ctg	cag	cac	aca	gat	cct	act	1056
Asn	His	Trp	Leu	Val	Ala	Ile	Thr	Phe	Leu	Gln	His	Thr	Asp	Pro	Thr	
			340					345					350			
ctt	ccc	cac	tac	acc	aac	gac	gag	tgg	aac	ttt	gtg	cga	ggt	gcc	gct	1104
Leu	Pro	His	Tyr	Thr	Asn	Asp	Glu	Trp	Asn	Phe	Val	Arg	Gly	Ala	Ala	
			355				360					365				
gca	acc	atc	gac	cga	gag	atg	ggc	ttc	att	gga	cgt	cat	ctg	ctc	cac	1152
Ala	Thr	Ile	Asp	Arg	Glu	Met	Gly	Phe	Ile	Gly	Arg	His	Leu	Leu	His	
	370					375				380						
ggc	att	atc	gag	act	cac	gtc	ctg	cat	cac	tac	gtc	tct	tcc	att	ccc	1200
Gly	Ile	Ile	Glu	Thr	His	Val	Leu	His	His	Tyr	Val	Ser	Ser	Ile	Pro	
385				390						395					400	
ttc	tac	aat	gcg	gac	gaa	gct	acc	gag	gcc	atc	aaa	cct	atc	atg	ggc	1248
Phe	Tyr	Asu	Ala	Asp	Glu	Ala	Thr	Glu	Ala	Ile	Lys	Pro	Ile	Met	Gly	
			405						410					415		
aag	cac	tat	cga	gct	gat	gtc	cag	gac	ggg	cct	cga	gga	ttc	att	cga	1296
Lys	His	Tyr	Arg	Ala	Asp	Val	Gln	Asp	Gly	Pro	Arg	Gly	Phe	Ile	Arg	
			420					425					430			
gcc	atg	tac	cga	tct	gca	cga	atg	tgc	cag	tgg	gtt	gaa	ccc	tcc	gct	1344
Ala	Met	Tyr	Arg	Ser	Ala	Arg	Met	Cys	Gln	Trp	Val	Glu	Pro	Ser	Ala	
			435				440					445				
ggt	gcc	gag	gga	gct	ggc	aag	ggt	gtc	ctg	ttc	ttt	cga	aac	cga	aac	1392
Gly	Ala	Glu	Gly	Ala	Gly	Lys	Gly	Val	Leu	Phe	Phe	Arg	Asn	Arg	Asn	
	450					455				460						
aat	gtg	ggc	act	cct	ccc	gct	gtc	atc	aag	ccc	gtt	gcc	taa			1434
Asn	Val	Gly	Thr	Pro	Pro	Ala	Val	Ile	Lys	Pro	Val	Ala				
465				470						475						
<210>57																
<211>477																
<212>PRT																
<213> 串珠镰刀菌																
<400>57																
Met	Ala	Ser	Thr	Ser	Ala	Leu	Pro	Lys	Gln	Asn	Pro	Ala	Leu	Arg	Arg	
1				5					10					15		

Thr Val Thr Ser Thr Thr Val Thr Asp Ser Glu Ser Ala Ala Val Ser			
20	25	30	
Pro Ser Asp Ser Pro Arg His Ser Ala Ser Ser Thr Ser Leu Ser Ser			
35	40	45	
Met Ser Glu Val Asp Ile Ala Lys Pro Lys Ser Glu Tyr Gly Val Met			
50	55	60	
Leu Asp Thr Tyr Gly Asn Gln Phe Glu Val Pro Asp Phe Thr Ile Lys			
65	70	75	80
Asp Ile Tyr Asn Ala Ile Pro Lys His Cys Phe Lys Arg Ser Ala Leu			
85	90	95	
Lys Gly Tyr Gly Tyr Ile Leu Arg Asp Ile Val Leu Leu Thr Thr Thr			
100	105	110	
Phe Ser Ile Trp Tyr Asn Phe Val Thr Pro Glu Tyr Ile Pro Ser Thr			
115	120	125	
Pro Ala Arg Ala Gly Leu Trp Ala Val Tyr Thr Val Leu Gln Gly Leu			
130	135	140	
Phe Gly Thr Gly Leu Trp Val Ile Ala His Glu Cys Gly His Gly Ala			
145	150	155	160
Phe Ser Asp Ser Arg Ile Ile Asn Asp Ile Thr Gly Trp Val Leu His			
165	170	175	
Ser Ser Leu Leu Val Pro Tyr Phe Ser Trp Gln Ile Ser His Arg Lys			
180	185	190	
His His Lys Ala Thr Gly Asn Met Glu Arg Asp Met Val Phe Val Pro			
195	200	205	
Arg Thr Arg Glu Gln Gln Ala Thr Arg Leu Gly Lys Met Thr His Glu			
210	215	220	
Leu Ala His Leu Thr Glu Glu Thr Pro Ala Phe Thr Leu Leu Met Leu			
225	230	235	240
Val Leu Gln Gln Leu Val Gly Trp Pro Asn Tyr Leu Ile Thr Asn Val			
245	250	255	
Thr Gly His Asn Tyr His Glu Arg Gln Arg Glu Gly Arg Gly Lys Gly			
260	265	270	
Lys His Asn Gly Leu Gly Gly Gly Val Asn His Phe Asp Pro Arg Ser			
275	280	285	
Pro Leu Tyr Glu Asn Ser Asp Ala Lys Leu Ile Val Leu Ser Asp Ile			
290	295	300	
Gly Ile Gly Leu Met Ala Thr Ala Leu Tyr Phe Leu Val Gln Lys Phe			
305	310	315	320
Gly Phe Tyr Asn Met Ala Ile Trp Tyr Phe Val Pro Tyr Leu Trp Val			

325	330	335
Asn His Trp Leu Val Ala Ile Thr Phe Leu Gln His Thr Asp Pro Thr		
340	345	350
Leu Pro His Tyr Thr Asn Asp Glu Trp Asn Phe Val Arg Gly Ala Ala		
355	360	365
Ala Thr Ile Asp Arg Glu Met Gly Phe Ile Gly Arg His Leu Leu His		
370	375	380
Gly Ile Ile Glu Thr His Val Leu His His Tyr Val Ser Ser Ile Pro		
385	390	395
Phe Tyr Asn Ala Asp Glu Ala Thr Glu Ala Ile Lys Pro Ile Met Gly		
405	410	415
Lys His Tyr Arg Ala Asp Val Gln Asp Gly Pro Arg Gly Phe Ile Arg		
420	425	430
Ala Met Tyr Arg Ser Ala Arg Met Cys Gln Trp Val Glu Pro Ser Ala		
435	440	445
Gly Ala Glu Gly Ala Gly Lys Gly Val Leu Phe Phe Arg Asn Arg Asn		
450	455	460
Asn Val Gly Thr Pro Pro Ala Val Ile Lys Pro Val Ala		
465	470	475

<210>58

<211>6303

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pZKUE3S

<400>58

```

ggccgcaagt gtggatgggg aagtgagtgc ccggttctgt gtgcacaatt ggcaatccaa      60
gatggatgga ttcaacacag ggatatagcg agctacgtgg tgggtgcgagg atatagcaac      120
ggatatattat gtttgacact tgagaatgta cgatacaagc actgtccaag tacaatacta      180
aacatactgt acatactcat actcgtaccc gggcaacggt ttcacttgag tgcagtggct      240
agtgtctctta ctcgtacagt gtgcaatact gcgtatcata gtctttgatg tatatcgtat      300
tcattcatgt tagttgcgta cgaggaaact gtctctgaac agaagaagga ggacgtctct      360
gactacgaga actcccagta caaggagttc ctagtccccct ctcccaacga gaagctggcc      420
agaggtctgc tcatgctggc cgagctgtct tgcaagggct ctctggccac tggcgagtac      480
tccaagcaga ccattgagct tgcccgatcc gaccccgagt ttgtgggttg cttcattgcc      540
cagaaccgac ctaagggcga ctctgaggac tggtttattc tgacccccgg ggtgggtctt      600
gacgacaagg gagacgtct cggacagcag taccgaactg ttgaggatgt catgtctacc      660
ggaacggata tcataattgt cggccgaggt ctgtacggcc agaaccgaga tcctattgag      720
gaggccaagc gataccagaa ggctggctgg gaggcctacc agaagattaa ctgttagagg      780

```

ttagactatg gatatgtaat ttaactgtgt atatagagag cgtgcaagta tggagcgctt	840
gttcagcttg tatgatggtc agacgacctg tctgatcgag tatgtatgat actgcacaac	900
ctgtgtatcc gcatgatctg tccaatgggg catgttgttg tgtttctcga tacggagatg	960
ctgggtacag tgctaatacag ttgaactact tatacttata tgaggctcga agaaagctga	1020
cttgtgtatg acttaattaa tegagcttgg cgtaatcatg gtcatactg tttcctgtgt	1080
gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag	1140
cctgggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgtt	1200
tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag	1260
gcggtttgcg tattgggcgc tcttccgctt cctcgtcac tgactcgtg cgctcggtcg	1320
ttcggctgcg gcgagcggta tcagctcact caaaggcggg aatacggta tccacagaat	1380
caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta	1440
aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa	1500
atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac caggcgtttc	1560
cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc ggataacctgt	1620
ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt aggtatctca	1680
gttcgggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc gttcagcccc	1740
accgctgcgc cttatccggg aactatcgtc ttgagtccaa cccggtaaga cacgacttat	1800
cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta ggcggtgcta	1860
cagagtctct gaagtgggtg cctaactacg gctacactag aaggacagta tttggatatct	1920
gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga tccggcaaac	1980
aaaccaccgc tggtagcggg ggtttttttg ttgcaagca gcagattacg cgcagaaaaa	2040
aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa	2100
actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc tagatccttt	2160
taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact tggctctgaca	2220
gttaccaatg cttaatcagt gaggcacctt tctcagcgat ctgtctattt cgttcatcca	2280
tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggtta ccatctggcc	2340
ccagtgtgc aatgataccg cgagaccac gctaccggc tccagattta tcagcaataa	2400
accagccagc cggaagggcc gagegcagaa gtggtcctgc aactttatcc gcctccatcc	2460
agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat agtttgcgca	2520
acgttgttgc cattgctaca ggcacgtgg tgtcacgctc gtcgtttggg atggcttcat	2580
tcagctccgg ttcceaacga tcaaggcgag ttacatgac ccccatgttg tgcaaaaaag	2640
cggttagctc ctccggctc ccatcggttgc tcagaagtaa gttggccgca gtgttatcac	2700
tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta agatgctttt	2760
ctgtgactgg tgagtactca accaagtcac tctgagaata gtgtatgcgg cgaccgagtt	2820
gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact ttaaaagtgc	2880
tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg ctgttgagat	2940
ccagttcgat gtaaccact cgtgcacca actgatcttc agcatctttt actttcacca	3000
gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga ataaggcgca	3060
cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc atttatcagg	3120

gttattgtct	catgagcggg	tacatatttg	aatgtattta	gaaaaataaa	caaatagggg	3180
ttccgcgcac	atttccccga	aaagtgccac	ctgacgcgcc	ctgtagcggc	gcattaagcg	3240
cggcgggtgt	ggtgggttacg	cgcagcgtga	ccgtacact	tgccagcgcc	ctagcgccccg	3300
ctcctttcgc	tttcttccct	tcctttctcg	ccacgttcgc	cggctttccc	cgtcaagctc	3360
taaatcgggg	gctcccttta	gggttccgat	ttagtgcttt	acggcacctc	gacccccaaa	3420
aacttgatta	gggtgatggg	tcacgtagt	ggccatcgcc	ctgatagacg	gtttttcgcc	3480
ctttgacgtt	ggagtccacg	ttctttaata	gtggactctt	gttccaaact	ggaacaacac	3540
tcaaccctat	ctcgggtctat	tcttttgatt	tataagggat	tttgccgatt	tcggcctatt	3600
ggttaaaaaa	tgagctgatt	taacaaaaat	ttaacgcgaa	ttttaacaaa	atattaacgc	3660
ttacaatttc	cattcgccat	tcaggctgcg	caactgttgg	gaagggcgat	cgggtgcgggc	3720
ctcttcgcta	ttacgccagc	tggcgaaagg	gggatgtgct	gcaaggcgat	taagttgggt	3780
aacgccaggg	ttttcccagt	cacgacgttg	taaaacgacg	gccagtgaat	tgtaatacga	3840
ctcactatag	ggcgaattgg	gtaccggggc	ccacctcgag	gtcgacgagt	atctgtctga	3900
ctcgtcattg	catgcctttg	gagtacgact	ccaactatga	gtgtgcttgg	atcactttga	3960
cgatacattc	ttcgttggag	gctgtgggtc	tgacagctgc	gttttcggcg	cggttggccg	4020
acaacaatat	cagctgcaac	gtcattgctg	gctttcatca	tgatcacatt	tttgtcggca	4080
aaggcgacgc	ccagagagcc	attgacgttc	tttctaattt	ggaccgatag	ccgtatagtc	4140
cagtctatct	ataagttcaa	ctaactcgta	actattacca	taacatatac	ttcactgccc	4200
cagataaggt	tccgataaaa	agttctgcag	actaaattta	tttcagtctc	ctcttcacca	4260
ccaaaatgcc	ctcctacgaa	gctcgagtgc	tcaagctcgt	ggcagccaag	aaaaccaacc	4320
tgtgtgcttc	tctggatggt	accaccacca	aggagctcat	tgagcttgcc	gataaggtcg	4380
gaccttatgt	gtgcatgata	aaaacccata	tcgacatcat	tgacgacttc	acctacgccg	4440
gcaactgtgt	ccccctcaag	gaacttgctc	ttaagcacgg	tttcttcctg	ttcgaggaca	4500
gaaagttcgc	agatattggc	aacactgtca	agcaccagta	ccggtgtcac	cgaatcgccg	4560
agtggtcgga	tatcaccaac	gcccacgggtg	tttaaacccg	gaaccggaat	cgataagctt	4620
gatatacgaat	tcatgctgtt	catcggtggt	aatgctgctg	tgtgctgtgt	gtgtgtgttg	4680
tttggcgctc	attgtttgcgt	tatgcagcgt	acaccacaat	attggaagct	tattagcctt	4740
tctatttttt	cgttttgcaag	gcttaacaac	attgctgtgg	agagggatgg	ggatatggag	4800
gccgctggag	ggagtccggag	aggcgttttg	gagcggcttg	gcctggcgcc	cagctcgcga	4860
aacgcaccta	ggaccctttg	gcacgccgaa	atgtgccact	tttcagtcta	gtaacgcctt	4920
acctacgtca	ttccatgcgt	gcatgtttgc	gccttttttc	ccttgccctt	gatcgccaca	4980
cagtacagtg	cactgtacag	tggaggtttt	gggggggtct	tagatgggag	ctaaaagcgg	5040
cctagcggta	cactagtggg	attgtatgga	gtggcatgga	gcctaggtgg	agcctgacag	5100
gacgcacgac	cggctagccc	gtgacagacg	atgggtggct	cctgttgtcc	accgcgtaca	5160
aatgtttggg	ccaaagtctt	gtcagccttg	cttgcgaaac	taattcccaa	ttttgtcact	5220
tcgcaccccc	attgategag	ccctaacccc	tgcccatcag	gcaatccaat	taagctcgca	5280
ttgtctgcct	tgttttagttt	ggctcctgcc	cgtttcggcg	tccacttgca	caaacacaaa	5340
caagcattat	atataaggct	cgtctctccc	tcccaaccac	actcactttt	ttgcccgtct	5400
tcccttgcta	acacaaaagt	caagaacaca	aacaaccacc	ccaacccctt	tacacacaag	5460

acatatctac accatggagt ctggacceat gcctgctggc attcccttcc ctgagtacta 5520
 tgacttcttt atggactgga agactccctt ggccatcgct gccacctaca ctgctgccgt 5580
 cggctctcttc aaccccaagg ttggcaaggt ctcccgagtg gttgccaagt cggctaacgc 5640
 aaagcctgcc gagegaaccc agtccggage tgccatgact gccttcgtct ttgtgcacaa 5700
 cctcattctg tgtgtctact ctggcateac cttctactac atgtttcctg ctatgggtcaa 5760
 gaacttccga acccacacac tgcacgaage ctactgcgac acggatcagt ccctctggaa 5820
 caacgcactt ggctactggg gttacctctt ctacctgtcc aagttctacg aggtcattga 5880
 caccatcate atcatectga agggacgacg gtcctcgctg cttcagacct accaccatgc 5940
 tggagccatg attaccatgt ggtctggcat caactaccaa gccactccca tttggatctt 6000
 tgttggtcttc aactccttca ttcacacat catgtactgt tactatgcct tcacctctat 6060
 cggattccat cctcctggca aaaagtacct gacttcgatg cagattactc agtttctggt 6120
 cggtatcacc attgccgtgt cctacctctt cgttcctggc tgcacccgaa cacccggtgc 6180
 tcagatggct gtctggatca acgtcggcta cctgtttccc ttgacctatc tgttcgtgga 6240
 ctttgccaag cgaacctact ccaagegate tgccattgcc gctcagaaaa aggctcagta 6300
 agc 6303

<210>59

<211>21

<212>DNA

<213> 人工序列

<220>

<223> 引物 pZP-GW-5-1

<400>59

cgacaagatg gaatgagaat g

21

<210>60

<211>22

<212>DNA

<213> 人工序列

<220>

<223> 引物 pZP-GW-5-2

<400>60

ctggtttttc aactacttct ac

22

<210>61

<211>21

<212>DNA

<213> 人工序列

<220>

<223> 引物 pZP-GW-5-3

<400>61

gtactgtcct gtgtctgttc c

21

<210>62	
<211>22	
<212>DNA	
<213> 人工序列	
<220>	
<223> 引物 pZP-GW-5-4	
<400>62	
ctacatcgtc cgaaagcaca ag	22
<210>63	
<211>24	
<212>DNA	
<213> 人工序列	
<220>	
<223> 引物 pZP-GW-3-1	
<400>63	
ctaccagatc gagcaccatc tctg	24
<210>64	
<211>21	
<212>DNA	
<213> 人工序列	
<220>	
<223> 引物 pZP-GW-3-2	
<400>64	
ctaccaggtg gaacagctgt g	21
<210>65	
<211>22	
<212>DNA	
<213> 人工序列	
<220>	
<223> 引物 pZP-GW-3-3	
<400>65	
tctgccccat gaaggtctcg tc	22
<210>66	
<211>22	
<212>DNA	
<213> 人工序列	
<220>	
<223> 引物 pZP-GW-3-4	
<400>66	

cctgtcccag ttcgctcgaa tg	22
<210>67	
<211>44	
<212>DNA	
<213> 人工序列	
<220>	
<223>Genome Walker 衔接子 -1	
<400>67	
gtaatacgac tatagggcac gcgtggtcga cggcccgggc tggt	44
<210>68	
<211>8	
<212>DNA	
<213> 人工序列	
<220>	
<223>Genome Walker 衔接子 -2	
<220>	
<221>misc_feature	
<222>(1).. (1)	
<223>5' 末端与 -P04 基附连	
<220>	
<221>misc_feature	
<222>(8).. (8)	
<223>3' 末端与 -H2N 基附连	
<400>68	
accagccc	8
<210>69	
<211>22	
<212>DNA	
<213> 人工序列	
<220>	
<223> 巢式衔接子引物	
<400>69	
gtaatacgac tcactatagg gc	22
<210>70	
<211>36	
<212>DNA	
<213> 人工序列	
<220>	
<223> 引物 Per10F1	

<400>70	
gatcaacat ggggggaagt tcacatgcat tcgctg	36
<210>71	
<211>29	
<212>DNA	
<213> 人工序列	
<220>	
<223> 引物 ZPGW-5-5	
<400>71	
gttatagttt tcatgtgaaa taccgagag	29
<210>72	
<211>37	
<212>DNA	
<213> 人工序列	
<220>	
<223> 引物 Per10R	
<400>72	
gatcaagcgg ccgccagacc tcgtcattat ctgatag	37
<210>73	
<211>7222	
<212>DNA	
<213> 人工序列	
<220>	
<223> 质粒 pFBAIn-MOD-1	
<400>73	
catggatcca ggcctgttaa cggccattac ggcctgcagg atccgaaaaa acctcccaca	60
cctccccctg aacctgaaac ataaaatgaa tgcaattgtt gttgttaact tgtttattgc	120
agcttataat ggttacaaat aaagcaatag catcacaaat ttcacaaata aagcattttt	180
ttcactgcat tctagttgtg gtttgtccaa actcatcaat gtatcttata atgtctgcgg	240
ccgcaagtgt ggatggggaa gtgagtgccg ggttctgtgt gcacaattgg caatccaaga	300
tggatggatt caacacaggg atatagcgag ctacgtggtg gtgcgaggat atagcaacgg	360
atatttatgt ttgacacttg agaatgtacg atacaagcac tgtccaagta caatactaaa	420
catactgtac atactcatac tcgtacccgg gcaacggttt cacttgagtg cagtggctag	480
tgctcttact cgtacagtgt gcaatactgc gtatcatagt ctttgatgta tatcgtattc	540
attcatgtta gttgcgtacg agccggaagc ataaagtgt aagcctgggg tgcctaataa	600
gtgagctaac tcacattaat tgcgttgcgc tctactgccc ctttccagtc gggaaacctg	660
tcgtgccage tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg	720
cgtctttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg	780
gtatcagctc actcaaagge ggtaatacgg ttatccacag aatcagggga taacgcagga	840

aagaacatgt	gagcaaaagg	ccagcaaaag	gccaggaacc	gtaaaaaggc	cgcgttgctg	900
gcgtttttcc	ataggtccg	ccccctgac	gagcatcaca	aaaatcgacg	ctcaagtcag	960
aggtggcgaa	acccgacagg	actataaaga	taccaggcgt	ttccccctgg	aagctccctc	1020
gtgcgctctc	ctgttccgac	cctgcccgtt	accggatacc	tgtccgcctt	tctcccttcg	1080
ggaagcgtgg	cgttttctca	tagctcacgc	tgtaggtatc	tcagttcggg	gtaggtcgtt	1140
cgctccaage	tgggctgtgt	gcacgaacce	cccgttcagc	ccgaccgctg	cgctttatcc	1200
ggtaactatc	gtcttgagtc	caaccgggta	agacacgact	tatcgccact	ggcagcagcc	1260
actggtaaca	ggattagcag	agcgagggtat	gtaggcgggtg	ctacagagtt	cttgaagtgg	1320
tggcctaact	acggctacac	tagaaggaca	gtatttggta	tctgcgctct	gctgaagcca	1380
gttaccttcg	gaaaaagagt	tggtagctct	tgatccggca	aacaaaccac	cgctggtagc	1440
ggtgggtttt	ttgtttgcaa	gcagcagatt	acgcgcagaa	aaaaaggatc	tcaagaagat	1500
cctttgatct	tttctacggg	gtctgacgct	cagtggaaacg	aaaactcacg	ttaagggtat	1560
ttggatcatga	gattatcaaa	aaggatcttc	acctagatcc	ttttaaatta	aaaatgaagt	1620
tttaaatcaa	tctaaagtat	atatgagtaa	acttggctctg	acagttacca	atgcttaatc	1680
agtgaggcac	ctatctcagc	gatctgtcta	tttcgttcat	ccatagttgc	ctgactcccc	1740
gtcgtgtaga	taactacgat	acgggaggggc	ttaccatctg	gccccagtcg	tgcaatgata	1800
ccgcgagacc	cacgctcacc	ggctccagat	ttatcagcaa	taaaccagcc	agccggaagg	1860
gccgagcgca	gaagtgggtcc	tgcaacttta	tccgcctcca	tccagtctat	taattgttgc	1920
cgggaagcta	gagtaagtag	ttcgccagtt	aatagtttgc	gcaacgttgt	tgccattgct	1980
acaggcatcg	tgggtgtcag	ctcgctggtt	ggtaggtcct	cattcagctc	cggttcccaa	2040
cgatcaaggc	gagttacatg	atcccccatg	ttgtgcaaaa	aagcggttag	ctccttcggg	2100
cctccgatcg	ttgtcagaag	taagttggcc	gcagtggtat	cactcatggt	tatggcagca	2160
ctgcataatt	ctcttactgt	catgccatcc	gtaagatgct	tttctgtgac	tgggtgagtac	2220
tcaaccaagt	cattctgaga	atagtgtatg	cggcgaccga	gttgcctctg	cccggcgctca	2280
atacgggata	ataccgcgcc	acatagcaga	actttaaaaag	tgctcatcat	tggaaaacgt	2340
tcttcggggc	gaaaactctc	aaggatctta	ccgctgttga	gatccagttc	gatgtaaccc	2400
actcgtgcac	ccaactgac	ttcagcatct	tttactttca	ccagcgtttc	tgggtgagca	2460
aaaacaggaa	ggcaaaatgc	cgcaaaaaag	ggaataaggg	cgacacggaa	atgttgaata	2520
ctcatactct	tcctttttca	atattattga	agcatttata	agggttattg	tctcatgagc	2580
ggatacatat	ttgaatgtat	ttagaaaaat	aaacaaatag	gggttccgcg	cacatttccc	2640
cgaaaagtgc	cacctgacgc	gccctgtagc	ggcgcatata	gcgcggcggg	tgtgggtggt	2700
acgcgcagcg	tgaccgctac	acttgccagc	gccctagcgc	ccgctccttt	cgttttcttc	2760
ccttcctttc	tcgccacgtt	cgccggcctt	ccccgtcaag	ctctaaatcg	ggggctccct	2820
ttagggttcc	gatttagtgc	tttacggcac	ctcgacccca	aaaaacttga	ttagggtgat	2880
ggttcacgta	gtgggccaac	gccctgatag	acggtttttc	gccctttgac	gttggagtcc	2940
acgttcttta	atagtggact	cttgttccaa	actggaacaa	cactcaaccc	tatctcggtc	3000
tattcttttg	atttataagg	gattttgccg	atttcggcct	attgggttaaa	aaatgagctg	3060
atttaacaaa	aatttaacgc	gaattttaac	aaaatattaa	cgcttacaat	ttccattcgc	3120
cattcaggct	gcgcaactgt	tgggaagggc	gatcggtgcg	ggcctcttcg	ctattacgcc	3180

agctggcgaa	agggggatgt	gctgcaaggc	gattaagttg	ggtaacgcca	gggttttccc	3240
agtcacgacg	ttgtaaaacg	acggccagtg	aattgtaata	cgactcacta	tagggcgaaat	3300
tgggtaccgg	gccccccctc	gaggtegatg	gtgtcgataa	gcttgataatc	gaattcatgt	3360
cacacaaacc	gatcttegcc	tcaaggaaac	ctaattctac	atccgagaga	ctgccgagat	3420
ccagtctaca	ctgattaatt	ttegggceaa	taatttaaaa	aaatcgtggt	atataatatt	3480
atatgtatta	tatatataca	tcatgatgat	actgacagtc	atgtcccat	gctaaataga	3540
cagactccat	ctgccgcctc	caactgatgt	tctcaatatt	taaggggtca	tctcgcattg	3600
tttaataata	aacagactcc	atctaccgcc	tccaaatgat	gttctcaaaa	tatattgtat	3660
gaacttattt	ttattactta	gtattattag	acaacttact	tgttttatga	aaaacacttc	3720
ctatttagga	aacaatttat	aatggcagtt	cgttcattta	acaatttatg	tagaataaat	3780
gttataaatg	cgtatgggaa	atcttaaata	tggatagcat	aatgatatac	tgcattgcct	3840
aattcgaaat	caacagcaac	gaaaaaaatc	ccttgtacaa	cataaatagt	catcgagaaa	3900
tatcaactat	caaagaacag	ctattcacac	gttactattg	agattattat	tggacgagaa	3960
tcacacactc	aactgtcttt	ctctcttcta	gaaatacagg	tacaagtatg	tactattctc	4020
attgttcata	cttctagtca	tttcatccca	catattcctt	ggattttctct	ccaatgaatg	4080
acattctatc	ttgcaaattc	aacaattata	ataagatata	ccaaagtagc	ggtatagtgg	4140
caatcaaaaa	gcttctctgg	tgtgtctctc	gtatttattt	ttattctaata	gatccattaa	4200
aggtatata	ttatttcttg	ttatataatc	cttttgttta	ttacatgggc	tggatacata	4260
aaggtatttt	gatttaattt	tttgcctaaa	ttcaatcccc	cctcgttcag	tgtcaactgt	4320
aatggtagga	aattaccata	cttttgaaga	agcaaaaaaa	atgaaagaaa	aaaaaaatcg	4380
tatttccagg	ttagacgttc	cgcagaatct	agaatgcggt	atgcggtaca	ttgttcttcg	4440
aacgtaaaag	ttgcgcctcc	tgagatatig	tacatttttg	cttttacaag	tacaagtaca	4500
tcgtacaact	atgtactact	gttgatgcat	ccacaacagt	ttgttttggt	tttttttggt	4560
tttttttttt	ctaattgattc	attaccgcta	tgtataccta	cttgtacttg	tagtaagccg	4620
ggttattggc	gttcaattaa	tcatagactt	atgaatctgc	acgggtgtgcg	ctgcgagtta	4680
cttttagctt	atgcatgcta	cttgggtgta	atattgggat	ctgttcggaa	atcaacggat	4740
gctcaatcga	tttcgacagt	aatttaattaa	gtcatacaca	agtcagcttt	cttcgagcct	4800
catataagta	taagtagttc	aacgtattag	cactgtaccc	agcatctccg	tatcgagaaa	4860
cacaacaaca	tgccccattg	gacagatcat	gcggatacac	aggttgtgca	gtatcataca	4920
tactcgatca	gacaggctgt	ctgaccatca	tacaagctga	acaagcgctc	catacttgca	4980
cgctctctat	atacacagtt	aaattacata	tccatagtct	aacctctaac	agttaatctt	5040
ctggtaagcc	tcccagccag	ccttctggta	tcgcttggcc	tcctcaatag	gatctcggtt	5100
ctggccgtac	agacctcggc	cgacaattat	gatatccggt	ccggtagaca	tgacatcctc	5160
aacagttcgg	tactgctgtc	cgagagcgctc	tcccttgctg	tcaagaccca	ccccgggggt	5220
cagaataagc	cagtcctcag	agtcgccctt	aggctcggtt	tgggcaatga	agccaaccac	5280
aaactcgggg	tcggatcggg	caagctcaat	ggtctgcttg	gagtactcgc	cagtggccag	5340
agagcccttg	caagacagct	cggccagcat	gagcagacct	ctggccagct	tctcgttggg	5400
agaggggact	aggaactcct	tgtactggga	gttctcgtag	tcagagacgt	cctccttctt	5460
ctgttcagag	acagtttctt	cggcaccagc	tcgcaggcca	gcaatgattc	cggttccggg	5520

tacaccgtgg	gcgttgggtga	tateggacca	ctcggcgatt	cgggtgacacc	ggtactgggtg	5580
cttgacagt	ttgccaatat	ctgcgaactt	tctgtcctcg	aacaggaaga	aaccgtgctt	5640
aagagcaagt	tccttgaggg	ggagcacagt	gccggcgtag	gtgaagtcgt	caatgatgtc	5700
gatatgggtt	ttgatcatgc	acacataagg	tccgacctta	tccgcaagct	caatgagctc	5760
cttgggtggtg	gtaacatcca	gagaagcaca	caggttgggtt	ttcttggctg	ccacgagctt	5820
gagcactcga	gcggcaaagg	cggacttgtg	gacgttagct	cgagcttcgt	aggaggggcat	5880
tttgggtggtg	aagaggagac	tgaaataaat	ttagtctgca	gaacttttta	tccgaacctt	5940
atctggggca	gtgaagtata	tgttatggta	atagttacga	gtagttgaa	cttatagata	6000
gactggacta	tacggctatc	ggtccaaatt	agaaagaacg	tcaatggctc	tctgggcgtc	6060
gcctttgccg	acaaaaatgt	gatcatgatg	aaagccagca	atgacgttgc	agctgatatt	6120
gttgtcggcc	aaccgcgccg	aaaacgcagc	tgtcagaccc	acagcctcca	acgaagaatg	6180
tatcgtcaaa	gtgatccaag	cacactcata	gttggagtcg	tactccaaag	gcggcaatga	6240
cgagtcagac	agatactcgt	cgaaaacagt	gtacgcagat	ctactataga	ggaacattta	6300
aattgccccg	gagaagacgg	ccaggccgcc	tagatgacaa	attcaacaac	tcacagctga	6360
ctttctgcca	ttgccactag	gggggggcct	ttttatatgg	ccaagccaag	ctctccacgt	6420
cggttgggct	gcaccaaca	ataaatgggt	agggttgcac	caacaaagg	atgggatggg	6480
gggtagaaga	tacgaggata	acggggctca	atggcacaaa	taagaacgaa	tactgccatt	6540
aagactcgtg	atccagegac	tgacaccatt	gcatcatcta	agggcctcaa	aactacctcg	6600
gaactgctgc	gctgatctgg	acaccacaga	ggttccgagc	actttagggt	gcaccaaatg	6660
ttccaccagg	tgcaggcaga	aaacgctgga	acagcgtgta	cagtttgtct	taacaaaaag	6720
tgagggcgct	gaggtcgagc	agggtgggtg	gacttggtat	agcctttaga	gctgcgaaag	6780
cgcgtatgga	tttggctcat	caggccagat	tgagggtctg	tggacacatg	tcatgttagt	6840
gtacttcaat	cgccccctgg	atatagcccc	gacaataggc	cgtggcctca	tttttttgcc	6900
ttccgcacat	ttccattgct	cggtagccac	accttgcttc	tcctgcactt	gccaacctta	6960
atactggttt	acattgacca	acatcttaca	agcggggggc	ttgtctaggg	tatatataaa	7020
cagtggctct	cccaatcggt	tgccagtcct	ttttttccct	tccttcccca	cagattcgaa	7080
atctaaacta	cacatcacag	aattccgagc	cgtgagtatc	cacgacaaga	tcagtgtcga	7140
gacgacgcgt	tttgtgtaat	gacacaatcc	gaaagtcgct	agcaacacac	actctctaca	7200
caaactaacc	cagctctggt	ac				7222

<210>74

<211>8133

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pFBA IN-Pex10

<400>74

ggccgcaagt	gtggatgggg	aagtgagtgc	ccggttctgt	gtgcacaatt	ggcaatccaa	60
gatggatgga	ttcaacacag	ggatatagcg	agctacgtgg	tggtgagagg	atatagcaac	120
ggatatattat	gtttgacact	tgagaatgta	cgatacaagc	actgtccaag	tacaatacta	180

aacatactgt	acatactcat	actcgtaccc	gggcaacggt	ttcacttgag	tgcagtggct	240
agtgtcttta	ctcgtacagt	gtgcaatact	gcgtatcata	gtctttgatg	tatatcgtat	300
tcattcatgt	tagttgcgta	cgagccggaa	gcataaagtg	taaagcctgg	ggtgcctaata	360
gagttagcta	actcacatta	attgcgttgc	gctcactgcc	cgttttccag	tcgggaaacc	420
tgtcgtgcca	gctgcattaa	tgaatcggcc	aacgcgcggg	gagaggcggt	ttgcgtattg	480
ggcgctcttc	cgttctctcg	ctcactgact	cgtcgcgtc	ggtcgttcgg	ctgcggcgag	540
cggatcagc	tcactcaaag	gcggtaatac	ggttatccac	agaatcaggg	gataacgcag	600
gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcgttgc	660
tggcgttttt	ccataggctc	cgtccccctg	acgagcatca	caaaaatcga	cgtcaagtc	720
agaggtggcg	aaacccgaca	ggactataaa	gataccaggc	gtttccccct	ggaagctccc	780
tcgtgcgctc	tcctgttccg	accctgccgc	ttaccggata	cctgtccgcc	tttctccctt	840
cgggaagcgt	ggcgctttct	catagctcac	gctgtaggta	tctcagttcg	gtgtaggctg	900
ttcgctccaa	gctgggctgt	gtgcacgaac	ccccggttca	gcccgaaccgc	tgcgccttat	960
ccggtaaacta	tcgtcttgag	tccaacccgg	taagacacga	cttatcgcca	ctggcagcag	1020
ccactggtaa	caggattagc	agagcgaggt	atgtaggcgg	tgtacagag	ttcttgaagt	1080
ggtggcctaa	ctacggctac	actagaagga	cagtatttgg	tatctgcgct	ctgctgaagc	1140
cagttacctt	cggaaaaaga	gttggtagct	cttgatccgg	caaacaaacc	accgctggta	1200
gcggtgggtt	ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	1260
atcctttgat	cttttctacg	gggtctgacg	ctcagtggaa	cgaaaactca	cgttaaggga	1320
ttttggctcat	gagattatca	aaaaggatct	tcacctagat	ccttttaaat	taaaaatgaa	1380
gttttaaatc	aatctaaagt	atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	1440
tcagttaggc	acctatctca	gcgatctgtc	tatttcgttc	atccatagtt	gcctgactcc	1500
ccgtcgtgta	gataactacg	atacgggagg	gcttaccatc	tggccccagt	gctgcaatga	1560
taccgcgaga	cccacgtcga	ccggctccag	atttatcagc	aataaaccag	ccagccggaa	1620
gggccgagcg	cagaagtggg	cctgcaactt	tatccgcctc	catccagtct	attaattggt	1680
gccgggaagc	tagagtaagt	agttcgccag	ttaatagttt	gcgcaacggt	gttgccattg	1740
ctacaggcat	cgtgggtgtca	cgctcgtcgt	ttggtatggc	ttcattcagc	tccgggtccc	1800
aacgatcaag	gcgagttaca	tgatccccc	tgttgtgcaa	aaaagcggtt	agctccttcg	1860
gtcctccgat	cgttgtcaga	agtaagttgg	ccgcagtggt	atcactcatg	gttatggcag	1920
cactgcataa	ttctcttact	gtcatgccat	ccgtaagatg	cttttctgtg	actggtgagt	1980
actcaaccaa	gtcattctga	gaatagtgtg	tgcggcgacc	gagttgctct	tgcccggcgt	2040
caatacggga	taataccgcg	ccacatagca	gaactttaaa	agtgtcatc	atttgaaaaac	2100
gttcttcggg	gcgaaaactc	tcaaggatct	taccgctggt	gagatccagt	tcgatgtaac	2160
ccactcgtgc	acccaactga	tcttcagcat	cttttacttt	caccagcggt	tctgggtgag	2220
caaaaacagg	aaggcaaaat	gccgcaaaaa	agggaataag	ggcgacacgg	aaatgttgaa	2280
tactcatact	cttccctttt	caatattatt	gaagcattta	tcagggttat	tgtctcatga	2340
gcggatacat	atttgaatgt	atttagaaaa	ataaacaat	aggggttccg	cgcacatttc	2400
cccgaaggt	gccacctgac	gcgccctgta	gcggcgcat	aagcgcgcg	ggtgtgggtg	2460
ttacgcgcag	cgtgaccgct	acacttgcca	gcgccctagc	gcccgtctct	ttcgctttct	2520

tcccttcctt	tctcgccacg	ttegcgggt	ttccccgtca	agctctaaat	cgggggctcc	2580
ctttagggtt	ccgatttagt	gctttacggc	acctcgaccc	caaaaaactt	gattaggggtg	2640
atggttcacg	tagtgggcca	tcgcccgtgat	agacggtttt	tcgccccttg	acgttggagt	2700
ccacgttctt	taatagtgga	ctcttggttc	aaactggaac	aacactcaac	cctatctcgg	2760
tctattcttt	tgatttataa	gggattttgc	cgatttcggc	ctattgggta	aaaaatgagc	2820
tgattttaaca	aaaattttaac	gcgaatttta	acaaaatatt	aacgcttaca	atttccattc	2880
gccattcagg	ctgcgcaact	gttgggaagg	gcgatcggtg	cgggcctctt	cgctattacg	2940
ccagctggcg	aaagggggat	gtgctgcaag	gcgattaagt	tgggtaacgc	cagggttttc	3000
ccagtcacga	cgttgtaaaa	cgacggccag	tgaattgtaa	tacgactcac	tatagggcga	3060
attgggtacc	gggccccccc	tcgaggctga	tgggtgctgat	aagcttgata	tcgaattcat	3120
gtcacacaaa	ccgatcttcg	cctcaaggaa	acctaatctt	acatccgaga	gactgccgag	3180
atccagtcta	cactgattaa	ttttcgggcc	aataatttaa	aaaaatcgtg	ttatataata	3240
ttatatgtat	tatatatata	catcatgatg	atactgacag	tcatgtccca	ttgctaaata	3300
gacagactcc	atctgcgcgc	tccaactgat	gttctcaata	tttaaggggt	catctcgcgt	3360
tgtttaataa	taaacagact	ccatctaccg	cctccaaatg	atgttctcaa	aatatatgtt	3420
atgaacttat	ttttattact	tagtattatt	agacaactta	cttgctttat	gaaaaacact	3480
tcctatttag	gaaacaattt	ataatggcag	ttcgttcatt	taacaattta	tgtagaataa	3540
atgttataaa	tgcgtatggg	aaatcttaaa	tatggatagc	ataaatgata	tctgcattgc	3600
ctaattcgaa	atcaacagca	acgaaaaaaaa	tcccttgtag	aacataaata	gtcatcgaga	3660
aatatcaact	atcaaagaac	agctattcac	acgttactat	tgagattatt	attggacgag	3720
aatcacacac	tcaactgtct	ttctctcttc	tagaaataca	ggtacaagta	tgtactattc	3780
tcatgtttca	tacttctagt	catttcatcc	cacatatctc	ttggattttc	ctccaatgaa	3840
tgacattcta	tcttgcaaat	tcaacaatta	taataagata	taccaaaagta	gcggtatagt	3900
ggcaatcaaa	aagcttctct	ggtgtgcttc	tcgtatttat	ttttattcta	atgatccatt	3960
aaaggtatat	atttattttc	tgttatataa	tccttttggt	tattacatgg	gctggataca	4020
taaaggtatt	ttgatttaat	tttttgctta	aattcaatcc	cccctcgttc	agtgtaact	4080
gtaatggtag	gaaattacca	tacttttgaa	gaagcaaaaa	aaatgaaaga	aaaaaaaaaat	4140
cgtattttcca	ggttagacgt	tcgcgagaat	ctagaatgcg	gtatgcggta	cattgttctt	4200
cgaacgtaaa	agttgcgctc	cctgagatat	tgtacatttt	tgctttttaca	agtacaagta	4260
catcgtaaaa	ctatgtacta	ctgttgatgc	atccacaaca	gtttgttttg	tttttttttg	4320
tttttttttt	ttctaattgat	tcattaccgc	tatgtatacc	tacttgtact	tgtagtaagc	4380
cgggttattg	gcgttcaatt	aatcatagac	ttatgaatct	gcacgggtgtg	cgctgcgagt	4440
tactttttage	ttatgcatgc	tacttgggtg	taatattggg	atctgttcgg	aaatcaacgg	4500
atgctcaate	gatttcgaca	gtaattaatt	aagtcataca	caagtcagct	ttcttcgagc	4560
ctcatataag	tataagtagt	tcaacgtatt	agcactgtac	ccagcatctc	cgtatcgaga	4620
aacacaacaa	catgccccat	tggacagatc	atgcggatac	acaggttgtg	cagtatcata	4680
catactcgat	cagacaggtc	gtctgaccat	catacaagct	gaacaagcgc	tccataactg	4740
cacgctctct	atatacacag	ttaaattaca	tatccatagt	ctaacctcta	acagttaate	4800
ttctggtaag	cctcccagcc	agccttctgg	tatcgcttgg	cctcctcaat	aggatctcgg	4860

ttctggccgt	acagacctcg	gccgacaatt	atgatatccg	ttccggtaga	catgacatcc	4920
tcaacagttc	ggtactgctg	tccgagagcg	tctcccttgt	cgtcaagacc	caccccgggg	4980
gtcagaataa	gccagtcctc	agagtcgccc	ttaggtcggt	tctgggcaat	gaagccaacc	5040
acaaactcgg	ggtcggatcg	ggcaagctca	atggtctgct	tggagtactc	gccagtggcc	5100
agagagccct	tgcaagacag	ctcggccagc	atgagcagac	ctctggccag	cttctcgttg	5160
ggagagggga	ctaggaactc	cttgtagctg	gagttctcgt	agtcagagac	gtcctccttc	5220
ttctgttcag	agacagtttc	ctcggcacca	gctcgcaggc	cagcaatgat	tccggttccg	5280
ggtacaccgt	gggcgttggg	gatateggac	cactcggcga	ttcggtgaca	ccggtactgg	5340
tgcttgacag	tgttgccaat	atctgcgaac	tttctgtcct	cgaacaggaa	gaaaccgtgc	5400
ttaagagcaa	gttccttgag	ggggagcaca	gtgccggcgt	aggtgaagtc	gtcaatgatg	5460
tcgatatggg	ttttgatcat	gcacacataa	ggtccgacct	tatcggcaag	ctcaatgagc	5520
tccttggtgg	tggtaacatc	cagagaagca	cacaggttgg	ttttcttggc	tgccacgagc	5580
ttgagcactc	gagcggcaaa	ggcggacttg	tggacgttag	ctcgagcttc	gtaggagggc	5640
attttggtgg	tgaagaggag	actgaaataa	atttagtctg	cagaactttt	tatcggaacc	5700
ttatctgggg	cagtgaagta	tatgttatgg	taatagttac	gagttagttg	aacttataga	5760
tagactggac	tatacggcta	tcggtcctaa	ttagaaagaa	cgtcaatggc	tctctggggc	5820
tcgcctttgc	cgacaaaaat	gtgatcatga	tgaaagccag	caatgacggt	gcagctgata	5880
ttgtttgtcg	ccaaccgcgc	cgaaaacgca	gctgtcagac	ccacagcctc	caacgaagaa	5940
tgtatcgtca	aagtgatcca	agcacactca	tagttggagt	cgtactccaa	aggcggcaat	6000
gacgagtcag	acagatactc	gtcgaaaaca	gtgtacgcag	atctactata	gaggaacatt	6060
taaattgccc	cggagaagac	ggccaggccg	cctagatgac	aaattcaaca	actcacagct	6120
gactttctgc	cattgccact	aggggggggg	ctttttatat	ggccaagcca	agctctccac	6180
gtcggttggg	ctgcacccaa	caataaatgg	gtagggttgc	accaacaaag	ggatgggatg	6240
gggggtagaa	gatacgagga	taacggggct	caatggcaca	aataagaacg	aatactgcca	6300
ttaagactcg	tgatccagcg	actgacacca	ttgcatcctc	taagggcctc	aaaactacct	6360
cggaaactgt	gcgctgatct	ggacaccaca	gaggttccga	gcactttagg	ttgcacccaa	6420
tgtcccacca	ggtgcaggca	gaaaacgctg	gaacagcgtg	tacagtttgt	cttaacaaaa	6480
agtgagggcg	ctgaggtcga	gcagggttgt	gtgacttggt	atagccttta	gagctgcgaa	6540
agcgcgtatg	gattttggctc	atcaggccag	attgagggtc	tgtggacaca	tgtcatgtta	6600
gtgtacttca	atcgccccct	ggatatagcc	ccgacaatag	gccgtggcct	catttttttg	6660
ccttccgcac	atttccattg	ctcggtagcc	acaccttgct	tctcctgcac	ttgccaacct	6720
taatactggg	ttacattgac	caacatctta	caagcggggg	gcttgtctag	ggtatatata	6780
aacagtggct	ctcccaatcg	gttgccagtc	tcttttttcc	tttctttccc	cacagattcg	6840
aaatctaaac	tacacatcac	agaattccga	gccgtgagta	tccacgacaa	gatcagtgtc	6900
gagacgacgc	gtttttgtgt	atgacacaat	ccgaaagtcg	ctagcaacac	acactctcta	6960
cacaaactaa	cccagctctg	gtaccatggg	gggaagttca	catgcattcg	ctggtgaatc	7020
tgatctgaca	ctacaactac	acaccaggtc	caacatgagc	gacaatacga	caatcaaaaa	7080
gccgatccga	cccaaaccga	tccggacgga	acgcctgcct	tacgctgggg	ccgcagaaat	7140
catccgagcc	aaccagaaag	accactactt	tgagtccgtg	cttgaacagc	atctcgtcac	7200


```

gtttctgcag aaatggaagg gagtacgatt tatccaccag tacaaggagg agctggagac 7260
ggcgtccaag ttgcatatc tcggtttgtg tacgcttgtg ggctccaaga ctctcggaga 7320
agagtacacc aatctcatgt acactatcag agaccgaaca gctctaccgg gggtggtgag 7380
acggtttggc tacgtgcttt ccaacactct gtttccatac ctgtttgtgc gctacatggg 7440
caagttgcgc gccaaactga tgcgcgagta tccccatctg gtggagtacg acgaagatga 7500
gcctgtgccc agcccggaaa catggaagga gcgggtcatc aagacgtttg tgaacaagtt 7560
tgacaagttc acggcgctgg aggggtttac cgcgatccac ttggcgattt tctacgtcta 7620
cggctcgtac taccagctca gtaagecgat ctggggcatg cgttatgtat ttggacaccg 7680
actggacaag aatgagcctc gaatcggtta cgagatgctc ggtctgctga ttttcgcccc 7740
gtttgccacg tcatttgtgc agacgggaag agagtacctc ggagcgctgc tggaaaagag 7800
cgtggagaaa gaggcagggg agaaggaaga tgaaaaggaa gcggttgtgc cgaaaaagaa 7860
gtcgtcaatt ccgttcattg aggatacaga aggggagacg gaagacaaga tcgatctgga 7920
ggaccctcga cagctcaagt tcattcctga ggcgccaga gcgtgcactc tgtgtctgtc 7980
atacattagt gcgcggcat gtacgccatg tggacacttt ttctgttggg actgtatttc 8040
cgaatgggtg agagagaagc ccgagtgtcc cttgtgtcgg cagggtgtga gagagcagaa 8100
cttgttgctt atcagataat gacgaggtct ggc 8133

```

<210>75

<211>35

<212>DNA

<213> 人工序列

<220>

<223> 引物 PEX10-R-BsiWI

<400>75

```

gatcaacgta cgcttcagca gtaactgtat tgctc 35

```

<210>76

<211>35

<212>DNA

<213> 人工序列

<220>

<223> 引物 PEX10-F1-SalI

<400>76

```

gatcaagtcg acattgtaac tagtcctgga gggtc 35

```

<210>77

<211>36

<212>DNA

<213> 人工序列

<220>

<223> 引物 PEX10-F2-SalI

<400>77

gatcaagtcg acgtcttagc gtcattgtatt ctcaag	36
<210>78	
<211>7277	
<212>DNA	
<213> 人工序列	
<220>	
<223> 质粒 pEXP-MOD1	
<400>78	
catggatcca ggcctgttaa cggccattac ggcctgcagg atccgaaaaa acctcccaca	60
cctccccctg aacctgaaac ataaaatgaa tgcaattggt gttgttaact tgtttattgc	120
agcttataat ggttacaaat aaagcaatag catcacaaat ttcacaaata aagcattttt	180
ttcactgcat tctagttgtg gtttgtccaa actcatcaat gtatcttata atgtctgcgg	240
ccgcaagtgt ggatggggaa gtgagtgcc gggtctgtgt gcacaattgg caatccaaga	300
tggatggatt caacacaggg atatagcgag ctacgtgggt gtgcgaggat atagcaacgg	360
atatttatgt ttgacacttg agaatgtacg atacaagcac tgtccaagta caatactaaa	420
catactgtac atactcatac tcgtaccgg gcaacgggtt cacttgagt cagtggctag	480
tgctcttact cgtacagtgt gcaatactgc gtatcatagt ctttgatgta tatcgtattc	540
attcatgtta gttgcgtacg agccgggaagc ataaagtgt aagcctgggg tgcctaata	600
gtgagctaac tcacattaat tgcgttgcgc tctactgccc ctttccagtc gggaaacctg	660
tcgtgccagc tgcattaatg aatcgccaa cgcgcgggga gagcggttt gcgtattggg	720
cgctcttccg ctctctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg	780
gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga	840
aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg	900
gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag	960
aggtggcgaa acccgacagg actataaaga taccaggcgt tccccctgg aagctccctc	1020
gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg	1080
ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc tcagttcggg ttaggtcgtt	1140
cgctccaagc tgggctgtgt gcacgaacct ccggttcagc ccgaccgctg cgccttatcc	1200
ggtaactatc gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc	1260
actggtaaca ggattagcag agcgaggtat gtaggcgggt ctacagagtt cttgaagtgg	1320
tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca	1380
gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc	1440
ggtgggtttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat	1500
cctttgatct tttctacggg gtctgacgct cagtggaaacg aaaactcacg ttaagggtat	1560
ttggatcatga gattatcaaa aaggatcttc acctagatcc ttttaaatga aaaaatgaagt	1620
tttaaatcaa tctaaagtat atatgagtaa acttgggtctg acagttacca atgcttaate	1680
agtgaggcac ctatctcagc gatctgtcta tttcggttcat ccatagttgc ctgactcccc	1740
gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtc tgcaatgata	1800
ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg	1860

gccgagcgca	gaagtgggtcc	tgcaacttta	tccgcctcca	tccagtctat	taattgtttgc	1920
cgggaagcta	gagtaagtag	ttcgccagtt	aatagtttgc	gcaacgttgt	tgccattgct	1980
acaggcatcg	tgggtgtcacg	ctcgtcgttt	ggtatggctt	cattcagctc	cggttcccaa	2040
cgatcaagge	gagttacatg	atcccccatg	ttgtgcaaaa	aagcggttag	ctccttcggt	2100
cctccgateg	ttgtcagaag	taagttggcc	gcagtgttat	cactcatggt	tatggcagca	2160
ctgcataatt	ctcttactgt	catgccatcc	gtaagatgct	tttctgtgac	tgggtgagtac	2220
tcaaccaagt	cattctgaga	atagtgtatg	cggcgaccga	gttgcctctg	cccggcgtca	2280
atacgggata	ataccgcgcc	acatagcaga	actttaaaag	tgctcatcat	tggaaaacgt	2340
tcttcggggc	gaaaactctc	aaggatctta	ccgctgttga	gatccagttc	gatgtaaccc	2400
actcgtgcac	ccaactgate	ttcagcatct	tttactttca	ccagcgtttc	tgggtgagca	2460
aaaacaggaa	ggcaaaatgc	cgcaaaaaag	ggaataaggg	cgacacggaa	atgttgaata	2520
ctcatactct	tcctttttca	atattattga	agcattttatc	agggttattg	tctcatgagc	2580
ggatacatat	ttgaatgtat	ttagaaaaat	aaacaaatag	gggttccgcg	cacatttccc	2640
cgaaaagtgc	cacctgacgc	gccctgtage	ggcgcattaa	gcgcggcggg	tgtggtggtt	2700
acgcgcagcg	tgaccgctac	acttgccagc	gccctagcgc	ccgctccttt	cgttttcttc	2760
ccttcctttc	tcgccacgtt	cgcggctttt	ccccgtcaag	ctctaaatcg	ggggctccct	2820
ttagggttcc	gatttagtgc	tttacggcac	ctcgacccca	aaaaacttga	ttagggtgat	2880
ggttcacgta	gtgggccatc	gccctgatag	acggtttttc	gccctttgac	gttggagtcc	2940
acgttcttta	atagtggact	cttgtttcaa	actggaacaa	cactcaaccc	tatctcggtc	3000
tattcttttg	atttataagg	gattttgccg	atttcggcct	attgggttaa	aaatgagctg	3060
atttaacaaa	aatttaacgc	gaattttaac	aaaatattaa	cgtttacaat	ttccattcgc	3120
cattcaggct	gcgcaactgt	tgggaagggc	gatcgggtgcg	ggcctcttcg	ctattacgcc	3180
agctggcgaa	agggggatgt	gctgcaaggc	gattaagttg	ggtaacgcca	gggttttccc	3240
agtcacgacg	ttgtaaaacg	acggccagtg	aattgtaata	cgactcacta	tagggcgaa	3300
tgggtaccgg	gccccccctc	gaggctgatg	gtgtcgataa	gcttgatata	gaattcatgt	3360
cacacaaacc	gatcttcgcc	tcaaggaaac	ctaattctac	atccgagaga	ctgccgagat	3420
ccagtctaca	ctgattaatt	ttcgggccaa	taatttaaaa	aaatcgtgtt	atataatatt	3480
atatgtatta	tatatataca	tcatgatgat	actgacagtc	atgtcccatt	gctaaataga	3540
cagactccat	ctgccgcctc	caactgatgt	tctcaatatt	taaggggtca	tctcgcattg	3600
tttaataata	aacagactcc	atctaccgcc	tccaaatgat	gttctcaaaa	tatattgtat	3660
gaacttattt	ttattactta	gtattattag	acaacttact	tgctttatga	aaaacacttc	3720
ctatttagga	aacaatttat	aatggcagtt	cgttcattta	acaatttatg	tagaataaat	3780
gttataaatg	cgtatgggaa	atcttaaata	tggatagcat	aaatgatata	tgcatcgcct	3840
aattcgaaat	caacagcaac	gaaaaaaatc	ccttgtacaa	cataaatagt	catcgagaaa	3900
tatcaactat	caaagaacag	ctattcacac	gttactattg	agattattat	tggacgagaa	3960
tcacacactc	aactgtcttt	ctctcttcta	gaaatacagg	tacaagtatg	tactattctc	4020
attgttcata	cttctagtc	tttcatecca	catattcctt	ggattttctt	ccaatgaatg	4080
acattctata	ttgcaaattc	aacaattata	ataagatata	ccaaagtagc	ggtatagtgg	4140
caatcaaaaa	gcttctctgg	tgtgcttctc	gtatttattt	ttattctaat	gatccattaa	4200

aggtatatat	ttatttccttg	ttatataatc	cttttgttta	ttacatgggc	tggatacata	4260
aaggtatfff	gatttaatff	tttgcftaaa	ttcaatcccc	cctcgttcag	tgtcaactgt	4320
aatggtagga	aattaccata	cttttgaaga	agcaaaaaaa	atgaaagaaa	aaaaaaatcg	4380
tatttccagg	ttagacgttc	cgcagaatct	agaatgcggt	atgcggtaca	ttgttcttcg	4440
aacgtaaaag	ttgcgtccc	tgagatatig	tacatfffft	ctttttacaag	tacaagtaca	4500
tcgtacaact	atgtactact	gttgatgcat	ccacaacagt	ttgtttttgtt	ttttttttgtt	4560
tttttttttt	ctaattgattc	attaccgcta	tgtataccta	cttgtacttg	tagtaagccg	4620
ggttatitggc	gttcaattaa	tcatagactt	atgaatctgc	acgggtgtgcg	ctgcgagtta	4680
cttttagctt	atgcatgcta	cttgggtgta	atattgggat	ctgttcggaa	atcaacggat	4740
gctcaatcga	tttcgacagt	aatttaattaa	gtcatacaca	agtcagcttt	cttcgagcct	4800
catataagta	taagtagttc	aacgtattag	cactgtaccc	agcatctccg	tatcgagaaa	4860
cacaacaaca	tgccccattg	gacagatcat	gcggatacac	aggtttgtgca	gtatcataca	4920
tactcgatca	gacaggtcgt	ctgaccatca	tacaagctga	acaagcgctc	catacttgca	4980
cgctctctat	atacacagtt	aaattacata	tccatagtct	aacctctaac	agttaatctt	5040
ctggtaagcc	tcccagccag	ccttctggta	tcgcttggcc	tcctcaatag	gatctcggtt	5100
ctggccgtac	agacctcggc	cgacaattat	gatatccgtt	ccggtagaca	tgacatcctc	5160
aacagttcgg	tactgtctgc	cgagagcgtc	tcccttgtcg	tcaagaccca	ccccgggggt	5220
cagaataagc	cagtcctcag	agtcgccctt	aggtcggttc	tgggcaatga	agccaaccac	5280
aaactcgggg	tcggatcggg	caagctcaat	ggtctgcttg	gagtactcgc	cagtggccag	5340
agagcccttg	caagacagct	cggccagcat	gagcagacct	ctggccagct	tctcgttggg	5400
agaggggact	aggaactcct	tgtactggga	gttctcgtag	tcagagacgt	cctccttctt	5460
ctgttcagag	acagtttctt	cggcaccagc	tcgcaggcca	gcaatgattc	cggttccggg	5520
tacaccgtgg	gcgttgggtga	tatcggacca	ctcggcgatt	cggtgacacc	ggtactgggtg	5580
cttgacagtg	ttgccaatat	ctgcgaactt	tctgtcctcg	aacaggaaga	aaccgtgctt	5640
aagagcaagt	tccttgaggg	ggagcacagt	gccggcgtag	gtgaagtcgt	caatgatgtc	5700
gatatgggtt	ttgatcatgc	acacataagg	tccgacctta	tcggcaagct	caatgagctc	5760
cttgggtggtg	gtaacatcca	gagaagcaca	caggttgggt	ttcttggctg	ccacgagctt	5820
gagcactcga	gcggcaaagg	cggacttgtg	gacgttagct	cgagcttcgt	aggagggcat	5880
tttgggtggtg	aagaggagac	tgaataaat	ttagtctgca	gaacttttta	tcggaacctt	5940
atctggggca	gtgaagtata	tgttatggta	atagttacga	gttagttgaa	cttatagata	6000
gactggacta	tacggctatc	ggtccaaatt	agaaagaacg	tcaatggctc	tctgggcgtc	6060
gcctttgccg	acaaaaatgt	gatcatgatg	aaagccagca	atgacgttgc	agctgatatt	6120
gttgtcggcc	aaccgcgccg	aaaacgcagc	tgtcagaccc	acagcctcca	acgaagaatg	6180
tatcgtcaaa	gtgatccaag	cacactcata	gttggagtcg	tactccaaag	gcggcaatga	6240
cgagtcagac	agatactcgt	cgaccgtacg	gggagttitg	cgccccgtttt	ttcgagcccc	6300
acacgtttcg	gtgagtatga	gcggcggcag	attcgagcgt	ttccggtttc	cgcggctgga	6360
cgagagccca	tgatgggggc	tcccaccacc	agcaatcagg	gccctgatta	cacaccacc	6420
tgtaatgtca	tgctgttcat	cgatggttaa	tgctgtctgt	tgctgtgtgt	gtgtgttgtt	6480
tggcgctcat	tgttgcgtta	tgcagcgtac	accacaatat	tggaagctta	ttagcctttc	6540

tat	tttttttcg	tttgcaaggc	ttaacaacat	tgctgtggag	agggatgggg	atatggaggc	6600
cgct	ggaggg	agtcggagag	gcgttttggga	gcggcttggc	ctggcgccca	gctcgcgaaa	6660
cgcac	tagg	accctttggc	acgccgaaat	gtgccacttt	tcagtctagt	aacgccttac	6720
ctacgt	catt	ccatgcgtgc	atgtttgcgc	cttttttccc	ttgcccttga	tcgccacaca	6780
gtacagt	gca	ctgtacagt	gaggttttgg	gggggtctta	gatgggagct	aaaagcggcc	6840
tagcgg	taca	ctagtgggat	tgtatggagt	ggcatggagc	ctaggtggag	cctgacagga	6900
cgcacg	accg	gctagcccgt	gacagacgat	gggtggctcc	tgttgtccac	cgcgtaaaa	6960
tgttt	gggcc	aaagtcttgt	cagccttgtc	tgcgaacct	attcccaatt	ttgtcacttc	7020
gcaccccc	cat	tgategagcc	ctaaccctg	cccatcaggc	aatccaatta	agctcgcatt	7080
gtctgc	cttg	tttagtttgg	ctcctgccc	tttcggcgtc	cacttgcaca	aacacaaaca	7140
agcatt	tatat	ataaggctcg	tctctccctc	ccaaccacac	tcactttttt	gcccgtcttc	7200
ccttgcta	aac	acaaaagtca	agaacacaaa	caaccacccc	aacccccctta	cacacaagac	7260
atatctac	ag	caatggc					7277

<210>79

<211>7559

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pPEX10-1

<400>79

gtacgagccg	gaagcataaa	gtgtaaagcc	tgggggtgcct	aatgagtgag	ctaactcaca	60
ttaattgcgt	tgcgctcact	gcccgccttc	cagtcgggaa	acctgtcgtg	ccagctgcat	120
taatgaatcg	gccaacgcgc	ggggagaggg	ggtttgcgta	ttgggcgctc	ttccgcttcc	180
tcgctcactg	actcgctgcg	ctcggtcggt	cggctgcggc	gagcgggtatc	agctcactca	240
aaggcggtaa	tacggttatc	cacagaatca	ggggataacg	caggaaagaa	catgtgagca	300
aaaggccagc	aaaaggccag	gaaccgtaaa	aaggccgcgt	tgctggcggt	tttccatagg	360
ctccgcccc	ctgacgagca	tcacaaaaat	cgacgctcaa	gtcagagggtg	gcgaaacccg	420
acaggactat	aaagatacca	ggcgcttccc	cctggaagct	ccctcgctgcg	ctctcctggt	480
ccgaccctgc	cgcttaccgg	atacctgtcc	gcctttctcc	cttcgggaag	cgtggcgctt	540
tctcatagct	cacgctgtag	gtatctcagt	tcgggtgtagg	tcgttcgctc	caagctgggc	600
tgtgtgcacg	aaccccccg	tcagcccagc	cgctgcgcct	tatccggtaa	ctatcgtctt	660
gagtccaacc	cggtaagaca	cgacttatcg	ccactggcag	cagccactgg	taacaggatt	720
agcagagcga	ggtatgtagg	cggtgctaca	gagtctctga	agtgggtggc	taactacggc	780
tacactagaa	ggacagtatt	tggtatctgc	gctctgctga	agccagttac	cttcggaaaa	840
agagttagta	gctcttgatc	cggcaaacaa	accaccgctg	gtagcgggtgg	tttttttgtt	900
tgcaagcagc	agattacgcg	cagaaaaaaa	ggatctcaag	aagatccctt	gatcttttct	960
acgggggtctg	acgctcagt	gaacgaaaac	tcacgttaag	ggatttttgt	catgagatta	1020
tcaaaaagga	tcttcacct	gatcccttta	aattaaaaat	gaagttttta	atcaatctaa	1080
agtatatatg	agtaaacctg	gtctgacagt	taccaatgct	taatcagtga	ggcacctatc	1140

tcagegatct	gtctatttctg	ttcatecata	gttgccctgac	tccccgtcgt	gtagataact	1200
acgatacggg	agggttacc	atctggcccc	agtgtctgca	tgataccgcg	agaccacgc	1260
tcaccggctc	cagatttate	agcaataaac	cagccagccg	gaagggccga	gcgcagaagt	1320
ggctctgcaa	ctttatecgc	ctccatccag	tctattaatt	gttgccggga	agctagagta	1380
agtagttcgc	cagttaatag	tttgcgcaac	gttggttgcca	ttgctacagg	catcgtgggtg	1440
tcacgctcgt	cgttttggtat	ggcttcatte	agctccgggt	cccaacgate	aaggcgagtt	1500
acatgatccc	ccatgtttgtg	caaaaaagcg	gttagctcct	tcggctctcc	gatcgtttgtc	1560
agaagtaagt	tggccgcagt	gttatcactc	atgggttatgg	cagcactgca	taattctctt	1620
actgtcatgc	catccgtaag	atgcttttct	gtgactgggtg	agtactcaac	caagtcattc	1680
tgagaatagt	gtatgcggcg	accgagttgc	tcttgccccg	cgtcaatacg	ggataatacc	1740
gcgccacata	gcagaacttt	aaaagtgcct	atcattggaa	aacgttcttc	ggggcgaaaa	1800
ctctcaagga	tcttaccgct	gttgagatcc	agttcgatgt	aaccactcgc	tgcacccaac	1860
tgatcttcag	catcttttac	tttcaccagc	gtttctgggt	gagcaaaaac	aggaaggcaa	1920
aatgccgcaa	aaaagggaat	aagggcgaca	cggaaatgtt	gaatactcat	actcttcctt	1980
tttcaatatt	attgaagcat	ttatcagggt	tattgtctca	tgagcggata	catatttgaa	2040
tgtattttaga	aaaataaaca	aataggggtt	ccgcgcacat	ttccccgaaa	agtgccacct	2100
gacgcgcct	gtagcggcgc	attaagcgcg	gcgggtgtgg	tggttacgcg	cagcgtgacc	2160
gctacacttg	ccagcgcct	agcgcgcct	cctttcgctt	tcttcccttc	ctttctcgcc	2220
acgttcgccg	gctttccccg	tcaagctcta	aatcgggggc	tccctttagg	gttccgattt	2280
agtgtcttac	ggcacctcga	ccccaaaaaa	cttgattagg	gtgatgggtc	acgtagtggg	2340
ccatcgccct	gatagacggt	ttttcgccct	ttgacgttgg	agtccacgtt	ctttaatagt	2400
ggactcttgt	tccaaacttg	aacaacactc	aaccctatct	cggctctattc	ttttgattta	2460
taagggattt	tgccgatttc	ggcctatttg	ttaaaaaatg	agctgattta	acaaaaattt	2520
aacgcgaatt	ttaacaaaat	attaacgctt	acaatttcca	ttcgccattc	aggctgcgca	2580
actgtttgga	agggcgatcg	gtgcgggcct	cttcgctatt	acgccagctg	gcgaaagggg	2640
gatgtgctgc	aaggcgatta	agttgggtta	cgccagggtt	ttcccagtc	cgacgttgta	2700
aaacgacggc	cagtgaattg	taatacgact	cactataggg	cgaattgggt	accgggcccc	2760
ccctcgaggt	cgatggtgtc	gataagcttg	atatcgattt	catgtcacac	aaaccgatct	2820
tcgcctcaag	gaaacctaat	tctacatccg	agagactgcc	gagatccagt	ctacactgat	2880
taatttttcgg	gccaataatt	taaaaaaatc	gtgttatata	atattatatg	tattatatat	2940
atacatcatg	atgatactga	cagtcatgtc	ccattgctaa	atagacagac	tccatctgcc	3000
gcctccaact	gatgttctca	atatttaagg	ggtcactctc	cattgtttta	taataaacag	3060
actccatcta	ccgcctccaa	atgatgttct	caaaatatat	tgtatgaact	tatttttatt	3120
acttagtatt	attagacaac	ttacttgctt	tatgaaaaac	acttcctatt	taggaaacaa	3180
tttataatgg	cagttcgttc	atttaacaat	ttatgtagaa	taaatgttat	aaatgcgtat	3240
gggaaatctt	aaatatggat	agcataaatg	atatctgcat	tgccctaattc	gaaatcaaca	3300
gcaacgaaaa	aaatcccttg	tacaacataa	atagtcacgc	agaaatatca	actatcaaag	3360
aacagctatt	cacacgttac	tattgagatt	attattggac	gagaatcaca	cactcaactg	3420
tctttctctc	ttctagaaat	acaggtacaa	gtatgtacta	ttctcattgt	tcatacttct	3480

agtcatttca	tcccacatat	tccttggatt	tctctccaat	gaatgacatt	ctatcttgca	3540
aattcaacaa	ttataataag	atataccaaa	gtagcggtat	agtggaacatc	aaaaagcttc	3600
tctgggtg	ttctcgtatt	tatTTTTatt	ctaatgatcc	attaaaggta	tatatTTatt	3660
tcttggtata	taatcctttt	gtttattaca	tgggctggat	acataaagg	attttgattt	3720
aattttttgc	ttaaattcaa	teccccctcg	ttcagtgtea	actgtaattg	taggaaatta	3780
ccatactttt	gaagaagcaa	aaaaaatgaa	agaaaaaaa	aatcgtaatt	ccaggttaga	3840
cgttccgcag	aatctagaat	gcggtatgcg	gtacattggt	cttcgaacgt	aaaagttgcg	3900
ctccctgaga	tattgtacat	ttttgctttt	acaagtacaa	gtacatcgta	caactatgta	3960
ctactgttga	tgcateccaca	acagtttggt	ttgtTTTTTT	ttgtTTTTTT	TTTTtcta	4020
gattcattac	cgctatgtat	acctacttgt	acttgtagta	agccgggtta	ttggcggtca	4080
attaatcata	gacttatgaa	tctgcacgg	gtgcgctgcg	agttactttt	agcttatgca	4140
tgctacttgg	gtgtaatat	gggatctgtt	cggaaatcaa	cggatgctca	atcgatttcg	4200
acagtaatta	attaagtc	acacaagtca	gctttcttcg	agcctcatat	aagtataagt	4260
agttcaacgt	attagcactg	taccagcat	ctccgtatcg	agaaacacaa	caacatgccc	4320
cattggacag	atcatgcgga	tacacagggt	gtgcagtatc	atacatactc	gatcagacag	4380
gtcgtctgac	catcatacaa	gctgaacaag	cgctccatac	ttgcacgctc	tctatataca	4440
cagttaaatt	acatatccat	agtctaacct	ctaacagtta	atcttcttgt	aagcctccca	4500
gccagccttc	tgggtatcgct	tggcctcctc	aataggatct	cggttctggc	cgtacagacc	4560
tcgccgaca	attatgatat	ccgttccggt	agacatgaca	tcctcaacag	ttcgggtactg	4620
ctgtccgaga	gcgtctccct	tgctgtcaag	accacccccg	ggggtcagaa	taagccagtc	4680
ctcagagtcg	cccttaggtc	ggttctgggc	aatgaagcca	accacaaact	cggggtcgga	4740
tcgggcaagc	tcaatgggtct	gcttggagta	ctcgccagtg	gccagagagc	ccttgaaga	4800
cagctcggcc	agcatgagca	gacctctggc	cagcttctcg	ttgggagagg	ggactaggaa	4860
ctccttgtac	tgggagttct	cgtagtcaga	gacgtcctcc	ttcttctgtt	cagagacagt	4920
ttcctcggca	ccagctcgca	ggccagcaat	gattccgggt	ccgggtacac	cgtgggcgtt	4980
ggtgatatcg	gaccactcgg	cgattcgggtg	acaccgggtac	tgggtgcttga	cagtgttgcc	5040
aatatctgcg	aactttctgt	cctcgaacag	gaagaaaccg	tgcttaagag	caagttcctt	5100
gagggggagc	acagtgcggg	cgtaggtgaa	gtcgtcaatg	atgtcgatat	gggttttgat	5160
catgcacaca	taaggtccga	ccttatcggc	aagctcaatg	agctccttgg	tgggtgtaac	5220
atccagagaa	gcacacaggt	tggTTTTctt	ggctgccacg	agcttgagca	ctcgagcggc	5280
aaaggcggac	ttgtggacgt	tagctcgagc	ttcgtaggag	ggcatttttg	tgggtgaagag	5340
gagactgaaa	taaatttagt	ctgcagaact	ttttatcgga	accttatctg	gggcagtgaa	5400
gtatatgtta	tggtaatatg	tacaggttag	ttgaacttat	agatagactg	gactatacgg	5460
ctatcgggtc	aaattagaaa	gaacgtcaat	ggctctcttg	gcgtcgccct	tgccgacaaa	5520
aatgtgatca	tgatgaaagc	cagcaatgac	gttgcagctg	atattgttgt	cggccaaccg	5580
cgccgaaaac	gcagctgtca	gaccacagc	ctccaacgaa	gaatgtatcg	tcaaagtgat	5640
ccaagcacac	tcatagttag	agtcgtactc	caaaggcggc	aatgacgagt	cagacagata	5700
ctcgtcgaca	ttgtaactag	tcctggagg	tcTTTTTTat	ggataacctc	catgtacgat	5760
gtatccaaga	tctccacgta	ctgtgttctg	tttcctaagt	aatacccaac	aacctctcca	5820

acaaacactt	gggaagatgc	acttgtgctg	agatgtcaag	atgttagtac	tgtactggat	5880
ggagagaata	ttaataaata	attgttaccc	aactacatct	tgtcgattga	aagagatacc	5940
cctaagacag	ataggatata	tgcaaccgga	ggaatgaacc	ccccagcacc	ggcacccttt	6000
ctattaacaa	aatgccaact	gaaatttgaa	aagtccaact	aaacttatatt	gaccacaaaa	6060
aactcgtcaa	aagtggcggc	gaaagctggc	aatgatgac	atccccttgg	aactatgata	6120
tcccctcgga	atcttcgtcc	ccatttgcca	catctacttg	caacgccacg	tctgcttact	6180
aagcaaccca	aatctgcctc	ggctcaaaat	gtggggaagt	tcacatgcat	tcgctgggtga	6240
atctgatctg	acactacaac	tacacaccag	gtccaacatg	agcgacaata	cgacaatcaa	6300
aaagccgata	cgacccaaac	cgatccggac	ggaacgcctg	ccttacgctg	gggccgcaga	6360
aatcatccga	gccaaccaga	aagaccacta	ctttgagtc	gtgcttgaac	agcatctcgt	6420
cacgtttctg	cagaaatgga	aggaggtacg	atttatccac	cagtacaagg	aggagctgga	6480
gacggcgctc	aagtttgcac	atctcggttt	gtgtacgctt	gtgggctcca	agactctcgg	6540
agaagagtac	accaatctca	tgtacactat	cagagaccga	acagctctac	cgggggtggt	6600
gagacggttt	ggctacgtgc	tttccaacac	tctgtttcca	tacctgtttg	tgcgctacat	6660
gggcaagtgt	cgcgccaaac	tgatgcgcga	gtatcccat	ctggtggagt	acgacgaaga	6720
tgagcctgtg	cccagcccgg	aaacatggaa	ggagcgggtc	atcaagacgt	ttgtgaacaa	6780
gtttgacaag	ttcacggcgc	tggaggggtt	taccgcgata	cacttggcga	ttttctacgt	6840
ctacggctcg	tactaccagc	tcagtaagcg	gatctggggc	atgcgttatg	tatttggaca	6900
ccgactggac	aagaatgagc	ctcgaatcgg	ttacgagatg	ctcggctctg	tgattttcgc	6960
ccggtttgcc	acgtcatttg	tgcagacggg	aagagagtac	ctcggagcgc	tgctggaaaa	7020
gagcgtggag	aaagaggcag	gggagaagga	agatgaaaag	gaagcggttg	tgccgaaaaa	7080
gaagtcgtca	attccgttca	ttgaggatac	agaaggggag	acggaagaca	agatcgatct	7140
ggaggaccct	cgacagctca	agttcattcc	tgaggcgctc	agagcgtgca	ctctgtgtct	7200
gtcatacatt	agtgcgcggg	catgtacgcc	atgtggacac	tttttctgtt	gggactgtat	7260
ttccgaatgg	gtgagagaga	agcccagatg	tcccttgtgt	cggcagggtg	tgagagagca	7320
gaacttggtg	cctatcagat	aatgacgagg	tctggatgga	aggactagtc	agcgagacac	7380
agagcatcag	ggaccagaca	cgaccaattc	aatcgacaac	actgtgctgc	atagcagtgc	7440
acagaggtcc	tgggcatgaa	tatatatttag	cattggagat	atgagtggta	gagcgtatac	7500
agtattaatt	gtggaggtat	ctcgtcgcat	tgatagagca	atacagttac	tgctgaagc	7559

<210>80

<211>8051

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pPEX10-2

<400>80

gtacgagccg	gaagcataaa	gtgtaaagcc	tggggtgcct	aatgagttag	ctaactcaca	60
ttaattgcgt	tgcgtcact	gcccgttttc	cagtcgggaa	acctgtcgtg	ccagctgcat	120
taatgaatcg	gccaacgcgc	ggggagaggc	ggtttgcgta	ttgggcgctc	ttccgcttcc	180

tcgctcactg	actcgctgcg	ctcggtcggt	cggctgcggc	gagcgggtatc	agctcactca	240
aaggcggtaa	tacggttatc	cacagaatca	ggggataacg	caggaaagaa	catgtgagca	300
aaaggccagc	aaaaggccag	gaaccgtaaa	aaggccgcgt	tgctggcggt	tttccatagg	360
ctccgcccc	ctgacgagca	tcacaaaaat	cgacgctcaa	gtcagagggtg	gcgaaacccg	420
acaggactat	aaagatacca	ggcggtttccc	cctggaagct	ccctcgtgcg	ctctcctggt	480
ccgaccctgc	cgtttaccgg	atacctgtcc	gcctttctcc	cttcgggaag	cgtggcgctt	540
tctcatagct	cacgctgtag	gtatctcagt	tcgggtgtagg	tcgttcgctc	caagctgggc	600
tgtgtgcacg	aaccccccg	tcagcccgac	cgtgcgcct	tatccggtaa	ctatcgtctt	660
gagtccaacc	cggtaagaca	cgacttatcg	ccactggcag	cagccactgg	taacaggatt	720
agcagagcga	ggtatgtagg	cgggtgctaca	gagttcttga	agtgggtggc	taactacggc	780
tacactagaa	ggacagtatt	tggtatctgc	gctctgctga	agccagttac	cttcggaaaa	840
agagttggta	gctcttgatc	cggcaaaaa	accaccgctg	gtagcgggtg	tttttttgtt	900
tgcaagcagc	agattacgcg	cagaaaaaaa	ggatctcaag	aagatccttt	gatcttttct	960
acggggtctg	acgctcagt	gaacgaaaac	tcacgttaag	ggatttttgt	catgagatta	1020
tcaaaaagga	tcttcaccta	gaccttttta	aattaaaaat	gaagttttta	atcaatctaa	1080
agtatatatg	agtaaacttg	gtctgacagt	taccaatgct	taatcagtga	ggcacctatc	1140
tcagcgatct	gtctatttgc	ttcatccata	gttgccctgac	tccccgctcg	gtagataact	1200
acgatacggg	agggttacc	atctggcccc	agtgcctgca	tgataccgcg	agaccacgc	1260
tcaccggctc	cagatttate	agcaataaac	cagccagccg	gaagggccga	gcgcagaagt	1320
ggctctgcaa	ctttatccgc	ctccatccag	tctattaatt	gttgccggga	agctagagta	1380
agtagttcgc	cagttaatag	tttgcgcaac	gttggttgcca	ttgctacagg	catcgtgggtg	1440
tcacgctcgt	cgttttggtat	ggcttcattc	agctccgggt	cccaacgata	aaggcgagtt	1500
acatgatccc	ccatgtttgt	caaaaaagcg	gttagctcct	tcggctcctcc	gatcgtttgc	1560
agaagtaagt	tggccgcagt	gttatcactc	atgggttatgg	cagcactgca	taattctctt	1620
actgtcatgc	catccgtaag	atgcttttct	gtgactgggtg	agtactcaac	caagtcattc	1680
tgagaatagt	gtatgcggcg	accgagttgc	tcttgcccgg	cgtcaataacg	ggataataacc	1740
gcgccacata	gcagaacttt	aaaagtgtc	atcattggaa	aacgtttctc	ggggcgaaaa	1800
ctctcaagga	tcttaccgct	gttgagatcc	agttcgatgt	aaccactcgc	tgcacccaac	1860
tgatcttcag	catctttttac	tttcaccagc	gtttctgggt	gagcaaaaaac	aggaaggcaa	1920
aatgccgcaa	aaaagggaat	aagggcgaca	cggaaatgtt	gaatactcat	actcttcctt	1980
tttcaatatt	attgaagcat	ttatcagggt	tattgtctca	tgagcggata	catatttgaa	2040
tgtatttaga	aaaataaaca	aataggggtt	ccgcgcacat	ttccccgaaa	agtgccacct	2100
gacgcgcct	gtagcggcgc	attaagcgcg	gcgggtgtgg	tggttacgcg	cagcgtgacc	2160
gctacacttg	ccagcgcct	agcgcgcct	cctttcgctt	tcttcccttc	ctttctcgcc	2220
acgttcgcgc	gctttccccg	tcaagctcta	aatcgggggc	tcccttttagg	gttccgattt	2280
agtgtcttac	ggcacctcga	ccccaaaaaa	cttgattagg	gtgatgggtc	acgtagtggg	2340
ccatgcgcct	gatagacggg	ttttgcgcct	ttgacgttgg	agtcacggtt	ctttaatagt	2400
ggactcttgt	tccaaactgg	aacaacactc	aaccctatct	cggctctattc	ttttgattta	2460
taagggattt	tgcgcgattt	ggcctatttg	ttaaaaaaatg	agctgattta	acaaaaattt	2520

aacgcgaatt	ttaacaaaat	attaacgctt	acaatttcca	ttcgccattc	aggctgcgca	2580
actgttggga	agggcgatcg	gtgcgggcct	cttcgctatt	acgccagctg	gcgaaagggg	2640
gatgtgctgc	aaggcgatta	agttgggtaa	cgccagggtt	ttcccagtc	cgacgttgta	2700
aaacgacggc	cagtgaattg	taatacgact	cactataggg	cgaattgggt	accggggccc	2760
ccctcgaggt	cgatggtgtc	gataagcttg	atatcgaatt	catgtcacac	aaaccgatct	2820
tcgcctcaag	gaaacctaat	tctacatccg	agagactgcc	gagatccagt	ctacactgat	2880
taattttcgg	gccaataatt	taaaaaaatc	gtgttatata	atattatatg	tattatatat	2940
atacatcatg	atgatactga	cagtcatgtc	ccattgctaa	atagacagac	tccatctgcc	3000
gcctccaact	gatgttctca	atatttaagg	ggtcactctg	cattgtttta	taataaacag	3060
actccatcta	ccgcctccaa	atgatgttct	caaaatatat	tgtatgaact	tatttttatt	3120
acttagtatt	attagacaac	ttacttgctt	tatgaaaaac	acttcctatt	taggaaacaa	3180
tttataatgg	cagttcgttc	atttaacaat	ttatgtagaa	taaagtgtat	aatgcgtat	3240
gggaaatctt	aaatatggat	agcataaatg	atatctgcat	tgccataatc	gaaatcaaca	3300
gcaacgaaaa	aaatcccttg	tacaacataa	atagtcatcg	agaaatatca	actatcaaag	3360
aacagctatt	cacacgttac	tattgagatt	attattggac	gagaatcaca	cactcaactg	3420
tctttctctc	ttctagaaat	acaggtacaa	gtatgtacta	ttctcattgt	tcatacttct	3480
agtcatttca	tcccacatat	tccttggatt	tctctccaat	gaatgacatt	ctatcttgca	3540
aattcaacaa	ttataataag	atataccaaa	gtagcggtat	agtggaatc	aaaaagcttc	3600
tctgggtgtc	ttctcgtatt	tatttttatt	ctaatgatcc	attaaaggta	tatatattat	3660
tcttggttata	taatcctttt	gtttattaca	tgggctggat	acataaagg	attttgattt	3720
aatTTTTTgc	ttaaattcaa	tccccctcgc	ttcagtgtea	actgtaatgg	taggaaatta	3780
ccatactttt	gaagaagcaa	aaaaaatgaa	agaaaaaaaa	aatcgatatt	ccaggttaga	3840
cgttccgcag	aatctagaat	gcggtatgcg	gtacattggt	cttcgaacgt	aaaagttgcg	3900
ctccctgaga	tattgtacat	ttttgctttt	acaagtacaa	gtacatcgta	caactatgta	3960
ctactgttga	tgcatccaca	acagtttggt	ttgttttttt	ttgttttttt	tttttcta	4020
gattcattac	cgctatgtat	acctacttgt	acttgtagta	agccgggtta	ttggcggtca	4080
attaatcata	gacttatgaa	tctgcacggg	gtgcgctgcg	agttactttt	agcttatgca	4140
tgctacttgg	gtgtaatat	gggatctggt	cggaaatcaa	cggatgctca	atcgatttcg	4200
acagtaatta	attaagtc	acacaagtca	gctttcttcg	agcctcatat	aagtataagt	4260
agttcaacgt	attagcactg	taccagcat	ctccgtatcg	agaaacacaa	caacatgccc	4320
cattggacag	atcatgcgga	tacacagggt	gtgcagtatc	atacatactc	gatcagacag	4380
gtcgtctgac	catcatacaa	gctgaacaag	cgctccatac	ttgcacgctc	tctatataca	4440
cagttaaatt	acatatccat	agtctaacct	ctaacagtta	atcttctggg	aagcctccca	4500
gccagccttc	tggtatcgct	tggcctcctc	aataggatct	cggttctggc	cgtacagacc	4560
tcggccgaca	attatgatat	ccgttccggg	agacatgaca	tcctcaacag	ttcgggtactg	4620
ctgtccgaga	gcgtctccct	tgctcgtcaag	accacccccg	ggggtcagaa	taagccagtc	4680
ctcagagtcg	cccttaggtc	ggttctgggc	aatgaagcca	accacaaact	cggggctcgga	4740
tcgggcaage	tcaatgggtc	gcttggagta	ctcgccagtg	gccagagagc	ccttgcaaga	4800
cagctcggcc	agcatgagca	gacctctggc	cagcttctcg	ttgggagagg	ggactaggaa	4860

ctccttgtac	tgggagttct	cgtagtcaga	gacgtcctcc	ttcttctgtt	cagagacagt	4920
ttcctcggca	ccagctcgca	ggccagcaat	gattccgggt	ccgggtacac	cgtgggcgtt	4980
ggtgatateg	gaccactcgg	cgattcgggtg	acaccgggtac	tgggtgcttga	cagtgttgcc	5040
aatatctgcg	aactttctgt	cctegaacag	gaagaaaccg	tgcttaagag	caagttcctt	5100
gagggggagc	acagtgccgg	cgtaggtgaa	gtcgtcaatg	atgtcgatat	gggttttgat	5160
catgcacaca	taaggtecga	ccttatecggc	aagctcaatg	agctccttgg	tgggtgtaac	5220
atccagagaa	gcacacaggt	tggttttctt	ggctgccacg	agcttgagca	ctcgagcggc	5280
aaaggcggac	ttgtggacgt	tagctcgagc	ttcgtaggag	ggcattttgg	tgggtgaagag	5340
gagactgaaa	taaatttagt	ctgcagaact	ttttatcgga	accttatctg	gggcagtga	5400
gtatatgtta	tggtaatagt	tacgagttag	ttgaacttat	agatagactg	gactatacgg	5460
ctatcgggcc	aaattagaaa	gaacgtcaat	ggctctctgg	gcgtcgccctt	tgccgacaaa	5520
aatgtgatca	tgatgaaagc	cagcaatgac	gttcgacgtg	atattgttgt	cggccaaccg	5580
cgccgaaaac	gcagctgtca	gaccacagc	ctccaacgaa	gaatgtatcg	tcaaagtgat	5640
ccaagcacac	tcatagttgg	agtcgtactc	caaaggcggc	aatgacgagt	cagacagata	5700
ctcgtcgacg	tcttagcgtc	atgtattctc	aagcttagtc	agagagaagg	actatggagg	5760
agaaggggag	aattgagaag	ggtatttgaa	gggactttga	aggctcgcgtg	gaagaggtac	5820
ttgaagaggt	atttgaaggt	cacgtggaag	aggtatttga	agatcacgtg	gaagaagtac	5880
ttgttttaca	gagaatatcg	gggtgatttt	gacagtggga	ttgtctccca	agtcctaate	5940
gtttgacatg	ggagcagtga	aaagtcgggc	taaaaaagg	aatatcgga	atcggaaga	6000
cggaaagaat	tactggactc	atgttttagt	gatctgagca	cttcaaattt	gaaaatatct	6060
cttcaaacag	cagatcgggt	ggtcgtggag	gtaccatcaa	gggtaaaate	aaggctatca	6120
tcaagggcca	tatatcgcaa	gtttggggga	agataatatg	ttcatagtga	atcagggttg	6180
tggatttcct	catctaacgg	cattgttaact	agtcctggag	ggctcttttt	atggataacc	6240
tccatgtacg	atgtatccaa	gatctccacg	tactgtgttc	tgtttcctaa	gtaataacca	6300
acaacctctc	caacaaacac	ttgggaagat	gcacttgtgc	tgagatgtca	agatgttagt	6360
actgtactgg	atggagagaa	tattaataaa	taattgttac	ccaactacat	cttgtcgatt	6420
gaaagagata	cccctaagac	agataggata	tctgcaaccc	gaggaatgaa	ccccccagca	6480
ccggcaccct	ttctattaac	aaaatgccaa	ctgaaatttg	aaaagttcaa	ctaaacttat	6540
ttgaccacac	aaaactcgtc	aaaagtggcg	gcgaaagctg	gcaaatgatg	acatcccctt	6600
ggaactatga	tatcccctcg	gaatcttcgt	ccccatttgc	cacatctact	tgcaacgcca	6660
cgtctgctta	ctaagcaacc	caaactctgc	tcggctcaaa	atgtggggaa	gttcacatgc	6720
attcgtctgg	gaatctgate	tgacactaca	actacacacc	aggccaaca	tgagcgacaa	6780
tacgacaate	aaaaagccga	tccgacccaa	accgatccgg	acggaacgcc	tgccttacgc	6840
tggggccgca	gaaatcatcc	gagccaacca	gaaagaccac	tactttgagt	ccgtgcttga	6900
acagcatctc	gtcacgtttc	tgcagaaatg	gaaggagta	cgatttatcc	accagtacaa	6960
ggaggagctg	gagacggcgt	ccaagtttgc	atatctcggt	ttgtgtacgc	ttgtgggctc	7020
caagactctc	ggagaagagt	acaccaatct	catgtacact	atcagagacc	gaacagctct	7080
accgggggtg	gtgagacggg	ttggctacgt	gctttccaac	actctgtttc	catacctgtt	7140
tgtgcgctac	atgggcaagt	tgcgcgccaa	actgatgcgc	gagtatcccc	atctggtgga	7200

gtacgacgaa gatgagcctg tgcccagccc ggaaacatgg aaggagcggg tcatcaagac	7260
gtttgtgaac aagtttgaca agttcacggc gctggagggg tttaccgcga tccacttggc	7320
gattttctac gtctacggct cgtactacca gctcagtaag cggatctggg gcatgcgtta	7380
tgtatttggga caccgactgg acaagaatga gcctcgaatc ggttacgaga tgctcggctc	7440
gctgattttc gcccggtttg ccacgtcatt tgtgcagacg ggaagagagt acctcggagc	7500
gctgctggaa aagagcgtgg agaaagaggc aggggagaag gaagatgaaa aggaagcgg	7560
tgtgccgaaa aagaagtcgt caattccgtt cattgaggat acagaagggg agacggaaga	7620
caagatcgat ctggaggacc ctcgacagct caagttcatt cctgaggcgt ccagagcgtg	7680
cactctgtgt ctgtcataca ttagtgcgcc ggcatgtacg ccatgtggac actttttctg	7740
ttgggactgt atttccgaat gggtgagaga gaagcccag tgtcccttgt gtcggcaggg	7800
tgtgagagag cagaacttgt tgcctatcag ataatgacga ggtctggatg gaaggactag	7860
tcagcgagac acagagcatt agggaccaga cacgaccaat tcaatcgaca aactgtgct	7920
gcatagcagt gcacagaggt cctgggcatg aatatatttt agcattggag atatgagtgg	7980
tagagcgtat acagtattaa ttgtggaggt atctcgtcgc attgatagag caatacagtt	8040
actgctgaag c	8051

<210>81

<211>15877

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pZKL1-2SP98C

<400>81

aaatgatgtc gacgcagtag gatgtcctgc acgggtcttt ttgtggggtg tggagaaagg	60
gggtgcttga tcgatggaag ccggtagaac cgggctgctt gtgcttggag atggaagccg	120
gtagaaccgg gctgcttggg gggatttggg gccgctgggc tccaaagagg ggtaggcatt	180
tcgttgggggt tacgtaattg cggcatttgg gtccctgcgcg catgtcccat tggtcagaat	240
tagtccggat aggagactta tcagccaatc acagcgccgg atccacctgt aggttgggtt	300
gggtgggagc acccctccac agagtagagt caaacagcag cagcaacatg atagttgggg	360
gtgtgcgtgt taaaggaaaa aaaagaagct tgggttatat tcccgtcta tttagaggtt	420
gcgggataga cgccgacgga gggcaatggc gctatggaac cttgcggata tccatacgcc	480
gcggcggact gcgtccgaac cagctccagc agcgtttttt ccgggccatt gagccgactg	540
cgaccccgcc aacgtgtctt ggcccacgca ctcattgtcat gttggtgttg ggaggccact	600
ttttaagtag cacaaggcac ctagctcgca gcaagggtgc cgaaccaaaag aagcggctgc	660
agtggtgcaa acggggcgga aacggcgga aaaagccacg ggggcacgaa ttgaggcacg	720
ccctcgaatt tgagacgagt cacggcccca ttcccccgcg caatggctcg ccaacgcccg	780
gtcttttgcg ccacatcagg ttaccccaag ccaaaccttt gtgttaaaaa gcttaacata	840
ttataccgaa cgtaggtttg ggcgggcttg ctccgtctgt ccaaggcaac atttatataa	900
gggtctgcat cgccggtcga attgaatctt ttttctctt ctcttctcta tattcattct	960
tgaattaaac acacatcaac catgggcgta ttcattaaac aggagcagct tccggctctc	1020

aagaagtaca	agtactccgc	cgaggatcac	tcgttcatct	ccaacaacat	tctgcgcccc	1080
ttctggcgac	agttttgtcaa	aattcttccct	ctgtggatgg	cccccaacat	ggtgactctg	1140
ctgggcttct	tctttgtcat	tgtgaacttc	atcaccatgc	tcattgttga	tcccaccac	1200
gaccgcgagc	ctcccagatg	ggtctacctc	acctacgctc	tgggtctgtt	cctttaccag	1260
acatttgatg	cctgtgacgg	atcccattgcc	cgacgaactg	gccagagtgg	accccttgga	1320
gagctgtttg	accactgtgt	cgacgccatg	aatacctctc	tgattctcac	ggtgggtggtg	1380
tccaccaccc	atatgggata	taacatgaag	ctactgattg	tgcagattgc	cgctctcgga	1440
aacttctacc	tgtcgacctg	ggagacctac	cataccggaa	ctctgtacct	ttctggcttc	1500
tctggtcctg	ttgaaggtat	cttgattctg	gtggctcttt	tcgtcctcac	cttcttcaact	1560
ggtcccaacg	tgtacgctct	gaccgtctac	gaggtctctc	ccgagtccat	cacttcgctg	1620
ctgcctgcca	gcttctctgga	cgtcaccatc	accagatct	acattggatt	cggagtgtctg	1680
ggcatggtgt	tcaacatcta	cggcgcctgc	ggaaacgtga	tcaagtacta	caacaacaag	1740
ggcaagagcg	ctctccccgc	cattctcgga	atcgccccct	ttggcatctt	ctacgtcggc	1800
gtctttgcct	gggcccattgt	tgtctctctg	cttctctcca	agtacgccat	cgtctatctg	1860
tttgccattg	gggctgcctt	tgccatgcaa	gtcggccaga	tgattcttgc	ccatctcgtg	1920
cttgctccct	tccccactg	gaacgtgctg	ctcttcttcc	cctttgtggg	actggcagtg	1980
cactacattg	caccctgtgt	tggctgggac	gccgatatcg	tgtcgggttaa	cactctcttc	2040
acctgttttg	gcgccaccct	ctccatttac	gccttctttg	tgcttgagat	catcgacgag	2100
atcaccaact	acctcgatat	ctgggtgtctg	cgaatcaagt	accctcagga	gaagaagacc	2160
gaataagcgg	cgcattggag	cgtgtgttct	gagtcgatgt	tttctatgga	gttgtgagtg	2220
ttagtagaca	tgatgggttt	atatatgatg	aatgaataga	tgtgattttg	atttgcacga	2280
tggaattgag	aactttgtaa	acgtacatgg	gaatgtatga	atgtgggggt	tttgtgactg	2340
gataactgac	ggtcagtgga	cgccgttgtt	caaatatcca	agagatgcga	gaaactttgg	2400
gtcaagtga	catgtcctct	ctgttcaagt	aaaccatcaa	ctatgggtag	tatatattagt	2460
aaggacaaga	gttgagattc	tttggagtcc	tagaaacgta	ttttcgcgtt	ccaagatcaa	2520
attagtagag	taatacgggc	acgggaatcc	attcatagtc	tcaatcctgc	aggtgagtta	2580
attaatcgag	cttggcgtaa	tcatggctcat	agctgtttcc	tgtgtgaaat	tgttatccgc	2640
tcacaattcc	acacaacgta	cgatagttag	tagacaacaa	tcagaacatc	tccctcctta	2700
tataatcaca	caggccagaa	cgcgctaaac	taaagcgctt	tggacactat	gttacattgg	2760
cattgattga	actgaaacca	cagtctccct	cgcctgaatc	gagcaatgga	tgttgtcgga	2820
agtcaacttc	actagaagag	cggttctatg	ccttgtcaag	atcatatcat	aaactcactc	2880
tgtattacce	catctataga	acacttgtaa	tgaatgggcg	gaaacattcc	gctatatgca	2940
cctttccaca	ctaattgcaa	gatgtgcata	ttcaacgggt	agtaagactg	gttccgactt	3000
ccgttgcatg	gagagcaatg	acctcgataa	tgcgaacatc	ccccacatat	acactcttac	3060
acaggccaat	ataatctgtg	catttactaa	atattttaagt	ctatgcacct	gcttgatgaa	3120
aagcggcacg	gatggatatca	tctagtcttc	gccaatccaa	gaaccaactg	tgttggcagt	3180
ggtgtagccc	atggcacaca	gaccaaagat	gaaaatacag	acatcggcgg	ttcgagccgt	3240
ggtgcctcga	gcaacaccct	tgtaatgcaa	aagaggaggg	taaatgtaca	ccagaggcac	3300
acatgcaaac	gatccggtga	gagcgacgaa	ccgatcgaga	tcgtcggcac	ctcccccattgc	3360

aacaaaggcg	gtgacaaaca	caaggaagaa	ccggaaaatg	ttcttctgcc	acttgatggt	3420
agagtgtgac	ttgcctgata	gggtgaagag	accattctcg	atgattcgga	tggcgcgcca	3480
gctgcattaa	tgaatcggcc	aacgcgcggg	gagaggcggt	ttgcgtattg	ggcgcctctc	3540
cgttctctcg	ctcactgact	cgtcgcgctc	ggtcgttcgg	ctgcggcgag	cggatatcagc	3600
tactcaaaag	gcggtaatac	ggttatecac	agaatcaggg	gataacgcag	gaaagaacat	3660
gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	gccgcgttgc	tggcgttttt	3720
ccataggctc	cgcctccctg	acgagcatca	caaaaatcga	cgtcaagtc	agagggtggcg	3780
aaacccgaca	ggactataaa	gataccaggc	gtttcccccct	ggaagctccc	tcgtgcgctc	3840
tcctgttccg	accctgccgc	ttaccggata	cctgtccgcc	tttctccctt	cgggaagcgt	3900
ggcgttttct	catagctcac	gctgtaggta	tctcagttcg	gtgtaggtcg	ttcgtctcaa	3960
gctgggctgt	gtgcacgaac	ccccgttca	gcccgaccgc	tgcgccttat	ccggtaacta	4020
tcgtcttgag	tccaacccgg	taagacacga	cttatcgcca	ctggcagcag	ccactggtaa	4080
caggattagc	agagcgaggt	atgtaggcgg	tgctacagag	ttcttgaagt	ggtggcctaa	4140
ctacggctac	actagaagaa	cagtatttgg	tatctgcgct	ctgctgaagc	cagttacctt	4200
cggaaaaaga	gttggtagct	cttgatccgg	caacaaacc	accgctggta	gcggtggttt	4260
ttttgtttgc	aagcagcaga	ttacgcgcag	aaaaaaagga	tctcaagaag	atcctttgat	4320
cttttctacg	gggtctgacg	ctcagttgaa	cgaaaactca	cgtaaaggga	ttttggtcac	4380
gagattatca	aaaaggatct	tcacctagat	ccttttaaat	taaaaatgaa	gttttaaatc	4440
aatctaaagt	atatatgagt	aaacttggtc	tgacagttac	caatgcttaa	tcagtgaggc	4500
acctatctca	gcgatctgtc	tatttcgttc	atccatagtt	gcctgactcc	ccgtcgtgta	4560
gataactacg	atacgggagg	gcttaccatc	tggccccagt	gctgcaatga	taccgcgaga	4620
cccacgtca	ccggctccag	atttatcagc	aataaaccag	ccagccggaa	gggccgagcg	4680
cagaagtggg	cctgcaactt	tatccgcctc	catccagtct	attaattggt	gccgggaagc	4740
tagagtaagt	agttcgccag	ttaatagttt	gcgcaacggt	gttgccattg	ctacaggcat	4800
cgtgggtgtca	cgctcgtcgt	ttgggtatggc	ttcattcagc	tccggttccc	aacgatcaag	4860
gcgagttaca	tgatcccca	tgttgtgcaa	aaaagcggtt	agctccttcg	gtcctccgat	4920
cgttggtcaga	agtaagttag	ccgcagtggt	atcactcatg	gttatggcag	cactgcataa	4980
ttctcttact	gtcatgccat	ccgtaagatg	cttttctgtg	actggtgagt	actcaaccaa	5040
gtcattctga	gaatagtgtg	tgcggcgacc	gagttgctct	tgcccggcgt	caatacggga	5100
taataccgcg	ccacatagca	gaactttaaa	agtgtctatc	attggaaaac	gttcttcggg	5160
gcgaaaactc	tcaaggatct	taccgctgtt	gagatccagt	tcgatgtaac	ccactcgtgc	5220
accaactga	tcttcagcat	cttttacttt	caccagcggt	tctgggtgag	caaaaacagg	5280
aaggcaaaat	gccgcaaaaa	agggaataag	ggcgacacgg	aaatgttgaa	tactcatact	5340
cttccttttt	caatattatt	gaagcattta	tcagggttat	tgtctcatga	gcggatacat	5400
atttgaatgt	atttagaaaa	ataaacaaat	aggggttccg	cgcacatttc	cccgaagaat	5460
gccacctgat	gcggtgtgaa	ataccgcaca	gatgcgtaag	gagaaaaatac	cgcacagga	5520
aattgtaage	gttaatatat	tgtaaaaatt	cgcgttaaat	ttttgttaaa	tcagctcatt	5580
ttttaaccaa	taggccgaaa	tcggcaaaaat	cccttataaa	tcaaaaagaat	agaccgagat	5640
agggttgagt	gttggttccag	tttggaaaca	gagtcacta	ttaaaagaacg	tggactccaa	5700

cg tcaaaggg	cg aaaaaccg	tctatcaggg	cgatggccca	ctacgtgaac	catcacccta	5760
atcaagtttt	ttggggtcga	ggtgccgtaa	agcactaaat	cggaacccta	aaggagagccc	5820
ccgatttaga	gcttgacggg	gaaagccggc	gaacgtggcg	agaaaggaag	ggaagaaagc	5880
gaaaggagcg	ggcgctaggg	cgctggcaag	tgtagcggtc	acgctgcgcg	taaccaccac	5940
acccgccgcg	cttaatgcgc	cgctacaggg	cgcgtccatt	cgccattcag	gctgcgcaac	6000
tgttgggaag	ggcgatcggt	gcgggcctct	tcgctattac	gccagctggc	gaaaggggga	6060
tgtgctgcaa	ggcgattaag	ttgggtaacg	ccagggtttt	cccagtcacg	acgttgtaaa	6120
acgacggcca	gtgaattgta	atacgactca	ctatagggcg	aattgggccc	gacgtcgcgt	6180
gcttagaagt	gaggattaca	agaagcctct	ggatatcaat	gatgaacgta	ctcagcggct	6240
ggtcaagcat	ttcgaccgtc	gaatcgacga	ggtgttcacc	tttgacaagc	gagggttccc	6300
aattgatcac	gttctcgagt	tgttcaaate	ttctctcaac	atctctctgc	atgaactatc	6360
tctgttgacg	aacgtgtcac	ccactgttcc	tcgaacgccc	ttctccgagt	ttggtctgaa	6420
catcttcgat	ctcaaaactga	ccccgcaggt	gatcaatagt	gccatgccac	tgccgatgcg	6480
gtgcgaacat	ccctggaggg	attctcggag	ctctacacaa	tgcagattct	gtcgtcgagt	6540
actctctacc	ttgctcgaat	gacttattgt	gctactactg	cactcatgct	tcgatcatgt	6600
gccctactgc	accccaaatt	tggatgatctg	attgagacag	agtaccctct	tcagctgatt	6660
cagaagatca	tcagcaacat	gaatgatgtg	gttgaccagg	caggctgttg	tagtcacgtc	6720
cttcacttca	agttcattct	tcatctgctt	ctgttttact	ttgacaggca	aatgaagaca	6780
tggtagcact	tgatggaggc	caagaacgcc	atttcacccc	gagacaccga	agtgccctgaa	6840
atcctggctg	ccccattga	taacatcgga	aactacggta	ttccggaaag	tgtatataga	6900
acctttcccc	agcttgtgtc	tgtggatatg	gatgggtgta	tccccctaat	taactcacct	6960
gcaggattga	gactatgaat	ggattcccgt	gcccgtatta	ctctactaat	ttgatcttgg	7020
aacgcgaaaa	tacgttttcta	ggactccaaa	gaatctcaac	tcttgtcctt	actaaatata	7080
ctacccatag	ttgatggttt	acttgaacag	agaggacatg	ttcacttgac	ccaaagtttc	7140
tcgcatctct	tggatatttg	aacaacggcg	tccactgacc	gtcagttatc	cagtcacaaa	7200
acccccacat	tcatacatte	ccatgtacgt	ttacaaagtt	ctcaattcca	tcgtgcaaat	7260
caaaatcaca	tctattcatt	catcatatat	aaacccatca	tgtctactaa	cactcacaaac	7320
tccatagaaa	acatcgactc	agaacacacg	ctccatgcgg	ccgcttaggc	aacgggcttg	7380
atgacagcgg	gaggagtgcc	cacattgttt	cggtttcgaa	agaacaggac	acccttgcca	7440
gctccctcgg	caccagcgga	gggttcaacc	cactggcaca	ttcgtgcaga	tcggtacatg	7500
gctcgaatga	atcctcgagg	accgtcctgg	acatcagctc	gatagtgcct	gcccattgata	7560
ggttttagtg	cctcggtagc	ttcgtccgca	ttgtagaagg	gaatggaaga	gacgtagtga	7620
tgcaggacgt	gagtctcgat	aatgccgtgg	agcagatgac	gtccaatgaa	gcccattctct	7680
cggtcgatgg	ttgcagcggc	acctcgcaca	aagtccact	cgctcgttgg	gtagtgggga	7740
agagtaggat	ctgtgtgctg	cagaaaggta	atggcgacga	gccagtgggt	aaccacacaag	7800
tagggaacga	agtaccagat	ggccatgttg	tagaatccga	acttctgaac	gagaaaagtac	7860
agagcggtag	ccataagacc	aatgccaatg	tcggagagca	cgatgagctt	ggcgtcgtctg	7920
ttctcgtaca	gaggagatcg	gggatcgaaa	tggttaaactc	caccgccaag	accgttgtgc	7980
tttcccttgc	ctcgaccctc	tcgctgccgc	tcatggtagt	tgtgtccagt	aacgttggta	8040

atgagatagt	tgggccaacc	gaccagttgc	tgaagcacia	gcatgagcag	ggtgaaagca	8100
ggagtttctt	cggtaagatg	ggcgagttcg	tgggtcatct	tgccgagtcg	agtagcttgc	8160
tgtctctggg	ttcgaggaac	gaagaccatg	tctcgtctca	tgtttccagt	ggccttgtga	8220
tgtttccggt	gggagatttg	ccagctgaag	tagggaacaa	gcagggaaga	gtgaagcacc	8280
cagccagtaa	tgtcgttgat	gattcgggaa	tcggagaaag	caccatgtcc	acactcgtgg	8340
gcaatgaccc	acagtccagt	accgaagagt	ccctgaagaa	cgggtgtacac	agcccacaga	8400
ccggctcgag	caggagtggg	gggaatgtac	tcgggtgtca	caaagttgta	ccagatgctg	8460
aaagtggtag	tcaggaggac	aatgtctcga	agaatgtagc	cgtatccctt	gagagcagat	8520
cgcttgaagc	agtgccttggg	aatagcgttg	tagatgtcct	tgatggtgaa	gtcgggaact	8580
tcgaactggt	tgccgtaggt	atccagcatg	acaccgtact	cggacttggg	cttggaatg	8640
tccacctcgg	acatggaaga	cagcgatgta	gaggaggccg	agtgtctggg	agaatcggag	8700
ggagagacgg	cagcagactc	cgagtcggtc	acagtgggtg	aagtgcgggt	tcgtcggagg	8760
gcagggttct	gcttgggcag	agccgaggtg	gaggccatgg	ccattgctgt	agatatgtct	8820
tgtgtgtaag	ggggttgggg	tggttgtttg	tgttcttgac	ttttgtgtta	gcaagggaag	8880
acgggcaaaa	aagtgagtgt	ggttgggagg	gagagacgag	ccttatatat	aatgcttgtt	8940
tgtgtttgtg	caagtggacg	ccgaaacggg	caggagccaa	actaaacaag	gcagacaatg	9000
cgagcttaat	tggattgcct	gatgggcagg	ggttagggct	cgatcaatgg	gggtgcgaag	9060
tgacaaaatt	gggaattagg	ttcgcaagca	aggctgacaa	gactttggcc	caaacatttg	9120
tacgcggtgg	acaacaggag	ccaccatcgc	tctgtcacgg	gctagccggt	cgtgcgtcct	9180
gtcaggctcc	acctaggctc	catgccactc	catacaatcc	cactagtgtg	ccgctaggcc	9240
gcttttagct	cccatctaag	acccccccaa	aacctccact	gtacagtgtg	ctgtactgtg	9300
tggcgatcaa	gggcaaggga	aaaaaggcgc	aaacatgcac	gcatggaatg	acgtaggtaa	9360
ggcgttacta	gactgaaaag	tggcacattt	cggcgtgcc	aagggtccta	ggtgcgtttc	9420
gcgagctggg	cgccaggcca	agccgctcca	aaacgcctct	ccgactccct	ccagcggcct	9480
ccatataccc	atccctctcc	acagcaatgt	tgttaagcct	tgcaaacgaa	aaaatagaaa	9540
ggctaataag	cttccaatat	tgtggtgtac	gctgcataac	gcaacaatga	gcgccaacaa	9600
acacacacac	acagcacaca	gcagcattaa	ccacgatgaa	cagcatgaat	tcctttacct	9660
gcaggataac	ttcgtataat	gtatgtctata	cgaagttagt	atctctctct	tgagcttttc	9720
cataacaagt	tcttctgcct	ccaggaagtc	catgggtggg	ttgatcatgg	ttttggtgta	9780
gtggtagtgc	agtgggtggt	ttgtgactgg	ggatgtagtt	gagaataagt	catacacaa	9840
tcagctttct	tcgagcctca	tataagtata	agtagttcaa	cgtattagca	ctgtaccacg	9900
catctccgta	tcgagaaaca	caacaacatg	ccccattgga	cagatcatgc	ggatacacag	9960
gttgtgcagt	atcatacata	ctcgatcaga	caggctcgtc	gaccatcata	caagctgaac	10020
aagcgtccca	tacttgacag	ctctctatat	acacagttaa	attacatatc	catagtctaa	10080
cctctaacag	ttaatcttct	ggtaagcctc	ccagccagcc	ttctgggtatc	gcttggcctc	10140
ctcaatagga	tctcggttct	ggccgtacag	acctcgcccg	acaattatga	tatccgttcc	10200
ggtagacatg	acatcctcaa	cagttcggta	ctgctgtccg	agagcgtctc	ccttgtcgtc	10260
aagaccaccc	ccgggggtca	gaataagcca	gtcctcagag	tcgcccttag	gtcggttctg	10320
ggcaatgaag	ccaaccacaa	actcgggggtc	ggatcgggca	agctcaatgg	tctgcttgga	10380

gtactcgcca	gtggccagag	agcccttgca	agacagctcg	gccagcatga	gcagacctct	10440
ggccagcttc	tcgttgggag	aggggactag	gaactccttg	tactgggagt	tctcgtagtc	10500
agagacgtcc	tccttcttct	gttcagagac	agtttcctcg	gcaccagctc	gcaggccagc	10560
aatgattccg	gttcegggtg	caccgtgggc	gttggtgata	tcggaccact	cggcgattcg	10620
gtgacaccgg	tactggtgct	tgacagtgtt	gccaatatct	gcgaactttc	tgtcctcgaa	10680
caggaagaaa	ccgtgcttaa	gagcaagttc	cttgaggggg	agcacagtgc	cggcgtaggt	10740
gaagtcgtca	atgatgtcga	tatgggtttt	gatcatgcac	acataagggtc	cgaccttate	10800
ggcaagctca	atgagctcct	tggtggtggt	aacatccaga	gaagcacaca	ggttggtttt	10860
cttggtgcc	acgagcttga	gcactcgagc	ggcaaaggcg	gacttgtgga	cgtagctcg	10920
agcttcgtag	gagggcattt	tggtggtgaa	gaggagactg	aaataaattt	agtctgcaga	10980
actttttatc	ggaaccttat	ctggggcagt	gaagtatatg	ttatggtaat	agttacgagt	11040
tagttgaact	tatagataga	ctggactata	cggctatcgg	tccaaattag	aaagaacgtc	11100
aatggctctc	tgggcgtcgc	ctttgccgac	aaaaatgtga	tcatgatgaa	agccagcaat	11160
gacgttgcag	ctgatattgt	tgtcggccaa	ccgcgccgaa	aacgcagctg	tcagaccac	11220
agcctccaac	gaagaatgta	tcgtcaaagt	gatccaagca	cactcatagt	tggagtcgta	11280
ctccaaaggc	ggcaatgacg	agtcagacag	atactcgctg	acgcgataac	ttcgtataat	11340
gtatgctata	cgaagttatc	gtacgatagt	tagtagacaa	caatcgatcg	aggaagagga	11400
caagcggctg	cttcttaagt	ttgtgacatc	agtatccaag	gcaccattgc	aaggattcaa	11460
ggctttgaac	ccgtcatttg	ccattcgtaa	cgttggtaga	caggttgac	ggttccttac	11520
ggcctccacc	tgtgtcaatc	ttctcaagct	gcctgactat	caggacattg	atcaacttcg	11580
gaagaaactt	ttgtatgcca	ttcgatcaca	tgtctggttc	gatttgtctt	agaggaacgc	11640
atatacagta	atcatagaga	ataaacgata	ttcattttatt	aaagtagata	gttgaggtag	11700
aagttgtaaa	gagtgataaa	tagcggccgc	tcactgaatc	tttttggtc	ccttgtgctt	11760
tcggacgatg	taggtctgca	cgtagaagtt	gaggaacaga	cacaggacag	taccaacgta	11820
gaagtagttg	aaaaaccagc	caaacattct	cattccatct	tgtcggtagc	agggaatgtt	11880
ccggtacttc	cagacgatgt	agaagccaac	gttgaactga	atgatctgca	tagaagtaat	11940
cagggacttg	ggcataggga	acttgagctt	gatcagtcgg	gtccaatagt	agccgtacat	12000
gatccagtga	atgaagccgt	tgagcagcac	aaagatccaa	acggcttcgt	ttcggtagtt	12060
gtagaacagc	cacatgtcca	taggagctcc	gagatggtga	aagaactgca	accaggtcag	12120
aggcttgccc	atgaggggca	gatagaagga	gtcaatgtac	tcgaggaact	tgtgaggta	12180
gaacagctga	gtggtgattc	ggaagacatt	gttgtcgaaa	gccttctcgc	agttgtcgga	12240
catgacacca	atggtgtaca	tggcgtaggc	catagagagg	aaggagccca	gcgagtagat	12300
ggacatgagc	aggttgtagt	tggtgaacac	aaacttcatt	cgagactgac	ccttggtgcc	12360
gagaggacca	agggtgaact	tcaggatgac	gaaggcgatg	gagaggtaga	gcacctcgca	12420
gtgcgaggca	tcagaccaga	gctgagcata	gtcgaccttg	ggaagaacct	cctggccaat	12480
ggagacgatt	tcgttcacga	cctccatggt	tgtgaattag	ggtggtgaga	atggttggtt	12540
gtagggaaga	atcaaaggcc	ggtctcggga	tccgtgggta	tatatatata	tatatatata	12600
tacgatecct	cgttacctcc	ctgttctcaa	aactgtgggt	tttcgttttt	cgttttttgc	12660
tttttttgat	tttttttaggg	ccaactaagc	ttccagattt	cgctaataac	ccttgtacta	12720

attacaagaa	aggaagaagc	tgattagagt	tgggcttttt	atgcaactgt	gctactcctt	12780
atctctgata	tgaaagtgtg	gacceaatca	catcatgtca	tttagagttg	gtaatactgg	12840
gaggatagat	aaggcacgaa	aacgagccat	agcagacatg	ctgggtgtag	ccaagcagaa	12900
gaaagtagat	gggagccaat	tgacgagcga	gggagctacg	ccaatccgac	atacgacacg	12960
ctgagatcgt	cttggecggg	gggtacctac	agatgtccaa	gggtaagtgc	ttgactgtaa	13020
ttgtatgtct	gaggacaaat	atgtagtcag	ccgtataaag	tcataccagg	caccagtgcc	13080
atcatcgaac	cactaactct	ctatgataca	tgccctccgg	attattgtac	catgcgtcgc	13140
tttgtttacat	acgtatcttg	cctttttctc	tcagaaactc	cagacttttg	ctattggctg	13200
agataagccc	ggaccatagt	gagtctttca	cactctgttt	aaacaccact	aaaaccccac	13260
aaaatatatc	ttaccgaata	tacagatcta	ctatagagga	acaattgccc	cggagaagac	13320
ggccaggccg	cctagatgac	aaattcaaca	actcacagct	gactttctgc	cattgccact	13380
aggggggggc	cTTTTtata	ggccaagcca	agctctccac	gtcggttggg	ctgcacccaa	13440
caataaatgg	gtagggttgc	accaacaaag	ggatgggatg	gggggtagaa	gatacgagga	13500
taacggggct	caatggcaca	aataagaacg	aatactgcca	ttaagactcg	tgatccagcg	13560
actgacacca	ttgcatcacc	taagggcctc	aaaactacct	cggaaactgt	gcgctgatct	13620
ggacaccaca	gaggttccga	gcactttagg	ttgcacaaaa	tgtcccacca	ggtgcaggca	13680
gaaaacgctg	gaacagcgtg	tacagtttgt	cTTaacaAAA	agtgaggggc	ctgaggtcga	13740
gcagggttgt	gtgactttgt	atagccttta	gagctgcgaa	agcgcgtatg	gattttggctc	13800
atcaggccag	attgagggtc	tgtggacaca	tgtcatgtta	gtgtacttca	atcgccccct	13860
ggatatagcc	ccgacaatag	gccgtggcct	catttttttg	ccttccgcac	atttccattg	13920
ctcggtagcc	acaccttgct	tctcctgcac	ttgccaacct	taatactggg	ttacattgac	13980
caacatctta	caagcggggg	gcttgtctag	ggtatatata	aacagtggct	ctcccaatcg	14040
gttgccagtc	tcttttttcc	tttctttccc	cacagattcg	aaatctaaac	tacacatcac	14100
acaatgcctg	ttactgacgt	ccttaagcga	aagtcgggtg	tcacgtcggg	cgacgatgtc	14160
cgagccgtga	gtatccacga	caagatcagt	gtcgagacga	cgcgttttgt	gtaatgacac	14220
aatccgaaag	tcgctagcaa	cacacactct	ctacacaaac	taaccagct	ctccatgggt	14280
aaggcttctc	gacaggctct	gccccctgtc	atcgacggaa	aggtgtacga	cgtctccgct	14340
tgggtgaact	tccaccctgg	tggagctgaa	atcattgaga	actaccaggg	acgagatgct	14400
actgacgcct	tcattggttat	gcactctcag	gaagccttcg	acaagctcaa	gcgaatgccc	14460
aagatcaacc	aggcttccga	gctgcctccc	caggctgccg	tcaacgaagc	tcaggaggat	14520
ttccgaaagc	tccgagaaga	gctgategcc	actggcatgt	ttgacgcctc	tcccctctgg	14580
tactcgtaca	agatcttgac	caccctgggt	cttggcgtgc	ttgccttctt	catgctggtc	14640
cagtaccacc	tgtacttcat	tgggtgtctc	gtgtcgggta	tgcactacca	gcaaattggga	14700
tggctgtctc	atgacatctg	ccaccaccag	accttcaaga	accgaaactg	gaataacgtc	14760
ctgggtcttg	tcttttgcaa	cggactccag	ggcttctccg	tgacctgggtg	gaaggacaga	14820
cacaacgecc	atcattctgc	taccaacgtt	cagggtcacg	atcccgcacat	tgataacctg	14880
cctctgtctg	cctgggtccga	ggacgatgtc	actcgagctt	ctcccatctc	ccgaaaagctc	14940
attcagttcc	aacagtacta	tttcttggtc	atctgtattc	tcctgcgatt	catctgggtg	15000
ttccagtctg	tgtgaccgt	tcgaccctc	aaggaccgag	acaaccagtt	ctaccgatct	15060

cagtacaaga aagaggccat tggactcgct ctgcactgga ctctcaagac cctgttccac 15120
 ctctttcttta tgccctccat cctgacctcg atgctgggtgt tctttgtttc cgagctcgtc 15180
 ggtggcttcg gaattgccat cgtggctctc atgaaccact accctctgga gaagatcggt 15240
 gattccgtct gggacggaca tggcttctct gtgggtcaga tccatgagac catgaacatt 15300
 cgacgaggca tcattactga ctggttcttt ggaggcctga actaccagat cgagcaccat 15360
 ctcttgccca cctgacctcg acacaacctc actgccgttt cctaccaggt ggaacagctg 15420
 tgccagaagc acaacctccc ctaccgaaac cctctgcccc atgaaggctct cgtcatcctg 15480
 ctccgatacc tgctccagtt cgctcgaatg gccgagaagc agcccgggtgc caaggctcag 15540
 taagcggccg catgagaaga taaatatata aatacattga gatattaaat gcgctagatt 15600
 agagagcctc atactgctcg gagagaagcc aagacgagta ctcaaagggg attacaccat 15660
 ccatatccac agacacaagc tggggaaagg ttctatatac actttccgga ataccgtagt 15720
 ttccgatggt atcaatgggg gcagccagga ttccaggcac ttcgggtgtct cgggggtgaaa 15780
 tggcgttctt ggcctccatc aagtcgtacc atgtcttcat ttgcctgtca aagtaaaaca 15840
 gaagcagatg aagaatgaac ttgaagtga ggaattt 15877

<210>82

<211>1185

<212>DNA

<213> 解脂耶氏酵母

<220>

<221>CDS

<222>(1).. (1185)

<223> 二酰基甘油胆碱磷酸转移酶 (Y1CPT1)

<300>

<302> 高二十碳五烯酸生产菌株

<310>W0 2006/052870

<311>2005-11-03

<312>2006-05-18

<313>(1).. (1185)

<300>

<302> 高二十碳五烯酸生产菌株

<310>US 2006-0115881-A1

<311>2005-11-02

<312>2006-06-01

<313>(1).. (1185)

<400>82

atg ggc gta ttc att aaa cag gag cag ctt ccg gct ctc aag aag tac 48
 Met Gly Val Phe Ile Lys Gln Glu Gln Leu Pro Ala Leu Lys Lys Tyr
 1 5 10 15
 aag tac tcc gcc gag gat cac tcg ttc atc tcc aac aac att ctg cgc 96

Lys	Tyr	Ser	Ala	Glu	Asp	His	Ser	Phe	Ile	Ser	Asn	Asn	Ile	Leu	Arg	
20				25				30								
ccc	ttc	tgg	cga	cag	ttt	gtc	aaa	atc	ttc	cct	ctg	tgg	atg	gcc	ccc	144
Pro	Phe	Trp	Arg	Gln	Phe	Val	Lys	Ile	Phe	Pro	Leu	Trp	Met	Ala	Pro	
35				40				45								
aac	atg	gtg	act	ctg	ttg	ggc	ttc	ttc	ttt	gtc	att	gtg	aac	ttc	atc	192
Asn	Met	Val	Thr	Leu	Leu	Gly	Phe	Phe	Phe	Val	Ile	Val	Asn	Phe	Ile	
50				55				60								
acc	atg	ctc	att	gtt	gat	ccc	acc	cac	gac	cgc	gag	cct	ccc	aga	tgg	240
Thr	Met	Leu	Ile	Val	Asp	Pro	Thr	His	Asp	Arg	Glu	Pro	Pro	Arg	Trp	
65				70				75				80				
gtc	tac	ctc	acc	tac	gct	ctg	ggt	ctg	ttc	ctt	tac	cag	aca	ttt	gat	288
Val	Tyr	Leu	Thr	Tyr	Ala	Leu	Gly	Leu	Phe	Leu	Tyr	Gln	Thr	Phe	Asp	
85				90				95								
gcc	tgt	gac	gga	tcc	cat	gcc	cga	cga	act	ggc	cag	agt	gga	ccc	ctt	336
Ala	Cys	Asp	Gly	Ser	His	Ala	Arg	Arg	Thr	Gly	Gln	Ser	Gly	Pro	Leu	
100				105				110								
gga	gag	ctg	ttt	gac	cac	tgt	gtc	gac	gcc	atg	aat	acc	tct	ctg	att	384
Gly	Glu	Leu	Phe	Asp	His	Cys	Val	Asp	Ala	Met	Asn	Thr	Ser	Leu	Ile	
115				120				125								
ctc	acg	gtg	gtg	gtg	tcc	acc	acc	cat	atg	gga	tat	aac	atg	aag	ctg	432
Leu	Thr	Val	Val	Val	Ser	Thr	Thr	His	Met	Gly	Tyr	Asn	Met	Lys	Leu	
130				135				140								
ctg	att	gtg	cag	att	gcc	gct	ctc	gga	aac	ttc	tac	ctg	tcg	acc	tgg	480
Leu	Ile	Val	Gln	Ile	Ala	Ala	Leu	Gly	Asn	Phe	Tyr	Leu	Ser	Thr	Trp	
145				150				155				160				
gag	acc	tac	cat	acc	gga	act	ctg	tac	ctt	tct	ggc	ttc	tct	ggt	cct	528
Glu	Thr	Tyr	His	Thr	Gly	Thr	Leu	Tyr	Leu	Ser	Gly	Phe	Ser	Gly	Pro	
165				170				175								
gtt	gaa	ggt	atc	ttg	att	ctg	gtg	gct	ctt	ttc	gtc	ctc	acc	ttc	ttc	576
Val	Glu	Gly	Ile	Leu	Ile	Leu	Val	Ala	Leu	Phe	Val	Leu	Thr	Phe	Phe	
180				185				190								
act	ggt	ccc	aac	gtg	tac	gct	ctg	acc	gtc	tac	gag	gct	ctt	ccc	gaa	624
Thr	Gly	Pro	Asn	Val	Tyr	Ala	Leu	Thr	Val	Tyr	Glu	Ala	Leu	Pro	Glu	
195				200				205								
tcc	atc	act	tcg	ctg	ctg	cct	gcc	agc	ttc	ctg	gac	gtc	acc	atc	acc	672
Ser	Ile	Thr	Ser	Leu	Leu	Pro	Ala	Ser	Phe	Leu	Asp	Val	Thr	Ile	Thr	
210				215				220								
cag	atc	tac	att	gga	ttc	gga	gtg	ctg	ggc	atg	gtg	ttc	aac	atc	tac	720

Gln	Ile	Tyr	Ile	Gly	Phe	Gly	Val	Leu	Gly	Met	Val	Phe	Asn	Ile	Tyr	
225					230				235						240	
ggc	gcc	tgc	gga	aac	gtg	atc	aag	tac	tac	aac	aac	aag	ggc	aag	agc	768
Gly	Ala	Cys	Gly	Asn	Val	Ile	Lys	Tyr	Tyr	Asn	Asn	Lys	Gly	Lys	Ser	
				245					250						255	
gct	ctc	ccc	gcc	att	ctc	gga	atc	gcc	ccc	ttt	ggc	atc	ttc	tac	gtc	816
Ala	Leu	Pro	Ala	Ile	Leu	Gly	Ile	Ala	Pro	Phe	Gly	Ile	Phe	Tyr	Val	
			260					265					270			
ggc	gtc	ttt	gcc	tgg	gcc	cat	gtt	gct	cct	ctg	ctt	ctc	tcc	aag	tac	864
Gly	Val	Phe	Ala	Trp	Ala	His	Val	Ala	Pro	Leu	Leu	Leu	Ser	Lys	Tyr	
			275				280						285			
gcc	atc	gtc	tat	ctg	ttt	gcc	att	ggg	gct	gcc	ttt	gcc	atg	caa	gtc	912
Ala	Ile	Val	Tyr	Leu	Phe	Ala	Ile	Gly	Ala	Ala	Phe	Ala	Met	Gln	Val	
	290					295				300						
ggc	cag	atg	att	ctt	gcc	cat	ctc	gtg	ctt	gct	ccc	ttc	ccc	cac	tgg	960
Gly	Gln	Met	Ile	Leu	Ala	His	Leu	Val	Leu	Ala	Pro	Phe	Pro	His	Trp	
305				310					315					320		
aac	gtg	ctg	ctc	ttc	ttc	ccc	ttt	gtg	gga	ctg	gca	gtg	cac	tac	att	1008
Asn	Val	Leu	Leu	Phe	Phe	Pro	Phe	Val	Gly	Leu	Ala	Val	His	Tyr	Ile	
				325				330					335			
gca	ccc	gtg	ttt	ggc	tgg	gac	gcc	gat	atc	gtg	tcg	gtt	aac	act	ctc	1056
Ala	Pro	Val	Phe	Gly	Trp	Asp	Ala	Asp	Ile	Val	Ser	Val	Asn	Thr	Leu	
			340					345					350			
ttc	acc	tgt	ttt	ggc	gcc	acc	ctc	tcc	att	tac	gcc	ttc	ttt	gtg	ctt	1104
Phe	Thr	Cys	Phe	Gly	Ala	Thr	Leu	Ser	Ile	Tyr	Ala	Phe	Phe	Val	Leu	
			355				360						365			
gag	atc	atc	gac	gag	atc	acc	aac	tac	ctc	gat	atc	tgg	tgt	ctg	cga	1152
Glu	Ile	Ile	Asp	Glu	Ile	Thr	Asn	Tyr	Leu	Asp	Ile	Trp	Cys	Leu	Arg	
	370					375				380						
atc	aag	tac	cct	cag	gag	aag	aag	act	gag	taa						1185
Ile	Lys	Tyr	Pro	Gln	Glu	Lys	Lys	Thr	Glu							
385				390												
<210>83																
<211>394																
<212>PRT																
<213> 解脂耶氏酵母																
<400>83																
Met	Gly	Val	Phe	Ile	Lys	Gln	Glu	Gln	Leu	Pro	Ala	Leu	Lys	Lys	Tyr	
1				5					10					15		

Lys Tyr Ser Ala Glu Asp His Ser Phe Ile Ser Asn Asn Ile Leu Arg		
20	25	30
Pro Phe Trp Arg Gln Phe Val Lys Ile Phe Pro Leu Trp Met Ala Pro		
35	40	45
Asn Met Val Thr Leu Leu Gly Phe Phe Phe Val Ile Val Asn Phe Ile		
50	55	60
Thr Met Leu Ile Val Asp Pro Thr His Asp Arg Glu Pro Pro Arg Trp		
65	70	75
Val Tyr Leu Thr Tyr Ala Leu Gly Leu Phe Leu Tyr Gln Thr Phe Asp		
85	90	95
Ala Cys Asp Gly Ser His Ala Arg Arg Thr Gly Gln Ser Gly Pro Leu		
100	105	110
Gly Glu Leu Phe Asp His Cys Val Asp Ala Met Asn Thr Ser Leu Ile		
115	120	125
Leu Thr Val Val Val Ser Thr Thr His Met Gly Tyr Asn Met Lys Leu		
130	135	140
Leu Ile Val Gln Ile Ala Ala Leu Gly Asn Phe Tyr Leu Ser Thr Trp		
145	150	155
Glu Thr Tyr His Thr Gly Thr Leu Tyr Leu Ser Gly Phe Ser Gly Pro		
165	170	175
Val Glu Gly Ile Leu Ile Leu Val Ala Leu Phe Val Leu Thr Phe Phe		
180	185	190
Thr Gly Pro Asn Val Tyr Ala Leu Thr Val Tyr Glu Ala Leu Pro Glu		
195	200	205
Ser Ile Thr Ser Leu Leu Pro Ala Ser Phe Leu Asp Val Thr Ile Thr		
210	215	220
Gln Ile Tyr Ile Gly Phe Gly Val Leu Gly Met Val Phe Asn Ile Tyr		
225	230	235
Gly Ala Cys Gly Asn Val Ile Lys Tyr Tyr Asn Asn Lys Gly Lys Ser		
245	250	255
Ala Leu Pro Ala Ile Leu Gly Ile Ala Pro Phe Gly Ile Phe Tyr Val		
260	265	270
Gly Val Phe Ala Trp Ala His Val Ala Pro Leu Leu Leu Ser Lys Tyr		
275	280	285
Ala Ile Val Tyr Leu Phe Ala Ile Gly Ala Ala Phe Ala Met Gln Val		
290	295	300
Gly Gln Met Ile Leu Ala His Leu Val Leu Ala Pro Phe Pro His Trp		
305	310	315
Asn Val Leu Leu Phe Phe Pro Phe Val Gly Leu Ala Val His Tyr Ile		

325	330	335
Ala Pro Val Phe Gly Trp Asp Ala Asp Ile Val Ser Val Asn Thr Leu		
340	345	350
Phe Thr Cys Phe Gly Ala Thr Leu Ser Ile Tyr Ala Phe Phe Val Leu		
355	360	365
Glu Ile Ile Asp Glu Ile Thr Asn Tyr Leu Asp Ile Trp Cys Leu Arg		
370	375	380
Ile Lys Tyr Pro Gln Glu Lys Lys Thr Glu		
385	390	

<210>84

<211>15812

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pZKL2-5U89Gc

<400>84

```

gtacgttatac atttgaacag tgaaaggcta cagtaacaga agcagttgta aacttcattc      60
cgttgattct gtactacagt accccactac gccgcttcg ctgacactgt tcaacccaaa      120
aactacatct gcgtgcgctg tgtaaggcta tcatcagata catactgtag attctgtaga      180
tgcgaaacctg cttgtatcat atacatcccc ctccccctga cctgcacaag caagcaatgt      240
gacattgata ttgctgctta tctagtgcg aggatgtgaa agccgagact caaacatttc      300
ttttactctc ttgttccctga ccagacctgg cggagattac gccagtatga ttcttgagg      360
tctgagacaa gcctggaaca gccaacattt atttttcgaa gcgagaaaca tgccacaccc      420
cggcacgttc agagatgcat atgatttggt tttcgagtaa cagtaccccc cccccccccc      480
ccaatgaaac cagtattact cacaccatcc tcattcaaag cgttacactg attacgcgcc      540
catcaacgac agcatgaggg gactgctgat ctgatctaata caaatgacta caaaaatcgc      600
aataatgaag agcaaacgac aaaaaagaaa cagggttaacc aatccccgctt caatgtctca      660
ccacaatcca gcaactgtttc tcattacctc ctccctctaa ttccagagtt gcatcagggt      720
ccttgatggc gcgccagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc      780
gtattgggcg ctcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggtgc      840
ggcgagcggt atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata      900
acgcaggaaa gaacatgtga gcaaaaaggcc agcaaaaaggc caggaaccgt aaaaaggccg      960
cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct      1020
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccctggaa      1080
gtccctcgt gcgctctcct gtccgaccc tgccgcttac cggatacctg tccgccttc      1140
tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt      1200
aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg      1260
ccttatecgg taactatcgt cttgagtcca acccggttaag acacgactta tcgccactgg      1320
cagcagccac tggtaacagg attagcagag cgaggatatgt aggcggtgct acagagttct      1380

```

tgaagtgggtg	gcctaactac	ggctacacta	gaagaacagt	atttgggtatc	tgcgctctgc	1440
tgaagccagt	taccttcgga	aaaagagttg	gtagctcttg	atccggcaaa	caaaccaccg	1500
ctggtagcgg	tggttttttt	gtttgcaagc	agcagattac	gcgcagaaaa	aaaggatctc	1560
aagaagatcc	tttgatcttt	tctacggggt	ctgacgctca	gtggaacgaa	aactcacgtt	1620
aagggatttt	ggtcattgaga	ttatcaaaaa	ggatcttcac	ctagatcctt	ttaaattaaa	1680
aatgaagttt	taaatcaatc	taaagtatat	atgagtaaac	ttggctctgac	agttaccaat	1740
gcttaatcag	tgaggcacct	atctcagcga	tctgtctatt	tcgttcatcc	atagttgcct	1800
gactccccgt	cgtgtagata	actacgatac	gggagggcct	accatctggc	cccagtgcgtg	1860
caatgatacc	gcgagaccca	cgtcaccggg	ctccagattt	atcagcaata	aaccagccag	1920
ccggaagggc	cgagcgcaga	agtggctcctg	caactttatc	cgcttccatc	cagtctatta	1980
attgttgccg	ggaagctaga	gtaagtagtt	cgccagttaa	tagtttgccg	aacgttggtg	2040
ccattgctac	aggcatcggtg	gtgtcacgct	cgtcgtttgg	tatggcttca	ttcagctccg	2100
gttcccaacg	atcaaggcga	gttacctgat	cccccatggt	gtgcaaaaaa	gcggttagct	2160
ccttcgggtc	tccgategtt	gtcagaagta	agttggccgc	agtgttatca	ctcatggtta	2220
tggcagcact	gcataattct	cttactgtca	tgccatccgt	aagatgcttt	tctgtgactg	2280
gtgagtactc	aaccaagtca	ttctgagaat	agtgtatgcg	gcgaccgagt	tgctcttgcc	2340
cggcgtcaat	acgggataat	accgcgccac	atagcagaac	tttaaaagtg	ctcatcattg	2400
gaaaacgttc	ttcggggcga	aaactctcaa	ggatcttacc	gctgttgaga	tccagttcga	2460
tgtaaccac	tcgtgcaccc	aactgatctt	cagcatcttt	tactttcacc	agcgtttctg	2520
ggtgagcaaa	aacaggaagg	caaatgccg	caaaaaaggg	aataagggcg	acacggaaat	2580
gttgaatact	catactcttc	ctttttcaat	attattgaag	catttatcag	ggttattgtc	2640
tcatgagcgg	atacatattt	gaatgtattt	agaaaaataa	acaaataggg	gttccgcgca	2700
catttccccg	aaaagtgcc	cctgatgcgg	tgtgaaatac	cgcacagatg	cgtaaggaga	2760
aaataccgca	tcaggaaatt	gtaagcgta	atattttgtt	aaaattcgcg	ttaaattttt	2820
gttaaatacag	ctcatttttt	aaccaatagg	ccgaaatcgg	caaaatccct	tataaatcaa	2880
aagaatagac	cgagataggg	ttgagtgttg	ttccagtttg	gaacaagagt	ccactattaa	2940
agaacgtgga	ctccaacgtc	aaagggcgaa	aaaccgtcta	tcagggcgat	ggcccactac	3000
gtgaaccatc	accctaatac	agtttttttg	ggtcgaggtg	ccgtaaagca	ctaaatcgga	3060
accctaaagg	gagcccccg	tttagagctt	gacggggaaa	gccggcgaa	gtggcgagaa	3120
aggaagggaa	gaaagcgaaa	ggagcgggcg	ctagggcgct	ggcaagtgt	gcggtcacgc	3180
tgcgcgtaac	caccacaccc	gccgcgctta	atgcgccgct	acagggcgcg	tccattcgcc	3240
attcaggctg	cgcaactgtt	gggaagggcg	atcggtgcgg	gcctcttcgc	tattacgcca	3300
gctggcgaaa	gggggatgtg	ctgcaaggcg	attaagtgtg	gtaacgccag	ggttttccca	3360
gtcacgacgt	tgtaaaacga	cggccagtga	attgtaatac	gactcactat	agggcgaaat	3420
gggcccgaac	tcgcatgctg	gtttcgattt	gtcttagagg	aacgcatata	cagtaatcat	3480
agagaataaa	cgatattcat	ttattaaagt	agatagtgtg	ggtagaagtt	gtaaagagtg	3540
ataaatagct	tagataccac	agacaccctc	ggtgacgaag	tactgcagat	ggtttccaat	3600
cacattgacc	tgctggagca	gagtgttacc	ggcagagcac	tgttttattg	tctggccctg	3660
gcacatgaca	acgttggaga	gaggagggtg	gatcaggggc	cagtcaataa	agacctcacc	3720

agagcagtgc	tggtaacctg	cccagaaggg	cacttgaggg	acgatatctc	ctcgggtgggt	3780
gattcggtag	agctttcggt	ctttggacac	cttggagaca	tcggggttct	cctggccaaa	3840
gaagagttta	tcgaccaggt	tagcaaagcc	agcgttaccg	acaatgggct	gaccaagagt	3900
aacaacgagg	ggatcgtagc	cgtaaacctt	gaggttgatt	ccgaacagaa	gggctgcagc	3960
tcctccgaga	gagtgaccgg	tgacagcaat	ctggtagtcg	ggatactgct	caatcacaga	4020
gtcgagcttg	gggcccgatc	gattgtaggt	gttggttag	gactggatga	agccattgtg	4080
gacaagacag	tcatacacaag	tagcagtaga	agagatgtta	gcagcaagat	caaagttaat	4140
taactcacct	gcaggattga	gactatgaat	ggattcccgt	gcccgtatta	ctctactaat	4200
ttgatcttgg	aacgcgaaaa	tacgtttcta	ggactccaaa	gaatctcaac	tcttgtcctt	4260
actaaatata	ctacccatag	ttgatggttt	acttgaacag	agaggacatg	ttcacttgac	4320
ccaaagtttc	tcgcatctct	tggatatttg	aacaacggcg	tccactgacc	gtcagttatc	4380
cagtcacaaa	acccccacat	tcatacattc	ccatgtacgt	ttacaaagtt	ctcaattcca	4440
tcgtgcaaat	caaaatcaca	tctattcatt	catcatatat	aaacccatca	tgtctactaa	4500
cactcacaac	tccatagaaa	acatcgactc	agaacacacg	ctccatgcgg	ccgcttagga	4560
atcctgagcg	tccttgacac	agtgaaccac	accgactttg	tgcatgtact	tgagggtgga	4620
aatgatgttg	cccacaatgg	tagggtagaa	gacgtaccga	actccgtgtc	gttcgcaaca	4680
ctctcggaca	gcttgctgca	cgaagggata	gtgccaagac	gacattcgag	gaaagagggtg	4740
atgctcgatc	tggaagttga	gaccgccagt	aaagaacatg	gcaatgggtc	caccgtaggt	4800
ggaagaggtc	tccacctgag	ctctgtacca	gtcgatctga	tcggcttcaa	cgctcttctc	4860
ggagctcttg	accttgcaat	tcttgcgggg	gattcgctcc	gagccatcga	agttgtgaga	4920
caagatgaaa	aagaaggtga	ggaaggcacc	ggtagcagtg	ggcaccagag	gaatgggtgat	4980
gagcagggag	gttccagtga	gataccaggg	caagaaggcg	gttcgaaaaga	tgaagaaagc	5040
tcgcataacg	aatgcaaggg	ttcggtaccg	tcgcagaaaag	ccgttctctc	gcatggctgt	5100
gacagactcg	ggaatgggtg	cggtgtgctg	cattcggaag	atgtagagag	ggtgttacac	5160
cagcgaaaacg	ccgtaggctc	caagcacgag	gtacatgtac	caggcctgga	atcgggtgaaa	5220
ccactttcga	gcagtgttgg	cagcagggta	gttggtggaac	acaaggaatg	gttctgcgga	5280
ctcggcatcc	aggctcgagac	catgctgatt	ggtgtaggtg	tgatgtcgca	tgatgtgaga	5340
ctgcagccag	atccatctgg	acgatecaat	gacgtcgatg	ccgtaggcaa	agagagcggt	5400
gaccaggggc	tttttgctga	tggcaccatg	agaggcatcg	tgctgaatgg	acaggccgat	5460
ctgcatgtgc	atgaatccag	tcaagagacc	ccacagcacc	attccggtag	tagcccagtg	5520
ccactcgcaa	aaggcgggtg	cagcaatgat	gccaacgggt	cgcagccaga	atccagggtg	5580
ggcataccag	ttccgacctt	tcatgacctc	tcgcatagtt	cgcttgacgt	cctgtgcaaa	5640
gggagagtcg	taggtgtaga	caatgtcctt	ggaggttcgg	tcgtgcttgc	ctcgcacgaa	5700
ctgttgaagc	agcttcgagt	tctcgggctt	gacgtaaggg	tgcatggagt	agaacagagg	5760
agaagcatcg	gaggcaccag	aagcgaggat	caagtcgcct	ccgggatgga	ccttggcaag	5820
accttccaga	tcgtagagaa	tgccgtcgat	ggcaaccagg	tcgggtcgct	cgagcagctg	5880
ctcggtagta	aggagagag	ccatggttgt	gaattaggg	ggtgagaatg	gttgggttga	5940
gggaagaatc	aaaggccggt	ctcgggatcc	gtgggtatat	atatatatat	atatatatat	6000
gaccttctgt	tacctccctg	ttctcaaaac	tgtgggtttt	cgtttttctg	tttttgcttt	6060

ttttgatttt	tttagggcca	actaagcttc	cagatttcgc	taatcacctt	tgtactaatt	6120
acaagaaagg	aagaagctga	ttagagttgg	gctttttatg	caactgtgct	actccttate	6180
tctgatatga	aagtgtagac	ccaatcacat	catgtcattt	agagttggta	atactgggag	6240
gatagataag	gcacgaaaac	gagccatagc	agacatgctg	ggtgtagcca	agcagaagaa	6300
agtagatggg	agccaattga	cgagcgaggg	agctacgcca	atccgacata	cgacacgctg	6360
agatcgtctt	ggccgggggg	tacctacaga	tgtccaaggg	taagtgcttg	actgtaattg	6420
tatgtctgag	gacaaatatg	tagtcagccg	tataaagtca	taccaggcac	cagtgccatc	6480
atcgaaccac	taactctcta	tgatacatgc	ctccggtatt	attgtaccat	gcgtcgcttt	6540
gttacatacg	tatcttgcct	ttttctctca	gaaactccag	aattctctct	cttgagcttt	6600
tccataacaa	gttcttctgc	ctccaggaag	tccatgggtg	gtttgatcat	ggttttggtg	6660
tagtggtagt	gcagtgggtg	tattgtgact	ggggatgtag	ttgagaataa	gtcatacaca	6720
agtcagcttt	cttcgagcct	catataagta	taagtagttc	aacgtattag	cactgtaccc	6780
agcatctccg	tatcgagaaa	cacaacaaca	tgccccattg	gacagatcat	gcggatacac	6840
aggttgtgca	gtatcataca	tactcgatca	gacaggctgt	ctgaccatca	tacaagctga	6900
acaagcgctc	catacttgca	cgctctctat	atacacagtt	aaattacata	tccatagtct	6960
aacctctaac	agttaatctt	ctggtaagcc	tcccagccag	ccttctggta	tcgcttggcc	7020
tcctcaatag	gatctcggtt	ctggccgtac	agacctcggc	cgacaattat	gatatccgtt	7080
ccggtagaca	tgacatcctc	aacagttcgg	tactgtctgc	cgagagcgtc	tcccttgtcg	7140
tcaagacca	ccccgggggt	cagaataagc	cagtcctcag	agtcgccctt	aggtcggttc	7200
tgggcaatga	agccaaccac	aaactcgggg	tcggatcggg	caagctcaat	ggtctgcttg	7260
gagtactcgc	cagtggccag	agagcccttg	caagacagct	cggccagcat	gagcagacct	7320
ctggccagct	tctcgttggg	agaggggact	aggaactcct	tgtactggga	gttctcgtag	7380
tcagagacgt	cctccttctt	ctgttcagag	acagtttctt	cggcaccagc	tcgcaggcca	7440
gcaatgattc	cggttccggg	tacaccgtgg	gcgttgggtg	tatcggaaca	ctcggcgatt	7500
cggtgacacc	ggtactgggtg	cttgacagtg	ttgccaatat	ctgcgaactt	tctgtcctcg	7560
aacaggaaga	aaccgtgctt	aagagcaagt	tccttgaggg	ggagcacagt	gccggcgtag	7620
gtgaagtcgt	caatgatgtc	gatatgggtt	ttgatcatgc	acacataagg	tccgacctta	7680
tcggcaagct	caatgagctc	cttgggtggg	gtaacatcca	gagaagcaca	caggttgggt	7740
ttcttggctg	ccacgagctt	gagcactcga	gcggcaaagg	cggacttgtg	gacgttagct	7800
cgagcttcgt	aggagggcat	tttgggtggg	aagaggagac	tgaaataaat	ttagtctgca	7860
gaacttttta	tcggaacctt	atctggggca	gtgaagtata	tgttatggta	atagttacga	7920
gttagttgaa	cttatagata	gactggacta	tacggctatc	ggtccaaatt	agaaagaacg	7980
tcaatggctc	tctgggcgtc	gcctttgccg	acaaaaatgt	gatcatgatg	aaagccagca	8040
atgacgttgc	agctgatatt	gttgctggcc	aaccgcgccg	aaaacgcagc	tgtcagaccc	8100
acagcctcca	acgaagaatg	tatcgtcaaa	gtgatccaag	cacactcata	gttggagtcg	8160
tactccaaag	gcggcaatga	cgagtcagac	agatactcgt	cgaccttttc	cttgggaacc	8220
accacgtca	gcccttctga	ctcacgtatt	gtagccaccg	acacaggcaa	cagtccgtgg	8280
atagcagaat	atgtcttgtc	ggtccatttc	tcaccaactt	taggcgtcaa	gtgaatgttg	8340
cagaagaagt	atgtgccttc	attgagaatc	ggtgttgctg	atttcaataa	agtcttgaga	8400

tcagtttggc	cagtcattgtt	gtggggggta	attggattga	gttatcgctt	acagtctgta	8460
caggtatact	cgctgcccac	tttatacttt	ttgattccgc	tgcacttgaa	gcaatgtcgt	8520
ttacccaaag	tgagaatgct	ccacagaaca	caccccaggg	tatggttgag	caaaaaataa	8580
acactccgat	acggggaate	gaaccccggg	ctccacgggt	ctcaagaagt	attcttcatg	8640
agagcgtatc	gategaggaa	gaggacaagc	ggctgcttct	taagtttgtg	acatcagtat	8700
ccaaggcacc	attgcaagga	ttcaaggctt	tgaacccgtc	atttgccatt	cgtaacgctg	8760
gtagacaggt	tgategggtc	cctacggcct	ccacctgtgt	caatcttctc	aagctgcctg	8820
actatcagga	cattgatcaa	cttcggaaga	aacttttgta	tgccattcga	tcacatgctg	8880
gtttcgattt	gtcttagagg	aacgcatata	cagtaatcat	agagaataaa	cgatattcat	8940
ttattaaagt	agatagttga	ggtagaagtt	gtaaagagtg	ataaatagcg	gccgctcact	9000
gaatcttttt	ggctcccttg	tgctttcgga	cgatgtaggt	ctgcacgtag	aagttgagga	9060
acagacacag	gacagtacca	acgtagaagt	agttgaaaaa	ccagccaaac	attctcattc	9120
catcttgtcg	gtagcaggga	atgttccggg	acttccagac	gatgtagaag	ccaacgttga	9180
actgaatgat	ctgcatagaa	gtaatcaggg	acttgggcat	agggaacttg	agcttgatca	9240
gtcgggtcca	atagtagccg	tacatgatcc	agtgaatgaa	gccgttgagc	agcacaaga	9300
tccaaacggc	ttcgtttcgg	tagttgtaga	acagccacat	gtccatagga	gctccgagat	9360
ggtgaaagaa	ctgcaaccag	gtcagaggct	tgcccatgag	gggcagatag	aaggagtcaa	9420
tgtactcgag	gaacttgctg	aggtagaaca	gctgagtggg	gattcggaag	acattgttgt	9480
cgaaagcctt	ctcgcagttg	tcggacatga	caccaatggg	gtacatggcg	taggccatag	9540
agaggaagga	gcccagcgag	tagatggaca	tgagcaggtt	gtagttgggtg	aacacaaact	9600
tcattcgaga	ctgacccttg	ggtcgcagag	gaccaagggt	gaacttcagg	atgacgaagg	9660
cgatggagag	gtacagcacc	tcgcagtgcg	aggcatcaga	ccagagctga	gcatagtcga	9720
ccttgggaag	aacctccttg	ccaatggaga	cgatttcgtt	cacgacctcc	atggttgatg	9780
tgtgtttaat	tcaagaatga	atatagagaa	gagaagaaga	aaaaagattc	aattgagccg	9840
gcgatgcaga	cccttatata	aatgttgcct	tggacagacg	gagcaagccc	gccccaaacct	9900
acgttcggta	taatattgta	agctttttta	cacaaagggt	tggcttgggg	taacctgatg	9960
tggtgcaaaa	gaccgggctg	tggcgagcca	ttgcgcgggc	gaatggggcc	gtgactcgtc	10020
tcaaattcga	gggcgtgcct	caattcgtgc	ccccgtggct	ttttcccgcc	gtttccgccc	10080
cgtttgcacc	actgcagccg	cttcttttgt	tcggacacct	tgctgcgagc	taggtgcctt	10140
gtgctactta	aaaagtggcc	tcccaacacc	aacatgacat	gagtgcgtgg	gccaagacac	10200
gttggcgggg	tcgcagtcgg	ctcaatggcc	cggaaaaaac	gctgctggag	ctggttcgga	10260
cgcagtcgcg	cgcggcgtat	ggatatccgc	aaggttccat	agcgccattg	ccctccgctg	10320
gcgtctatcc	cgcaacctct	aaatagagcg	ggaatataac	ccaagcttct	tttttttccct	10380
ttaacacgea	cacccccaac	tatcatgttg	ctgctgctgt	ttgactctac	tctgtggagg	10440
ggtgtctcca	cccaacccaa	cctacagggt	gatccggcgc	tgtgattggc	tgataagtct	10500
cctatccgga	ctaattctga	ccaatgggac	atgcgcgcag	gacccaaatg	ccgcaattac	10560
gtaaccccaa	cgaaatgcct	acccctcttt	ggagcccagc	ggccccaaat	ccccccaage	10620
agccccgttc	taccggcttc	catctccaag	cacaagcagc	ccggttctac	cggttcccat	10680
ctccaagcac	ccctttctcc	acaccccaca	aaaagacccg	tgcaggacat	cctactgcgt	10740

gtttaaacac	cactaaaacc	ccacaaaata	tatcttaccg	aatatacaga	tctactatag	10800
aggaacaatt	gccccggaga	agacggccag	gccgcctaga	tgacaaattc	aacaactcac	10860
agctgacttt	ctgccattgc	cactagggggg	gggccttttt	atatggccaa	gccaagctct	10920
ccacgtcggt	tgggctgcac	ccaacaataa	atgggtaggg	ttgcaccaac	aaagggatgg	10980
gatggggggt	agaagatacg	aggataacgg	ggctcaatgg	cacaaataag	aacgaatact	11040
gccattaaga	ctcgtgatec	agcgactgac	accattgcat	catctaaggg	cctcaaaact	11100
acctcggaac	tgctgcgctg	atctggacac	cacagagggt	ccgagcactt	taggttgcac	11160
caaatgtccc	accaggtgca	ggcagaaaac	gctggaacag	cgtgtacagt	ttgtcttaac	11220
aaaaagtgag	ggcgctgagg	tcgagcaggg	tgggtgtgact	tgttatagcc	tttagagctg	11280
cgaagcgcg	tatggatttg	gctcatcagg	ccagattgag	ggctctgtga	cacatgtcat	11340
gttagtgtac	ttcaatcgcc	ccctggatat	agccccgaca	ataggccgtg	gcctcatttt	11400
tttgccttcc	gcacatttcc	attgctcggt	accacacact	tgtctctcct	gcacttgcca	11460
accttaatac	tggttttacat	tgaccaacat	cttacaagcg	gggggcttgt	ctaggggtata	11520
tataaacagt	ggctctccca	atcggttgcc	agtctctttt	ttcctttctt	tccccacaga	11580
ttcgaaatct	aaactacaca	tcacacaatg	cctgttactg	acgtccttaa	gcgaaagtcc	11640
ggtgtcatcg	tcggcgacga	tgtccgagcc	gtgagtatcc	acgacaagat	cagtgtcgag	11700
acgacgcgtt	ttgtgtaatg	acacaatccg	aaagtcgcta	gcaacacaca	ctctctacac	11760
aaactaaccc	agctctccat	ggtgaaggct	tctcgacagg	ctctgccccct	cgtcatcgac	11820
ggaaagggtg	acgacgtctc	cgcttgggtg	aacttccacc	ctgggtggagc	tgaaatcatt	11880
gagaactacc	agggacgaga	tgctactgac	gccttcatgg	ttatgcactc	tcaggaagcc	11940
ttcgacaagc	tcaagcgaat	gcccagatc	aaccaggctt	ccgagctgcc	tccccaggct	12000
gccgtcaacg	aagctcagga	ggatttccga	aagctccgag	aagagctgat	cgccactggc	12060
atgtttgacg	cctctcccct	ctggctactg	tacaagatct	tgaccaccct	gggtcttggc	12120
gtgcttgcc	tcttcatgct	ggctccagtac	cacctgtact	tcattgggtg	tctcgtgctc	12180
ggtatgcact	accagcaaat	gggatggctg	tctcatgaca	tctgccacca	ccagaccttc	12240
aagaaccgaa	actggaataa	cgctcctgggt	ctggctctttg	gcaacggact	ccagggcttc	12300
tccgtgacct	ggtggaagga	cagacacaa	gcccacatt	ctgctaccaa	cgttcagggt	12360
cacgatcccg	acattgataa	cctgcctctg	ctcgccctgg	ccgaggacga	tgtcactcga	12420
gcttctccca	tctcccga	gctcattcag	ttccaacagt	actatttcct	ggctcatctgt	12480
attctcctgc	gattcatctg	gtgtttccag	tctgtgctga	ccgttcgata	cctcaaggac	12540
cgagacaacc	agttctaccg	atctcagtac	aagaaagagg	ccattggact	cgctctgcac	12600
tggactctca	agaccctgtt	ccacctcttc	tttatgccct	ccatcctgac	ctcgatgctg	12660
gtgttctttg	tttccgagct	cgctcggtgg	ttcggaattg	ccatcggtgg	cttcatgaac	12720
cactaccctc	tggagaagat	cggtgattcc	gtctgggacg	gacatggctt	ctctgtgggt	12780
cagatccatg	agaccatgaa	cattcgacga	ggcatcatta	ctgactgggt	ctttggaggc	12840
ctgaactacc	agategagca	ccatctctgg	cccaccctgc	ctcgacacaa	cctcactgcc	12900
gtttcctacc	aggtggaaca	gctgtgccag	aagcacaacc	tccccctaccg	aaacctctg	12960
ccccatgaag	gtctcgtcat	cctgctccga	tacctgtccc	agttcgctcg	aatggccgag	13020
aagcagcccc	gtgccaaggc	tcagtaagcg	gccgcatgag	aagataaata	tataaataca	13080

ttgagatatt	aaatgcgcta	gattagagag	cctcatactg	ctcggagaga	agccaagacg	13140
agtactcaaa	ggggattaca	ccatccatat	ccacagacac	aagctgggga	aaggttctat	13200
atacactttc	cggaataccg	tagtttccga	tgttatcaat	gggggcagcc	aggatttcag	13260
gcacttcggt	gtctcggggt	gaaatggcgt	tcttggcctc	catcaagtcg	taccatgtct	13320
tcatttgcct	gtcaaagtaa	aacagaagca	gatgaagaat	gaacttgaag	tgaaggaatt	13380
taaatagtgt	gagcaaggga	gaaatgtaga	gtgtgaaaga	ctcactatgg	tccgggctta	13440
tctcgaccaa	tagccaaagt	ctggagtttc	tgagagaaaa	aggcaagata	cgtatgtaac	13500
aaagcgacgc	atggtacaat	aataccggag	gcatgtatca	tagagagtta	gtggttcgat	13560
gatggcactg	gtgcctggta	tgactttata	cggtcgacta	catatttgtc	ctcagacata	13620
caattacagt	caagcactta	cccttggaca	tctgtaggta	ccccccggcc	aagacgatct	13680
cagcgtgtcg	tatgtcggat	tggcgtagct	ccctcgctcg	tcaattggct	cccatctact	13740
ttcttctgct	tggctacacc	cagcatgtct	gctatggctc	gttttcgtgc	cttatctatc	13800
ctcccagtat	taccaactct	aaatgacatg	atgtgattgg	gtctacactt	tcatatcaga	13860
gataaggagt	agcacagttg	cataaaaagc	ccaactctaa	tcagcttctt	cctttcttgt	13920
aattagtaca	aaggtgatta	gcgaaatctg	gaagcttagt	tggccctaaa	aaaatcaaaa	13980
aaagcaaaaa	acgaaaaacg	aaaaaccaca	gttttgagaa	cagggaggta	acgaaggatc	14040
gtatatatat	atatatatat	atatacccac	ggatcccgag	accggccttt	gattcttccc	14100
tacaaccaac	cattctcacc	accctaattc	acaaccatgg	gcgtattcat	taaacaggag	14160
cagcttccgg	ctctcaagaa	gtacaagtac	tccgccgagg	atcactcggt	catctccaac	14220
aacattctgc	gccccttctg	gcgacagttt	gtcaaaaatct	tccctctgtg	gatggccccc	14280
aacatggtga	ctctgctggg	cttcttcttt	gtcattgtga	acttcatcac	catgctcatt	14340
gttgatccca	cccacgaccg	cgagcctccc	agatgggtct	acctcaccta	cgctctgggt	14400
ctgttccctt	accagacatt	tgatgcctgt	gacggatccc	atgcccgcg	aactggccag	14460
agtggacccc	tggagagct	gtttgaccac	tgtgtcgacg	ccatgaatac	ctctctgatt	14520
ctcacggtgg	tgggtgtccac	cacccatatg	ggatataaca	tgaagctact	gattgtgcag	14580
attgccgctc	tcggaaactt	ctacctgtcg	acctgggaga	cctaccatac	cggaactctg	14640
tacctttctg	gcttctctgg	tctgtttgaa	ggatatctga	ttctgggtggc	tcttttcgtc	14700
ctcaccttct	tactgggtcc	caacgtgtac	gctctgaccg	tctacgaggc	tcttcccag	14760
tccatcactt	cgctgctgcc	tgccagcttc	ctggacgtca	ccatcaccca	gatctacatt	14820
ggattcggag	tgtctgggcat	ggtgttcaac	atctacggcg	cctgcggaaa	cgtgatcaag	14880
tactacaaca	acaagggcaa	gagcgctctc	cccgccattc	tcggaatcgc	cccccttggc	14940
atctttctacg	tcggcgctct	tgcctgggcc	catgttgctc	ctctgcttct	ctccaagtac	15000
gccatcgctc	atctgtttgc	cattgggggt	gcctttgcca	tgcaagtcgg	ccagatgatt	15060
cttgcceate	tcgtgcttgc	tccctttccc	cactggaacg	tgctgctctt	cttccccctt	15120
gtgggactgg	cagtgcacta	cattgcaccc	gtgtttggct	gggacgccga	tatcgtgtcg	15180
gttaacacte	tcttcacctg	ttttggcgcc	accctctcca	tttacgcctt	ctttgtgctt	15240
gagatcateg	acgagatcac	caactacctc	gatatctggt	gtctgcgaat	caagtaccct	15300
caggagaaga	agaccgaata	agcggccgca	tggagcgtgt	gttctgagtc	gatgttttct	15360
atggagttgt	gagtgttagt	agacatgatg	ggtttatata	tgatgaatga	atagatgtga	15420

```

ttttgatttg cacgatggaa ttgagaactt tgtaaacgta catgggaatg tatgaatgtg 15480
ggggttttgt gactggataa ctgacgggtca gtggacgccg ttgttcaaat atccaagaga 15540
tgcgagaaac tttgggtcaa gtgaacatgt cctctctgtt caagtaaacc atcaactatg 15600
ggtagtatat ttagtaagga caagagttag gattcttttg agtcctagaa acgtattttc 15660
gcgttccaag atcaaattag tagagtaata cgggcacggg aatccattca tagtctcaat 15720
cctgcagggtg agttaattaa tcgagcttgg cgtaatcatg gtcatactgt tttcctgtgt 15780
gaaattgtta tccgctcaca attccacaca ac 15812

```

<210>85

<211>4313

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pZKUM

<400>85

```

taatcgagct tggcgtaatc atggtcatac ctgtttcctg tgtgaaattg ttatccgctc 60
acaattccac acaacatacg agccggaagc ataaagtgtg aagcctgggg tgcctaata 120
gtgagctaac tcacattaat tgcgttgcgc tctactgccg ctttccagtc gggaaacctg 180
tcgtgccagc tgcattaatg aatcgcccaa cgcgcgggga gaggcgggtt gcgtattggg 240
cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcgggt gcggcgagcg 300
gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga 360
aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg 420
gcgtttttcc ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag 480
aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg aagctccctc 540
gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg 600
ggaagcgtgg cgctttctca tagctcacgc ttaggtatc tcagttcggg ttaggtcgtt 660
cgctccaagc tgggctgtgt gcacgaacct ccggttcagc ccgaccgctg cgccttatcc 720
ggtaactatc gtcttgagtc caaccggta agacacgact tatcgccact ggcagcagcc 780
actggtaaca ggattagcag agcgaggtat gtaggcgggt ctacagagtt cttgaagtgg 840
tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca 900
gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc 960
ggtgggtttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 1020
cctttgatct tttctacggg gtctgacgct cagtggaaacg aaaactcacg ttaagggtat 1080
ttggatcatg gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt 1140
tttaaataca tctaaagtat atatgagtaa acttgggtctg acagttacca atgcttaate 1200
agttaggcac ctatctcagc gatctgtcta tttcgttcat ccatagtgtc ctgactcccc 1260
gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtc tgcaatgata 1320
ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg 1380
gccgagcgca gaagtgggtc tgcaacttta tccgcctcca tccagtctat taattgttgc 1440
cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct 1500

```

acaggcatcg	tgggtgtcacg	ctcgctgttt	ggtatggctt	cattcagctc	cggttcccaa	1560
cgatcaagge	gagttacatg	atccccatg	ttgtgcaaaa	aagcggttag	ctccttcggt	1620
cctccgateg	ttgtcagaag	taagttggcc	gcagtgttat	cactcatggt	tatggcagca	1680
ctgcataatt	ctcttactgt	catgccaatc	gtaagatgct	tttctgtgac	tgggtgagtac	1740
tcaaccaagt	cattctgaga	atagtgtatg	cggcgaccga	gttgcctctg	cccggcgtca	1800
atacgggata	ataccgcgcc	acatagcaga	actttaaaag	tgctcatcat	tggaaaacgt	1860
tcttcggggc	gaaaactctc	aaggatctta	ccgctgttga	gatccagttc	gatgtaaccc	1920
actcgtgcac	ccaactgate	ttcagcatct	tttactttca	ccagcgtttc	tgggtgagca	1980
aaaacaggaa	ggcaaaatgc	cgcaaaaaag	ggaataaggg	cgacacggaa	atgttgaata	2040
ctcatactct	tcctttttca	atattattga	agcattttatc	agggttattg	tctcatgagc	2100
ggatacatat	ttgaatgtat	ttagaaaaat	aaacaaatag	gggttccgcg	cacatttccc	2160
cgaaaagtgc	cacctgacgc	gccctgtagc	ggcgcattaa	gcgcggcggg	tgtgggtggtt	2220
acgcgcagcg	tgaccgctac	acttgccagc	gccctagcgc	ccgctccttt	cgctttcttc	2280
ccttcctttc	tcgccacgtt	cgccggcttt	ccccgtcaag	ctctaaatcg	ggggctccct	2340
ttagggttcc	gatttagtgc	tttacggcac	ctcgacccca	aaaaacttga	ttagggtgat	2400
ggttcacgta	gtgggccatc	gccctgatag	acggtttttc	gccctttgac	gttggagtcc	2460
acgttcttta	atagtggact	cttgttccaa	actggaacaa	cactcaaccc	tatctcggtc	2520
tattcttttg	atttataagg	gattttgccg	atttcggcct	attgggttaa	aaatgagctg	2580
atttaacaaa	aatttaacgc	gaattttaac	aaaatattaa	cgcttacaat	ttccattcgc	2640
cattcaggct	gcgcaactgt	tgggaagggc	gatcgggtgcg	ggcctcttcg	ctattacgcc	2700
agctggcgaa	agggggatgt	gctgcaaggc	gattaagttg	ggtaacgcca	gggttttccc	2760
agtcacgacg	ttgtaaaacg	acggccagtg	aattgtaata	cgactcacta	tagggcgaa	2820
tgggtaccgg	gccccccctc	gaggctgacg	agtatctgtc	tgactcgtca	ttgccgcctt	2880
tggagtacga	ctccaactat	gagtgtgctt	ggatcacttt	gacgatacat	tcttcgttgg	2940
aggctgtggg	tctgacagct	gcgttttcgg	cgcggttggc	cgacaacaat	atcagctgca	3000
acgtcattgc	tggctttcat	catgatcaca	tttttgcgg	caaaggcgac	gccagagag	3060
ccattgacgt	tctttcta	ttggaccgat	agccgtatag	tccagtctat	ctataagttc	3120
aactaactcg	taactattac	cataacatat	acttcaactgc	cccagataag	gttccgataa	3180
aaagttctgc	agactaaatt	tatttcagtc	tcctcttcac	caccaaaatg	ccctcctacg	3240
aagctcgagt	gctcaagctc	gtggcagcca	agaaaaccaa	cctgtgtgct	tctctggatg	3300
ttaccaccac	caaggagctc	attgagcttg	ccgataaggt	cggaccttat	gtgtgcatga	3360
tcaaaaccca	tatcgacatc	attgacgact	tcacctacgc	cggcactgtg	ctccccctca	3420
aggaacttgc	tcttaagcac	ggtttcttcc	tgttcgagga	cagaaagttc	gcagatattg	3480
gcaacactgt	caagcaccag	taccgggtgc	accgaatcgc	cgagtgggtc	gatataacca	3540
acgcccacgg	tgtaccggga	accggaatcg	attgctggcc	tgcgagctgg	tgcgtacgag	3600
gaaactgtct	ctgaacagaa	gaaggaggac	gtctctgact	acgagaactc	ccagtacaag	3660
gagttcctag	tccccctctc	caacgagaag	ctggccagag	gtctgctcat	gctggccgag	3720
ctgtcttgca	agggtctctc	ggccactggc	gagtactcca	agcagaccat	tgagcttgcc	3780
cgatccgacc	ccgagtttgt	ggttggcttc	attgcccaga	accgacctaa	gggcgactct	3840

```

gaggactggc ttattctgac ccccggggtg ggtcttgacg acaagggaga cgctctcgga 3900
cagcagtacc gaactgttga ggatgtcatg tctaccggaa cggatatcat aattgtcggc 3960
cgaggtctgt acggccagaa ccgagatcct attgaggagg ccaagcgata ccagaaggct 4020
ggctgggagg cttaccagaa gattaactgt tagaggttag actatggata tgtaatttaa 4080
ctgtgtatat agagagecgt caagtatgga gcgcttggtc agcttgtatg atggtcagac 4140
gacctgtctg atcgagtatg tatgatactg cacaacctgt gtatccgcat gatctgtcca 4200
atggggcatg ttgttgtgtt tctcgatacg gagatgctgg gtacagtgtc aatacgttga 4260
actacttata cttatatgag gctcgaagaa agctgacttg tgtatgactt aat 4313

```

<210>86

<211>1459

<212>DNA

<213> 人工序列

<220>

<223> 合成突变型 Ura3 基因, 包含对耶氏酵母属 Ura3 编码区 (GenBank 保藏号 AJ306421) 的一个从 +21 至 +53 的 33bp 的缺失, 一个在 +376 的 1bp 缺失, 和一个从 +400 至 403 的 3bp 缺失

<400>86

```

gagtatctgt ctgactcgtc attgccgcct ttggagtacg actccaacta tgagtgtgct 60
tggatcactt tgacgataca ttcttcgttg gaggtgttgg gtctgacagc tgcgttttcg 120
gcgcggttgg ccgacaacaa tatcagctgc aacgtcattg ctggctttca tcatgatcac 180
atTTTTgtcg gcaaaggcga cgcccagaga gccattgacg ttctttctaa tttggaccga 240
tagccgtata gtccagtcta tctataagtt caactaactc gtaactatta ccataacata 300
tacttcactg ccccagataa ggttccgata aaaagtcttg cagactaaat ttatttcagt 360
ctcctcttca ccacaaaat gccctcctac gaagctcgag tgctcaagct cgtggcagcc 420
aagaaaacca acctgtgtgc ttctctggat gttaccacca ccaaggagct cattgagctt 480
gccgataagg tcggacctta tgtgtgcatg atcaaaaccc atatcgacat cattgacgac 540
ttcacctacg ccggcactgt gtccccctc aaggaacttg ctcttaagca cggtttcttc 600
ctgttcgagg acagaaagt ttgcagatatt ggcaaacactg tcaagcacca gtaccggtgt 660
caccgaatcg ccgagtggtc cgatatcacc aacgcccacg gtgtaccggg aaccggaatc 720
gattgctggc ctgcgagctg gtgcgtacga ggaaactgtc tctgaacaga agaaggagga 780
cgtctctgac tacgagaact ccagttacaa ggagttccta gtcccccttc ccaacgagaa 840
gctggccaga ggtctgctca tgctggccga gctgtcttgc aagggctctc tggccactgg 900
cgagtactcc aagcagacca ttgagcttgc ccgatccgac cccgagtttg tggttggctt 960
cattgcccag aaccgacctt agggcgactc tgaggactgg cttattctga ccccggggtt 1020
gggtcttgac gacaaggag acgctctcgg acagcagtac cgaactgttg aggatgtcat 1080
gtctaccgga acgatataca taattgtcgg ccgaggtctg tacggccaga accgagatcc 1140
tattgaggag gccaagcgat accagaaggc tggctgggag gcttaccaga agattaactg 1200
ttagaggtta gactatggat atgtaattta actgtgtata tagagagcgt gcaagtatgg 1260
agcgcttgtt cagcttgtat gatggtcaga cgacctgtct gatcgagtat gtatgatact 1320

```


gcacaacctg tgtatccgca tgatctgtcc aatggggcat gttgttgtgt ttctcgatac 1380
 ggagatgctg ggtacagtgc taatacgttg aactacttat acttatatga ggctcgaaga 1440
 aagctgactt gtgtatgac 1459

<210>87

<211>7966

<212>DNA

<213> 人工序列

<220>

<223> 质粒 pYPS161

<400>87

aaatgtaacg aaactgaaat ttgaccagat attgtgtccg cgggtggagct ccagcttttg 60
 ttcccttttag tgagggttaa ttctgagctt ggcgtaatca tggtcatagc tgtttcctgt 120
 gtgaaattgt tatecgctca caagcttcca cacaacgtac gttctggttg gctcggatga 180
 ttctctgcggc ccagcgtaa ggcaggcggt cgcgccgat cggtttgggt cggatcggct 240
 ttttgattgt cgtattgtcg ctcatgttgg acctggtgtg tagttgtagt gtcagatcag 300
 attcaccagc gaatgcatgt gaacttcccc acattttgag ccgaggcaga tttgggttgc 360
 ttagtaagca gacgtggcgt tgcaagtaga tgtggcaaatt ggggacgaag attccgaggg 420
 gatatcatag ttccaagggg atgtcatcat ttgccagctt tcgccgccac ttttgacgag 480
 tttttgtggg tcaaataagt ttagttgaac ttttcaaatt tcagttggca ttttgttaat 540
 agaaagggtg cgggtgctgg ggggttcatt cctcgggttg cagatatact atctgtctta 600
 ggggtatctc ttccaatcga caagatgtag ttgggtaaca attatttatt aatattctct 660
 ccatccagta cagtactaac atcttgacat ctccagcaca gtgcatcttc ccaagtgttt 720
 gttggagagg ttgttgggta ttacttagga aacagaacac agtacgtgga gatcttggat 780
 acatcgtaca tggaggttat ccataaaaaa gaccctccag gactagttac aatgccgtta 840
 gatgaggaaa tccacaaccc tgattcacta tgaacatatt atcttcccc aaacttgcga 900
 tatatggccc ttgatgatag ctttgatatt acccttgatg gtacctccac gaccaaccga 960
 tctgctgttt gaagagatat ttccaatttt gaagtgtcga gatctactaa acatgagtcc 1020
 agtaattctt tccgtctttc cgatttccga tattcccttt ttagcccgga cttttcactg 1080
 ctcccatgtc aaacgattag gacttgggag acaatcccac tgtcaaaatc accccgatat 1140
 tctctgtaaa acaagtactt ctccacgtg atcttcaaatt acctcttcca cgtgaccttc 1200
 aaatacctct tcaagtacct ctccacgcg accttcaaag tcccttcaaa tacccttctc 1260
 aattctcccc ttctctcca tagtcttct ctctgactaa gcttgagaat acatgacgct 1320
 aagacgaaaa cacactagag accctgagag cctgaacatg catccactct gcagttgcgc 1380
 acgtgcctac agcaactatc ggggccagtg ctggatctga cactgcgtct ccctatgaag 1440
 aaactgataa acagatctgc actcataaca atgatctgag cgatgaaaac gtgacctcca 1500
 cagccacaag tcataatcgg cgcgccagct gcattaatga atcggccaac gcgcggggag 1560
 aggcggtttg cgtattgggc gctcttccgc ttctcgtct actgactcgc tgcgctcggt 1620
 cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggg tatccacaga 1680
 atcaggggat aacgcaggaa agaacatgtg agcaaaaagg cagcaaaaagg ccaggaaccg 1740

taaaaaggcc	gcgttgctgg	cgtttttcca	taggctccgc	ccccctgacg	agcatcacia	1800
aaatcgacgc	tcaagtcaga	ggtggcgaaa	cccgacagga	ctataaagat	accaggcggtt	1860
tccccctgga	agctccctcg	tgcgctctcc	tgttccgacc	ctgccgctta	ccggataacct	1920
gtccgccttt	ctcccttcgg	gaagcgtggc	gctttctcat	agctcacgct	gtaggtatct	1980
cagttcgggtg	taggtcggtt	gctccaagct	gggctgtgtg	cacgaacccc	ccgttcagcc	2040
cgaccgctgc	gccttatecg	gtaactatcg	tcttgagtcc	aacccggtaa	gacacgactt	2100
atcgccactg	gcagcagcca	ctggtaacag	gattagcaga	gcgaggatatg	taggcgggtgc	2160
tacagagttc	ttgaagtggg	ggcctaacta	cggctacact	agaagaacag	tatttgggtat	2220
ctgcgctctg	ctgaagccag	ttaccttcgg	aaaaagagtt	ggtagctctt	gatccggcaa	2280
acaaccacc	gctggtagcg	gtggtttttt	tgtttgcaag	cagcagatta	cgcgcagaaa	2340
aaaaggatct	caagaagatc	ctttgatctt	ttctacgggg	tctgacgctc	agtggaacga	2400
aaactcacgt	taagggatct	tggatcatgag	attatcaaaa	aggatcttca	cctagatcct	2460
tttaaattaa	aatgaagtt	ttaaataaat	ctaaagtata	tatgagtaaa	cttgggtctga	2520
cagttaccaa	tgcttaatac	gtgaggcacc	tatctcagcg	atctgtctat	ttcgttcac	2580
catagttagc	tgactccccg	tcgtgtagat	aactacgata	cgggagggtc	taccatctgg	2640
ccccagtgt	gcaatgatac	cgcgagaccc	acgtcaccg	gctccagatt	tatcagcaat	2700
aaaccagcca	gccggaaggg	ccgagcgcag	aagtggctct	gcaactttat	ccgcctccat	2760
ccagtctatt	aattgtttgc	gggaagctag	agtaagtagt	tcgccagtta	atagtttgcg	2820
caacgttggt	gccattgcta	caggcatcgt	ggtgtcacgc	tcgtcgtttg	gtatggcttc	2880
attcagctcc	ggttcccaac	gatcaaggcg	agttacatga	tcccccatgt	tgtgcaaaaa	2940
agcggttagc	tccttcggtc	ctccgatcgt	tgtcagaagt	aagttggccg	cagtgttatc	3000
actcatgggt	atggcagcac	tgcataattc	tcttactgtc	atgccatccg	taagatgctt	3060
ttctgtgact	ggtgagtact	caaccaagtc	attctgagaa	tagtgtatgc	ggcgaccgag	3120
ttgctcttgc	ccggcgctca	tacgggataa	taccgcgcca	catagcagaa	ctttaaaagt	3180
gctcatcatt	ggaaaacgtt	cttcggggcg	aaaactctca	aggatcttac	cgctgttgag	3240
atccagttcg	atgtaacca	ctcgtgcacc	caactgatct	tcagcatctt	ttactttcac	3300
cagcgtttct	gggtgagcaa	aaacaggaag	gcaaaatgcc	gcaaaaaagg	gaataagggc	3360
gacacggaaa	tgttgaatac	tcatactctt	cttttttcaa	tattattgaa	gcatttatca	3420
gggttattgt	ctcatgagcg	gatacatatt	tgaatgtatt	tagaaaaata	aacaaatagg	3480
ggttccgcgc	acatttcccc	gaaaagtgcc	acctgatgcg	gtgtgaaata	ccgcacagat	3540
gcgtaaggag	aaaataccgc	atcaggaaat	tgtaaagcgt	aatattttgt	taaaattcgc	3600
gttaaatttt	tgttaaatac	gctcatTTTT	taaccaatag	gccgaaatcg	gcaaaatccc	3660
ttataaatca	aaagaataga	ccgagatagg	gttgagtgtt	gttccagttt	ggaacaagag	3720
tccactatta	aagaacgtgg	actccaacgt	caaagggcga	aaaaccgtct	atcagggcga	3780
tggcccaacta	cgtgaaccat	caccctaata	aagttttttg	gggtcgaggt	gccgtaaaagc	3840
actaaatcgg	aaccctaaag	ggagcccccg	atttagagct	tgacggggaa	agccggcgaa	3900
cgtggcgaga	aaggaaggga	agaaagcgaa	aggagcgggc	gctagggcgc	tggcaagtgt	3960
agcggtcacg	ctgcgcgtaa	ccaccacacc	cgccgcgctt	aatgcgccgc	tacagggcgc	4020
gtccattcgc	cattcaggct	gcgcaactgt	tgggaagggc	gatcgggtgcg	ggcctcttcg	4080

ctattacgcc	agctggcgaa	aggggggatgt	gctgcaaggc	gattaagttg	ggtaacgcca	4140
gggttttccc	agtcacgacg	ttgtaaaacg	acggccagtg	aattgtaata	cgactcacta	4200
tagggcgaa	tgggcccga	gtcgcatgca	actattagtg	aggcttcggg	agtggttgtc	4260
tcggttgtct	cattcagact	cgttgtgttg	tatctatata	tatataaaca	ctcttgtccc	4320
tcaatccca	tgccatcttt	tgtataactt	gccgccaata	tgaactcat	ctccctcatc	4380
accgtcgcta	ccaccgtctt	ggcggtgtgc	ggagacaagt	acaagctgac	ctataaccaga	4440
tcagacgccc	aatcggtcga	atctctgccc	gtcacctacc	aagatgacct	gatcaccgcc	4500
tccaccgacg	gcgaacccat	caccatcacc	gagggcgagg	gcaacacctt	ctctgttaac	4560
gacatgccc	tcgctatct	ggagctgcag	gctttgttct	ggaccggcga	ctacggctac	4620
aagctccagg	gtcggtcttt	tgacattgcc	gccgatggaa	cctttgagct	gagagacggc	4680
cccaaggagt	actactattg	cactcctcac	cctgagcgaa	acgtcatcta	cgtcatcaac	4740
agccccgact	actccaagtg	tcggttcaag	cgtaccatca	agttccacgc	tgaagaagtc	4800
taagtggtaa	tcgaccgact	aaccattttt	agctgacaaa	cacttgctaa	ctcctataac	4860
gaatgaatga	ctaacttggc	atattgttac	caagtattac	ttgggatata	gttgagtgt	4920
accattgcta	agaatccaaa	ctggagcttc	taaaggctctg	ggagtcgccg	tatgtgttca	4980
tatcgaaatc	aaagaaatca	taatcgcaac	agaattcaaa	atcaagcaga	ttaatatcca	5040
ttattgtact	cggatcgatga	catatctgat	atgatctcgg	atatgatctc	tgactgttta	5100
ctgggagatt	tgttgaagat	ttgttgagg	tatctgaaaa	gtagacaata	gagacaaaat	5160
gacgatata	agaactgaat	cgggcccga	tactcggtat	cattcccttc	agcagtaact	5220
gtattgctct	atcaatgcga	cgagatacct	ccacaattaa	tactgtatac	gctctaccac	5280
tcatactctc	aatgctaaaa	tatattcatg	cccaggacct	ctgtgcactg	ctatgcagca	5340
cagtgttgct	gattgaattg	gtcgtgtctg	gtccctgatg	ctctgtgtct	cgctgactag	5400
tccttccatc	cagacctcgt	cattatctga	taggcaacaa	gttctgtctct	ctcacaccct	5460
gccgacacaa	gggacactcg	ggcttctctc	tcacccattc	ggaaatacag	tccttaatta	5520
agttgcgaca	catgtcttga	tagtatcttg	aattctctct	cttgagcttt	tccataacaa	5580
gttcttctgc	ctccaggaag	tccatgggtg	gtttgatcat	ggttttgggtg	tagtggtagt	5640
gcagtgggtg	tattgtgact	ggggatgtag	ttgagaataa	gtcatacaca	agtcagcttt	5700
cttcgagcct	catataagta	taagtagttc	aacgtattag	cactgtaccc	agcatctccg	5760
tatcgagaaa	cacaacaaca	tgccccattg	gacagatcat	gcggatacac	aggttgtgca	5820
gtatcataca	tactcgatca	gacaggctgt	ctgaccatca	tacaagctga	acaagcgctc	5880
catacttgca	cgctctctat	atacacagtt	aaattacata	tccatagtct	aacctctaac	5940
agttaatctt	ctggtaagcc	tcccagccag	ccttctggta	tcgcttggcc	tcctcaatag	6000
gatctcggtt	ctggccgtac	agacctcggc	cgacaattat	gatatccgtt	ccggtagaca	6060
tgacatcctc	aacagttcgg	tactgtgtgc	cgagagcgtc	tccttgtcgc	tcaagaccca	6120
ccccgggggt	cagaataagc	cagtcctcag	agtcgccctt	aggtcggttc	tgggcaatga	6180
agccaaccac	aaactcgggg	tcggatcggg	caagctcaat	ggtctgcttg	gagtactcgc	6240
cagtggccag	agagcccttg	caagacagct	cggccagcat	gagcagacct	ctggccagct	6300
tctcgttggg	agaggggact	aggaactcct	tgtactggga	gttctcgtag	tcagagacgt	6360
cctccttctt	ctgttcagag	acagtttctt	cggcaccagc	tcgcaggcca	gcaatgattc	6420

cggttccggg tacaccgtgg gcgttgggtga tatecggacca ctcggcgatt cggtgacacc 6480
 ggtactggtg cttgacagtg ttgccaatat ctgcgaactt tctgtcctcg aacaggaaga 6540
 aaccgtgctt aagagcaagt tccttgaggg ggagcacagt gccggcgtag gtgaagtcgt 6600
 caatgatgtc gatatgggtt ttgatcatgc acacataagg tccgacctta tcggcaagct 6660
 caatgagctc cttgggtggtg gtaacatcca gagaagcaca caggttgggtt ttcttggctg 6720
 ccacgagctt gagcactcga gcggcaaagg cggacttgtg gacgttagct cgagcttcgt 6780
 aggagggcat tttgggtggtg aagaggagac tgaaataaat ttagtctgca gaacttttta 6840
 tcggaacctt atctggggca gtgaagtata tgttatggta atagttacga gttagttgaa 6900
 cttatagata gactggacta tacggctatc ggtccaaatt agaaagaacg tcaatggctc 6960
 tctgggcgtc gcctttgccg acaaaaatgt gatcatgatg aaagccagca atgacgttgc 7020
 agctgatatt gttgtcggcc aaccgcgccg aaaacgcagc tgtcagaccc acagcctcca 7080
 acgaagaatg tategtcaaa gtgatccaag cacactcata gttggagtcg tactccaaag 7140
 gcggcaatga cgagtcagac agatactcgt cgaccttttc cttgggaacc accaccgtca 7200
 gcccttctga ctacgtatt gtagccaccg acacaggcaa cagtcctgtg atagcagaat 7260
 atgtcttgtc ggtccatttc tcaccaactt taggcgtcaa gtgaatgttg cagaagaagt 7320
 atgtgccttc attgagaatc ggtgttgcgt atttcaataa agtcttgaga tcagtttggc 7380
 cagtcatggt gtggggggta attggattga gttatgcct acagtctgta caggtatact 7440
 cgctgcccac tttatacttt ttgattccgc tgcacttgaa gcaatgtcgt ttaccaaaag 7500
 tgagaatgct ccacagaaca caccacaggg tatggttgag caaaaaataa aactccgat 7560
 acggggaate gaaccccggt ctccacggtt ctcaagaagt attcttgatg agagcgtatc 7620
 gatgagccta aaatgaaccc gagtatatct cataaaattc tcggtgagag gtctgtgact 7680
 gtcagtacaa ggtgccttca ttatgccctc aaccttacca tacctcactg aatgtagtgt 7740
 acctctaaaa atgaaataca gtgccaaaag ccaaggcact gagctcgtct aacggacttg 7800
 atatacaacc aattaaaaca atgaaaaga aatacagttc tttgtatcat ttgtaacaat 7860
 taccctgtac aaactaaggt attgaaatcc cacaatatc ccaaagtcca cccctttcca 7920
 aattgtcatg cctacaactc atataccaag cactaaccta ccgttt 7966

<210>88

<211>20

<212>DNA

<213> 人工序列

<220>

<223> 引物 Pex-10del13' . 正向

<400>88

ccaacatgag cgacaatacg

20

<210>89

<211>20

<212>DNA

<213> 人工序列

<220>

<223> 引物 Pex-10del25' . 反向

<400>89

caagttctgc tctctcacac

20

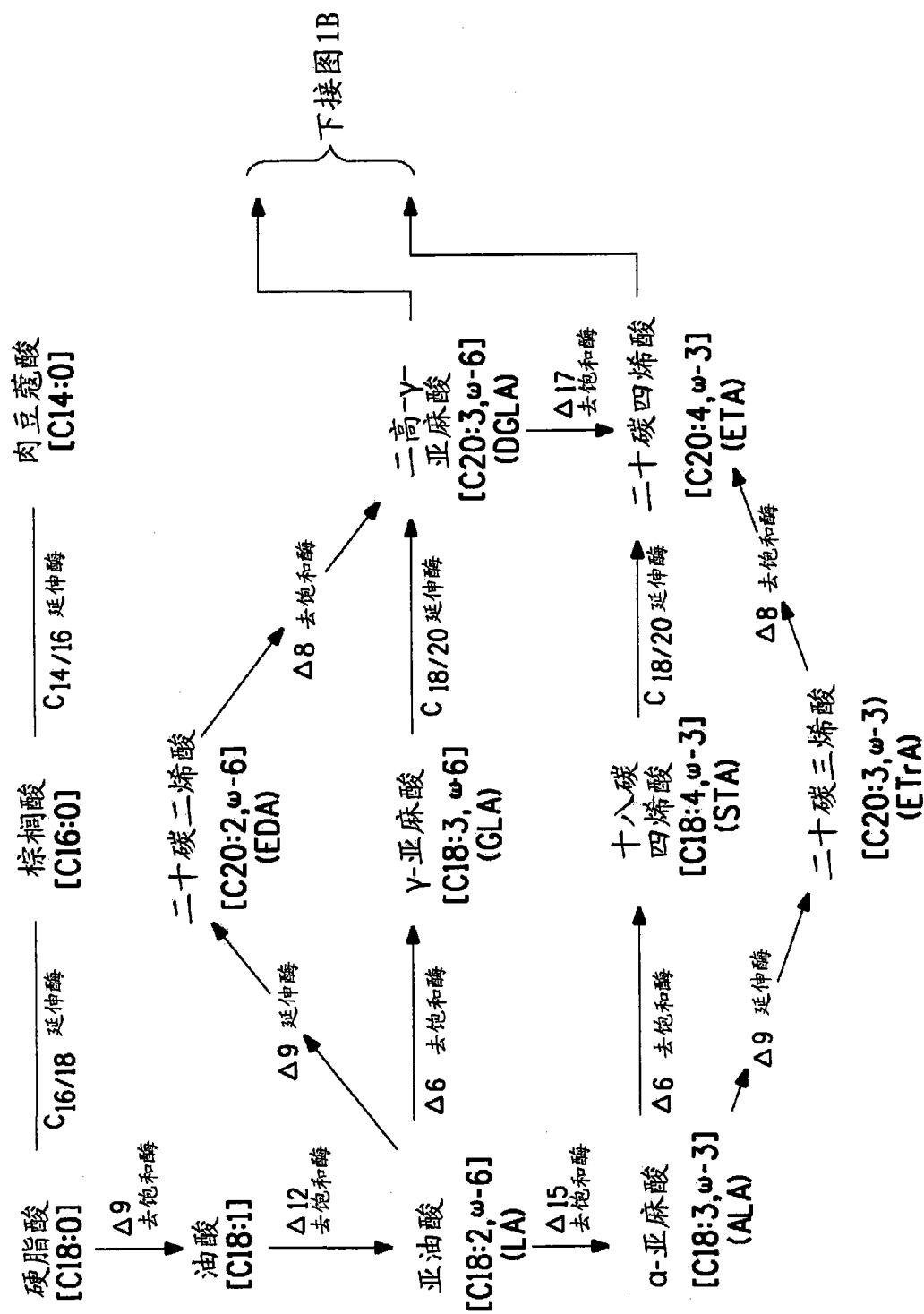


图 1A

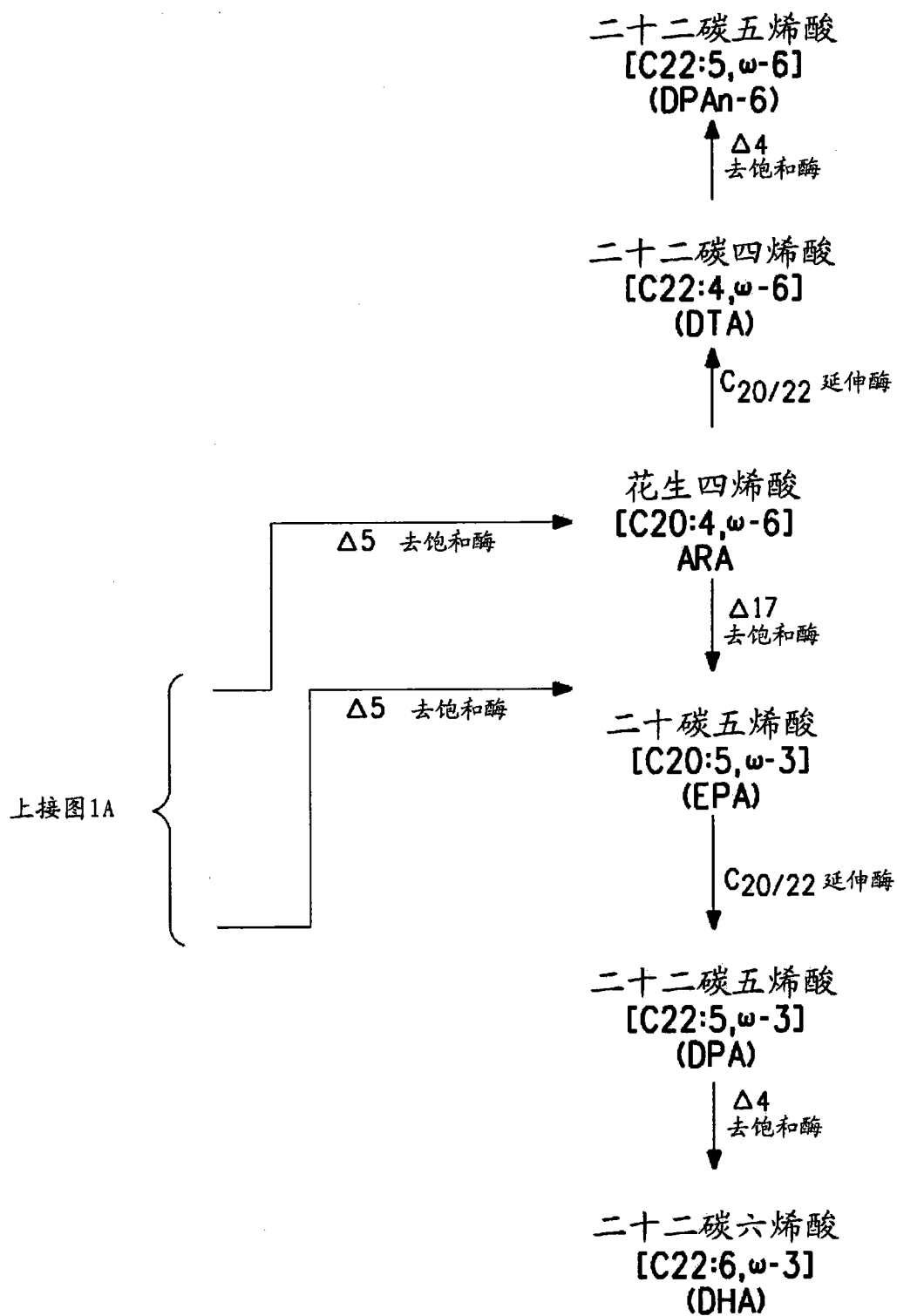


图 1B

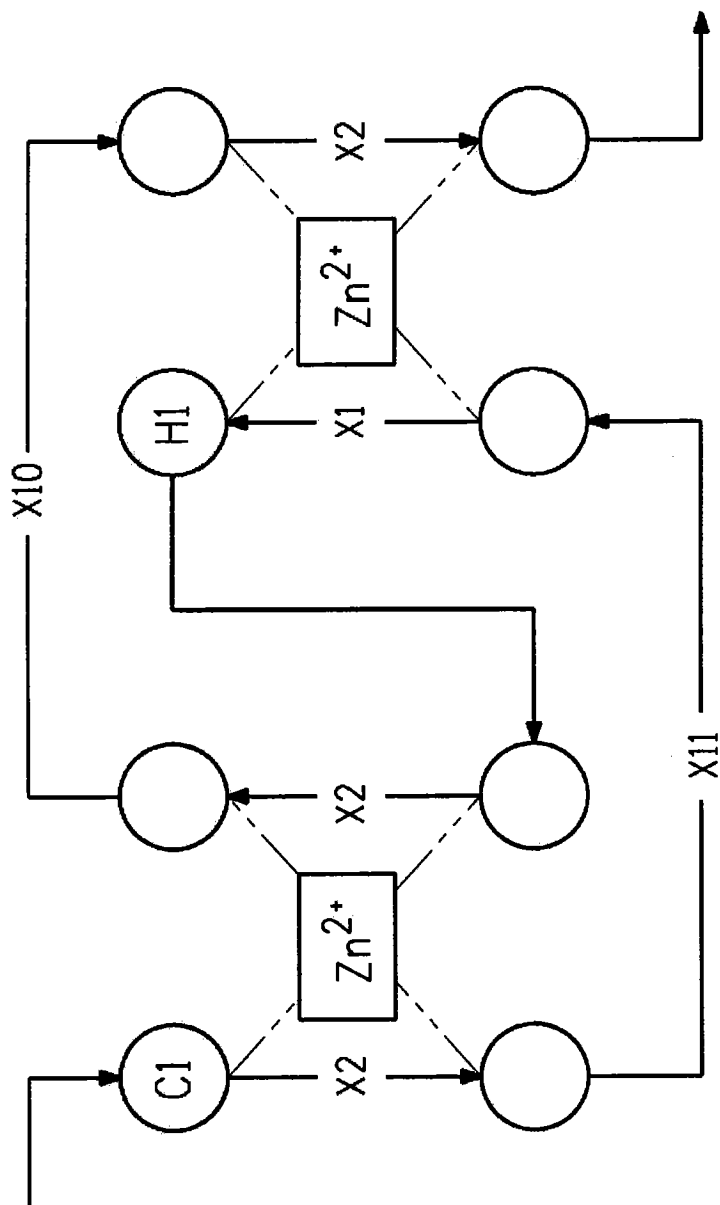


图 2B

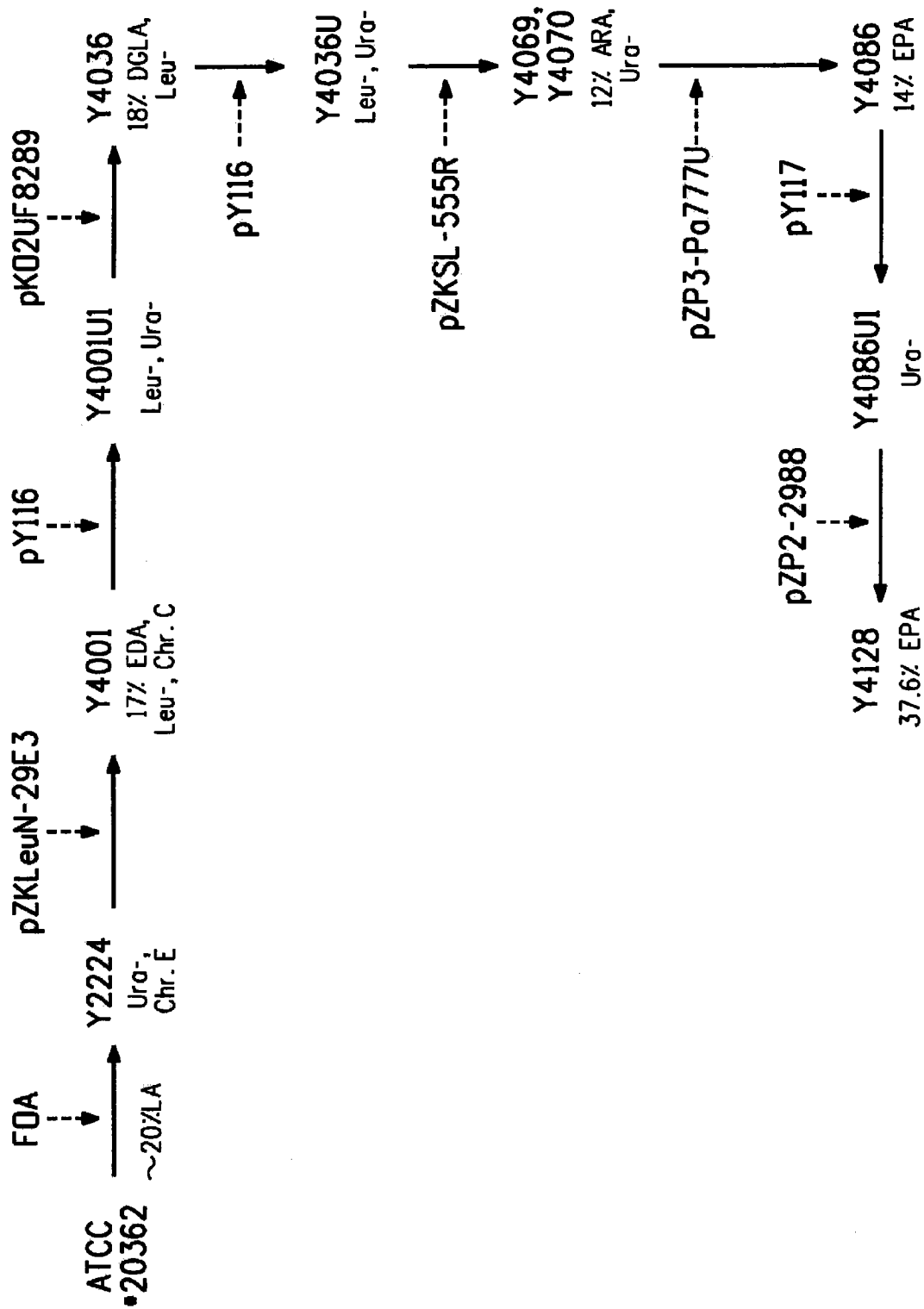


图 3A

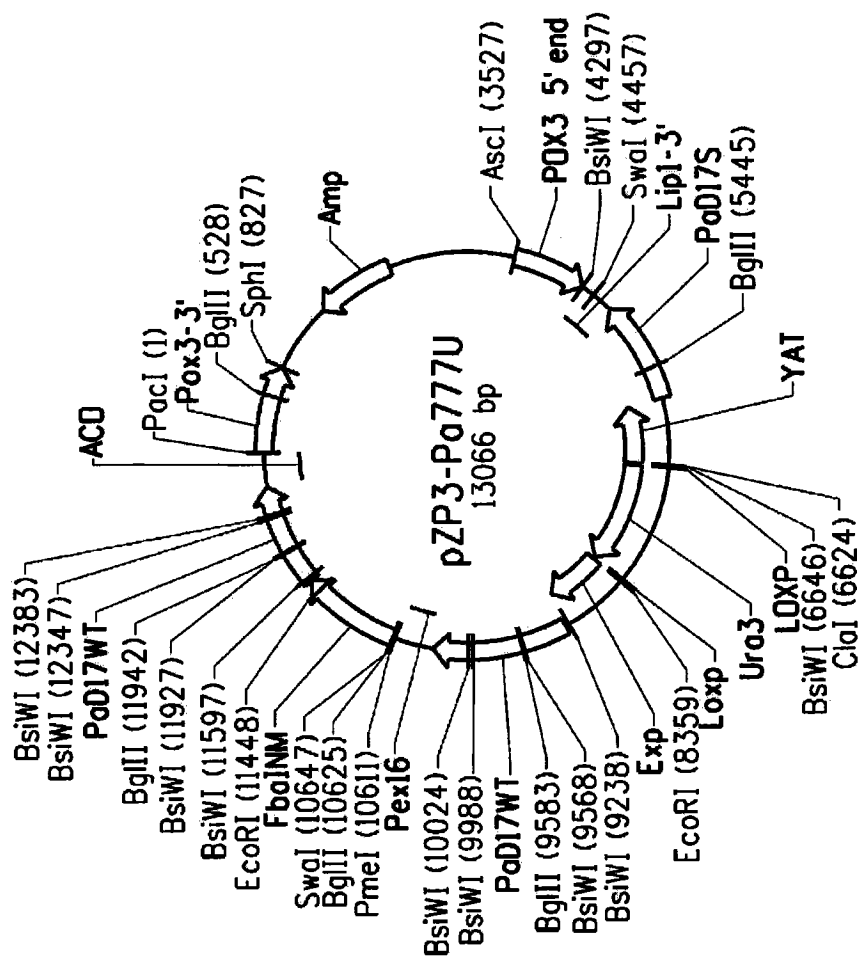


图 3B

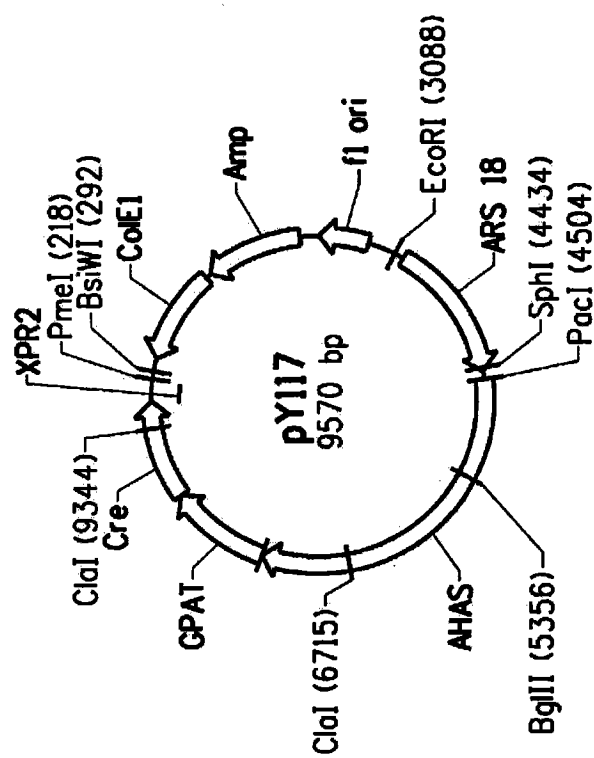


图 4A

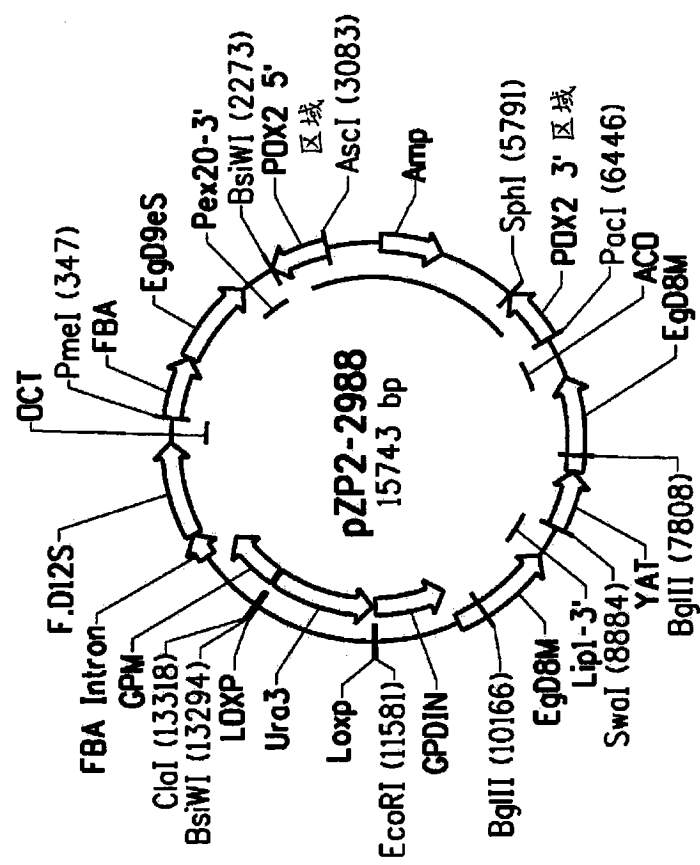


图 4B

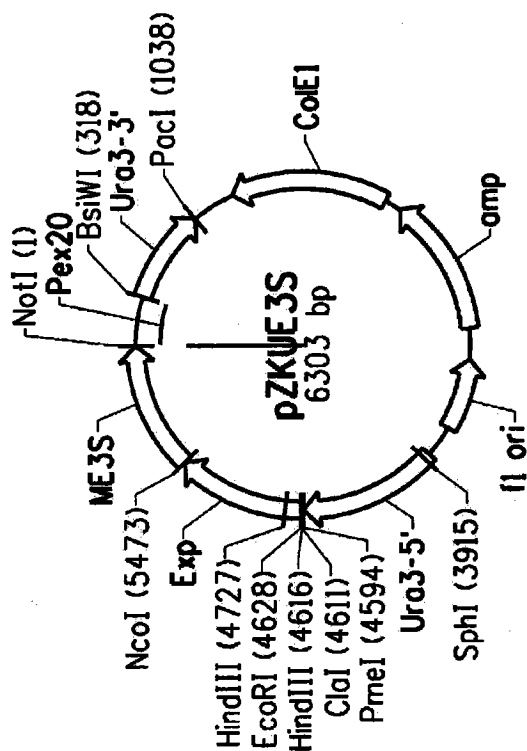


图 5A

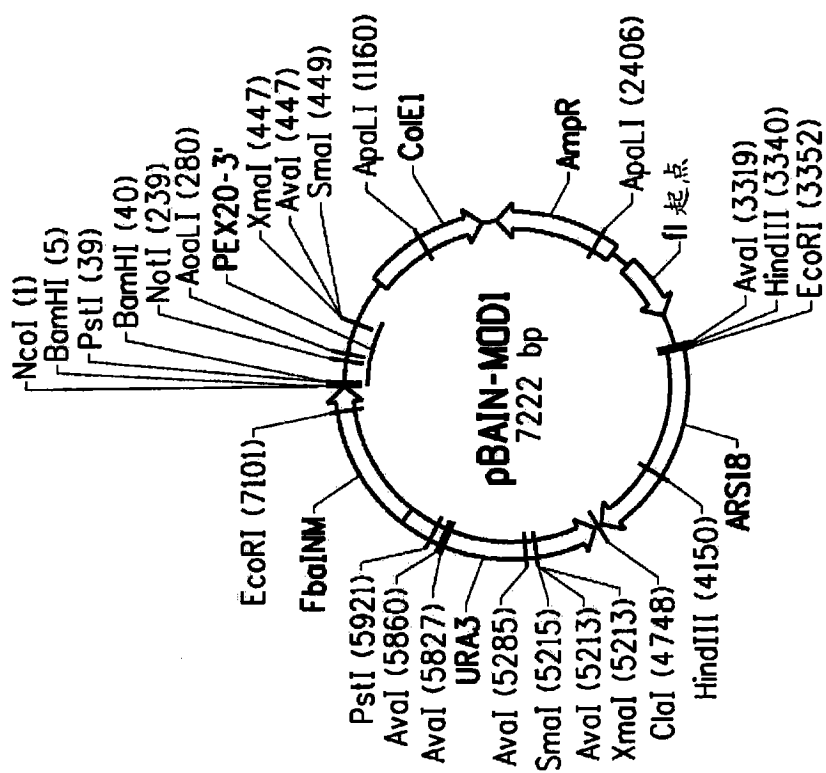


图 5B

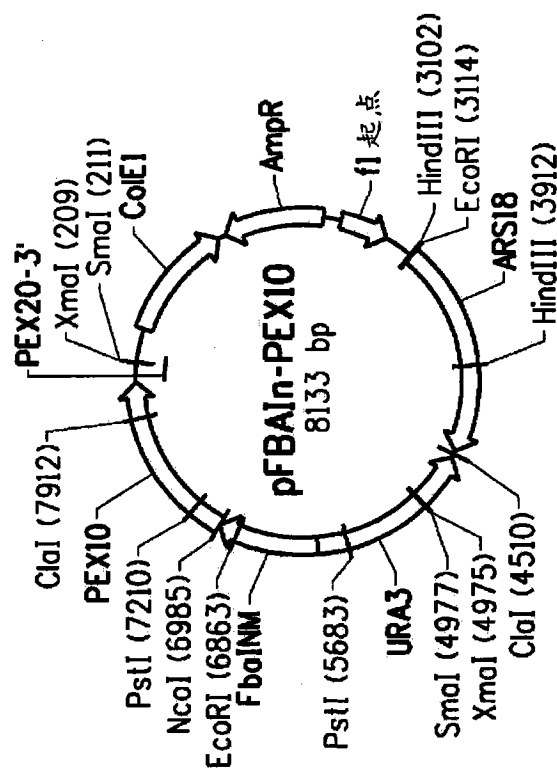


图 6A

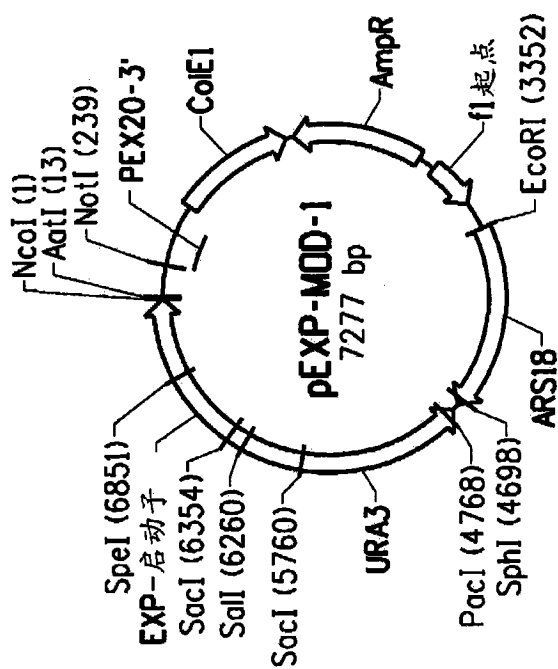


图 6B

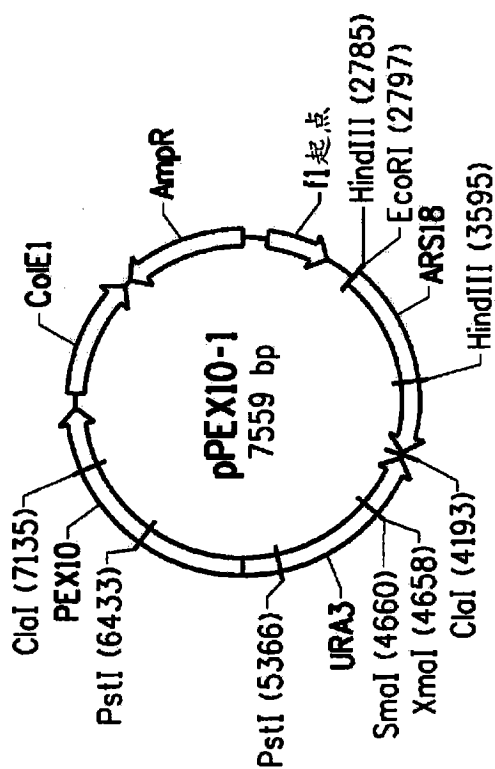


图 7A

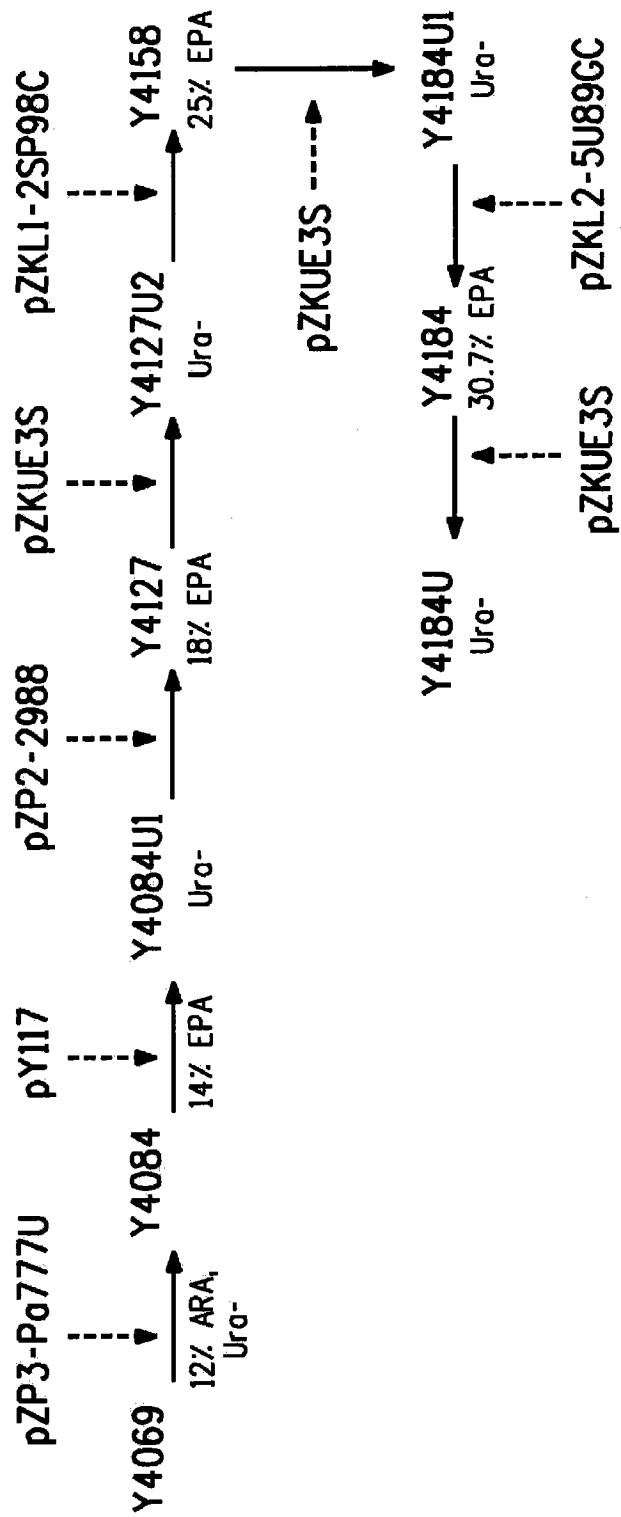


图 7B

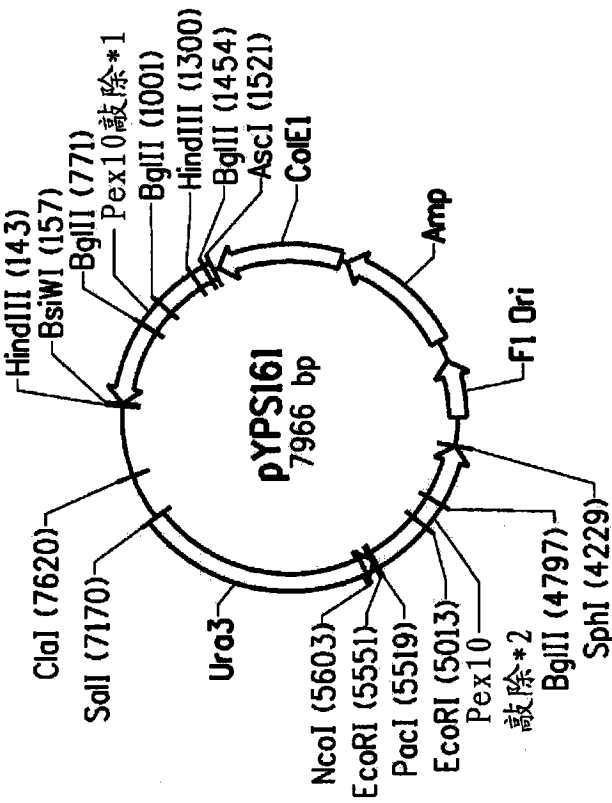


图 9A

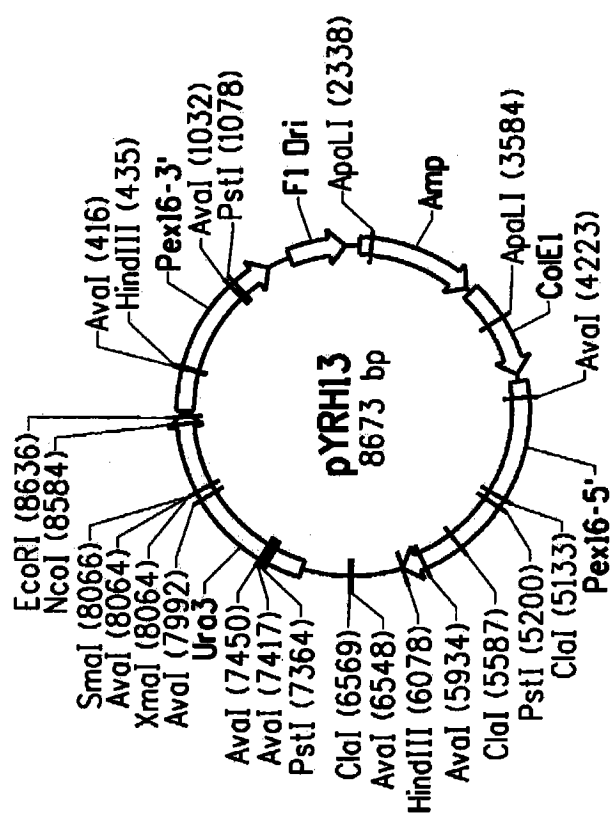


图 9B

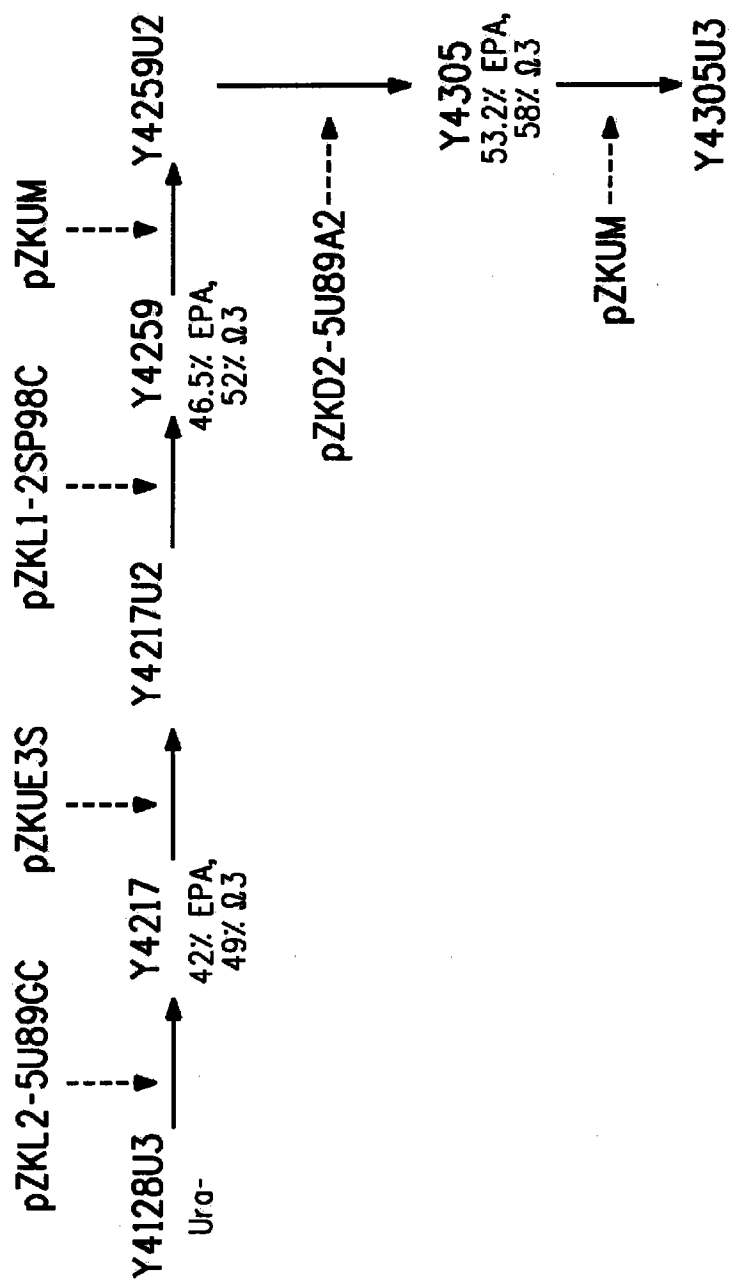


图 10

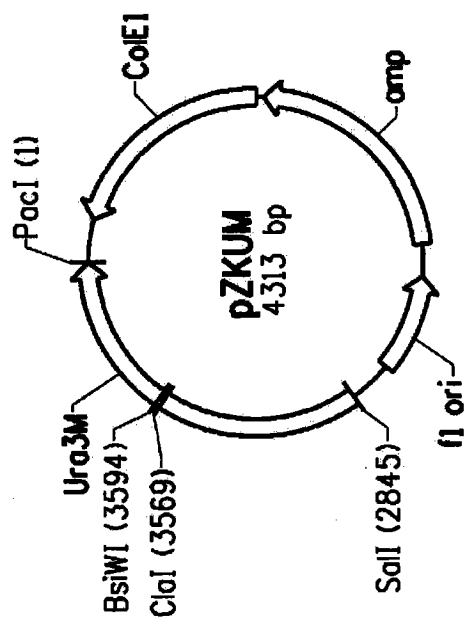


图 11A

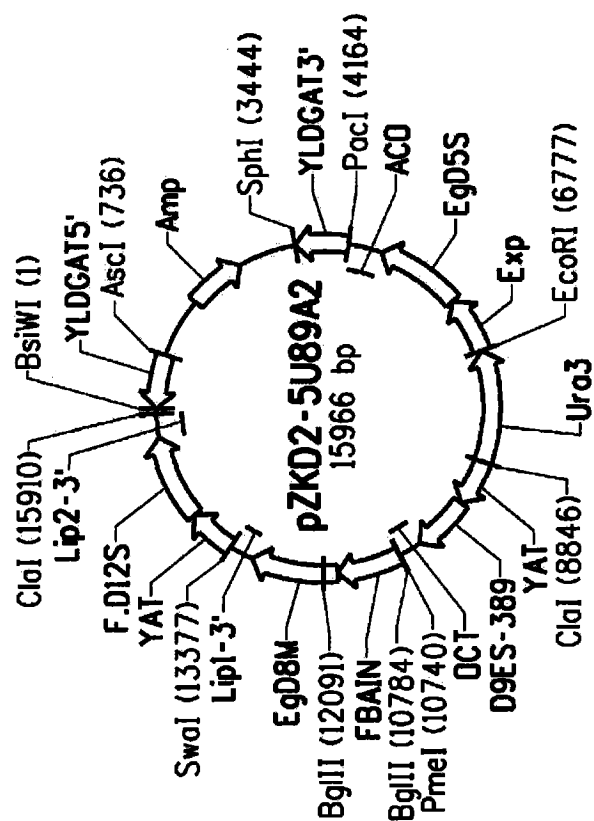


图 11B

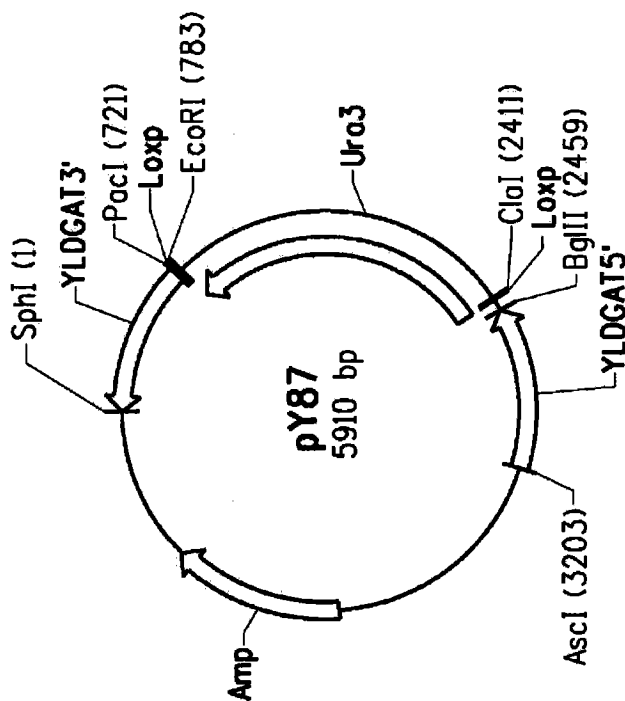


图 12A

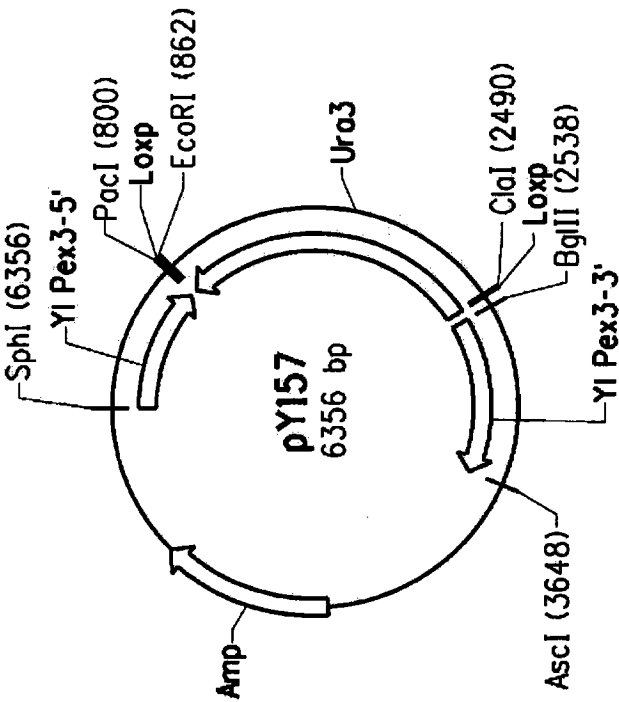


图 12B

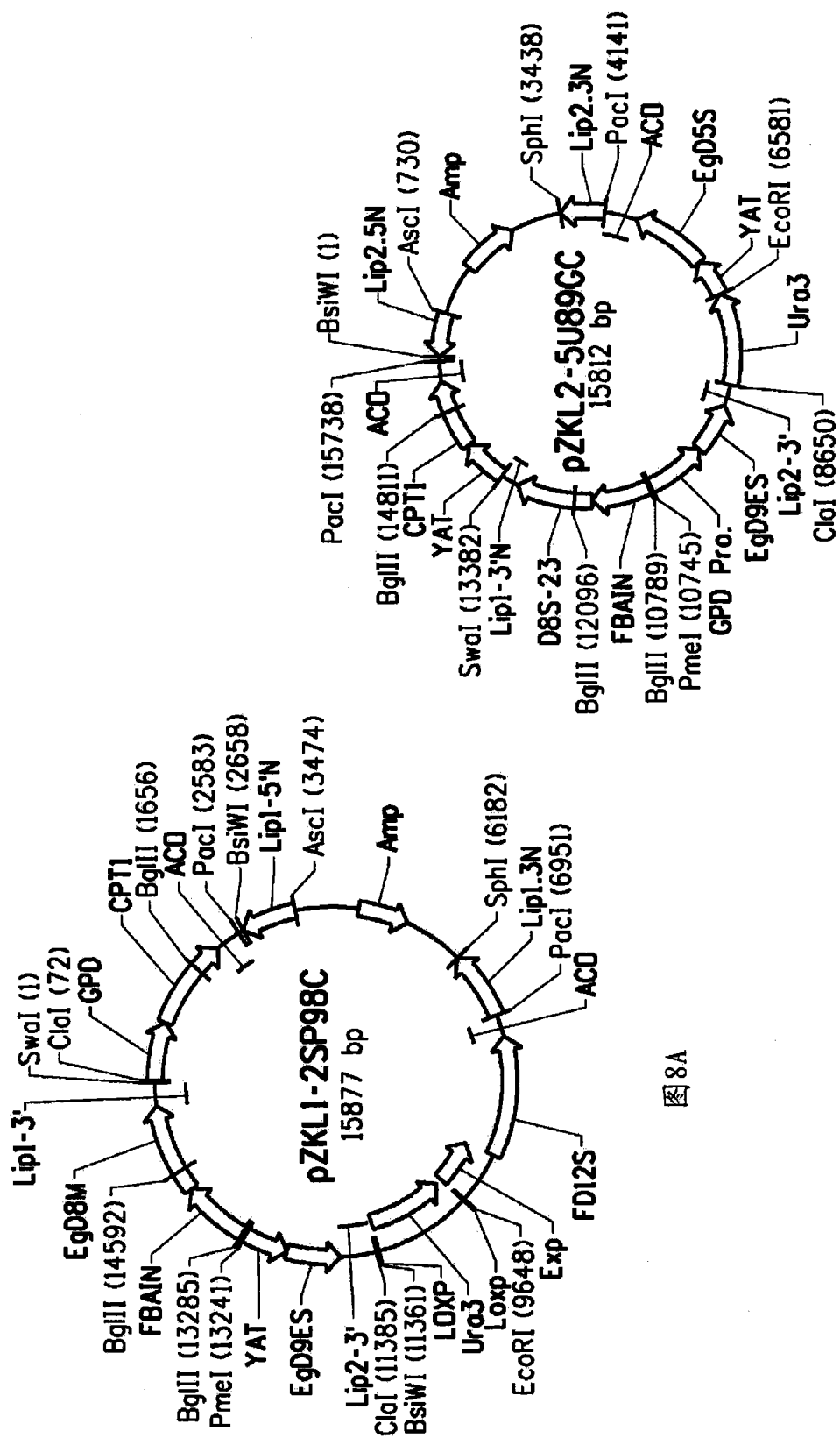


图 8A

图 8B