

(19) **DANMARK**

(10) **DK/EP 2475679 T3**



(12) **Oversættelse af
europæisk patentskrift**

Patent- og
Varemærkestyrelsen

-
- (51) Int.Cl.: **C 07 K 14/195 (2006.01)** **C 12 N 15/82 (2006.01)**
- (45) Oversættelsen bekendtgjort den: **2016-05-09**
- (80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2016-02-10**
- (86) Europæisk ansøgning nr.: **10790007.8**
- (86) Europæisk indleveringsdag: **2010-06-14**
- (87) Den europæiske ansøgnings publiceringsdag: **2012-07-18**
- (86) International ansøgning nr.: **US2010038535**
- (87) Internationalt publikationsnr.: **WO2010147904**
- (30) Prioritet: **2009-06-16 US 187366 P**
- (84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO SE SI SK SM TR**
- (73) Patenthaver: **E. I. du Pont de Nemours and Company, Chestnut Run Plaza , 974 Center Road , P.O. Box 2915, Wilmington, DE 19805, USA**
- (72) Opfinder: **HONG, Seung-Pyo, 9 Pierson's Ridge, Hockessin, Delaware 19707, USA**
SHARPE, Pamela, L., 232 Cayman Court, Wilmington, Delaware 19808, USA
XUE, Zhixiong, 111 Harvey Lane, Chadds Ford, Pennsylvania 19317-9728, USA
YADAV, Narendra, S., 126 Tuckaway Drive, Wilmington, Delaware 19803, USA
ZHU, Qun, 544 Revere Road, West Chester, Pennsylvania 19382, USA
- (74) Fuldmægtig i Danmark: **Marks & Clerk (Luxembourg) LLP, 44 rue de la Vallée, B.P. 1775, L-1017 Luxembourg, Luxembourg**
- (54) Benævnelse: **FORBEDREDE, OPTIMEREDE STAMMER AF YARROWIA LIPOLYTICA TIL FREMSTILLING AF HØJKONCENTRERET EICOSAPENTAENSYRE**
- (56) Fremdragne publikationer:
WO-A1-2009/046231
US-A1- 2006 094 092
US-A1- 2006 168 687
US-A1- 2008 254 191
US-A1- 2009 082 286
US-A1- 2009 093 543
DATABASE UniProt [Online] 16 August 2004 (2004-08-16), "SubName: Full=YALIOF19514p;", XP002692015, retrieved from EBI accession no. UNIPROT:Q6C138 Database accession no. Q6C138
YOUNG ET AL.: 'UniProtKB_Q1 MKL6: Malonyl CoA synthetase' RHIZOBIUM LEGUMINOSARUM BV. VICIAE 3841, [Online] 30 May 2006, Retrieved from the Internet: <URL:http://www.uniprot.org/uniprot/Q1MKL6> [retrieved on 2010-09-20]

Fortsættes ...

DESCRIPTION

[0001] This application claims the benefit of U.S. Provisional Applications No. 61/187366, No. 61/187368 and No. 61/187359.

FIELD OF THE INVENTION

[0002] This invention is in the field of biotechnology. More specifically, this invention pertains to an engineered strain of the oleaginous yeast *Yarrowia lipolytica* that is capable of efficiently producing eicosapentaenoic acid, an ω -3 polyunsaturated fatty acid ["PUFA"], in high concentrations.

BACKGROUND OF THE INVENTION

[0003] The clinical and pharmaceutical value of eicosapentaenoic acid ["EPA"; *cis*-5, 8, 11, 14, 17-eicosapentaenoic acid; ω -3] are well known (U.S. Pat. Appl. Pub. No. 2009-0093543-A1). Similarly, the advantages of producing EPA in microbes using recombinant means, as opposed to producing EPA from natural microbial sources or via isolation from fish oil and marine plankton, are also well recognized.

[0004] Although the literature reports a number of recent examples whereby various portions of the ω -3/ ω -6 polyunsaturated fatty acid ["PUFA"] biosynthetic pathway, responsible for EPA production, have been introduced into plants and non-oleaginous yeast, significant efforts by the Applicants' Assignee has focused on the use of the oleaginous yeast, *Yarrowia lipolytica* (U.S. Patent 7,238,482; U.S. Pat. Appl. Pub. No. 2006-0115881-A1; U.S. Pat. Appl. Pub. No. 2009-0093543-A1). Oleaginous yeast are defined as those yeast that are naturally capable of oil synthesis and accumulation, wherein oil accumulation is at least 25% of the cellular dry weight.

[0005] More specifically, U.S. Pat. Appl. Pub. No. 2006-0115881-A1 demonstrated production of 9% EPA of total fatty acids in a recombinant *Yarrowia lipolytica* strain without co-synthesis of γ -linolenic acid ["GLA"; ω -6], by expression of the following genes: $\Delta 9$ elongase, $\Delta 8$ desaturase, $\Delta 5$ desaturase, $\Delta 17$ desaturase, $\Delta 12$ desaturase and $C_{16/18}$ elongase.

[0006] U.S. Pat. Appl. Pub. No. 2009-0093543-A1 describes optimized recombinant *Yarrowia lipolytica* strains for EPA production and demonstrated production of up to 55.6% EPA of total fatty acids in a recombinant *Y. lipolytica* strain by expression of the following genes: $\Delta 9$ elongase, $\Delta 8$ desaturase, $\Delta 5$ desaturase, $\Delta 17$ desaturase, $\Delta 12$ desaturase, $C_{16/18}$ elongase and diacylglycerol cholinephosphotransferase, within a host cell comprising a disruption in the native peroxisome biogenesis factor 10 protein (PEX10).

[0007] U.S. Pat. Appl. Pub. No. 2008-254191-A1 describes a multizyme comprising for example a delta-8 desaturase and a delta-9 elongase connected by a linker for enhancing the production of PUFA. A recombinant *Yarrowia lipolytica* comprising a multizyme is disclosed.

[0008] U.S. Pat. Appl. Pub. No. 2006-0094092-A1 describes engineered recombinant *Yarrowia lipolytica* strains for ARA production, wherein the cells comprise a variety of heterologous desaturases, and/or acyltransferases.

[0009] U.S. Pat. Appl. Pub. No. 2009-0082286-A1 describes methods for enhancing the production of resveratrol in recombinant oleaginous organisms. Described therein is a recombinant *Yarrowia lipolytica* comprising a heterologous malonyl-CoA synthetase enzyme.

[0010] Despite the disclosures cited above, strain improvements are necessary for commercial production of EPA that will permit production of high EPA as a weight percent of the total fatty acids in addition to high total lipid content, while minimizing production of intermediate fatty acids, such as linoleic acid ["LA"; ω -6], and byproduct fatty acids in the final oil product. Applicants have solved the stated problem by engineering improved optimized strains of *Yarrowia lipolytica*, wherein the improvement enables at least one of the following: production of 61.8% EPA in the total oil fraction, production of 39.6% total fatty acids as a percent of the dry cell weight, or production of lipids having an EPA to LA ratio of 6.1.

SUMMARY OF THE INVENTION

[0011] In a first embodiment, the invention concerns a recombinant *Yarrowia* sp. host cell that produces an oil comprising:

1. (a) at least 50 weight percent of eicosapentaenoic acid measured as a weight percent of total fatty acids; and
2. (b) said oil has a ratio of at least 3.5 of eicosapentaenoic acid, measured as a weight percent of total fatty acids, to linoleic acid, measured as a weight percent of total fatty acids, wherein said host cell comprises:
 1. (a) at least one multizyme which comprises a polypeptide having at least one delta-9 elongase linked to at least one delta-8 desaturase;
 2. (b) at least one peroxisome biogenesis factor protein whose expression has been down-regulated; and
 3. (c) at least one recombinant construct comprising a non-native chimeric gene comprising a nucleotide sequence encoding a heterologous malonyl-CoA synthetase, and at least one recombinant construct comprising a non-native chimeric gene comprising a nucleotide sequence encoding an acyl-CoA lysophospholipid acyltransferase (LPLAT) such that the LPLAT is overexpressed in the host cell, wherein said chimeric genes comprise regulatory and coding sequences not found together in nature.

[0012] Preferably, the nucleotide sequence encoding the malonyl-CoA synthetase encodes a polypeptide consisting essentially of an amino acid sequence selected from the group consisting of SEQ ID NO:40 and SEQ ID NO:42.

[0013] Preferably, the nucleotide sequence encoding the acyl-CoA lysophospholipid acyltransferase encodes a polypeptide consisting essentially of an amino acid sequence selected from the group consisting of SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:25, SEQ ID NO:29, SEQ ID NO:31 and SEQ ID NO:32.

[0014] Preferably, the multizyme comprises a linker between the delta-9 elongase and delta-8 desaturase, which linker is selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

[0015] Preferably, the multizyme consists essentially of a sequence selected from the group consisting of: SEQ ID NO:9, SEQ ID NO:11 and SEQ ID NO:13.

[0016] In a second embodiment, the invention concerns a method for making a microbial oil comprising eicosapentaenoic acid comprising:

1. a) culturing the host cell of any of the invention wherein a microbial oil comprising eicosapentaenoic acid is produced; and,
2. b) optionally recovering the microbial oil of step (a).

BIOLOGICAL DEPOSITS

[0017] The following biological materials have been deposited with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, and bear the following designations, accession numbers and dates of deposit.

Biological Material	Accession No.	Date of Deposit
<i>Yarrowia lipolytica</i> Y8406	ATCC PTA-10025	May 14, 2009
<i>Yarrowia lipolytica</i> Y8412	ATCC PTA-10026	May 14, 2009
<i>Yarrowia lipolytica</i> Y8259	ATCC PTA-10027	May 14, 2009

[0018] The biological materials listed above were deposited under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. The listed deposit will be maintained in the indicated international depository for at least 30 years and will be made available to the public upon the grant of a patent disclosing it. The availability of a deposit does not constitute a license to practice the subject invention in derogation of patent rights granted by government action.

BRIEF DESCRIPTION OF THE DRAWINGS AND**SEQUENCE DESCRIPTIONS**

[0019]

FIG. 1A and FIG. 1B illustrate the ω -3/ ω -6 fatty acid biosynthetic pathway, and should be viewed together when considering the description of this pathway below.

FIG. 2 diagrams the development of *Yarrowia lipolytica* strains Y9481, Y9497 and Y9502, producing greater than 60.9% EPA in the total lipid fraction.

FIG. 3 provides a plasmid map for pY116.

FIG. 4 provides plasmid maps for the following: (A) pZKSL-5S5A5; and, (B) pZP3-Pa777U.

FIG. 5 provides plasmid maps for the following: (A) pZKUM; and, (B) pZKL2-5mB89C.

FIG. 6 provides plasmid maps for the following: (A) pZKL1-2SR9G85; and, (B) pZSCP-Ma83.

FIG. 7 provides plasmid maps for the following: (A) pZKL4-398F2; and, (B) pZP2-85m88F.

FIG. 8 provides plasmid maps for the following: (A) pZK16-ML8N; and, (B) pZK16-ML.

FIG. 9 diagrams the development of *Yarrowia lipolytica* strain Y8672, producing greater than 61.8% EPA in the total lipid fraction.

FIG. 10 provides plasmid maps for the following: (A) pZKL2-5m89C; and, (B) pY201, comprising a chimeric YAT1::ScAle1S::Lip1 gene.

FIG. 11 provides plasmid maps for the following: (A) pY168, comprising a chimeric YAT1::YAla1::Lip1 gene; and, (B) pY208, comprising a chimeric YAT1::MaLPAAT1S::Lip1 gene.

FIG. 12 provides plasmid maps for the following: (A) pY207, comprising a chimeric YAT1::YILPAAT1::Lip1 gene; and, (B) pY175, comprising a chimeric YAT1::CeLPCATS::Lip1 gene.

FIG. 13 provides a comparison of EPA % TFAs, LA % TFAs and the ratio of EPA % TFAs to LA % TFAs in each of the strains described in the Examples.

FIG. 14 provides plasmid maps for the following: (A) pY222, comprising a chimeric YAT1::ScLPAATS::Lip1 gene; and, (B) pY177, comprising a chimeric YAT1::YILPAAT1::Lip1 gene.

[0020] The invention can be more fully understood from the following detailed description and the accompanying sequence descriptions, which form a part of this application.

[0021] The following sequences comply with 37 C.F.R. §1.821-1.825 ("Requirements for Patent Applications Containing Nucleotide Sequences and/or Amino Acid Sequence Disclosures - the Sequence Rules") and are consistent with World Intellectual Property Organization (WIPO) Standard ST.25 (1998) and the sequence listing requirements of the EPO and PCT (Rules 5.2 and 49.5(a-bis), and Section 208 and Annex C of the Administrative Instructions). The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

[0022] SEQ ID NOs:1-156 are ORFs encoding promoters, genes or proteins (or fragments thereof) or plasmids, as identified in Table 1.

Table 1: Summary of Gene and Protein SEQ ID Numbers

Description	Nucleic acid SEQ ID NO.	Protein SEQ ID NO.
Multizyme linker GAGPARPAGLPPATYYDSLAVMGS	--	1
Multizyme linker GPARPAGLPPATYYDSLAV	--	2
Multizyme linker PARPAGLPPATYYDSLAV	--	3
Multizyme linker PTRPAGPPPATYYDSLAV	--	4
Multizyme linker PGGPGKPSIASLPPPIRPVGNPPAAYDALAT	--	5
Multizyme linker PARPAGLPPATYYDSLAVSGRT	--	6
Multizyme linker PGGPGKPSIASLPPPIRPVGNPPAAYDALATGR T	--	7
DGLA synthase, comprising EgD9eS/EgD8M gene fusion	8 (2112 bp)	9 (703 AA)
DGLA synthase, comprising EaD9eS/EaD8S gene fusion	10 (2109 bp)	11 (702 AA)
DGLA synthase, comprising E389D9eS/EgD8M gene fusion	12 (2127 bp)	13 (708 AA)
<i>Saccharomyces cerevisiae</i> Ale1 ("ScAle1"; also ORF "YOR175C")	14 (1860 bp)	15 (619 AA)
<i>Yarrowia lipolytica</i> Ale1 ("YAla1")	16 (1539 bp)	17 (512 AA)
membrane bound O-acyltransferase motif M(V/I)LxxKL	--	18
membrane bound O-acyltransferase motif FxxYxxW	--	19
membrane bound O-acyltransferase motif SAxWHG	--	20
U.S. Pat. Appl. Pub. No. 2008-0145867-A1 motif EX ₁ W _{NX} ₂ [T/V]-X ₂ W	--	21
Synthetic Ale1 derived from <i>Saccharomyces cerevisiae</i> , codon-optimized for expression in <i>Yarrowia lipolytica</i> ("ScAle1S")	22 (1870 bp)	23 (619 AA)
<i>Caenorhabditis elegans</i> LPCAT ("CeLPCAT")	24 (849 bp)	25 (282 AA)
Synthetic LPCAT derived from <i>Caenorhabditis elegans</i> , codon-optimized for expression in <i>Yarrowia lipolytica</i> ("CeLPCATS")	26 (859 bp)	27 (282 AA)
<i>Mortierella alpina</i> LPAAT1 ("MaLPAAT1")	28 (945 bp)	29 (314 AA)
<i>Yarrowia lipolytica</i> LPAAT1 ("YILPAAT1")	30 (1549 bp)	31 (282 AA)
<i>Saccharomyces cerevisiae</i> LPAAT ("ScLPAAT"; also ORF "YDL052C")	--	32 (303 AA)
1-acyl-sn-glycerol-3-phosphate acyltransferase motif NHxxxD	--	33
1-acyl-sn-glycerol-3-phosphate acyltransferase motif EGTR	--	34
Synthetic LPAAT1 derived from <i>Mortierella alpina</i> , codon-optimized for expression in <i>Yarrowia lipolytica</i> ("MaLPAAT1S")	35 (955 bp)	36 (314 AA)
<i>Yarrowia lipolytica</i> diacylglycerol cholinephosphotransferase gene ("YCPT1")	37 (1185 bp)	38 (394 AA)
<i>Rhizobium leguminosarum</i> bv. viciae 3841 malonyl-CoA synthetase (GenBank Accession No. YP_766603) ("MCS")	39 (1515 bp)	40 (504 AA)
Synthetic malonyl-CoA synthetase derived from <i>Rhizobium leguminosarum</i> bv. viciae 3841 (GenBank Accession No. YP_766603), codon-optimized for expression in <i>Yarrowia lipolytica</i> ("MCS")	41 (1518 bp)	42 (505 AA)
<i>Euglena gracilis</i> Δ 9 elongase ("EgD9e")	43 (777 bp)	44 (258 AA)
Synthetic Δ 9 elongase derived from <i>Euglena gracilis</i> , codon-optimized for expression in <i>Yarrowia lipolytica</i> ("EgD9eS")	45 (777 bp)	46 (258 AA)
<i>Eutrepitiella</i> sp. CCMP389 Δ 9 elongase ("E389D9e")	47 (792 bp)	48 (263 AA)
Synthetic Δ 9 elongase derived from <i>Eutrepitiella</i> sp. CCMP389 codon-optimized for expression in <i>Yarrowia lipolytica</i> ("E389D9eS")	49 (792 bp)	50 (263 AA)
<i>Euglena anabaena</i> UTEX 373 Δ 9 elongase ("EaD9Elo1")	51 (774 bp)	52 (258 AA)
Synthetic Δ 9 elongase derived from <i>Euglena anabaena</i> UTEX 373, codon-optimized for expression in <i>Yarrowia lipolytica</i> ("EaD9eS")	53 (774 bp)	54 (258 AA)
<i>Euglena gracilis</i> Δ 8 desaturase ("Eg5" or "EgD8")	55 (1271 bp)	56 (421 AA)

Description	Nucleic acid SEQ ID NO.	Protein SEQ ID NO.
Synthetic $\Delta 8$ desaturase derived from <i>Euglena gracilis</i> , codon-optimized for expression in <i>Yarrowia lipolytica</i> ("D8SF" or "EgD8S")	57 (1272 bp)	58 (422 AA)
Synthetic mutant $\Delta 8$ desaturase ("EgD8M"), derived from <i>Euglena gracilis</i> ("EgD8S")	59 (1272 bp)	60 (422 AA)
<i>Euglena anabaena</i> UTEX 373 $\Delta 8$ desaturase ("EaD8es3")	61 (1260 bp)	62 (420 AA)
Synthetic $\Delta 8$ desaturase derived from <i>Euglena anabaena</i> UTEX 373, codon-optimized for expression in <i>Yarrowia lipolytica</i> ("EaD8S")	63 (1260 bp)	64 (420 AA)
<i>Euglena gracilis</i> $\Delta 5$ desaturase ("EgD5")	65 (1350 bp)	66 (449 AA)
Synthetic $\Delta 5$ desaturase derived from <i>Euglena gracilis</i> , codon-optimized for expression in <i>Yarrowia lipolytica</i> ("EgD5S")	67 (1350 bp)	68 (449 AA)
Mutant $\Delta 5$ desaturase ("EgD5M"), derived from	69	70
<i>Euglena gracilis</i> ("EgD5") (U.S. Pat. Pub. No. 2010-0075386-A1)	(1350 bp)	(449 AA)
Synthetic mutant $\Delta 5$ desaturase ("EgD5SM"), derived from <i>Euglena gracilis</i> ("EgD5S") (U.S. Pat. Pub. No. 2010-0075386-A1)	71 (1350 bp)	72 (449 AA)
<i>Peridinium</i> sp. CCMP626 $\Delta 5$ desaturase ("RD5")	73 (1392 bp)	74 (463 AA)
Synthetic $\Delta 5$ desaturase derived from <i>Peridinium</i> sp. CCMP626, codon-optimized for expression in <i>Yarrowia lipolytica</i> ("RD5S")	75 (1392 bp)	76 (463 AA)
<i>Euglena anabaena</i> UTEX 373 $\Delta 5$ desaturase ("EaD5")	77 (1362 bp)	78 (454 AA)
Synthetic $\Delta 5$ desaturase derived from <i>Euglena anabaena</i> UTEX 373, codon-optimized for expression in <i>Yarrowia lipolytica</i> ("EaD5S")	79 (1362 bp)	80 (454 AA)
Synthetic mutant $\Delta 5$ desaturase ("EaD5SM"), derived from <i>Euglena anabaena</i> ("EaD5S") (U.S. Pat. Pub. No. 2010-0075386-A1)	81 (1365 bp)	82 (454 AA)
<i>Phytophthora ramorum</i> $\Delta 17$ desaturase ("PrD17")	83 (1086 bp)	84 (361 AA)
Synthetic $\Delta 17$ desaturase derived from <i>Phytophthora ramorum</i> , codon-optimized for expression in <i>Yarrowia lipolytica</i> ("PrD17S")	85 (1086 bp)	86 (361 AA)
<i>Pythium aphanidematum</i> $\Delta 17$ desaturase ("PaD17")	87 (1080 bp)	88 (359 AA)
Synthetic $\Delta 17$ desaturase derived from <i>Pythium aphanidematum</i> , codon-optimized for expression in <i>Yarrowia lipolytica</i> ("PaD17S")	89 (1080 bp)	90 (359 AA)
<i>Fusarium moniliforme</i> $\Delta 12$ desaturase ("FmD12")	91 (1434 bp)	92 (477 AA)
Synthetic $\Delta 12$ desaturase derived from <i>Fusarium moniliforme</i> , codon-optimized for expression in <i>Yarrowia lipolytica</i> ("FmD12S")	93 (1434 bp)	94 (477 AA)
<i>Mortierella alpina</i> C _{16:1a} elongase	95 (828 bp)	96 (275 AA)
Synthetic C _{16:1a} elongase derived from <i>Mortierella alpina</i> ELO3, codon-optimized for expression in <i>Yarrowia lipolytica</i> ("ME3S")	97 (828 bp)	98 (275 AA)
Shindou et al. membrane bound O-acyltransferase motif WHGxxxGYxxxF	--	99
Shindou et al. membrane bound O-acyltransferase motif YxxxF	--	100
Shindou et al. membrane bound O-acyltransferase motif YxxxYFxxdH	--	101
U.S. Pat. Appl. Pub. No. 2008-0145867-A1 motif M-[V/I]-[L/I]-xxK-[L/V/I]-xxxxxxDGG	--	102
U.S. Pat. Appl. Pub. No. 2008-0145867-A1 motif RxKYxxWxxx-[E/D]-[A/G]xxxxGxG-[F/Y]-xG	--	103
U.S. Pat. Appl. Pub. No. 2008-0145867-A1 motif SAxWHGxxPGYxx-[T/F]-F	--	104
Lewin, T.W. et al. & Yamashita et al. 1-acyl-sn-glycerol-3-phosphate acyltransferase motif GxxF-[D/R]-R	--	105
Lewin, T.W. et al. 1-acyl-sn-glycerol-3-phosphate acyltransferase motif [V/I]-[P/X]-[V/L]-[V]-P-[V/I]	--	106
Yamashita et al. 1-acyl-sn-glycerol-3-phosphate acyltransferase motif NPIVM	--	107
<i>Yarrowia lipolytica</i> Pex1p (GenBank Accession No. CAG82178)	--	108 (1024 AA)
<i>Yarrowia lipolytica</i> Pex2p (GenBank Accession No. CAG77647)	--	109 (381 AA)
<i>Yarrowia lipolytica</i> Pex3p (GenBank Accession No. CAG78565)	--	110 (431 AA)
<i>Yarrowia lipolytica</i> Pex3Bp (GenBank Accession No. CAG83356)	--	111 (395 AA)
<i>Yarrowia lipolytica</i> Pex4p (GenBank Accession No. CAG79130)	--	112 (153 AA)
<i>Yarrowia lipolytica</i> Pex5p (GenBank Accession No. CAG78803)	--	113 (598 AA)
<i>Yarrowia lipolytica</i> Pex6p (GenBank Accession No. CAG82306)	--	114 (1024 AA)
<i>Yarrowia lipolytica</i> Pex7p (GenBank Accession No. CAG78389)	--	115 (356 AA)
<i>Yarrowia lipolytica</i> Pex8p (GenBank Accession No. CAG80447)	--	116 (671 AA)
<i>Yarrowia lipolytica</i> Pex10p (GenBank Accession No. CAG81606)	--	117 (377 AA)
<i>Yarrowia lipolytica</i> Pex12p (GenBank Accession No. CAG81532)	--	118 (408 AA)
<i>Yarrowia lipolytica</i> Pex13p (GenBank Accession No. CAG81789)	--	119 (412 AA)
<i>Yarrowia lipolytica</i> Pex14p (GenBank Accession No. CAG79323)	--	120 (380 AA)
<i>Yarrowia lipolytica</i> Pex16p (GenBank Accession No. CAG79622)	--	121 (391 AA)
<i>Yarrowia lipolytica</i> Pex17p (GenBank Accession No. CAG84025)	--	122 (225 AA)
<i>Yarrowia lipolytica</i> Pex19p (GenBank Accession No. AAK84827)	--	123 (324 AA)
<i>Yarrowia lipolytica</i> Pex20p (GenBank Accession No. CAG79226)	--	124 (417 AA)
<i>Yarrowia lipolytica</i> Pex22p (GenBank Accession No. CAG77876)	--	125 (195 AA)
<i>Yarrowia lipolytica</i> Pex26p (GenBank Accession No. NC_006072, antisense translation of nucleotides 117230-118387)	--	126 (386 AA)
Plasmid pY116	127 (8739 bp)	--
Plasmid pZKSL-5S5A5	128 (13975 bp)	--
Plasmid pZP3-Pa777U	129 (13066 bp)	--
Plasmid pZKUM	130 (4313 bp)	--
Plasmid pZKL2-5mB89C	131 (15991 bp)	--
Plasmid pZKL1-2SR9G85	132 (14554 bp)	--
Plasmid pZSCP-Ma83	133 (15119 bp)	--
Plasmid pZKL4-398F2	134 (14623 bp)	--
Plasmid pZP2-85m98F	135 (14619 bp)	--
Plasmid pZK16-ML8N	136 (15262 bp)	--
Plasmid pZK16-ML	137 (13075bp)	--
Plasmid pZKL2-5m89C	138 (15799 bp)	--
Plasmid pY201	139 (9641 bp)	--
<i>Escherichia coli</i> LoxP recombination site, recognized by a Cre recombinase enzyme	140 (34 bp)	--
Primer 798	141	--
Primer 799	142	--
Primer 800	143	--
Primer 801	144	--

Description	Nucleic acid SEQ ID NO.	Protein SEQ ID NO.
Plasmid pY168	145 (9320 bp)	—
Plasmid pY208	146 (8726 bp)	—
Primer 856	147	—
Primer 857	148	—
Plasmid pY207	149 (8630 bp)	—
Plasmid pY175	150 (8630 bp)	—
Synthetic LPAAT derived from <i>Saccharomyces cerevisiae</i> , codon-optimized for expression in <i>Yarrowia lipolytica</i> ("ScLPAATS")	151 (926 bp)	152 (303 AA)
Plasmid pY222	153 (7891 bp)	—
Primer 869	154	—
Primer 870	155	—
Plasmid pY177	156 (9598 bp)	—

DETAILED DESCRIPTION OF THE INVENTION

[0023] Described herein are production host strains of *Yarrowia lipolytica* that are capable of producing greater than 50% eicosapentaenoic acid ["EPA"; 20:5 ω -3]. Accumulation of this particular polyunsaturated fatty acid ["PUFA"] is accomplished by introduction of a functional ω -3/ ω -6 fatty acid biosynthetic pathway comprising proteins with $\Delta 9$ elongase, $\Delta 8$ desaturase, $\Delta 5$ desaturase, $\Delta 17$ desaturase, $\Delta 12$ desaturase and C_{16/18} elongase activities, which thereby enables production of an EPA oil with minimal γ -linolenic acid ["GLA"]. Thus, this disclosure demonstrates that *Y. lipolytica* can be engineered to enable commercial production of EPA and derivatives thereof. Methods of production, are also claimed.

[0024] PUFAs, such as EPA (or derivatives thereof), are used as dietary substitutes, or supplements, particularly infant formulas, for patients undergoing intravenous feeding or for preventing or treating malnutrition. Alternatively, the purified PUFAs (or derivatives thereof) may be incorporated into cooking oils, fats or margarines formulated so that in normal use the recipient would receive the desired amount for dietary supplementation. The PUFAs may also be incorporated into infant formulas, nutritional supplements or other food products and may find use as anti-inflammatory or cholesterol lowering agents. Optionally, the compositions may be used for pharmaceutical use, either human or veterinary.

[0025] Supplementation of humans or animals with PUFAs produced by recombinant means can result in increased levels of the added PUFAs, as well as their metabolic progeny. For example, treatment with EPA can result not only in increased levels of EPA, but also downstream products of EPA such as eicosanoids (i.e., prostaglandins, leukotrienes, thromboxanes), docosapentaenoic acid ["DPA"; *cis*-7, 10, 13, 16, 19-docosapentaenoic; 22:5 ω -3] and docosahexaenoic acid ["DHA"; *cis*-4, 7, 10, 13, 16, 19-docosahexaenoic acid; 22:6 ω -3]. Complex regulatory mechanisms can make it desirable to combine various PUFAs, or add different conjugates of PUFAs, in order to prevent, control or overcome such mechanisms to achieve the desired levels of specific PUFAs in an individual.

[0026] Alternately, PUFAs, or derivatives thereof, made by the methodology disclosed herein can be utilized in the synthesis of animal and aquaculture feeds, such as dry feeds, semi-moist and wet feeds, since these formulations generally require at least 1-2% of the nutrient composition to be ω -3 and/or ω -6 PUFAs.

[0027] In this disclosure, a number of terms and abbreviations are used. The following definitions are provided.

"Open reading frame" is abbreviated as "ORF".

"Polymerase chain reaction" is abbreviated as "PCR".

"American Type Culture Collection" is abbreviated as "ATCC".

"Polyunsaturated fatty acid(s)" is abbreviated as "PUFA(s)".

"Triacylglycerols" are abbreviated as "TAGs".

"Co-enzyme A" is abbreviated as "CoA".

"Total fatty acids" are abbreviated as "TFAs".

"Fatty acid methyl esters" are abbreviated as "FAMES".

"Dry cell weight" is abbreviated as "DCW".

[0028] As used herein the term "invention" or "present invention" is intended to refer to all aspects and embodiments of the invention as described in the claims herein and should not be read so as to be limited to any particular embodiment or aspect.

[0029] The term "food product" refers to any food generally suitable for human consumption. Typical food products include, but are not limited to: meat products, cereal products, baked foods, snack foods, dairy products, beverages and the like. The terms "food analog", "functional food", "medical food" and "medical nutritional" are defined as in U.S. Pat. Appl. Pub. No. 2006-0115881-A1.

[0030] The term "pharmaceutical" as used herein means a compound or substance which if sold in the United States would be controlled by Section 503 or 505 of the Federal Food, Drug and Cosmetic Act.

[0031] The term "infant formula" means a food which is designed exclusively for consumption by the human infant by reason of its simulation of human breast milk. Typical commercial examples of infant formula include, but are not limited to: Similac® and Isomil®.

[0032] The term "dietary supplement" refers to a product that: (i) is intended to supplement the diet and thus is not represented for use as a conventional food or as a sole item of a meal or the diet; (ii) contains one or more dietary ingredients (including, e.g., vitamins, minerals, herbs or other botanicals, amino acids, enzymes and glandulars) or their constituents; (iii) is intended to be taken by mouth as a pill, capsule, tablet, or liquid; and, (iv) is labeled as being a dietary supplement.

[0033] The term "animal feed" refers to feeds intended exclusively for consumption by animals, including domestic animals such as pets, farm animals, etc. or for animals raised for the production of food, such as for e.g., fish farming. The terms "aquaculture feed", "aquafeed" and "feed nutrient" are as defined in U.S. Pat. Appl. Pub. No. 2006-0115881-A1.

[0034] As used herein the term "biomass" refers specifically to spent or used yeast cellular material from the fermentation of a recombinant production host producing EPA in commercially significant amounts, wherein the preferred production host is a recombinant strain of the oleaginous yeast, *Yarrowia lipolytica*. The biomass may be in the form of whole cells, whole cell lysates, homogenized cells, partially hydrolyzed cellular material, and/or partially purified cellular material (e.g., microbially produced oil).

[0035] The term "lipids" refer to any fat-soluble (i.e., lipophilic), naturally-occurring molecule. A general overview of lipids is provided in U.S. Pat. Appl. Pub. No. 2009-0093543-A1 (see Table 2 therein).

[0036] The term "glycerophospholipids" refers to a broad class of molecules, having a glycerol core with fatty acids at the sn-1 position and sn-2 position, and a polar head group (e.g., phosphate, choline, ethanolamine, glycerol, inositol, serine, cardiolipin) joined at the sn-3 position via a phosphodiester bond. Glycerophospholipids thus include phosphatidic acid ["PA"], phosphatidylcholines ["PC"], phosphatidylethanolamines ["PE"], phosphatidylglycerols ["PG"], phosphatidylinositols ["PI"], phosphatidylserines ["PS"] and cardiolipins ["CL"]. Glycerophospholipids possess tremendous diversity, not only resulting from variable phosphoryl head groups, but also as a result of differing chain lengths and degrees of saturation of their fatty acids. Generally, saturated and monounsaturated fatty acids are esterified at the sn-1 position, while polyunsaturated fatty acids are esterified at the sn-2 position.

[0037] "Lysophospholipids" are derived from glycerophospholipids, by deacylation of the sn-2 position fatty acid. Lysophospholipids include, e.g., lysophosphatidic acid ["LPA"], lysophosphatidylcholine ["LPC"], lysophosphatidylethanolamine ["LPE"], lysophosphatidylserine ["LPS"], lysophosphatidylglycerol ["LPG"] and lysophosphatidylinositol ["LPI"].

[0038] The term "oil" refers to a lipid substance that is liquid at 25 °C and usually polyunsaturated. In oleaginous organisms, oil constitutes a major part of the total lipid. "Oil" is composed primarily of

triacylglycerols ["TAGs"] but may also contain other neutral lipids, phospholipids and free fatty acids. The fatty acid composition in the oil and the fatty acid composition of the total lipid are generally similar; thus, an increase or decrease in the concentration of PUFAs in the total lipid will correspond with an increase or decrease in the concentration of PUFAs in the oil, and vice versa.

[0039] "Neutral lipids" refer to those lipids commonly found in cells in lipid bodies as storage fats and are so called because at cellular pH, the lipids bear no charged groups. Generally, they are completely non-polar with no affinity for water. Neutral lipids generally refer to mono-, di-, and/or triesters of glycerol with fatty acids, also called monoacylglycerol, diacylglycerol or triacylglycerol, respectively, or collectively, acylglycerols. A hydrolysis reaction must occur to release free fatty acids from acylglycerols.

[0040] The term "triacylglycerols" ["TAGs"] refers to neutral lipids composed of three fatty acyl residues esterified to a glycerol molecule. TAGs can contain long chain PUFAs and saturated fatty acids, as well as shorter chain saturated and unsaturated fatty acids.

[0041] The term "total fatty acids" ["TFAs"] herein refer to the sum of all cellular fatty acids that can be derivitized to fatty acid methyl esters ["FAMES"] by the base transesterification method (as known in the art) in a given sample, which may be the biomass or oil, for example. Thus, total fatty acids include fatty acids from neutral lipid fractions (including diacylglycerols, monoacylglycerols and TAGs) and from polar lipid fractions (including, e.g., the PC and the PE fractions) but not free fatty acids.

[0042] The term "total lipid content" of cells is a measure of TFAs as a percent of the dry cell weight ["DCW"], although total lipid content can be approximated as a measure of FAMES as a percent of the DCW ["FAMES % DCW"]. Thus, total lipid content ["TFAs % DCW"] is equivalent to, e.g., milligrams of total fatty acids per 100 milligrams of DCW.

[0043] The concentration of a fatty acid in the total lipid is expressed herein as a weight percent of TFAs ["% TFAs"], e.g., milligrams of the given fatty acid per 100 milligrams of TFAs. Unless otherwise specifically stated in the disclosure herein, reference to the percent of a given fatty acid with respect to total lipids is equivalent to concentration of the fatty acid as % TFAs (e.g., % EPA of total lipids is equivalent to EPA % TFAs).

[0044] In some cases, it is useful to express the content of a given fatty acid(s) in a cell as its weight percent of the dry cell weight ["% DCW"]. Thus, for example, EPA % DCW would be determined according to the following formula: (EPA % TFAs) * (TFAs % DCW)/100. The content of a given fatty acid(s) in a cell as its weight percent of the dry cell weight ["% DCW"] can be approximated, however, as: (EPA % TFAs) * (FAMES % DCW)/100.

[0045] The terms "lipid profile" and "lipid composition" are interchangeable and refer to the amount of individual fatty acids contained in a particular lipid fraction, such as in the total lipid or the oil, wherein the amount is expressed as a weight percent of TFAs. The sum of each individual fatty acid present in the mixture should be 100.

[0046] The term "extracted oil" refers to an oil that has been separated from other cellular materials, such as the microorganism in which the oil was synthesized. Extracted oils are obtained through a wide variety of methods, the simplest of which involves physical means alone. For example, mechanical crushing using various press configurations (e.g., screw, expeller, piston, bead beaters, etc.) can separate oil from cellular materials. Alternately, oil extraction can occur via treatment with various organic solvents (e.g., hexane), via enzymatic extraction, via osmotic shock, via ultrasonic extraction, via supercritical fluid extraction (e.g., CO₂ extraction), via saponification and via combinations of these methods. An extracted oil does not require that it is not necessarily purified or further concentrated. The extracted oils described herein will comprise at least 50 EPA % TFAs.

[0047] The term "blended oil" refers to an oil that is obtained by admixing, or blending, the extracted oil described herein with any combination of, or individual, oil to obtain a desired composition. Thus, for example, types of oils from different microbes can be mixed together to obtain a desired PUFA composition. Alternatively, or additionally, the PUFA-containing oils disclosed herein can be blended with fish oil, vegetable oil or a mixture of both to obtain a desired composition.

[0048] The term "fatty acids" refers to long chain aliphatic acids (alkanoic acids) of varying chain lengths, from about C₁₂ to C₂₂, although both longer and shorter chain-length acids are known. The predominant chain lengths are between C₁₆ and C₂₂. The structure of a fatty acid is represented by a simple notation system of "X:Y", where X is the total number of carbon ["C"] atoms in the particular fatty acid and Y is the number of double bonds. Additional details concerning the differentiation between "saturated fatty acids" versus "unsaturated fatty acids", "monounsaturated fatty acids" versus "polyunsaturated fatty acids" ["PUFAs"], and "omega-6 fatty acids" ["ω-6" or "n-6"] versus "omega-3 fatty acids" ["ω-3" or "n-3"] are provided in U.S. Patent 7,238,482.

[0049] Nomenclature used to describe PUFAs herein is given in Table 2. In the column titled "Shorthand Notation", the omega-reference system is used to indicate the number of carbons, the number of double bonds and the position of the double bond closest to the omega carbon, counting from the omega carbon, which is numbered 1 for this purpose. The remainder of the Table summarizes the common names of ω-3 and ω-6 fatty acids and their precursors, the abbreviations that will be used throughout the specification and the chemical name of each compound.

Table 2: Nomenclature of Polyunsaturated Fatty Acids And Precursors

Common Name	Abbreviation	Chemical Name	Shorthand Notation
Myristic	--	tetradecanoic	14:0
Palmitic	Palmitate	hexadecanoic	16:0
Palmitoleic	--	9-hexadecenoic	16:1
Stearic	--	octadecanoic	18:0
Oleic	--	cis-9-octadecenoic	18:1
Linoleic	LA	cis-9, 12-octadecadienoic	18:2 ω-6
γ-Linolenic	GLA	cis-6, 9, 12-octadecatrienoic	18:3 ω-6
Eicosadienoic	EDA	cis-11, 14-eicosadienoic	20:2 ω-6
Dihomo-γ-Linolenic	DGLA	cis-8, 11, 14-eicosatrienoic	20:3 ω-6
Arachidonic	ARA	cis-5, 8, 11, 14-eicosatetraenoic	20:4 ω-6
α-Linolenic	ALA	cis-9, 12, 15-octadecatrienoic	18:3 ω-3
Stearidonic	STA	cis-6, 9, 12, 15-octadecatetraenoic	18:4 ω-3
Eicosatrienoic	ETA	cis-11, 14, 17-eicosatrienoic	20:3 ω-3
Sciadonic	SCI	cis-5, 11, 14-eicosatrienoic	20:3b ω-6
Juniperonic	JUP	cis-5, 11, 14, 17-eicosatetraenoic	20:4b ω-3
Eicosatetraenoic	ETA	cis-8, 11, 14, 17-eicosatetraenoic	20:4 ω-3
Eicosapentaenoic	EPA	cis-5, 8, 11, 14, 17-eicosapentaenoic	20:5 ω-3
Docosatetraenoic	DTA	cis-7, 10, 13, 16-docosatetraenoic	22:4 ω-6
Docosapentaenoic	DPAn-6	cis-4, 7, 10, 13, 16-docosapentaenoic	22:5 ω-6
Docosapentaenoic	DPA	cis-7, 10, 13, 16, 19-docosapentaenoic	22:5 ω-3
Docosahexaenoic	DHA	cis-4, 7, 10, 13, 16, 19-docosahexaenoic	22:6 ω-3

[0050] The term "PUFA biosynthetic pathway" refers to a metabolic process that converts oleic acid to ω-6 fatty acids such as LA, EDA, GLA, DGLA, ARA, DTA and DPAn-6 and ω-3 fatty acids such as ALA, STA, ETrA, ETA, EPA, DPA and DHA. This process is well described in the literature (e.g., see U.S. Pat. Appl. Pub. No. 2006-0115881-A1). Briefly, this process involves elongation of the carbon chain through the addition of carbon atoms and desaturation of the molecule through the addition of double bonds, via a series of special elongation and desaturation enzymes termed "PUFA biosynthetic pathway enzymes" that are present in the endoplasmic reticulum membrane. More specifically, "PUFA biosynthetic pathway enzymes" refer to any of the following enzymes (and genes which encode said enzymes) associated with the biosynthesis of a PUFA, including: Δ4 desaturase, Δ5 desaturase, Δ6 desaturase, Δ12 desaturase, Δ15 desaturase, Δ17 desaturase, Δ9 desaturase, Δ8 desaturase, Δ9 elongase, C₁₄/16 elongase, C₁₈/20 elongase, C₁₈/20 elongase and/or C₂₀/22 elongase.

[0051] The term "Δ9 elongase/Δ8 desaturase pathway" will refer to a PUFA biosynthetic pathway that minimally includes at least one Δ9 elongase and at least one Δ8 desaturase, thereby enabling biosynthesis of DGLA and/or ETA from LA and ALA, respectively, with EDA and/or ETrA as intermediate fatty acids. With expression of other desaturases and elongases, ARA, DTA, DPAn-6, EPA, DPA and DHA may also be synthesized.

[0052] The term "desaturase" refers to a polypeptide that can desaturate, i.e., introduce a double bond, in one or more fatty acids to produce a fatty acid or precursor of interest. Despite use of the omega-reference system throughout the specification to refer to specific fatty acids, it is more convenient to indicate the activity of a desaturase by counting from the carboxyl end of the substrate using the delta-

system. Of particular interest herein are: $\Delta 8$ desaturases, $\Delta 5$ desaturases, $\Delta 17$ desaturases and $\Delta 12$ desaturases. Other useful desaturases can include $\Delta 4$ desaturases, $\Delta 6$ desaturases, $\Delta 15$ desaturases and $\Delta 9$ desaturases.

[0053] The term "elongase" refers to a polypeptide that can elongate a fatty acid carbon chain to produce an acid 2 carbons longer than the fatty acid substrate that the elongase acts upon. This process of elongation occurs in a multi-step mechanism in association with fatty acid synthase, as described in Intl. App. Pub. No. WO 2005/047480. Examples of reactions catalyzed by elongase systems are the conversion of GLA to DGLA, STA to ETA, ARA to DTA and EPA to DPA. In general, the substrate selectivity of elongases is somewhat broad but segregated by both chain length and the degree and type of unsaturation. For example, a $C_{14/16}$ elongase will utilize a C_{14} substrate (e.g., myristic acid), a $C_{16/18}$ elongase will utilize a C_{16} substrate (e.g., palmitate), a $C_{18/20}$ elongase will utilize a C_{18} substrate (e.g., GLA, STA) and a $C_{20/22}$ elongase [also referred to as a $\Delta 5$ elongase or C_{20} elongase] will utilize a C_{20} substrate (e.g., ARA, EPA). For the purposes herein, two distinct types of $C_{18/20}$ elongases can be defined: a $\Delta 6$ elongase will catalyze conversion of GLA and STA to DGLA and ETA, respectively, while a $\Delta 9$ elongase is able to catalyze the conversion of LA and ALA to EDA and ETrA, respectively.

[0054] The term "multizyme" or "fusion protein" refers to a single polypeptide having at least two independent and separable enzymatic activities, wherein the first enzymatic activity is preferably linked to the second enzymatic activity (U.S. Pat. Appl. Pub. No. 2008-0254191-A1). The "link" or "bond" between the at least two independent and separable enzymatic activities is minimally comprised of a single polypeptide bond, although the link may also be comprised of one amino acid residue, such as proline, or a polypeptide comprising at least one proline amino acid residue. Preferred linkers are selected from the group consisting of: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

[0055] The term "DGLA synthase" refers to a multizyme, wherein a $\Delta 9$ elongase is linked to a $\Delta 8$ desaturase. The term "EgD9eS/EgD8M" refers to a DGLA synthase (SEQ ID NOs:8 and 9) created by linking the $\Delta 9$ elongase "EgD9eS" (U.S. Patent 7,645,604) to the $\Delta 8$ desaturase "EgD8M" (U.S. Patent 7,709,239) with a linker sequence (i.e., SEQ ID NO:1 [GAGPARPAGLPPATYYDSLAVMGSG]; U.S. Pat. Appl. Pub. No. 2008-0254191-A1). Similarly, the term "EaD9eS/EaD8S" refers to a DGLA synthase (SEQ ID NOs:10 and 11) created by linking the $\Delta 9$ elongase "EaD9eS" (U.S. Pat. Appl. Pub. No. 2008-0254522-A1) to the $\Delta 8$ desaturase "EaD8S" (U.S. Pat. Appl. Pub. No. 2008-0254521-A1) with the linker sequence set forth as SEQ ID NO:1. And, the term "E389D9eS/EgD8M" refers to a DGLA synthase (SEQ ID NOs:12 and 13) created by linking the $\Delta 9$ elongase "E389D9eS" (U.S. Patent 7,645,604) to the $\Delta 8$ desaturase "EgD8M" (*supra*) with the linker sequence set forth as SEQ ID NO:1.

[0056] The terms "conversion efficiency" and "percent substrate conversion" refer to the efficiency by which a particular enzyme, such as a desaturase, elongase or multizyme, can convert substrate to product. The conversion efficiency is measured according to the following formula: $(\text{product}/[\text{substrate}+\text{product}]) \times 100$, where "product" includes the immediate product and all products in the pathway derived from it.

[0057] The term " C_{18} to C_{20} elongation conversion efficiency" refers to the efficiency by which $C_{18/20}$ elongases can convert C_{18} substrates (i.e., LA, ALA, GLA, STA) to C_{20} products (i.e., EDA, ETrA, DGLA, ETA). These $C_{18/20}$ elongases can be either $\Delta 9$ elongases or $\Delta 6$ elongases.

[0058] The term " $\Delta 9$ elongation conversion efficiency" refers to the efficiency by which $\Delta 9$ elongase can convert C_{18} substrates (i.e., LA, ALA) to C_{20} products (i.e., EDA, ETrA).

[0059] The term "acyltransferase" refers to an enzyme responsible for transferring an acyl group from a donor lipid to an acceptor lipid molecule.

[0060] The term "acyl-CoA:lysophospholipid acyltransferase" ["LPLAT"] refers to a broad class of acyltransferases, having the ability to acylate a variety of lysophospholipid substrates at the sn-2 position. More specifically, LPLATs include LPA acyltransferases ["LPAATs"] having the ability to catalyze conversion of LPA to PA, LPC acyltransferases ["LPCATs"] having the ability to catalyze conversion of LPC to PC, LPE acyltransferases ["LPEATs"] having the ability to catalyze conversion of LPE to PE, LPS acyltransferases ["LPSATs"] having the ability to catalyze conversion of LPS to PS, LPG acyltransferases ["LPGATs"] having the ability to catalyze conversion of LPG to PG, and LPI acyltransferases ["LPIATs"] having the ability to catalyze conversion of LPI to PI. Standardization of LPLAT nomenclature has not been formalized, so various other designations are used in the art (for example, LPAATs have also been referred to as acyl-CoA:1-acyl- sn-glycerol-3-phosphate 2-O-acyltransferases, 1-acyl- sn-glycerol-3-phosphate acyltransferases and/or 1-acylglycerolphosphate acyltransferases ["AGPATs"] and LPCATs are often referred to as acyl-CoA:1-acyl lysophosphatidyl-choline acyltransferases). Additionally, it is important to note that some LPLATs, such as the *Saccharomyces cerevisiae* Ale1 (ORF YOR175C; SEQ ID NO:15), have broad specificity and thus a single enzyme may be capable of catalyzing several LPLAT reactions, including LPAAT, LPCAT and LPEAT reactions (Tamaki, H. et al., J. Biol. Chem., 282:34288-34298 (2007); Ståhl, U. et al., FEBS Letters, 582:305-309 (2008); Chen, Q. et al., FEBS Letters, 581:5511-5516 (2007); Benghezal, M. et al., J. Biol. Chem., 282:30845-30855 (2007); Riekhof, et al., J. Biol. Chem., 282:28344-28352 (2007)).

[0061] More specifically, the term "polypeptide having at least lysophosphatidylcholine acyltransferase ["LPCAT"] activity" will refer to those enzymes capable of catalyzing the reaction: acyl-CoA + 1-acyl- sn-glycerol-3-phosphocholine $\rightarrow$ CoA + 1,2-diacyl- sn-glycerol-3-phosphocholine (EC 2.3.1.23). LPCAT activity has been described in two structurally distinct protein families, i.e., the LPAAT protein family (Hshikawa, et al., Proc. Natl. Acad. Sci. U.S.A., 105:2830-2835 (2008); Intl. App. Pub. No. WO 2004/076617) and the ALE1 protein family (Tamaki, H. et al., *supra*; Ståhl, U. et al., *supra*; Chen, Q. et al., *supra*; Benghezal, M. et al., *supra*; Riekhof, et al., *supra*).

[0062] The term "LPCAT" refers to a protein of the ALE1 protein family that: 1) has LPCAT activity (EC 2.3.1.23) and shares at least about 45% amino acid identity, based on the Clustal W method of alignment, when compared to an amino acid sequence selected from the group consisting of SEQ ID NO:15 (ScAle1) and SEQ ID NO:17 (YIAle1); and/or 2) has LPCAT activity (EC 2.3.1.23) and has at least one membrane bound O-acyltransferase ["MBOAT"] protein family motif selected from the group consisting of: M(V/I)LxxL (SEQ ID NO:18), RxxYxxW (SEQ ID NO:19), SxxWHG (SEQ ID NO:20) and EX₁₁WNx₂-[T/V]-X₂W (SEQ ID NO:21). Examples of ALE1 polypeptides include ScAle1 and YIAle1.

[0063] The term "ScAle1" refers to a LPCAT (SEQ ID NO:15) isolated from *Saccharomyces cerevisiae* (ORF "YOR175C"), encoded by the nucleotide sequence set forth as SEQ ID NO:14. In contrast, the term "ScAle1S" refers to a synthetic LPCAT derived from *S. cerevisiae* that is codon-optimized for expression in *Yarrowia lipolytica* (i.e., SEQ ID NOs:22 and 23).

[0064] The term "YIAle1" refers to a LPCAT (SEQ ID NO:17) isolated from *Yarrowia lipolytica*, encoded by the nucleotide sequence set forth as SEQ ID NO:16.

[0065] The term "LPCAT" also refers to a protein that has LPCAT activity (EC 2.3.1.23) and shares at least about 90% amino acid identity, based on the Clustal W method of alignment, when compared to an amino acid sequence as set forth in SEQ ID NO:25 (CeLPCAT).

[0066] The term "CeLPCAT" refers to a LPCAT enzyme (SEQ ID NO:25) isolated from *Caenorhabditis elegans*, encoded by the nucleotide sequence set forth as SEQ ID NO:24. In contrast, the term "CeLPCATs" refers to a synthetic LPCAT derived from *C. elegans* that is codon-optimized for expression in *Yarrowia lipolytica* (i.e., SEQ ID NOs:26 and 27).

[0067] The term "polypeptide having at least lysophosphatidic acid acyltransferase ["LPAAT"] activity" will refer to those enzymes capable of catalyzing the reaction: acyl-CoA + 1-acyl- sn-glycerol 3-phosphate $\rightarrow$ CoA + 1,2-diacyl- sn-glycerol 3-phosphate (EC 2.3.1.51).

[0068] The term "LPAAT" refers to a protein that: 1) has LPAAT activity and shares at least about 43.9% amino acid identity, based on the Clustal W method of alignment, when compared to an amino acid sequence selected from the group consisting of SEQ ID NO:29 (MaLPAAT1), SEQ ID NO:31 (YILPAAT1) and SEQ ID NO:32 (ScLPAAT1); and/or 2) has LPAAT activity and has at least one 1-acyl- sn-glycerol-3-phosphate acyltransferase family motif selected from the group consisting of: N⁺xxxD (SEQ ID NO:33) and EGTR (SEQ ID NO:34). Examples of LPAAT polypeptides include ScLPAAT, MaLPAAT1 and YILPAAT1.

[0069] The term "ScLPAAT" refers to a LPAAT (SEQ ID NO:32) isolated from *Saccharomyces cerevisiae* (ORF "YDL052C").

[0070] The term "MaLPAAT1" refers to a LPAAT (SEQ ID NO:29) isolated from *Mortierella alpina*, encoded by the nucleotide sequence set forth as SEQ ID NO:28. In contrast, the term "MaLPAAT1 S" refers to a synthetic LPAAT derived from *M. alpina* that is codon-optimized for expression in *Yarrowia lipolytica* (i.e., SEQ ID NOs:35 and 36).

[0071] The term "YILPAAT1" refers to a LPAAT (SEQ ID NO:31) isolated from *Yarrowia lipolytica*, encoded by the nucleotide sequence set forth as SEQ ID NO:30.

[0072] The term "ortholog" refers to a homologous protein from a different species that evolved from a common ancestor protein as evidenced by being in one clade of phylogenetic tree analysis and that catalyzes the same enzymatic reaction.

[0073] The term "diacylglycerol cholinephosphotransferase" refers to an enzyme (EC 2.7.8.2) that catalyzes the synthesis of phosphatidylcholines from CDP-choline and 1,2-diacylglycerols. This enzyme is part of the CDP-choline pathway, responsible for phosphatidylcholine ["PC"] biosynthesis.

[0074] The term "YICPT1" refers to a diacylglycerol cholinephosphotransferase enzyme (SEQ ID NO:38) isolated from *Yarrowia lipolytica*, encoded by SEQ ID NO:37. YICPT1 is described in Intl. App. Pub. No. WO 2006/052870 (see also GenBank Accession No. XM_501703 (YALIO10989g)).

[0075] The term "malonic acid", also referred to as propanedioic acid according to International Union of Pure and Applied Chemistry ["IUPAC"] systematic nomenclature, refers to a dicarboxylic acid having the chemical structure set forth as CH₂(COOH)₂. The malonate or propanedioate ion is derived from malonic acid by loss of two hydrogen ions (i.e., CH₂(COO)₂²⁻). Salts and esters of malonic acid include, but are not limited to, diethyl malonate [(C₂H₅)₂CH₂(C₂H₃O₄)], dimethyl malonate [(CH₃)₂CH₂(C₂H₃O₄)] and disodium malonate [Na₂(C₂H₂O₄)].

[0076] "Malonates" refer to the ionised form of malonic acid, as well as its esters and salts. All of these are referred to herein collectively as "malonates".

[0077] "Malonyl-CoA" [CAS Registry No. 524-14-1] refers to an acyl thioester that can be formed by the carboxylation of acetyl-CoA to malonyl-CoA. Alternatively, malonyl-CoA is produced enzymatically from the substrate malonate, via a malonyl-CoA synthetase.

[0078] "Malonyl-CoA synthetase" [EC 6.2.1.-] catalyzes the following enzymatic reaction: malonate + ATP + CoA → malonyl-CoA + AMP + pyrophosphate (PPi). The enzyme was first purified from malonate-grown *Pseudomonas fluorescens* (Kim, Y.S. and S.K. Bang, J. Biol. Chem., 260:5098-5104 (1985)), although various *Rhizobia* homologs have since been isolated from bacteroids within legume nodules (see, for example, Kim, Y.S. and H.Z. Chae, Biochem. J., 273:511-516 (1991) and Kim, Y.S. and S.W. Kang, Biochem. J., 297:327-333 (1994)).

[0079] As used herein, the term "rMCS" refers to a gene (SEQ ID NO:39) encoding malonyl-CoA synthetase (SEQ ID NO:40) isolated from *Rhizobium leguminosarum* bv. viciae 3841 (GenBank Accession No. YP_766603). Similarly, the term "MCS" refers to a synthetic gene encoding malonyl-CoA synthetase derived from *Rhizobium leguminosarum* bv. viciae 3841 that is codon-optimized for expression in *Yarrowia lipolytica* (i.e., SEQ ID NOs:41 and 42).

[0080] The term "peroxisomes" refers to ubiquitous organelles found in all eukaryotic cells. They have a single lipid bilayer membrane that separates their contents from the cytosol and that contains various membrane proteins essential to the functions described below. Peroxisomes selectively import proteins via an "extended shuttle mechanism". More specifically, there are at least 32 known peroxisomal proteins, called peroxins, which participate in the process of importing proteins by means of ATP hydrolysis through the peroxisomal membrane. Once cellular proteins are imported into the peroxisome, they are typically subjected to some means of degradation. For example, peroxisomes contain oxidative enzymes, such as e.g., catalase, D-amino acid oxidase and uric acid oxidase, that enable degradation of substances that are toxic to the cell. Alternatively, peroxisomes breakdown fatty acid molecules to produce free molecules of acetyl-CoA which are exported back to the cytosol, in a process called β -oxidation.

[0081] The terms "peroxisome biogenesis factor protein", "peroxin" and "Pex protein" are interchangeable and refer to proteins involved in peroxisome biogenesis and/or that participate in the process of importing cellular proteins by means of ATP hydrolysis through the peroxisomal membrane. The acronym of a gene that encodes any of these proteins is "Pex gene". A system for nomenclature is described by Distel et al., J. Cell Biol., 135:1-3 (1996). At least 32 different Pex genes have been identified so far in various eukaryotic organisms. Many Pex genes have been isolated from the analysis of mutants that demonstrated abnormal peroxisomal functions or structures. Based on a review by Kiel, J. A. K. W., et al. (Traffic, 7:1291-1303 (2006)), wherein *in silico* analysis of the genomic sequences of 17 different fungal species was performed, the following Pex proteins were identified: 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, Pex21 p, Pex21 Bp, Pex22p, Pex22p-like and Pex26p. Collectively, these proteins will be referred to herein as "Pex proteins", encoded by "Pex genes".

[0082] The term "conserved domain" or "motif" means a set of amino acids conserved at specific positions along an aligned sequence of evolutionarily related proteins. While amino acids at other positions can vary between homologous proteins, amino acids that are highly conserved at specific positions indicate amino acids that are essential in the structure, the stability, or the activity of a protein. Because they are identified by their high degree of conservation in aligned sequences of a family of protein homologues, they can be used as identifiers, or "signatures", to determine if a protein with a newly determined sequence belongs to a previously identified protein family.

[0083] The term "down-regulated" in or in connection with at least one peroxisome biogenesis factor protein refers to reduction in, or abolishment of, the activity of a native peroxisome biogenesis factor protein, as compared to the activity of the wildtype protein. Down-regulation typically occurs when a native Pex gene has a "disruption", referring to an insertion, deletion, or targeted mutation within a portion of that gene, that results in either a complete gene knockout such that the gene is deleted from the genome and no protein is translated or a translated Pex protein having an insertion, deletion, amino acid substitution or other targeted mutation. The location of the disruption in the protein may be, for example, within the N-terminal portion of the protein or within the C-terminal portion of the protein. The disrupted Pex protein will have impaired activity with respect to the Pex protein that was not disrupted, and can be non-functional. Down-regulation that results in low or lack of expression of the Pex protein, could also result via manipulating the regulatory sequences, transcription and translation factors and/or signal transduction pathways or by use of sense, antisense or RNAi technology, etc.

[0084] The term "oleaginous" refers to those organisms that tend to store their energy source in the form of oil (Weete, In: Fungal Lipid Biochemistry, 2nd Ed., Plenum, 1980). The term "oleaginous yeast" refers to those microorganisms classified as yeasts that can make oil. Generally, the cellular oil content of oleaginous microorganisms follows a sigmoid curve, wherein the concentration of lipid increases until it reaches a maximum at the late logarithmic or early stationary growth phase and then gradually decreases during the late stationary and death phases (Yongmanitchai and Ward, Appl. Environ. Microbiol., 57:419-25 (1991)). It is not uncommon for oleaginous microorganisms to accumulate in excess of about 25% of their dry cell weight as oil. Examples of oleaginous yeast include, but are not limited to, the following genera: *Yarrowia*, *Candida*, *Rhodotorula*, *Rhodospiridium*, *Cryptococcus*, *Trichosporon* and *Lipomyces*.

[0085] The term "fermentable carbon source" means a carbon source that a microorganism will metabolize to derive energy. Typical carbon sources include, but are not limited to: monosaccharides, disaccharides, oligosaccharides, polysaccharides, alkanes, fatty acids, esters of fatty acids, monoglycerides, diglycerides, triglycerides, carbon dioxide, methanol, formaldehyde, formate and carbon-containing amines.

[0086] The terms "polynucleotide", "polynucleotide sequence", "nucleic acid sequence", "nucleic acid fragment" and "isolated nucleic acid fragment" are used interchangeably herein. These terms encompass nucleotide sequences and the like. A polynucleotide may be a polymer of RNA or DNA that is single- or double-stranded, that optionally contains synthetic, non-natural or altered nucleotide bases. A polynucleotide in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA, synthetic DNA, or mixtures thereof. Nucleotides (usually found in their 5'-monophosphate form) are referred to by a single letter designation as follows: "A" for adenylate or deoxyadenylate (for RNA or DNA, respectively), "C" for cytidylate or deoxycytidylate, "G" for guanylate or deoxyguanylate, "U" for uridylate, "T" for deoxythymidylate, "R" for purines (A or G), "Y" for pyrimidines (C or T), "K" for G or T, "H" for A or C or T, "I" for inosine, and "N" for any nucleotide.

[0087] A "substantial portion" of an amino acid or nucleotide sequence is that portion comprising enough of the amino acid sequence of a polypeptide or the nucleotide sequence of a gene to putatively identify that polypeptide or gene, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul, S. F., et al., J. Mol. Biol., 215:403-410 (1993)). In general, a sequence of ten or more contiguous amino acids or thirty or more nucleotides is necessary in order to identify putatively a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene-specific oligonucleotide probes comprising 20-30 contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation, such as *in situ* hybridization of bacterial colonies or bacteriophage plaques. In addition, short oligonucleotides of 12-15 bases may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises enough of the sequence to specifically identify and/or isolate a nucleic acid fragment comprising the sequence.

[0088] The term "complementary" is used to describe the relationship between nucleotide bases that are capable of hybridizing to one another. For example, with respect to DNA, adenosine is complementary to thymine and cytosine is complementary to guanine.

[0089] "Codon degeneracy" refers to the nature in the genetic code permitting variation of the nucleotide sequence without affecting the amino acid sequence of an encoded polypeptide. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a gene for improved expression in a host cell, it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

[0090] "Synthetic genes" can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These oligonucleotide building blocks are annealed and then ligated to form gene segments that are then enzymatically assembled to construct the entire gene. Accordingly, the genes can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell, where sequence information is available. For example, the codon usage profile for *Yarrowia lipolytica* is provided in U.S. Patent 7,125,672.

[0091] "Gene" refers to a nucleic acid fragment that expresses a specific protein, and which may refer to the coding region alone or may include regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers to any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. "Endogenous gene" refers to a native gene in its natural location in the genome of an organism. A "foreign" gene refers to a gene that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, native genes introduced into a new location within the native host, or chimeric genes. A "transgene" is a gene that has been introduced into the genome by a transformation procedure. A "codon-optimized gene" is a gene having its frequency of codon usage designed to mimic the frequency of preferred codon usage of the host cell.

[0092] "Coding sequence" refers to a DNA sequence that codes for a specific amino acid sequence. "Suitable regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, enhancers, silencers, 5' untranslated leader sequence (e.g., between the transcription start site and the translation initiation codon), introns, polyadenylation recognition sequences, RNA processing sites, effector binding sites and stem-loop structures.

[0093] "Promoter" refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental or physiological

conditions. Promoters that cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths may have identical promoter activity.

[0094] The terms "3' non-coding sequence" and "transcription terminator" refer to DNA sequences located downstream of a coding sequence. This includes polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The 3' region can influence the transcription, RNA processing or stability, or translation of the associated coding sequence.

[0095] The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence. That is, the coding sequence is under the transcriptional control of the promoter. Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

[0096] The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA. Expression may also refer to translation of mRNA into a polypeptide.

[0097] "Transformation" refers to the transfer of a nucleic acid molecule into a host organism. The nucleic acid molecule may be a plasmid that replicates autonomously, for example, or, it may integrate into the genome of the host organism. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" or "recombinant" or "transformed" organisms or "transformant".

[0098] "Stable transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, including both nuclear and organellar genomes, resulting in genetically stable inheritance (i.e., the nucleic acid fragment is "stably integrated"). In contrast, "transient transformation" refers to the transfer of a nucleic acid fragment into the nucleus, or DNA-containing organelle, of a host organism resulting in gene expression without integration or stable inheritance.

[0099] The terms "plasmid" and "vector" refer to an extra chromosomal element often carrying genes that are not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA fragments. Such elements may be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear or circular, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction that is capable of introducing an expression cassette(s) into a cell.

[0100] The term "expression cassette" refers to a fragment of DNA comprising the coding sequence of a selected gene and regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence that are required for expression of the selected gene product. Thus, an expression cassette is typically composed of: 1) a promoter sequence; 2) a coding sequence (i.e., ORF) and, 3) a 3' untranslated region (i.e., a terminator) that, in eukaryotes, usually contains a polyadenylation site. The expression cassette(s) is usually included within a vector, to facilitate cloning and transformation. Different expression cassettes can be transformed into different organisms including bacteria, yeast, plants and mammalian cells, as long as the correct regulatory sequences are used for each host.

[0101] The term "sequence analysis software" refers to any computer algorithm or software program that is useful for the analysis of nucleotide or amino acid sequences. "Sequence analysis software" may be commercially available or independently developed. Typical sequence analysis software will include, but is not limited to: 1) the GCG suite of programs (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI); 2) BLASTP, BLASTN, BLASTX (Altschul et al., J. Mol. Biol., 215:403-410 (1990)); 3) DNASTAR (DNASTAR, Inc. Madison, WI); 4) Sequencher (Gene Codes Corporation, Ann Arbor, MI); and, 5) the FASTA program incorporating the Smith-Waterman algorithm (W. R. Pearson, Comput. Methods Genome Res., [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor(s): Suhai, Sandor. Plenum: New York, NY). Within the context of this application it will be understood that where sequence analysis software is used for analysis, that the results of the analysis will be based on the "default values" of the program referenced, unless otherwise specified. As used herein "default values" will mean any set of values or parameters that originally load with the software when first initialized.

[0102] Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989) (hereinafter "Maniatis"); by Silhavy, T. J., Brennan, M. L. and Enquist, L. W., Experiments with Gene Fusions, Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1984); and by Ausubel, F. M. et al., Current Protocols in Molecular Biology, published by Greene Publishing Assoc. and Wiley-Interscience, Hoboken, NJ (1987).

[0103] In general, lipid accumulation in oleaginous microorganisms is triggered in response to the overall carbon to nitrogen ratio present in the growth medium. This process, leading to the *de novo* synthesis of free palmitate (16:0) in oleaginous microorganisms, is described in detail in U.S. Patent 7,238,482. Palmitate is the precursor of longer-chain saturated and unsaturated fatty acid derivatives, which are formed through the action of elongases and desaturases (FIG. 1).

[0104] A wide spectrum of fatty acids (including saturated and unsaturated fatty acids and short-chain and long-chain fatty acids) can be incorporated into TAGs, the primary storage unit for fatty acids. In the methods and host cells described herein, incorporation of EPA into TAGs is most desirable, although the structural form of the EPA is not limiting (thus, for example, the EPA may exist in the total lipids as free fatty acids or in esterified forms such as acylglycerols, phospholipids, sulfolipids or glycolipids).

[0105] Although most PUFAs are incorporated into TAGs as neutral lipids and are stored in lipid bodies, it is important to note that a measurement of the total PUFAs within an oleaginous organism should minimally include those PUFAs that are located in the PC, PE and TAG fractions.

[0106] The metabolic process wherein oleic acid is converted to EPA involves elongation of the carbon chain through the addition of carbon atoms and desaturation of the molecule through the addition of double bonds. This requires a series of special desaturation and elongation enzymes present in the endoplasmic reticulum membrane. However, as seen in FIG. 1 and as described below, multiple alternate pathways exist for EPA production.

[0107] Specifically, FIG. 1 depicts the pathways described below. All pathways require the initial conversion of oleic acid to linoleic acid ["LA"], the first of the ω -6 fatty acids, by a Δ 12 desaturase. Then, using the " Δ 9 elongase/ Δ 8 desaturase pathway" and LA as substrate, long-chain ω -6 fatty acids are formed as follows: 1) LA is converted to eicosadienoic acid ["EDA"] by a Δ 9 elongase; 2) EDA is converted to dihomo- γ -linolenic acid ["DGLA"] by a Δ 8 desaturase; 3) DGLA is converted to arachidonic acid ["ARA"] by a Δ 5 desaturase; 4) ARA is converted to docosatetraenoic acid ["DTA"] by a C_{20/22} elongase; and, 5) DTA is converted to docosapentaenoic acid ["DPA-6"] by a Δ 4 desaturase.

[0108] The " Δ 9 elongase/ Δ 8 desaturase pathway" can also use α -linolenic acid ["ALA"] as substrate to produce long-chain ω -3 fatty acids as follows: 1) LA is converted to ALA, the first of the ω -3 fatty acids, by a Δ 15 desaturase; 2) ALA is converted to eicosatrienoic acid ["ETra"] by a Δ 9 elongase; 3) ETra is converted to eicosatetraenoic acid ["ETA"] by a Δ 8 desaturase; 4) ETA is converted to eicosapentaenoic acid ["EPA"] by a Δ 5 desaturase; 5) EPA is converted to docosapentaenoic acid ["DPA"] by a C_{20/22} elongase; and, 6) DPA is converted to docosahexaenoic acid ["DHA"] by a Δ 4 desaturase. Optionally, ω -6 fatty acids may be converted to ω -3 fatty acids. For example, ETA and EPA are produced from DGLA and ARA, respectively, by Δ 17 desaturase activity. Advantageously for the purposes herein, the Δ 9 elongase/ Δ 8 desaturase pathway enables production of an EPA oil that lacks significant amounts of γ -linolenic acid ["GLA"].

[0109] Alternate pathways for the biosynthesis of ω -3/ ω -6 fatty acids utilize a Δ 6 desaturase and C_{18/20} elongase, that is, the " Δ 6 desaturase/ Δ 6 elongase pathway". More specifically, LA and ALA may be converted to GLA and stearidonic acid ["STA"], respectively, by a Δ 6 desaturase; then, a C_{18/20} elongase converts GLA to DGLA and/or STA to ETA.

[0110] Economical commercial production of EPA in a recombinant *Yarrowia* sp. host cell requires consideration of a variety of variables, including the EPA concentration ["EPA % TFAs"] and total lipid content ["TFAs % DCW"]. Furthermore, it is desirable to reduce the production of intermediate fatty acids and byproduct fatty acids in the final oil product, in order to maximize production of the desired fatty acid, i.e., EPA.

[0111] Intermediate fatty acids are those fatty acids (e.g., oleic acid, LA, ALA, EDA, DGLA, ETA) that can be further converted to EPA by the action of other metabolic pathway enzymes. In contrast, by-product fatty acids (e.g., sciadonic acid, juniperonic acid) refer to any fatty acid produced that is neither EPA nor an intermediate fatty acid that can be further converted to EPA.

[0112] U.S. Pat. Appl. Pub. No. 2009-0093543-A1 describes optimized strains of recombinant *Yarrowia lipolytica* having the ability to produce microbial oils comprising at least about 43.3 EPA % TFAs, with less than about 23.6 LA % TFAs (an EPA:LA ratio of 1.83). The preferred strain was Y4305, whose maximum production was 55.6 EPA % TFAs, with an EPA:LA ratio of 3.03. Generally, the EPA strains of U.S. Pat. Appl. Pub. No. 2009-0093543-A1 comprised the following genes of the ω -3/ ω -6 fatty acid biosynthetic pathway:

1. a) at least one gene encoding Δ 9 elongase; and,
2. b) at least one gene encoding Δ 8 desaturase; and,
3. c) at least one gene encoding Δ 5 desaturase; and,
4. d) at least one gene encoding Δ 17 desaturase; and,
5. e) at least one gene encoding Δ 12 desaturase; and,
6. f) at least one gene encoding C_{16/18} elongase; and,
7. g) optionally, at least one gene encoding diacylglycerol cholinephosphotransferase (CPT1).

[0113] Examples of preferred genes having the enzymatic functionalities described above are set forth in Table 3 (although these genes are not intended to be limiting).

Table 3: Preferred Desaturases And Elongases For EPA Biosynthesis In *Yarrowia lipolytica*

ORF	Organism	Co-pending Patent Application References	Wildtype Abbreviation and SEQ ID NO	Codon-Optimized Abbreviation and SEQ ID NO	Mutant Abbreviation and SEQ ID NO
$\Delta 9$ elongase	<i>Euglena gracilis</i>	U.S. Patent 7,645,604	"EgD9e" (SEQ ID NOs:43 and 44)	"EgD9eS" (SEQ ID NOs:45 and 46)	--
	<i>Eutreptiella sp.</i> CCMP389	U.S. Patent 7,645,604	"E389D9e" (SEQ ID NOs:47 and 48)	"E389D9eS" (SEQ ID NOs:49 and 50)	--
	<i>Euglena anabaena</i> UTEX 373	U.S. Pat. Appl. Pub. No. 2008-0254522-A1; Intl. App. Pub. No. WO 2008/128241	"EaD9e" * (SEQ ID NOs:51 and 52)	"EaD9eS" (SEQ ID NOs:53 and 54)	--
$\Delta 8$ desaturase	<i>Euglena gracilis</i>	U.S. Patent 7,256,033; U.S. Patent 7,709,239	"EgD8" * (SEQ ID NOs:55 and 56)	"EgD8S" * (SEQ ID NOs:57 and 58)	"EgD8M" * (SEQ ID NOs:59 and 60)
	<i>Euglena anabaena</i> UTEX 373	U.S. Pat. Appl. Pub. No. 2008-0254521-A1; Intl. App. Pub. No. WO 2008/124194	"EaD8" * (SEQ ID NOs:61 and 62)	"EaD8S" (SEQ ID NOs:63 and 64)	--
$\Delta 5$ desaturase	<i>Euglena gracilis</i>	U.S. Patent 7,678,560; U.S. Pat. Pub. No. 2010-0075386-A1	"EgD5" (SEQ ID NOs:65 and 66)	"EgD5S" (SEQ ID NOs:67 and 68)	"EgD5M" (SEQ ID NOs:69 and 70); "EgD5SM" (SEQ ID NOs:71 and 72)
	<i>Peridinium sp.</i> CCMP626	U.S. Patent 7,695,950; U.S. Pat. Pub. No. 2010-0075386-A1	"RD5" (SEQ ID NOs:73 and 74)	"RD5S" (SEQ ID NOs:75 and 76)	--
	<i>Euglena anabaena</i> UTEX 373	U.S. Pat. Appl. Pub. No. 2008-0274521-A1; U.S. Pat. Pub. No. 2010-0075386-A1	"EaD5" * (SEQ ID NOs:77 and 78)	"EaD5S" * (SEQ ID NOs:79 and 80)	"EaD5SM" (SEQ ID NOs:81 and 82)
$\Delta 17$ desaturase	<i>Phytophthora ramorum</i>	U.S. Patent 7,465,793	"PrD17" (SEQ ID NOs:83 and 84)	"PrD17S" (SEQ ID NOs:85 and 86)	--
	<i>Pythium aphanidematum</i>	U.S. Patent 7,556,949	"PaD17" (SEQ ID NOs:87 and 88)	"PaD17S" (SEQ ID NOs:89 and 90)	--
$\Delta 12$ desaturase	<i>Fusarium moniliforme</i>	U.S. Patent 7,504,259	"FmD12" * (SEQ ID NOs:91 and 92)	"FmD12S" (SEQ ID NOs:93 and 94)	--
$C_{16/18}$ elongase	<i>Mortierella alpina</i>	U.S. Patent 7,470,532	"ELO3" (SEQ ID NOs:95 and 96)	"ME3S" (SEQ ID NOs:97 and 98)	--
Diacyl-glycerol choline-phosphotransferase	<i>Yarrowia lipolytica</i>	Intl. App. Pub. No. WO 2006/052870	"YCPT" (SEQ ID NOs:37 and 38)	--	--

* Notes: EaD9e was identified as "EaD9Elo1" in U.S. Pat. Appl. Pub. No. 2008-0254522-A1; EgD8 was identified as "Eg5" in U.S. Patent 7,256,033; EgD8S was identified as "D8SF" in U.S. Patent 7,256,033; EgD8M was identified as "EgD8S-23" in U.S. Patent 7,709,239; EaD8 was identified as "EaD8Des3" in U.S. Pat. Appl. Pub. No. 2008-0254521-A1; EaD5 was identified as "EaD5Des1" in U.S. Pat. Appl. Pub. No. 2008-0274521-A1; and, FmD12 was identified as "Fm2" in U.S. Patent 7,504,259.

[0114] Provided herein are optimized strains of recombinant *Yarrowia lipolytica* having the ability to produce improved microbial oils relative to those strains described in U.S. Pat. Appl. Pub. No. 2009-0093543-A1, based on the EPA % TFAs and the ratio of EPA:LA. In addition to expressing genes of the ω -3/ ω -6 fatty acid biosynthetic pathway as defined above and as detailed in U.S. Pat. Appl. Pub. No. 2009-0093543-A1, these improved strains are distinguished by:

- 1) comprising at least one multizyme, wherein said multizyme comprises a polypeptide having at least one fatty acid $\Delta 9$ elongase linked to at least one fatty acid $\Delta 8$ desaturase [a "DGLA synthase"];
- 2) comprising at least one polynucleotide encoding an enzyme selected from the group consisting of a malonyl CoA synthetase and an acyl-CoA lysophospholipid acyltransferase ["LPLAT"];
- 3) comprising at least one peroxisome biogenesis factor protein whose expression has been down-regulated;
- 4) producing at least about 50 EPA % TFAs; and,
- 5) having a ratio of EPA:LA of at least about 3.5.

[0115] U.S. Pat. Appl. Pub. No. 2008-0254191-A1, and especially Examples 55 and 56, describes DGLA synthases that possess improved enzymatic activity with respect to their individual $\Delta 9$ elongase and/or $\Delta 8$ desaturase counterparts, when heterologously expressed in *Yarrowia lipolytica*. Particularly relevant to the disclosure herein, a linker sequence (i.e., SEQ ID NO:1 [GAGPARPAGLPPATYYDSLAVMGS]) was used to fuse a $\Delta 9$ elongase (i.e., EgD9eS, EaD9eS or E389D9eS) to a $\Delta 8$ desaturase (i.e., EgD8M or EaD8S), thereby creating EgD9eS/EgD8M (SEQ ID NOs:8 and 9), EaD9eS/EaD8S (SEQ ID NOs:10 and 11) and E389D9eS/EgD8M (SEQ ID NOs:12 and 13). Surprisingly, fusing the two independent enzymes together as one fusion protein separated by a linker region increased flux from LA to DGLA, suggesting that the product of $\Delta 9$ elongase may be directly channeled as substrate of $\Delta 8$ desaturase in the fusion protein.

[0116] Table 4 below provides a summary of the improvements noted in conversion efficiency in U.S. Pat. Appl. Pub. No. 2008-0254191-A1, as a result of the gene fusion. Specifically, the number shown in bold text is the percent improvement in elongase or desaturase activity, while the details shown in parentheses provide the elongase or desaturase conversion efficiency in the gene fusion versus when the elongase or desaturase conversion efficiency when the gene was expressed alone.

Table 4: Improvement In $\Delta 9$ Elongase And $\Delta 8$ Desaturase Conversion As A Result Of Gene Fusion

Gene fusion	$\Delta 9$ Improvement	$\Delta 8$ Improvement
EgD9eS/EgD8M	5%	97%
(SEQ ID NOs:8 and 9)	(21% versus 20% conversion)	(73% versus 37% conversion)
EaD9eS/EaD8S	38%	32%
(SEQ ID NOs:10 and 11)	(18% versus 13% conversion)	(58% versus 41% conversion)
E389D9eS/EgD8M	50%	89%
(SEQ ID NOs:12 and 13)	(18% versus 12% conversion)	(70% versus 37% conversion)

[0117] Based on the results described above, expression of at least one DGLA synthase, such as the EgD9eS/EgD8M, EaD9eS/EaD8S and E389D9eS/EgD8M gene fusions described above, is preferred in improved optimized strains of recombinant *Yarrowia lipolytica* having the ability to produce improved EPA % TFAs. This gene fusion can be created using any combination of preferred $\Delta 9$ elongases and $\Delta 8$ desaturases suitable for expression in *Y. lipolytica*; and, the linker can be selected from the group consisting of: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6 and SEQ ID NO:7.

[0118] Previous studies have determined that many of the genetic mutations relating to engineering production of PUFAs in *Yarrowia lipolytica* result in increased by production of malonates during the fermentation (malonates accounted for ~45% of the total organic acids accumulated). Expression of a heterologous malonyl-CoA synthetase reversed this effect and resulted in substantially reduced by production of malonates.

[0119] More specifically, U.S. Patent Appl. Pub. No. US 2010-0159558 A1 describes generalized methods to avoid accumulation of organic acid (and in particular, malonate) "byproducts" that cannot be further utilized during a fermentation, during production of a product. This avoids carbon and energy waste within the organism, reduces the amount of base required to maintain an optimal pH range during the fermentation process, and reduces the amount of byproduct organic acids that require neutralization within the fermentation waste stream.

[0120] Malonyl-CoA synthetase [EC 6.2.1.-] catalyzes the following enzymatic reaction: malonate + ATP + CoA $\rightarrow$ malonyl-CoA + AMP + pyrophosphate (PP_i). By converting the byproduct (i.e., malonate) into malonyl-CoA, this substrate becomes available for use during the synthesis of fatty acids within the organism. Specifically, fatty acid synthesis can be summarized by the following equation (ignoring H⁺ and water): acetyl-CoA + 7 malonyl-CoA + 14 NADPH $\rightarrow$ palmitate + 7 CO₂ + 14 NADP⁺ + 8 CoA.

[0121] A codon-optimized malonyl-CoA synthetase was created and expressed in *Yarrowia lipolytica* in U.S. Patent Application No. 12/637877. Specifically, the codon-optimized malonyl-CoA synthetase gene ("MCS", SEQ ID NO:41) was designed based on the coding sequence of the malonyl-CoA synthetase gene from *Rhizobium leguminosarum* bv. viciae 3841 (rMCS; SEQ ID NOs:39 and 40, corresponding to GenBank Accession No. YP_766603). In addition to modification of the translation initiation site, 233 bp of the 1515 bp coding region (including the stop codon) were modified (15.4%), 219 codons were

optimized (43.4%), the GC content was reduced from 61.4% within the wild type gene (i.e., rMCS) to 55.6% within the synthetic gene (i.e., MCS) and the translation initiation codon "ATG" was added in front of the rMCS gene (SEQ ID NO:39) since *Yarrowia* cannot use the "GTG" codon for translation initiation. The codon-optimized MCS gene (SEQ ID NO:41) is 1518 bp encoding a peptide of 505 amino acids and a stop codon (SEQ ID NO:42).

[0122] Expression of MCS (SEQ ID NO:42) in *Yarrowia lipolytica* strain Y4305U, producing 49 EPA % TFAs, lowered the total amount of malonates (g/g DCW) -94% without impacting either the fatty acid profile or the total lipid yield (TFAs % DCW).

[0123] Based on the results described above, expression of at least one malonyl-CoA synthetase in improved optimized strains of recombinant *Yarrowia lipolytica* is desirable, as a means to reduce generation of unwanted byproducts and thereby decrease the cost of manufacture. Preferred malonyl-CoA synthetases are set forth as SEQ ID NOs:40 and 42, but these are should not be limiting to the disclosure herein. One skilled in the art could readily identify alternate heterologous malonyl-CoA synthetases suitable for expression in *Y. lipolytica*.

[0124] Glycerophospholipids, the main component of biological membranes, contain a glycerol core with fatty acids attached as R groups at the sn-1 position and sn-2 position, and a polar head group joined at the sn-3 position via a phosphodiester bond. Table 5 below summarizes the steps in the *de novo* biosynthetic pathway, originally described by Kennedy and Weiss (J. Biol. Chem., 222:193-214 (1956)):

Table 5. General Reactions Of *de Novo* Glycerophospholipid Biosynthesis

sn-Glycerol-3-Phosphate → Lysophosphatidic Acid (1-acyl-sn-glycerol 3-phosphate or "LPA")	Glycerol-3-phosphate acyltransferase (GPAT) [E.C. 2.3.1.15] esterifies 1 st acyl-CoA to sn-1 position of sn-glycerol 3-phosphate
LPA → Phosphatidic Acid (1,2-diacylglycerol phosphate or "PA")	Lysophosphatidic acid acyltransferase (LPAAT) [E.C. 2.3.1.51] esterifies 2nd acyl-CoA to sn-2 position of LPA
PA → 1,2-Diacylglycerol ("DAG")	Phosphatidic acid phosphatase [E.C. 3.1.3.4] removes a phosphate from PA; DAG can subsequently be converted to phosphatidylcholines ["PC"], phosphatidylethanolamines ["PE"] or TAG (TAG synthesis requires either a diacylglycerol acyltransferase (DGAT) [E.C. 2.3.1.20] or a phospholipid:diacylglycerol acyltransferase (PDAT) [E.C.2.3.1.158])
Or	
PA → Cytidine Diphosphate Diacylglycerol ("CDP-DG")	CDP-diacylglycerol synthase [EC 2.7.7.41] causes condensation of PA and cytidine triphosphate, with elimination of pyrophosphate; CDP-DG can subsequently be converted to phosphatidylglycerols ["PG"], phosphatidylinositols ["PI"], phosphatidylserines ["PS"] or cardiolipins ["CL"]

[0125] Following their *de novo* synthesis, glycerophospholipids can undergo rapid turnover of their fatty acyl composition at the sn-2 position. This "remodeling", or "acyl editing", has been attributed to deacylation of the glycerophospholipid and subsequent reacylation of the resulting lysophospholipid.

[0126] In the Lands' cycle (Lands, W.E., J. Biol. Chem., 231:883-888 (1958)), remodeling occurs through the concerted action of: 1) a phospholipase, such as phospholipase A₂, that releases fatty acids from the sn-2 position of phosphatidylcholine; and, 2) acyl-CoA:lysophospholipid acyltransferases ["LPLATs"], such as lysophosphatidylcholine acyltransferase ["LPCAT"] that reacylates the lysophosphatidylcholine ["LPC"] at the sn-2 position. Other glycerophospholipids can also be involved in the remodeling with their respective lysophospholipid acyltransferase activity, including LPLAT enzymes having lysophosphatidylethanolamine acyltransferase ["LPEAT"] activity, lysophosphatidylserine acyltransferase ["LPSAT"] activity, lysophosphatidylglycerol acyltransferase ["LPGAT"] activity and lysophosphatidylinositol acyltransferase ["LPIAT"] activity. In all cases, LPLATs are responsible for removing acyl-CoA fatty acids from the cellular acyl-CoA pool and acylating various lysophospholipid substrates at the sn-2 position in the phospholipid pool. Finally, LPLATs also include LPAAT enzymes that are involved in the *de novo* biosynthesis of PA from LPA.

[0127] In other cases, this sn-2 position remodeling has been attributed to the forward and reverse reactions of enzymes having LPCAT activity (Stymne S. and A.K. Stobart, Biochem J., 223(2):305-314 (1984)).

[0128] Several recent reviews by Shindou et al. provide an overview of glycerophospholipid biosynthesis and the role of LPLATs (J. Biol. Chem., 284(1):1-5 (2009); J. Lipid Res., 50:S46-S51 (2009)). And, numerous LPLATs have been reported in public and patent literature, based on the presence of conserved motifs.

[0129] More specifically, a variety of LPLAT motifs have been proposed, with slight variation based on the specific species that are included in analyzed alignments. For example, Shindou et al. (Biochem. Biophys. Res. Comm., 383:320-325 (2009)) proposed the following membrane bound O-acyltransferase ["MBOAT"] family motifs to be important for LPLAT activity, based on alignment of sequences from *Homo sapiens*, *Gallus gallus*, *Danio rerio* and *Caenorhabditis elegans*: WD, WHGxxGYYxxF (SEQ ID NO:98), YxxxF (SEQ ID NO:100) and YxxYFxxH (SEQ ID NO:101). Of these, the WD, WHGxxGYYxxF and YxxxF motifs are present in ScAe (SEQ ID NO:15) and YIAe1 (SEQ ID NO:17), but the YxxYFxxH motif is not. Alternate non-plant motifs for Ale1 homologs are also described in U.S. Pat. Appl. Pub. No. 2008-0145867-A1; specifically, these include: M-[V/I]-[L/I]-xxK-[L/V/I]-xxxxDG (SEQ ID NO:102), RxKYxxWxx-[E/D]-[A/G]xxxGxG-[F/Y]-xG (SEQ ID NO:103), EX11WNx₂-[T/V]-X₂W (SEQ ID NO:21) and SxWHGxxPGYxx-[T/F]-F (SEQ ID NO:104).

[0130] Similarly, Lewin, T.W. et al. (Biochemistry, 38:5764-5771 (1999)) and Yamashita et al. (Biochim. Biophys. Acta, 1771:1202-1215 (2007)) proposed the following 1-acyl-sn-glycerol-3-phosphate acyltransferase ["LPAAT"] family motifs to be important for LPLAT activity, based on alignment of sequences from bacteria, yeast, nematodes and mammals: NfxxxD (SEQ ID NO:33), GxH-[D/R]-R (SEQ ID NO:105), EGTR (SEQ ID NO:34) and either [V/I]-[P/X]-[V/L]-[V]-P-[V/I] (SEQ ID NO:106) or IPIVM (SEQ ID NO:107). The NfxxxD and EGTR motifs are present in MaLPAAT1 (SEQ ID NO:29), YILPAAT1 (SEQ ID NO:31) and CeLPCAT (SEQ ID NO:25), but the other LPAAT family motifs are not.

[0131] Based on publicly available Ale1, LPCAT and LPAAT protein sequences, including those described herein, LPLATs for inclusion in the improved optimized strains of recombinant *Yarrowia lipolytica* herein possess either MBOAT family motifs selected from the group consisting of: M([V/I])LxxKL (SEQ ID NO:18), RxKYxxW (SEQ ID NO:19), SxWHG (SEQ ID NO:20) and EX11WNx₂-[T/V]-X₂W (SEQ ID NO:21) or 1-acyl-sn-glycerol-3-phosphate acyltransferase family motifs selected from the group consisting of: NfxxxD (SEQ ID NO:33) and EGTR (SEQ ID NO:34).

[0132] The effect of LPLATs on PUFA production has been contemplated, since fatty acid biosynthesis requires rapid exchange of acyl groups between the acyl-CoA pool and the phospholipid pool. Specifically, desaturations occur mainly at the sn-2 position of phospholipids, while elongation occurs in the acyl-CoA pool.

[0133] More specifically, it has been previously hypothesized that LPCATs could be important in the accumulation of EPA in the TAG fraction of *Yarrowia lipolytica* (U.S. Pat. Appl. Pub. No. 2006-0115881-A1). As described therein, this hypothesis was based on the following studies: 1) Stymne S. and A.K. Stobart (Biochem. J., 223(2):305-314(1984)), who hypothesized that the exchange between the acyl-CoA pool and PC pool may be attributed to the forward and backward reaction of LPCAT; 2) Domergue, F. et al. (J. Bio. Chem., 278:35115 (2003)), who suggested that accumulation of GLA at the sn-2 position of PC and the inability to efficiently synthesize ARA in yeast was a result of the elongation step involved in PUFA biosynthesis occurring within the acyl-CoA pool, while the Δ5 and Δ6 desaturation steps occurred predominantly at the sn-2 position of PC; 3) Abbadi, A. et al. (The Plant Cell, 16:2734-2748 (2004)), who suggested that LPCAT plays a critical role in the successful reconstitution of a Δ6 desaturase/Δ6 elongase pathway, based on analysis on the constraints of PUFA accumulation in transgenic oilseed plants; and, 4) the work of Renz, A. et al. in Intl. App. Publications No. WO 2004/076617 A2 and No. WO 2004/087902 A2.

[0134] More specifically, Intl. App. Pub. No. WO 2004/076617 A2 describes the isolation of a LPCAT from *Caenorhabditis elegans* (clone T06E8.1) ["CeLPCAT"] and reports increase in the efficiency of Δ6 desaturation and Δ6 elongation, as well as an increase in biosynthesis of the long-chain PUFAs eicosadienoic acid ["EDA"; 20:2] and eicosatetraenoic acid ["ETA"; 20:4], respectively, when the LPCAT was expressed in an engineered strain of *Saccharomyces cerevisiae* that was fed exogenous 18:2 or α-linolenic ["ALA"; 18:3] fatty acids, respectively.

[0135] Intl. App. Pub. No. WO 2004/087902 A2 (Example 16) describes the isolation of *Mortierella alpina* LPAAT-like proteins (encoded by proteins having 417 amino acids in length or 389 amino acids in length, respectively, that are identical except for an N-terminal extension of 28 amino acid residues) and reports expression of one of these proteins using similar methods to those of Intl. App. Pub. No. WO 2004/076617 A2, which results in similar improvements in EDA and ETA biosynthesis.

[0136] Both Intl. App. Publications No. WO 2004/076617 and No. WO 2004/087902 teach that the improvement in EDA and ETA biosynthesis is due to reversible LPCAT activity in CeLPLAT and some LPAAT-like proteins, although not all LPAAT-like proteins have LPCAT activity. Furthermore, Renz, A. et al. concluded that LPCAT allowed efficient and continuous exchange of the newly synthesized fatty acids between phospholipids and the acyl-CoA pool, since desaturases catalyze the introduction of double bonds in PC-coupled fatty acids while elongases exclusively catalyze the elongation of CoA esterified fatty acids (acyl-CoAs).

[0137] Numerous other references generally describe benefits of co-expressing LPLATs with PUFA biosynthetic genes, to increase the amount of a desired fatty acid in the oil of a transgenic organism, increase total oil content or selectively increase the content of desired fatty acids (e.g., Intl. App. Publications No. WO 2006/069936, No. WO 2006/052870, No. WO 2009/001315, No. WO 2009/014140).

[0138] Herein, it is demonstrated that LPAAT and LPCAT are indeed important in the accumulation of EPA in the TAG fraction of *Yarrowia lipolytica*. Specifically, it was found that over-expression of LPLATs can result in an improvement in the Δ9 elongase conversion efficiency. As previously defined, conversion efficiency is a term that refers to the efficiency by which a particular enzyme, such as a Δ9 elongase, can convert substrate (e.g., LA) to product (e.g., EDA). Thus, in a strain engineered to produce EPA, improvement in Δ9 elongase conversion efficiency was demonstrated to result in increased EPA % TFAs and/or EPA % DCW.

[0139] These results, and additional supporting work, are the cornerstone of a method disclosed below for improving C₁₈ to C₂₀ elongation conversion efficiency in a LC-PUFA-producing recombinant

oleaginous microbial host cell, wherein said method comprises:

1. a) introducing into said LC-PUFA-producing recombinant host cell at least one isolated polynucleotide encoding a polypeptide having at least acyl-CoA:lysophospholipid acyltransferase activity wherein the polypeptide is selected from the group consisting of:
 1. (i) a polypeptide having at least 45% amino acid identity, based on the Clustal W method of alignment, when compared to an amino acid sequence selected from the group consisting of SEQ ID NO:15 (ScA1e1) and SEQ ID NO:17 (YIA1e1);
 2. (ii) a polypeptide having at least one membrane bound O-acyltransferase protein family motif selected from the group consisting of: M(V/I)LxxKL (SEQ ID NO:18), RxKYxxW (SEQ ID NO:19), SxWVHG (SEQ ID NO:20) and EX₁W_NX₂[T/V]-X₂W (SEQ ID NO:21);
 3. (iii) a polypeptide having at least 90% amino acid identity, based on the Clustal W method of alignment, when compared to an amino acid sequence as set forth in SEQ ID NO:25 (CeLPCAT);
 4. (iv) a polypeptide having at least 43.9% amino acid identity, based on the Clustal W method of alignment, when compared to an amino acid sequence selected from the group consisting of SEQ ID NO:29 (MaLPAAT1), SEQ ID NO:31 (YILPAAT1) and SEQ ID NO:32 (ScLPAAT1); and,
 5. (v) a polypeptide having at least one 1-acyl-sn-glycerol-3-phosphate acyltransferase protein family motif selected from the group consisting of: N_hxxxD (SEQ ID NO:33) and EGTR (SEQ ID NO:34);
 wherein the at least one isolated polynucleotide encoding a polypeptide having at least acyl-CoA:lysophospholipid acyltransferase activity is operably linked to at least one regulatory sequence, said regulatory sequence being the same or different; and,
2. b) growing the oleaginous microbial host cell;

wherein the C₁₈ to C₂₀ elongation conversion efficiency of the oleaginous microbial host cell is increased relative to the control host cell.

[0140] Preferably, the polynucleotide encoding a polypeptide having at least acyl-CoA:lysophospholipid acyltransferase activity is stably integrated and the an increase in C₁₈ to C₂₀ elongation conversion is at least about 4%,

[0141] More preferred, the increase in C₁₈ to C₂₀ elongation conversion efficiency is at least about 4-10%, more preferred at least about 10-20%, more preferred at least about 20-40%, and most preferred at least about 40-60% or greater in at least one LC-PUFA-producing oleaginous microbial host cell when compared to the control host cell.

[0142] Based on the improvement in C₁₈ to C₂₀ elongation conversion efficiency described above, optimized strains of recombinant *Yarrowia lipolytica* having the ability to produce improved EPA % TFAs, relative to those strains described in U.S. Pat. Appl. Pub. No. 2009-0093543-A1, will optionally comprise at least one acyl-CoA:lysophospholipid acyltransferase ["LPLAT"] as defined in the methods described above. As described herein, the amino acid sequence of the LPLAT is selected from the group consisting of: SEQ ID NO:15 (ScA1e1), SEQ ID NO:16 (YIA1e1), SEQ ID NO:25 (CeLPCAT), SEQ ID NO:29 (MaLPAAT1), SEQ ID NO:31 (YILPAAT1) and SEQ ID NO:32 (ScLPAAT1).

[0143] U.S. Pat. Appl. Pub. No. 2009-0093543-A1 describes a variety of knockouts useful in recombinant *Yarrowia* sp., including those useful for selection of transformants, those that diminish fatty acid degradation and TAG degradation and those that appear to result in a phenotypically "neutral" mutation (wherein the *Yarrowia* host cell seems unaffected). Most preferred, however, are those gene knockouts (e.g., diacylglycerol acyltransferase gene knockouts, peroxisome biogenesis factor protein ["PEX"] gene knockouts) that result in increases in the concentration of EPA relative to the total fatty acids ["EPA % TFAs"].

[0144] More specifically, U.S. Pat. Appl. Pub. No. 2009-0093543-A1 contemplates that in some preferred recombinant *Yarrowia* production hosts, the host is devoid of a native gene encoding a peroxisome biogenesis factor protein selected from the group consisting of: Pex1p (SEQ ID NO:108), Pex2p (SEQ ID NO:109), Pex3p (SEQ ID NO:110), Pex3Bp (SEQ ID NO:111), Pex4p (SEQ ID NO:112), Pex5p (SEQ ID NO:113), Pex6p (SEQ ID NO:114), Pex7p (SEQ ID NO:115), Pex8p (SEQ ID NO:116), Pex10p (SEQ ID NO:117), Pex12p (SEQ ID NO:118), Pex13p (SEQ ID NO:119), Pex14p (SEQ ID NO:120), Pex16p (SEQ ID NO:121), Pex17p (SEQ ID NO:122), Pex19p (SEQ ID NO:123), Pex20p (SEQ ID NO:124), Pex22p (SEQ ID NO:125) and Pex26p (SEQ ID NO:126). More preferred disrupted peroxisome biogenesis factor proteins are Pex2p, Pex3p, Pex10p, Pex12p and Pex16p, although data is provided only concerning Pex10p.

[0145] Intl. App. Pub. No. WO 2009/046248 confirms and expands the hypotheses and studies presented in U.S. Pat. Appl. Pub. No. 2009-0093543-A1, by comparing ΔPex16, ΔPex3 and ΔPex10 strains of *Yarrowia lipolytica*. Results therein demonstrated that Pex10 disruption was responsible for a 3.3 fold increase in EPA % TFAs and a 1.7 fold increase in the amount of C₂₀ PUFAs relative to the non-disrupted strain engineered for EPA production. Similarly, a 1.65 fold increase in DGLA % TFAs and a 1.3 fold increase in C₂₀ PUFAs % TFAs was observed in a ΔPex16 strain engineered for DGLA production. A 2.0 fold increase in DGLA % TFAs and a 1.7 fold increase in C₂₀ PUFAs % TFAs was observed in a ΔPex3 strain engineered for DGLA production.

[0146] These results, and additional supporting work, are the cornerstone of a method disclosed below for increasing the weight percent of one PUFA or a combination of PUFAs relative to the weight percent of total fatty acids in an oleaginous eukaryotic organism having a total lipid content, a total lipid fraction and an oil fraction, said method comprising:

1. a) providing an oleaginous eukaryotic organism comprising a disruption in a native gene encoding a peroxisome biogenesis factor protein, which creates a PEX-disruption organism; and genes encoding a functional PUFA biosynthetic pathway; and,
2. b) growing the PEX-disrupted organism under conditions wherein the weight percent of at least one PUFA is increased in the total lipid fraction and in the oil fraction relative to the weight percent of the total fatty acids, when compared with those weight percents in an oleaginous eukaryotic organism whose native peroxisome biogenesis factor protein has not been disrupted.

The amount of PUFAs that increases as a percent of total fatty acids can be: 1) the PUFA that is the desired end product of a functional PUFA biosynthetic pathway, as opposed to PUFA intermediates or by-products; 2) C₂₀ to C₂₂ PUFAs; and/or, 3) total PUFAs.

[0147] In addition to the increase in the weight percent of one or a combination of PUFAs relative to the weight percent of the total fatty acids, in some cases, the total lipid content (TFA % DCW) of the cell may be increased or decreased. What this means is that regardless of whether the disruption in the PEX gene causes the amount of total lipids in the PEX-disrupted cell to increase or decrease, the disruption always causes the weight percent of a PUFA or of a combination of PUFAs to increase.

[0148] Based on the above, optimized strains of recombinant *Yarrowia lipolytica* having the ability to produce improved EPA % TFAs, relative to those strains described in U.S. Pat. Appl. Pub. No. 2009-0093543-A1, will comprise at least one peroxisome biogenesis factor protein whose expression has been down-regulated (i.e., thereby producing a PEX-disrupted organism). In preferred strains, the down-regulated peroxisome biogenesis factor protein is Pex3p (SEQ ID NO:110), Pex10p (SEQ ID NO:117), or Pex16p (SEQ ID NO:121).

[0149] Although numerous techniques are available to one of skill in the art to achieve disruption of a native *Yarrowia* gene, generally the endogenous activity of a particular gene can be reduced or eliminated by the following techniques, for example: 1) disrupting the gene through insertion, substitution and/or deletion of all or part of the target gene; or, 2) manipulating the regulatory sequences controlling the expression of the protein. Both of these techniques are discussed in U.S. Pat. Appl. Pub. No. 2009-0093543-A1, as well as Intl. App. Pub. No. WO 2009/046248. One skilled in the art would appreciate that these and other methods are well described in the existing literature and are not limiting to the methods, host cells, and products described herein. One skilled in the art will also appreciate the most appropriate technique for use with any particular oleaginous yeast.

[0150] The optimized strains will produce at least about 40-50 EPA % TFAs, preferably at least about 50-55 EPA % TFAs, more preferably at least about 55-60 EPA % TFAs, more preferably at least 60-70 EPA % TFAs, and most preferably at least about 70-80 EPA % TFAs.

[0151] As will be clear to one of skill in the art, a multitude of different optimized *Yarrowia* strains producing at least about 50 EPA % TFAs could be engineered using the methodologies described herein. Selection of a preferred strain for commercial purposes will therefore also consider the total lipid content of the engineered strain, since both the concentration of EPA as a percent of the total fatty acids ["EPA % TFAs"] and total lipid content ["TFAs % DCW"] affect the cellular content of EPA as a percent of the dry cell weight ["EPA % DCW"]. That is, EPA % DCW is calculated as: (EPA % TFAs) * (TFA % DCW)/100. For example, a strain producing 50 EPA % TFAs and having 24 TFAs % DCW, a strain producing 55 EPA % TFAs and having 21.82 TFAs % DCW, a strain producing 60 EPA % TFAs and having 20 TFAs % DCW, a strain producing 65 EPA % TFAs and having 18.46 TFAs % DCW and a strain producing 70 EPA % TFAs and having 17.14 TFAs % DCW all produce 12 EPA % DCW. As described herein, an improved optimized strain of *Yarrowia lipolytica* may produce at least about 10-12 EPA % DCW, preferably at least about 12-14 EPA % DCW, more preferably at least about 14-16 EPA % DCW, more preferably at least about 16-18 EPA % DCW, more preferably at least about 18-20 EPA % DCW, more preferably at least about 20-22 EPA % DCW, more preferably at least about 22-24 EPA % DCW, and most preferably at least about 24-26 EPA % DCW.

[0152] In addition to possessing at least about 50 EPA % TFAs, the lipid profile within an improved optimized strain of *Yarrowia lipolytica*, or within extracted or unconcentrated oil therefrom, will have a ratio of EPA % TFAs to LA % TFAs of at least about 3:1. As demonstrated in U.S. Pat. Appl. Pub. No. 2009-0093543-A1 (Table 23), EPA, LA and oleic acid constituted approximately 76-80% of the fatty acids present in the lipid profile of a strain of *Y. lipolytica* producing greater than 40 EPA % TFAs. Of this, LA % TFAs was ca. three-fold greater than oleic acid % TFAs. Based on these observations, one of skill in the art will appreciate that minimizing the concentration of the intermediate fatty acid, LA (resulting in increased ratios of EPA:LA), will result in greater "pushing" of the carbon through the PUFA biosynthetic pathway and permit increased synthesis of EPA. As described herein, the ratio of EPA:LA will be at least about 3:1-3.5, more preferably at least about 3.5-4.5, more preferably at least about 4.5-5.5, and most preferably at least about 5.5-6.5.

[0153] Lipids produced by an improved optimized recombinant *Y. lipolytica* strains described herein will also be distinguished as having less than about 0.5% GLA or DHA (when measured by GC analysis using equipment having a detectable level down to about 0.1 %) and having a saturated fatty acid content of less than about 8%. This low percent of saturated fatty acids (i.e., 16:0 and 18:0) results in

substantial health benefits to humans and animals.

[0154] Microbial expression systems and expression vectors containing regulatory sequences that direct high-level expression of foreign proteins are well known to those skilled in the art. Any of these could be used to construct chimeric genes encoding the preferred desaturase, elongase, CPT1, DGLA synthase, malonyl CoA synthetase and acyl-CoA lysophospholipid acyltransferase proteins. These chimeric genes could then be introduced into *Yarrowia lipolytica* using standard methods of transformation to provide high-level expression of the encoded enzymes.

[0155] Vectors (e.g., constructs, plasmids) and DNA expression cassettes useful for the transformation of *Yarrowia* host cells are well known in the art. The specific choice of sequences present in the construct is dependent upon the desired expression products, the nature of the host cell, and the proposed means of separating transformed cells versus non-transformed cells. Typically, however, the vector contains at least one expression cassette, a selectable marker and sequences allowing autonomous replication or chromosomal integration. Suitable expression cassettes typically comprise a region 5' of the gene that controls transcriptional initiation (e.g., a promoter), the gene coding sequence, and a region 3' of the DNA fragment that controls transcriptional termination (i.e., a terminator). It is most preferred when both control regions are derived from genes from the transformed host cell, although they need not be derived from genes native to the production host (e.g., *Yarrowia lipolytica*).

[0156] Where two or more genes are expressed from separate replicating vectors, it is desirable that each vector has a different means of selection and should lack homology to the other constructs to maintain stable expression and prevent reassortment of elements among constructs. Judicious choice of regulatory regions, selection means and method of propagation of the introduced construct can be experimentally determined so that all introduced genes are expressed at the necessary levels to provide for synthesis of the desired products.

[0157] Constructs or vectors comprising the gene(s) of interest may be introduced into a host cell such as *Yarrowia* by any standard technique. These techniques include transformation (e.g., lithium acetate transformation [Methods in Enzymology, 194:186-187 (1991)]), bolistic impact, electroporation, microinjection, or any other method that introduces the gene(s) of interest into the host cell. More preferred herein for *Yarrowia lipolytica* are integration techniques based on linearized fragments of DNA, as described in U.S. Patent 4,880,741 and U.S. Patent 5,071,764 and Chen, D. C. et al. (Appl. Microbiol. Biotechnol., 48(2):232-235 (1997)).

[0158] For convenience, a host cell that has been manipulated by any method to take up a DNA sequence (e.g., an expression cassette) is referred to herein as "transformed", "transformant" or "recombinant". The transformed host will have at least one copy of the expression cassette and may have two or more, depending upon whether the expression cassette is integrated into the genome or is present on an extrachromosomal element having multiple copy numbers. The transformed host cell can be identified by various selection techniques, as described in U.S. Patent 7,238,482 and U.S. Patent 7,259,255.

[0159] Preferred selection methods for use herein are resistance to kanamycin, hygromycin and the amino glycoside G418, as well as ability to grow on media lacking uracil, leucine, lysine, tryptophan or histidine. Alternatively, 5-fluoroorotic acid (5-fluorouracil-6-carboxylic acid monohydrate; "5-FOA") is used for selection of yeast Ura⁻ mutants (U.S. Pat. Appl. Pub. No. 2009-0093543-A1), or a native acethydroxyacid synthase (or acetolactate synthase; E.C. 4.1.3.18) that confers sulfonyl urea herbicide resistance (Intl. App. Pub. No. WO 2006/052870) is utilized for selection of transformants. A unique method of "recycling" a pair of preferred selection markers for their use in multiple sequential transformations, by use of site-specific recombinase systems, is also taught in U.S. Pat. Appl. Pub. No. 2009-0093543-A1.

[0160] As is well known to one of skill in the art, merely inserting a gene (e.g., a desaturase, elongase, CPT1, DGLA synthase, malonyl CoA synthetase, acyl-CoA lysophospholipid acyltransferase) into a cloning vector does not ensure its expression at the desired rate, concentration, amount, etc. It may be desirable to manipulate a number of different genetic elements that control aspects of transcription, RNA stability, translation, protein stability and protein location, oxygen limitation and secretion from the host cell. More specifically, gene expression may be controlled by altering the following: the nature of the relevant transcriptional promoter and terminator sequences; the number of copies of the cloned gene; whether the gene is plasmid-borne or integrated into the genome of the host cell; the final cellular location of the synthesized foreign protein; the efficiency of translation in the host organism; the intrinsic stability of the cloned gene protein within the host cell; and, the codon usage within the cloned gene, such that its frequency approaches the frequency of preferred codon usage of the host cell. Several of these methods of overexpression will be discussed below, and are useful in recombinant *Yarrowia* host cells as a means to overexpress e.g., desaturases, elongases, CPT1 proteins, DGLA synthases, malonyl CoA synthetases and acyl-CoA lysophospholipid acyltransferases.

[0161] Expression of the desired gene(s) can be increased at the transcriptional level through the use of a stronger promoter (either regulated or constitutive) to cause increased expression, by removing/deleting destabilizing sequences from either the mRNA or the encoded protein, or by adding stabilizing sequences to the mRNA (U.S. Patent 4,910,141).

[0162] Transcription initiation control regions or promoters which are useful to drive expression of heterologous genes or portions thereof in *Yarrowia* host cells are numerous and known to those skilled in the art. Expression can be accomplished in an induced or constitutive fashion. Induced expression can be accomplished by inducing the activity of a regulatable promoter operably linked to the gene of interest, while constitutive expression can be achieved by the use of a constitutive promoter operably linked to the gene of interest. Virtually any promoter (i.e., native, synthetic, or chimeric) capable of directing expression of desaturase, elongase, CPT1, DGLA synthase, malonyl CoA synthetase and acyl-CoA lysophospholipid acyltransferase genes in *Yarrowia* will be suitable, although transcriptional and translational regions from the host species are particularly useful.

[0163] In general, the termination region can be derived from the 3' region of the gene from which the initiation region was obtained or from a different gene. A large number of termination regions are known and function satisfactorily in a variety of hosts when utilized both in the same and different genera and species from which they were derived. The termination region usually is selected more as a matter of convenience rather than because of any particular property. Preferably, the termination region is derived from a yeast gene. The 3'-region can also be synthetic, as one of skill in the art can utilize available information to design and synthesize a 3'-region sequence that functions as a transcription terminator. A termination site may be unnecessary, but it is highly preferred.

[0164] Although not intended to be limiting, preferred promoter regions and termination regions useful in the disclosure herein are those taught in U.S. Pat. Pub. No. 2009-0093543-A1.

[0165] Additional copies (i.e., more than one copy) of the PUFA biosynthetic pathway desaturase, elongase and DGLA synthase genes and/or CPT1, malonyl CoA synthetase and acyl-CoA lysophospholipid acyltransferase genes may be introduced into *Yarrowia lipolytica* to thereby increase EPA production and accumulation. Specifically, additional copies of genes may be cloned within a single expression construct; and/or, additional copies of the cloned gene(s) may be introduced into the host cell by increasing the plasmid copy number or by multiple integration of the cloned gene into the genome (*infra*).

[0166] It is important to note that the when preparing optimized strains of *Y. lipolytica* according to the methodology herein, copies of various desaturases, elongases, CPT1s, DGLA synthases, malonyl CoA synthetases and acyl-CoA lysophospholipid acyltransferases are often referred to. If, for example, 2 copies of a $\Delta 9$ elongase are required, this can refer to: 1) two copies of an identical coding sequence for a particular $\Delta 9$ elongase isolated from a single species; or, 2) one coding sequence for a $\Delta 9$ elongase isolated from a species "A" and one coding sequence for a $\Delta 9$ elongase isolated from a species "B", thus collectively resulting in two $\Delta 9$ elongases.

[0167] In general, once a DNA cassette (e.g., comprising a chimeric gene comprising a promoter, ORF and terminator) suitable for expression in an oleaginous yeast has been obtained, it is either placed in a plasmid vector capable of autonomous replication in a host cell or directly integrated into the genome of the host cell. Integration of expression cassettes can occur randomly within the host genome or can be targeted through the use of constructs containing regions of homology with the host genome sufficient to target recombination with the host locus. Although not relied on herein, all or some of the transcriptional and translational regulatory regions can be provided by the endogenous locus where constructs are targeted to an endogenous locus.

[0168] The preferred method of expressing genes in *Yarrowia lipolytica* is by integration of a linear DNA fragment into the genome of the host. Integration into multiple locations within the genome can be particularly useful when high level expression of genes are desired. Preferred loci include those taught in U.S. Pat. Pub. No. 2009-0093543-A1.

[0169] Juretzek et al. (Yeast, 18:97-113 (2001)) note that the stability of an integrated DNA fragment in *Yarrowia lipolytica* is dependent on the individual transformants, the recipient strain and the targeting platform used. Thus, the skilled artisan will recognize that multiple transformants must be screened in order to obtain a strain displaying the desired expression level and pattern. Such screening may be accomplished by Southern analysis of DNA blots (Southern, J. Mol. Biol., 98:503 (1975)), Northern analysis of mRNA expression (Kroczeck, J. Chromatogr. Biomed. Appl., 618 (1-2):133-145 (1993)), Western analysis of protein expression, phenotypic analysis or GC analysis of the PUFA products.

[0170] The transformed microbial host cell is grown under conditions that optimize expression of chimeric genes (e.g., encoding desaturases, elongases, CPT1, DGLA synthases, malonyl CoA synthetases, acyl-CoA lysophospholipid acyltransferases) and produce the greatest and the most economical yield of EPA. In general, media conditions may be optimized by modifying the type and amount of carbon source, the type and amount of nitrogen source, the carbon-to-nitrogen ratio, the amount of different mineral ions, the oxygen level, growth temperature, pH, length of the biomass production phase, length of the oil accumulation phase and the time and method of cell harvest. *Yarrowia lipolytica* are generally grown in a complex media such as yeast extract-peptone-dextrose broth ["YPD"] or a defined minimal media that lacks a component necessary for growth and thereby forces selection of the desired expression cassettes (e.g., Yeast Nitrogen Base (DIFCO Laboratories, Detroit, MI)).

[0171] Fermentation media for the methods and host cells described herein must contain a suitable carbon source, such as are taught in U.S. Patent 7,238,482 and U.S. Pat. Appl. No. 12/641,929 (filed December 19, 2009). Although it is contemplated that the source of carbon utilized in the presently disclosed methods may encompass a wide variety of carbon-containing sources, preferred carbon sources are sugars, glycerol and/or fatty acids. Most preferred is glucose, sucrose, invert sucrose, fructose and/or fatty acids containing between 10-22 carbons. For example, the fermentable carbon source can be selected from the group consisting of invert sucrose, glucose, fructose and combinations of these, provided that glucose is used in combination with invert sucrose and/or fructose.

[0172] The term "invert sucrose", also referred to herein as "invert sugar", refers to a mixture comprising equal parts of fructose and glucose resulting from the hydrolysis of sucrose. Invert sucrose may be a mixture comprising 25 to 50% glucose and 25 to 50% fructose. Invert sucrose may also comprise sucrose, the amount of which depends on the degree of hydrolysis.

[0173] Nitrogen may be supplied from an inorganic (e.g., $(\text{NH}_4)_2\text{SO}_4$) or organic (e.g., urea or glutamate) source. In addition to appropriate carbon and nitrogen sources, the fermentation media must also contain suitable minerals, salts, cofactors, buffers, vitamins and other components known to those skilled in the art suitable for the growth of the high EPA-producing oleaginous yeast and the promotion of the enzymatic pathways for EPA production. Particular attention is given to several metal ions, such as Fe^{+2} , Cu^{+2} , Mn^{+2} , Co^{+2} , Zn^{+2} and Mg^{+2} , that promote synthesis of lipids and PUFAs (Nakahara, T. et al., Ind. Appl. Single Cell Oils, D. J. Kyle and R. Colin, eds. pp 61-97 (1992)).

[0174] Preferred growth media for the methods and host cells described herein are common commercially prepared media, such as Yeast Nitrogen Base (DIFCO Laboratories, Detroit, MI). Other defined or synthetic growth media may also be used and the appropriate medium for growth of *Yarrowia lipolytica* will be known by one skilled in the art of microbiology or fermentation science. A suitable pH range for the fermentation is typically between about pH 4.0 to pH 8.0, wherein pH 5.5 to pH 7.5 is preferred as the range for the initial growth conditions. The fermentation may be conducted under aerobic or anaerobic conditions, wherein microaerobic conditions are preferred.

[0175] Typically, accumulation of high levels of PUFAs in oleaginous yeast cells requires a two-stage process, since the metabolic state must be "balanced" between growth and synthesis/storage of fats. Thus, most preferably, a two-stage fermentation process is necessary for the production of EPA in *Yarrowia lipolytica*. This approach is described in U.S. Patent 7,238,482, as are various suitable fermentation process designs (i.e., batch, fed-batch and continuous) and considerations during growth.

[0176] In some aspects herein, the primary product is oleaginous yeast biomass. As such, isolation and purification of the EPA-containing oils from the biomass may not be necessary (i.e., wherein the whole cell biomass is the product).

[0177] However, certain end uses and/or product forms may require partial and/or complete isolation/purification of the EPA-containing oil from the biomass, to result in partially purified biomass, purified oil, and/or purified EPA. PUFAs, including EPA, may be found in the host microorganism (e.g., *Yarrowia*) as free fatty acids or in esterified forms such as acylglycerols, phospholipids, sulfolipids or glycolipids. These fatty acids may be extracted from the host cell through a variety of means well-known in the art. One review of extraction techniques, quality analysis and acceptability standards for yeast lipids is that of Z. Jacobs (Critical Reviews in Biotechnology, 12(5/6):463-491 (1992)). A brief review of downstream processing is also available by A. Singh and O. Ward (Adv. Appl. Microbiol., 45:271-312 (1997)).

[0178] In general, means for the purification of EPA and other PUFAs from *Yarrowia* biomass may include extraction (e.g., U.S. Patent 6,797,303 and U.S. Patent 5,648,564) with organic solvents, sonication, supercritical fluid extraction (e.g., using carbon dioxide), saponification and physical means such as presses, bead beaters, or combinations thereof. One is referred to the teachings of U.S. Patent 7,238,482 for additional details.

[0179] Oils containing EPA that have been refined and/or purified can be hydrogenated, to thereby result in fats with various melting properties and textures. Many processed fats, including spreads, confectionary fats, hard butters, margarines, baking shortenings, etc., require varying degrees of solidity at room temperature and can only be produced through alteration of the source oil's physical properties. This is most commonly achieved through catalytic hydrogenation as described in WO 2006/052870.

[0180] Food products, infant formulas, functional foods, medical foods, medical nutritionals, dietary supplements, pharmaceutical compositions, animal feeds, and personal care products comprising oleaginous yeast biomass comprising EPA are provided herein. Similarly, also provided are food products, infant formulas, functional foods, medical foods, medical nutritionals, dietary supplements, pharmaceutical compositions, animal feeds, and personal care products comprising EPA or microbial oil comprising EPA isolated from the recombinant oleaginous yeast biomass.

[0181] One of skill in the art of processing and formulation will understand how the amount and composition of the biomass, partially purified biomass, purified oil, and/or purified EPA may be added to a particular product according to target species and/or end use. More specifically, an "effective" amount should be incorporated into a product formulation, although this amount will depend on the food or feed product, the diet that the product is intended to supplement or the medical condition that the medical food or medical nutritional is intended to correct or treat. Most desirably, the effective amount of EPA will be sufficient to provide the desirable health characteristics associated with ω -3/ ω -6 PUFA consumption. Typically, the amount of EPA incorporated into the product takes into account losses associated with processing conditions, typical handling and storage conditions, the stability of EPA in the product, and the bioavailability/ bioabsorption efficiency with the target species, to name a few.

[0182] One of skill in the art of processing and formulation will be familiar with processes to concentrate the oil produced from the recombinant *Yarrowia* production host cells described herein, to thereby increase the concentration of EPA in the total lipid fraction such that it comprises at least about 60%, at least about 70%, at least about 80% or even at least about 90% EPA. Means to blend the purified oils described herein with other purified fatty acids (e.g., LA, GLA, EDA, DGLA, ARA, DTA, DPA α -6, ALA, STA, ETrA, ETA, DPA and DHA), or oils containing alternate fatty acids in preferred concentrations, are also well known to one of skill in the art. These techniques readily permit the creation of an oil comprising a uniquely tailored fatty acid profile.

[0183] Personal Care Products: Within the context of personal care products, ω -3 fatty acids have particular application in skin formulations where they may be used to enhance the skin conditioning effect. The skilled person will understand how to provide an effective amount of the relevant ω -3 fatty acid(s) or oil comprising the same to a skin care composition. In addition to the PUFA oil or ω -3 fatty acid, the skin care composition may further comprise a cosmetically acceptable medium for skin care compositions, examples of which are described by Philippe et al. in U.S. Patent 6,280,747. For example, the cosmetically acceptable medium may be an anhydrous composition containing a fatty substance in a proportion generally from about 10% to about 90% by weight relative to the total weight of the composition, where the fatty phase contains at least one liquid, solid or semi-solid fatty substance. The fatty substance includes, but is not limited to, oils, waxes, gums, and so-called pasty fatty substances. Alternatively, the compositions may be in the form of a stable dispersion such as a water-in-oil or oil-in-water emulsion. Additionally, the compositions may contain one or more conventional cosmetic or dermatological additives or adjuvants including, but not limited to, antioxidants, preserving agents, fillers, surfactants, UVA and/or UVB sunscreens, fragrances, thickeners, wetting agents, anionic or nonionic or amphoteric polymers, and dyes.

[0184] Foodstuffs: The market place currently supports a large variety of food and feed products, incorporating ω -3 and/or ω -6 fatty acids (particularly LA, GLA, ARA, EPA, DPA and DHA). It is contemplated that the yeast biomass, partially purified biomass, purified oil, and/or purified EPA described herein will function in food products to impart the health benefits of current formulations.

[0185] Food products will include, but not be limited to: food analogs, drinks, meat products, cereal products, baked foods, snack foods and dairy products.

[0186] Food analogs can be made using processes well known to those skilled in the art. There can be mentioned meat analogs, cheese analogs, milk analogs and the like. Meat analogs made from soybeans contain soy protein or tofu and other ingredients mixed together to simulate various kinds of meats. These meat alternatives are sold as frozen, canned or dried foods. Usually, they can be used the same way as the foods they replace. Examples of meat analogs include, but are not limited to: ham analogs, sausage analogs, bacon analogs, and the like.

[0187] Food analogs can be classified as imitation or substitutes depending on their functional and compositional characteristics. For example, an imitation cheese need only resemble the cheese it is designed to replace. However, a product can generally be called a substitute cheese only if it is nutritionally equivalent to the cheese it is replacing and meets the minimum compositional requirements for that cheese. Thus, substitute cheese will often have higher protein levels than imitation cheeses and be fortified with vitamins and minerals.

[0188] Milk analogs or nondairy food products include, but are not limited to: imitation milks and nondairy frozen desserts (e.g., those made from soybeans and/or soy protein products).

[0189] Meat products encompass a broad variety of products. In the United States "meat" includes "red meats" produced from cattle, hogs and sheep. In addition to the red meats there are poultry items which include chickens, turkeys, geese, guineas, ducks and the fish and shellfish. There is a wide assortment of seasoned and processed meat products: fresh, cured and fried, and cured and cooked. Sausages and hot dogs are examples of processed meat products. Thus, the term "meat products" as used herein includes, but is not limited to, processed meat products.

[0190] A cereal food product is a food product derived from the processing of a cereal grain. A cereal grain includes any plant from the grass family that yields an edible grain (seed). The most popular grains are barley, corn, millet, oats, quinoa, rice, rye, sorghum, triticale, wheat and wild rice. Examples of cereal food products include, but are not limited to: whole grain, crushed grain, grits, flour, bran, germ, breakfast cereals, extruded foods, pastas, and the like.

[0191] A baked goods product comprises any of the cereal food products mentioned above and which has been baked or processed in a manner comparable to baking, i.e., to dry or harden by subjecting to heat. Examples of a baked good product include, but are not limited to: bread, cakes, doughnuts, bars, pastas, bread crumbs, baked snacks, mini-biscuits, mini-crackers, mini-cookies, and mini-pretzels. As was mentioned above, oils from the recombinant EPA production host cells can be used as an ingredient.

[0192] A snack food product comprises any of the above or below described food products.

[0193] A fried food product comprises any of the above or below described food products that has been fried.

[0194] The beverage can be in a liquid or in a dry powdered form. For example, there can be mentioned: non-carbonated drinks; fruit juices, fresh, frozen, canned or concentrate; flavored or plain milk drinks, etc. Adult and infant nutritional formulas are well known in the art and commercially available (e.g., Similac®, Ensure®, Jevity®, and Alimentum® from Ross Products Division, Abbott Laboratories).

[0195] A dairy product is a product derived from milk. A milk analog or nondairy product is derived from a source other than milk, for example, soymilk as was discussed above. These products include, but are not limited to: whole milk, skim milk, fermented milk products such as yogurt or sour milk, cream, butter, condensed milk, coffee whitener, coffee creamer, ice cream, cheese, etc.

[0196] Additional food products into which the *Yarrowia* biomass, partially purified biomass, purified oil, and/or purified EPA could be included are, for example: chewing gums, confections and frostings,

gelatins and puddings, hard and soft candies, jams and jellies, white granulated sugar, sugar substitutes, sweet sauces, toppings and syrups, and dry-blended powder mixes.

[0197] Infant Formulas: Infant formulas are liquids or reconstituted powders fed to infants and young children. "Infant formula" is defined herein as an enteral nutritional product which can be substituted for human breast milk in feeding infants and typically is composed of a desired percentage of fat mixed with desired percentages of carbohydrates and proteins in an aqueous solution (e.g., see U.S. Patent 4,670,285). Based on worldwide composition studies, as well as levels specified by expert groups, average human breast milk typically contains about 0.20% to 0.40% of total fatty acids (assuming about 50% of calories from fat); and, generally the ratio of DHA to ARA would range from about 1:1 to 1:2 (see, e.g., formulations of Enfamil LIPIL™ [Mead Johnson & Company] and Similac Advance™ [Ross Products Division, Abbott Laboratories]). Infant formulas have a special role to play in the diets of infants because they are often the only source of nutrients for infants. Although breast-feeding is still the best nourishment for infants, infant formula is a close enough second that babies not only survive but thrive.

[0198] Health Food Products And Pharmaceuticals: The present biomass, partially purified biomass, purified oil, and/or purified EPA may be used in formulations to impart health benefit in health food products, including functional foods, medical foods, medical nutritionals and dietary supplements. Additionally, *Yarrowia* biomass, partially purified biomass, purified oil, and/or purified EPA may be used in standard pharmaceutical compositions. The present engineered strains of *Yarrowia lipolytica* or the microbial oils produced therefrom comprising EPA could readily be incorporated into the any of the above mentioned food products, to thereby produce e.g., a functional or medical food. For example, more concentrated formulations comprising EPA include capsules, powders, tablets, softgels, gelpacs, liquid concentrates and emulsions which can be used as a dietary supplement in humans or animals other than humans.

[0199] Animal Feed Products: Animal feeds are generically defined herein as products intended for use as feed or for mixing in feed for animals other than humans. The *Yarrowia* biomass, partially purified biomass, purified oil, and/or purified EPA described herein can be used as an ingredient in various animal feeds.

[0200] More specifically, although not to be construed as limiting, it is expected that the EPA products from the recombinant *Yarrowia* host cells can be used within pet food products, ruminant and poultry food products and aquacultural food products. Pet food products are those products intended to be fed to a pet, such as a dog, cat, bird, reptile, rodent. These products can include the cereal and health food products above, as well as meat and meat byproducts, soy protein products, grass and hay products (such as alfalfa, timothy, oat or brome grass) and vegetables. Ruminant and poultry food products are those wherein the product is intended to be fed to e.g., turkeys, chickens, cattle and swine. As with the pet foods above, these products can include cereal and health food products, soy protein products, meat and meat byproducts, and grass and hay products as listed above. Aquacultural food products (or "aquafeeds") are those products intended to be used in aquafarming, which concerns the propagation, cultivation or farming of aquatic organisms and/or animals in fresh or marine waters.

[0201] It is contemplated that the present engineered strains of *Yarrowia lipolytica* that are producing high concentrations of EPA will be especially useful to include in most animal feed formulations. In addition to providing necessary ω -3 PUFAs, the yeast itself is a useful source of protein and other nutrients (e.g., vitamins, minerals, nucleic acids, complex carbohydrates, etc.) that can contribute to overall animal health and nutrition, as well as increase a formulation's palatability. Accordingly it is contemplated that the addition of yeast biomass comprising the recombinant *Yarrowia* production hosts will be an excellent additional source of feed nutrients in animal feed formulations (see U.S. Pat. Appl. Pub. No. 2009-0093543-A1 for additional details).

[0202] It is clear then that the present engineered strains of *Yarrowia lipolytica* that are producing high concentrations of EPA will be especially useful to include in most aquaculture feeds. In addition to providing necessary ω -3 and/or ω -6 PUFAs, the yeast itself is a useful source of protein that can increase the formulation's palatability. Alternatively, the oils produced by the present strains of *Y. lipolytica* could be introduced directly into the aquaculture feed formulations, following extraction and purification from the cell mass.

[0203] There is increasing awareness that EPA is an important ω -3 fatty acid in and of itself. As a result, it is expected herein that the EPA-enriched oils of the recombinant *Yarrowia* production hosts described herein will have very broad utility in a variety of therapeutic applications, e.g., inflammation, cardiovascular diseases, nutrient regulation of gene expression and dyslipidemia, and specifically in the treatment of clinical conditions including: coronary heart disease, high blood pressure, inflammatory disorders, Type II diabetes, ulcerative colitis, Crohn's disease, anorexia nervosa, burns, osteoarthritis, osteoporosis, depression, and attention deficit/hyperactivity disorder.

[0204] U.S. Pat. Appl. Pub. No. 2009-0093543-A1 describes additional clinical human studies, relating to EPA and inflammation, EPA and cardiovascular diseases, ω -3 PUFAs and nutrient regulation of gene expression, and ω -3 PUFAs and dyslipidemia and should be referred to therein. More recently, a randomized, double-blind placebo-controlled study was performed in 110 normal healthy subjects, wherein subjects were provided one of the following for 6 weeks, as a means to evaluate the effects of the oils on cardiovascular disease risk factors, adverse events and safety parameters: 600 mg/day EPA, 1800 mg/day EPA, 600 mg/day DHA or olive oil (placebo) (Gillies, P., "The New Science of Omega-3 Fatty Acids: Differential Nutritional Pharmacology" Texas Human Nutrition Conference, Texas A&M University, February 2010.. The 600 mg EPA per day supplement was found to maintain healthy cholesterol levels already in the normal range. Notably, the EPA oils of the study were derived from engineered strains of *Yarrowia lipolytica*.

EXAMPLES

[0205] The present invention is further defined in the following Examples. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions, as defined in the claims.

GENERAL METHODS

[0206] Standard recombinant DNA and molecular cloning techniques used in the Examples are well known in the art and are described by: 1) Sambrook, J., Fritsch, E. F. and Maniatis, T. Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989) (Maniatis); 2) T. J. Silhavy, M. L. Bannan, and L. W. Enquist, Experiments with Gene Fusions; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1984); and, 3) Ausubel, F. M. et al., Current Protocols in Molecular Biology, published by Greene Publishing Assoc. and Wiley-Interscience, Hoboken, NJ (1987).

[0207] Materials and methods suitable for the maintenance and growth of microbial cultures are well known in the art. Techniques suitable for use in the following examples may be found as set out in Manual of Methods for General Bacteriology (Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costlow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, Eds), American Society for Microbiology: Washington, D.C. (1994)); or by Thomas D. Brock in Biotechnology: A Textbook of Industrial Microbiology, 2nd ed., Sinauer Associates: Sunderland, MA (1989). All reagents, restriction enzymes and materials used for the growth and maintenance of microbial cells were obtained from Aldrich Chemicals (Milwaukee, WI), DIFCO Laboratories (Detroit, MI), New England Biolabs, Inc. (Beverly, MA), GIBCO/BRL (Gaithersburg, MD), or Sigma Chemical Company (St. Louis, MO), unless otherwise specified. *E. coli* strains were typically grown at 37 °C on Luria Bertani ["LB"] plates.

[0208] General molecular cloning was performed according to standard methods (Sambrook et al., *supra*). When PCR or site-directed mutagenesis was involved in subcloning, the constructs were sequenced to confirm that no errors had been introduced to the sequence. PCR products were cloned into Promega's pGEM-T-easy vector (Madison, WI).

[0209] The meaning of abbreviations is as follows: "sec" means second(s), "min" means minute(s), "h" means hour(s), "d" means day(s), "μL" means microliter(s), "mL" means milliliter(s), "L" means liter(s), "μM" means micromolar, "mM" means millimolar, "M" means molar, "mmol" means millimole(s), "μmole" mean micromole(s), "g" means gram(s), "μg" means microgram(s), "ng" means nanogram(s), "U" means unit(s), "bp" means base pair(s), "kB" means kilobase(s), "DCW" means dry cell weight, and "TFAs" means total fatty acids.

Nomenclature For Expression Cassettes

[0210] The structure of an expression cassette will be represented by a simple notation system of "X::Y::Z", wherein X describes the promoter fragment, Y describes the gene fragment, and Z describes the terminator fragment, which are all operably linked to one another.

Transformation And Cultivation Of *Yarrowia lipolytica*

[0211] *Yarrowia lipolytica* strain ATCC #20362 was purchased from the American Type Culture Collection (Rockville, MD). *Yarrowia lipolytica* strains were routinely grown at 28-30 °C in several media, according to the recipes shown below. Agar plates were prepared as required by addition of 20 g/L agar to each liquid media, according to standard methodology.

YPD agar medium (per liter): 10 g of yeast extract [Difco], 20 g of Bacto peptone [Difco], and 20 g of glucose.

Basic Minimal Media ["MM"] (per liter): 20 g glucose, 1.7 g yeast nitrogen base without amino acids, 1.0 g proline, and pH 6.1 (do not need to adjust).

Minimal Media + Uracil ["MM+uracil or MMU"] (per liter): Prepare MM media as above and add 0.1 g uracil and 0.1 g uridine.

Minimal Media + Uracil + Sulfonyleurea ["MMU+SU"] (per liter): Prepare MMU media as above and add 280 mg sulfonyleurea.

Minimal Media + Uracil + Lysine ["MMUraLys"] (per liter): Prepare MM media as above and add 0.1 g uracil, 0.1 g uridine, and 0.1 g lysine.

Minimal Media + 5-Fluoroorotic Acid ["MM + 5-FOA"] (per liter): 20 g glucose, 6.7 g Yeast Nitrogen base, 75 mg uracil, 75 mg uridine and appropriate amount of FOA (Zymo Research Corp., Orange, CA), based on FOA activity testing against a range of concentrations from 100 mg/L to 1000 mg/L (since variation occurs within each batch received from the supplier).

High Glucose Media ["HGM"] (per liter): 80 glucose, 2.58 g KH_2PO_4 and 5.36 g K_2HPO_4 , pH 7.5 (do not need to adjust).

Fermentation medium without Yeast Extract ["FM without YE"] (per liter): 6.70 g/L Yeast nitrogen base, 6.00 g KH_2PO_4 , 2.00 g K_2HPO_4 , 1.50 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, and 20 g glucose.

Fermentation medium ["FM"] (per liter): 6.70 g/L Yeast nitrogen base, 6.00 g KH_2PO_4 , 2.00 g K_2HPO_4 , 1.50 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 20 g glucose and 5.00 g Yeast extract (BBL).

[0212] Transformation of *Y. lipolytica* was performed as described in U.S. Pat. Appl. Pub. No. 2009-0093543-A1.

Fatty Acid Analysis Of *Yarrowia lipolytica*

[0213] For fatty acid ["FA"] analysis, cells were collected by centrifugation and lipids were extracted as described in Bligh, E. G. & Dyer, W. J. (Can. J. Biochem. Physiol., 37:911-917 (1959)). Fatty acid methyl esters ["FAMES"] were prepared by transesterification of the lipid extract with sodium methoxide (Roughan, G., and Nshida I., Arch Biochem Biophys., 276(1):38-46 (1990)) and subsequently analyzed with a Hewlett-Packard 6890 GC fitted with a 30-m X 0.25 mm (i.d.) HP-INNOWAX (Hewlett-Packard) column. The oven temperature was from 170 °C (25 min hold) to 185 °C at 3.5 °C/min.

[0214] For direct base transesterification, *Yarrowia* cells (0.5 mL culture) were harvested, washed once in distilled water, and dried under vacuum in a Speed-Vac for 5-10 min. Sodium methoxide (100 μl of 1 %) and a known amount of C15:0 triacylglycerol (C15:0 TAG; Cat. No. T-145, Nu-Check Prep, Elysian, MN) was added to the sample, and then the sample was vortexed and rocked for 30 min at 50 °C. After adding 3 drops of 1 M NaCl and 400 μl hexane, the sample was vortexed and spun. The upper layer was removed and analyzed by GC.

[0215] FAME peaks recorded via GC analysis were identified by their retention times, when compared to that of known fatty acids, and quantitated by comparing the FAME peak areas with that of the internal standard (C15:0 TAG) of known amount. Thus, the approximate amount (μg) of any fatty acid FAME [" μg FAME"] is calculated according to the formula: (area of the FAME peak for the specified fatty acid/ area of the standard FAME peak) * (μg of the standard C15:0 TAG), while the amount (μg) of any fatty acid [" μg FA"] is calculated according to the formula: (area of the FAME peak for the specified fatty acid/area of the standard FAME peak) * (μg of the standard C15:0 TAG) * 0.9503, since 1 μg of C15:0 TAG is equal to 0.9503 μg fatty acids. Note that the 0.9503 conversion factor is an approximation of the value determined for most fatty acids, which range between 0.95 and 0.96.

[0216] The lipid profile, summarizing the amount of each individual fatty acid as a weight percent of TFAs, was determined by dividing the individual FAME peak area by the sum of all FAME peak areas and multiplying by 100.

Analysis Of Total Lipid Content And Composition In *Yarrowia lipolytica* By Flask Assay

[0217] For a detailed analysis of the total lipid content and composition in a particular strain of *Y. lipolytica*, flask assays were conducted as followed. Specifically, one loop of freshly streaked cells was inoculated into 3 mL FM medium and grown overnight at 250 rpm and 30 °C. The $\text{OD}_{600\text{nm}}$ was measured and an aliquot of the cells were added to a final $\text{OD}_{600\text{nm}}$ of 0.3 in 25 mL FM medium in a 125 mL flask. After 2 days in a shaker incubator at 250 rpm and at 30 °C, 6 mL of the culture was harvested by centrifugation and resuspended in 25 mL HGM in a 125 mL flask. After 5 days in a shaker incubator at 250 rpm and at 30 °C, a 1 mL aliquot was used for fatty acid analysis (*supra*) and 10 mL dried for dry cell weight ["DCW"] determination.

[0218] For DCW determination, 10 mL culture was harvested by centrifugation for 5 min at 4000 rpm in a Beckman GH-3.8 rotor in a Beckman GS-6R centrifuge. The pellet was resuspended in 25 mL of water and re-harvested as above. The washed pellet was re-suspended in 20 mL of water and transferred to a pre-weighed aluminum pan. The cell suspension was dried overnight in a vacuum oven at 80 °C. The weight of the cells was determined.

[0219] Total lipid content of cells ["TFAs % DCW"] is calculated and considered in conjunction with data tabulating the concentration of each fatty acid as a weight percent of TFAs ["% TFAs"] and the EPA content as a percent of the dry cell weight ["EPA % DCW"]. Data from flask assays will be presented as a table that summarizes the total dry cell weight of the cells ["DCW"], the total lipid content of cells ["TFAs % DCW"], the concentration of each fatty acid as a weight percent of TFAs ["% TFAs"] and the EPA content as a percent of the dry cell weight ["EPA % DCW"]. More specifically, fatty acids will be identified as 16:0 (palmitate), 16:1 (palmitoleic acid), 18:0 (stearic acid), 18:1 (oleic acid), 18:2 (LA), ALA, EDA, DGLA, ARA, ETRA, ETA, EPA and other.

EXAMPLE 1

Generation Of *Yarrowia lipolytica* Strain L135 (*Ura3+*, *Leu-*, *Apex3*) To Produce About 46%DGLA Of Total Fatty Acids

[0220] The present Example describes the construction of strain L135, derived from *Yarrowia lipolytica* ATCC #20362, capable of producing about 46% DGLA relative to the total lipids via expression of a $\Delta 9$ elongase/ $\Delta 8$ desaturase pathway.

[0221] Briefly, as diagrammed in FIG. 2, strain L135 was derived from *Yarrowia lipolytica* ATCC #20362 via construction of strain Y2224 (a FOA resistant mutant from an autonomous mutation of the *Ura3* gene of wildtype *Yarrowia* strain ATCC #20362), strain Y4001 (producing 17% EDA with a *Leu-* phenotype), strain Y4001 U1 (*Leu-* and *Ura-*), strain Y4036 (producing 18% DGLA with a *Leu-* phenotype) and strain Y4036U (*Leu-* and *Ura-*). Further details regarding the construction of strains Y2224, Y4001, Y4001 U, Y4036 and Y4036U are described in the General Methods of U.S. Pat. App. Pub. No. 2008-0254191.

[0222] The final genotype of strain Y4036U with respect to wild type *Yarrowia lipolytica* ATCC #20362 was *Ura3-*, YAT1::ME3S::Pex16, EXP1::EgD9eS::Lip1, FBAINm::EgD9eS::Lip2, GPAT::EgD9e::Lip2, FBAINm::EgDBM::Pex20, EXP1::EgDBM::Pex16, GPD::FmD12::Pex20, YAT1::FmD12::OCT (wherein FmD12 is a *Fusarium moniliforme* $\Delta 12$ desaturase gene [U.S. Patent 7,504,259]; ME3S is a codon-optimized C16:1/8 elongase gene, derived from *Mortierella alpina* [U.S. Patent 7,470,532]; EgD9e is a *Euglena gracilis* $\Delta 9$ elongase gene [U.S. Patent 7,645,604]; EgD9eS is a codon-optimized $\Delta 9$ elongase gene, derived from *Euglena gracilis* [U.S. Patent 7,645,604]; EgD8M is a synthetic mutant $\Delta 8$ desaturase [U.S. Patent 7,709,239], derived from *Euglena gracilis* [U.S. Patent 7,256,033]).

Generation Of L135 Strain With Chromosomal Deletion Of Pex3

[0223] Construction of strain L135 is described in Example 12 of Int. App. Pub. No. WO 2009/046248. Briefly, construct pY157 was used to knock out the chromosomal gene encoding the peroxisome biogenesis factor 3 protein [peroxisomal assembly protein Peroxin 3 or "Pex3p"] in strain Y4036U, thereby producing strain L135 (also referred to as strain Y4036 (Δpex3)). Knockout of the chromosomal Pex3 gene in strain L135, as compared to in strain Y4036 (whose native Pex3p had not been knocked out) resulted in the following: higher lipid content (TFAs % DCW) (ca. 6.0% versus 4.7%), higher DGLA % TFAs (46% versus 19%), higher DGLA % DCW (ca. 2.8% versus 0.9%) and reduced LA % TFAs (12% versus 30%). Additionally, the $\Delta 9$ elongase percent conversion efficiency was increased from ca. 48% in strain Y4036 to 83% in strain L135.

[0224] The final genotype of strain L135 with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura3+*, *Leu-*, *Pex3-*, *unknownn1-*, YAT1::ME3S::Pex16, EXP1::EgD9eS::Lip1, FBAINm::EgD9eS::Lip2, GPAT::EgD9e::Lip2, FBAINm::EgDBM::Pex20, EXP1::EgDBM::Pex16, GPD::FmD12::Pex20, YAT1::FmD12::OCT.

EXAMPLE 2

Generation Of *Yarrowia lipolytica* Strains Producing From About 18% To About 41 %ARA Of Total Fatty Acids ["TFAs"]

[0225] The present Example describes the construction of strain Y8006, derived from *Yarrowia lipolytica* ATCC #20362, capable of producing about 41 %ARA relative to the total lipids via expression of a $\Delta 9$ elongase/ $\Delta 8$ desaturase pathway.

[0226] The development of strain Y8006 (FIG. 2) required the construction of strains Y2224, Y4001, Y4001 U, Y4036, Y4036U and L135 (described in Example 1), as well as construction of strains L135U9 and Y8002.

Generation of L135U9 (*Leu*⁻, *Ura*3⁻) Strain

[0227] Strain L135U was created via temporary expression of the *Cre* recombinase enzyme in plasmid pY116 (FIG. 3; SEQ ID NO:127; described in Example 7 of Intl. App. Pub. No. WO 2008/073367) within strain L135 to produce a *Leu*⁻ and *Ura*⁻ phenotype. Plasmid pY116 was used for transformation of freshly grown L135 cells according to the General Methods. The transformant cells were plated onto MMLeuUra and maintained at 30 °C for 3 to 4 days. Three colonies were picked, inoculated into 3 mL liquid YPD media at 30 °C and shaken at 250 rpm/min for 1 day. The cultures were diluted to 1:50,000 with liquid MMLeuUra media, and 100 µL was plated onto new YPD plates and maintained at 30 °C for 2 days. Eight colonies were picked from each of three plates (24 colonies total) and streaked onto MMLeu and MMLeuUra selection plates. The colonies that could grow on MMLeuUra plates but not on MMLeu plates were selected and analyzed by GC to confirm the presence of C20:2 (EDA). One strain, having a *Leu*⁻ and *Ura*⁻ phenotype, was designated as L135U9.

Generation Of Y8002 Strain To Produce About 32%ARA Of TFAs

[0228] Construct pZKSL-5S5A5 (FIG. 4A; SEQ ID NO:128) was generated to integrate three Δ5 desaturase genes into the *Lys* loci of strain L135U9, to thereby enable production of ARA. The pZKSL-5S5A5 plasmid contained the following components:

Table 6: Description of Plasmid pZKSL-5S5A5 (SEQ ID NO:128)

RE Sites And Nucleotides Within SEQ ID NO:128	Description Of Fragment And Chimeric Gene Components
<i>AscI/Bst</i> WI (5925-6645)	720 bp 5' portion of <i>Yarrowia Lys5</i> gene (GenBank Accession No. M34929; labeled as "lys5 5' region" in Figure)
<i>PacI/SphI</i> (2536-3225)	689 bp 3' portion of <i>Yarrowia Lys5</i> gene (GenBank Accession No. M34929; labeled as "Lys5-3'" in Figure)
<i>EcoRI/Bst</i> WI (9338-6645)	FBAIN::EgD5SM::Pex20, comprising: • FBAIN: <i>Yarrowia lipolytica</i> FBAIN promoter (U.S. Patent 7,202,356); • EgD5SM: Synthetic mutant Δ5 desaturase (SEQ ID NO:71; U.S. Pat. Pub. No. 2010-0075386-A1), derived from <i>Euglena gracilis</i> (U.S. Patent 7,678,560) (labeled as "ED5S" in Figure); • Pex20: Pex20 terminator sequence from <i>Yarrowia Pex20</i> gene (GenBank Accession No. AF054613)
<i>PmeI/Clal</i> (11503-1)	YAT1::EaD5SM::OCT, comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • EaD5SM: Synthetic mutant Δ5 desaturase (SEQ ID NO:81; • U.S. Pat. Pub. No. 2010-0075386-A1), derived from <i>Euglena anaethaena</i> (U.S. Pat. Appl. Pub. No. 2008-0274521-A1) (labeled as "EaD5S" in Figure); • OCT: OCT terminator sequence of <i>Yarrowia OCT</i> gene (GenBank Accession No. X69988)
<i>Clal/PacI</i> (1-2536)	EXP1::EgD5M::Pex16, comprising: • EXP1: <i>Yarrowia lipolytica</i> export protein (EXP1) promoter (labeled as "EXP" in Figure; Intl. App. Pub. No. WO 2006/052870); • EgD5M: Mutant Δ5 desaturase (SEQ ID NO:69; U.S. Pat. Pub. No. 2010-0075386-A1) with elimination of internal <i>EcoRI</i> , <i>BglII</i> , <i>HindIII</i> and <i>NcoI</i> restriction enzyme sites, derived from <i>Euglena gracilis</i> (U.S. Patent 7,678,560) (labeled as "Euglena D5WT" in Figure); • Pex16: Pex16 terminator sequence from <i>Yarrowia Pex16</i> gene (GenBank Accession No. U75433)
<i>EcoRI/PmeI</i> (9360-11503)	<i>Yarrowia Leu2</i> gene (GenBank Accession No. M37309)

[0229] The pZKSL-5S5A5 plasmid was digested with *AscI/SphI*, and then used for transformation of strain L135U9 according to the General Methods. The transformant cells were plated onto MMUraLys plates and maintained at 30 °C for 2 to 3 days. Single colonies were then re-streaked onto MMUraLys plates, and then inoculated into liquid MMUraLys at 30 °C and shaken at 250 rpm/min for 2 days. The cells were subjected to fatty acid analysis, according to the General Methods.

[0230] GC analyses showed the presence of ARA in the transformants containing the 3 chimeric genes of pZKSL-5S5A5, but not in the parent L135U9 strain. Five strains (i.e., #28, #62, #73, #84 and #95) that produced about 32.2%, 32.9%, 34.4%, 32.1% and 38.6% ARA of TFAs were designated as strains Y8000, Y8001, Y8002, Y8003 and Y8004, respectively. Further analyses showed that the three chimeric genes of pZKSL-5S5A5 were not integrated into the *Lys5* site in the Y8000, Y8001, Y8002, Y8003 and Y8004 strains. All strains possessed a *Lys*⁺ phenotype.

[0231] The final genotype of strains Y8000, Y8001, Y8002, Y8003 and Y8004 with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura*⁻, *Pex3*⁻, *unknown 1*⁻, *unknown 2*⁻, *Leu*⁺, *Lys*⁺, YAT1::ME3S::Pex16, GPD::FmD12::Pex20, YAT1::FmD12::OCT, GPAT::EgD9e::Lip2, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, FBAINm::EgD8M::Pex20, EXP1::EgDBM::Pex16, FBAIN::EgD5SM::Pex20, EXP1::EgD5M::Pex16, YAT1::EaD5SM::OCT.

Generation Of Y8006 Strain To Produce About 41 %ARA Of TFAs

[0232] Construct pZP3-Pa777U (FIG. 4B; SEQ ID NO:129; described in Table 9 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was generated to integrate three Δ17 desaturase genes into the *Pox3* loci (GenBank Accession No. AJ001301) of strain Y8002.

[0233] The pZP3-Pa777U plasmid was digested with *AscI/SphI*, and then used for transformation of strain Y8002 according to the General Methods. The transformant cells were plated onto MM plates and maintained at 30 °C for 2 to 3 days. Single colonies were then re-streaked onto MM plates, and then inoculated into liquid MM at 30 °C and shaken at 250 rpm/min for 2 days. The cells were subjected to fatty acid analysis, according to the General Methods.

[0234] GC analyses showed the presence of 26% to 31 % EPA of TFAs in most of the selected 96 transformants containing the 3 chimeric genes of pZP3-Pa777U, but not in the parent Y8002 strain. Strain #69 produced about 38% EPA of TFAs and was designated as Y8007. There was one strain (i.e., strain #9) that did not produce EPA, but produced about 41% ARA of TFAs. This strain was designated as Y8006. Based on the lack of EPA production in strain Y8006, its genotype with respect to wildtype *Yarrowia lipolytica* ATCC #20362 is assumed to be *Pex3*⁻, *unknown 1*⁻, *unknown 2*⁻, *unknown 3*⁻, *Leu*⁺, *Lys*⁺, Ura⁺, YAT1::ME3S::Pex16, GPD::FmD12::Pex20, YAT1::FmD12::OCT, GPAT::EgD9e::Lip2, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, FBAINm::EgD8M::Pex20, EXP1::EgDBM::Pex16, FBAIN::EgD5SM::Pex20, EXP1::EgD5M::Pex16, YAT1::EaD5SM::OCT.

[0235] In contrast, the final genotype of strain Y8007 with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Pex3*⁻, *unknown 1*⁻, *unknown 2*⁻, *unknown 3*⁻, *Leu*⁺, *Lys*⁺, Ura⁺, YAT1::ME3S::Pex16, GPD::FmD12::Pex20, YAT1::FmD12::OCT, GPAT::EgD9e::Lip2, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, FBAINm::EgD8M::Pex20, EXP1::EgDBM::Pex16, FBAIN::EgD5SM::Pex20, EXP1::EgD5M::Pex16, YAT1::EaD5SM::OCT, YAT1::PaD17S::Lip1, EXP1::PaD17::Pex16, FBAINm::PaD17::Aco (wherein PaD17 is a *Pythium aphanidermatum* Δ17 desaturase [U.S. Patent 7,556,949] and PaD17S is a codon-optimized Δ17 desaturase, derived from *Pythium aphanidermatum* [U.S. Patent 7,556,949].

[0236] Integration of the 3 chimeric genes of pZP3-Pa777U into the *Pox3* loci (GenBank Accession No. AJ001301) in strains Y8006 and Y8007 was not confirmed.

EXAMPLE 3

Generation Of *Yarrowia lipolytica* Strains Producing From About 24% To About 56% EPA Of Total Fatty Acids ("TFAs")

[0237] The present Example describes the construction of strain Y8412, derived from *Yarrowia lipolytica* ATCC #20362, capable of producing about 56% EPA relative to the total lipids via expression of a Δ9

elongase/ $\Delta 8$ desaturase pathway.

[0238] The development of strain Y8412 (FIG. 2) required the construction of strains Y2224, Y4001, Y4001 U, Y4036, Y4036U and L135 (described in Example 1), strains L135U9 and Y8002 (described in Example 2), and strains Y8006U6, Y8069, Y8069U, Y8154, Y8154U, Y8269 and Y8269U.

Generation Of Strain Y8006U6 (Ura3-)

[0239] In order to disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:130; described in Table 15 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y8006.

[0240] Plasmid pZKUM was digested with *SaII*/*PacI*, and then used to transform strain Y8006 according to the General Methods. Following transformation, cells were plated onto MM + 5-FOA selection plates and maintained at 30 °C for 2 to 3 days.

[0241] A total of 8 transformants grown on MM + 5-FOA plates were picked and re-streaked onto MM plates and MM + 5-FOA plates, separately. All 8 strains had a *Ura*- phenotype (i.e., cells could grow on MM + 5-FOA plates, but not on MM plates). The cells were scraped from the MM + 5-FOA plates and subjected to fatty acid analysis, according to the General Methods.

[0242] GC analyses showed the presence of 22.9%, 25.5%, 23.6%, 21.6%, 21.6% and 25% ARA in the pZKUM-transformant strains #1, #2, #4, #5, #6 and #7, respectively, grown on MM + 5-FOA plates. These six strains were designated as strains Y8006U1, Y8006U2, Y8006U3, Y8006U4, Y8006U5 and Y8006U6, respectively (collectively, Y8006U).

Generation Of Y8069 Strain To Produce About 37.5% EPA Of TFAs

[0243] Construct pZP3-Pa777U (FIG. 4B; SEQ ID NO:129; described in Table 9 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was used to integrate three $\Delta 17$ desaturase genes into the *Pox3* loci (GenBank Accession No. AJ001301) of strain Y8006U6.

[0244] The pZP3-Pa777U plasmid was digested with *AscI*/*SphI*, and then used for transformation of strain Y8006U6 according to the General Methods. The transformant cells were plated onto MM plates and maintained at 30 °C for 2 to 3 days. Single colonies were then re-streaked onto MM plates, and then inoculated into liquid MM at 30 °C and shaken at 250 rpm/min for 2 days. The cells were subjected to fatty acid analysis, according to the General Methods.

[0245] GC analyses showed the presence of EPA in the transformants containing the 3 chimeric genes of pZP3-Pa777U, but not in the parent Y8006U6 strain. Most of the selected 24 strains produced 24-37% EPA of TFAs. Four strains (i.e., #1, #6, #11 and #14) that produced 37.5%, 43.7%, 37.9% and 37.5% EPA of TFAs were designated as Y8066, Y8067, Y8068 and Y8069, respectively. Integration of the 3 chimeric genes of pZP3-Pa777U into the *Pox3* loci (GenBank Accession No. AJ001301) of strains Y8066, Y8067, Y8068 and Y8069 was not confirmed.

[0246] The final genotype of strains Y8066, Y8067, Y8068 and Y8069 with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura*+, *Pex3*-, *unknown 1*-, *unknown 2*-, *unknown 3*-, *unknown 4*-, *Lou*+, *Lys*+, *YAT1*::ME3S::Pex16, *GPD*::Fmd12::Pex20, *YAT1*::Fmd12::Oct, *GPAT*::EgD9e::Lip2, *FBAINm*::EgD9eS::Lip2, *EXP1*::EgD9eS::Lip1, *FBAINm*::EgD8M::Pex20, *EXP1*::EgD8M::Pex16, *FBAIN*::EgD5SM::Pex20, *EXP1*::EgD5M::Pex16, *YAT1*::EaD5M::Oct, *YAT1*::PaD17S::Lip1, *EXP1*::PaD17::Pex16, *FBAINm*::PaD17::Aco.

Generation Of Strain Y8069U (Ura3-)

[0247] In order to disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:130; described in Table 15 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y8069, in a manner similar to that described for pZKUM transformation of strain Y8006 (*supra*). A total of 3 transformants were grown and identified to possess a *Ura*- phenotype.

[0248] GC analyses showed the presence of 22.4%, 21.9% and 21.5% EPA in the pZKUM-transformant strains #1, #2 and #3, respectively, grown on MM + 5-FOA plates. These three strains were designated as strains Y8069U1, Y8069U2, and Y8069U3, respectively (collectively, Y8069U).

Generation Of Strain Y8154 To Produce about 44.8% EPA Of TFAs

[0249] Construct pZKL2-5mB89C (FIG. 5B; SEQ ID NO:131) was generated to integrate one $\Delta 5$ desaturase gene, one $\Delta 9$ elongase gene, one $\Delta 8$ desaturase gene, and one *Yarrowia lipolytica* diacylglycerol cholinephosphotransferase gene (CPT1) into the *Lip2* loci (GenBank Accession No. AJ012632) of strain Y8069U3 to thereby enable higher level production of EPA. The pZKL2-5mB89C plasmid contained the following components:

Table 7: Description of Plasmid pZKL2-5mB89C (SEQ ID NO:131)

RE Sites And Nucleotides Within SEQ ID NO:131	Description Of Fragment And Chimeric Gene Components
<i>AscI</i> / <i>Bst</i> NI (730-1)	722 bp 5' portion of <i>Yarrowia Lip2</i> gene (labeled as "Lip2.5N" in Figure; GenBank Accession No. AJ012632)
<i>PacI</i> / <i>SphI</i> (4141-3438)	697 bp 3' portion of <i>Yarrowia Lip2</i> gene (labeled as "Lip2.3N" in Figure; GenBank Accession No. AJ012632)
<i>SwaI</i> / <i>Bst</i> WI (13561-1)	YAT1::YCPT1::Aco, comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • YCPT1: <i>Yarrowia lipolytica</i> diacylglycerol cholinephosphotransferase gene (SEQ ID NO:37) (labeled as "Y. lipolytica CPT1 cDNA" in Figure; Intl. App. Pub. No. WO 2006/052870); • Aco: Aco terminator sequence from <i>Yarrowia Aco</i> gene (GenBank Accession No. AJ001300)
<i>PmeI</i> / <i>SwaI</i> (10924-13561)	FBAIN::EgD8M::Lip1 comprising: • FBAIN: <i>Yarrowia lipolytica</i> FBAIN promoter (U.S. Patent 7,202,356); • EgD8M: Synthetic mutant $\Delta 8$ desaturase (SEQ ID NO:59; U.S. Patent 7,709,239), derived from <i>Euglena gracilis</i> ("EgD8S"; U.S. Patent 7,256,033) (labeled as "D8S-23" in Figure); • Lip1: Lip1 terminator sequence from <i>Yarrowia Lip1</i> gene (GenBank Accession No. Z50020)
<i>PmeI</i> / <i>ClaI</i> (10924-9068)	YAT1::EgD9eS::Lip2, comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • EgD9eS: codon-optimized $\Delta 9$ elongase (SEQ ID NO:45), derived from <i>Euglena gracilis</i> (U.S. Patent 7,645,604); • Lip2: Lip2 terminator sequence from <i>Yarrowia Lip2</i> gene (GenBank Accession No. AJ012632)
<i>ClaI</i> / <i>EcoRI</i> (9068-6999)	<i>Yarrowia Ura3</i> gene (GenBank Accession No. AJ306421)
<i>EcoRI</i> / <i>PacI</i> (6999-4141)	GPDI::EgD5SM::ACO, comprising: • GPDI: <i>Yarrowia lipolytica</i> GPDI promoter (U.S. Patent 7,459,546); • EgD5SM: Synthetic mutant $\Delta 5$ desaturase (SEQ ID NO:71; U.S. Pat. Pub. No. 2010-0075386-A1), derived from <i>Euglena gracilis</i> (U.S. Patent 7,678,560) (labeled as "EgD5S-HPGS" in Figure); • Aco: Aco terminator sequence from <i>Yarrowia Aco</i> gene (GenBank Accession No. AJ001300)

[0250] The pZKL2-5mB89C plasmid was digested with *AscI*/*SphI*, and then used for transformation of strain Y8069U3 according to the General Methods. The transformant cells were plated onto MM plates and maintained at 30 °C for 3 to 4 days. Single colonies were re-streaked onto MM plates, and then inoculated into liquid MM at 30 °C and shaken at 250 rpm/min for 2 days. The cells were collected by centrifugation, resuspended in HGM and then shaken at 250 rpm/min for 5 days. The cells were subjected to fatty acid analysis, according to the General Methods.

[0251] GC analyses showed that most of the selected 96 strains produced approximately 38-44% EPA of TFAs. Seven strains (i.e., #1, #39, #49, #62, #70, #85 and #92) that produced about 44.7%, 45.2%, 45.4%, 44.8%, 46.1 %, 48.6% and 45.9% EPA of TFAs were designated as strains Y8151, Y8152, Y8153, Y8154, Y8155, Y8156 and Y8157, respectively. Knockout of the Lip2 gene was not confirmed in

these EPA strains.

[0252] The final genotype of strains Y8151, Y8152, Y8153, Y8154, Y8155, Y8156 and Y8157 with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura+*, *Pex3-*, *unknown 1-*, *unknown 2-*, *unknown 3-*, *unknown 4-*, *unknown 5-*, *Leu+*, *Lys+*, YAT1::ME3S::Pex16, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9eS::Lip2, YAT1::EgD9eS::Lip2, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::EgD5M::Pex16, YAT1::EaD5SM::Oct, FBAIN::EgD5SM::Pex20, GPDIN::EgD5SM::Aco, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::YCPT::Aco.

Generation Of Strain Y8154U1 (*Ura3-*)

[0253] In order to disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:130; described in Table 15 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y8154, in a manner similar to that described for pZKUM transformation of strain Y8006 (*supra*). A total of 8 transformants were grown and identified to possess a *Ura-* phenotype.

[0254] GC analyses showed that there was 23.1 % EPA of TFAs in the pZKUM-transformant strain #7. This strain was designated as strain Y8154U1.

Generation Of Strain Y8269 To Produce About 45.3% EPA Of TFAs

[0255] Construct pZKL1-2SR9G85 (FIG. 6A; SEQ ID NO:132) was generated to integrate one DGLA synthase, one $\Delta 12$ desaturase gene and one $\Delta 5$ desaturase gene into the *Lip1* loci (GenBank Accession No. Z50020) of strain Y8154U1 to thereby enable higher level production of EPA. A DGLA synthase is a multizyme comprising a $\Delta 9$ elongase linked to a $\Delta 8$ desaturase.

[0256] The pZKL1-2SR9G85 plasmid contained the following components:

Table 8: Description of Plasmid pZKL1-2SR9G85 (SEQ ID NO:132)

RE Sites And Nucleotides Within SEQ ID NO:132	Description Of Fragment And Chimeric Gene Components
<i>AscI/BsNI</i> (4189-3373)	809 bp 5' portion of <i>Yarrowia Lip1</i> gene (labeled as "Lip1-5'N" in Figure; GenBank Accession No. Z50020)
<i>PacI/SphI</i> (7666-6879)	763 bp 3' portion of <i>Yarrowia Lip1</i> gene (labeled as "Lip1.3N" in Figure; GenBank Accession No. Z50020)
<i>ClaI/SwaI</i> (1-3217)	YAT1::E389D9eS/EgD8M::Lip1, comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • E389D9eS/EgD8M: gene fusion comprising a codon-optimized $\Delta 9$ elongase derived from <i>Eutryptella</i> sp. CCMP389 ("E389D9eS"), a linker, and the synthetic mutant $\Delta 8$ desaturase derived from <i>Euglena gracilis</i> ("EgD8M") (SEQ ID NO:12) (labeled as individually as "E389S", "Linker" and "EgD8M" in Figure; U.S. Pat. Appl. Pub. No. 2008-0254191-A1); • Lip1: Lip1 terminator sequence from <i>Yarrowia Lip1</i> gene (GenBank Accession No. Z50020)
<i>SalI/ClaI</i> (11982-1)	GPM::EgD5SM::Oct comprising: • GPM: <i>Yarrowia lipolytica</i> GPM promoter (labeled as "GPML" in Figure; U.S. Patent 7,202,356); • EgD5SM: Synthetic mutant $\Delta 5$ desaturase (SEQ ID NO:71; U.S. Pat. Pub. No. 2010-0075386-A1), derived from <i>Euglena gracilis</i> (U.S. Patent 7,678,560) (labeled as "ED5S" in Figure); • OCT: OCT terminator sequence of <i>Yarrowia</i> OCT gene (GenBank Accession No. X69988)
<i>SalI/EcoRI</i> (11982-10363)	<i>Yarrowia Ura3</i> gene (GenBank Accession No. AJ306421)
<i>EcoRI/PacI</i> (10363-7666)	EXP1::FmD12S::ACO, comprising: • EXP1: <i>Yarrowia lipolytica</i> export protein (EXP1) promoter (labeled as "Exp" in Figure; Intl. App. Pub. No. WO 2006/052870); • FmD12S: codon-optimized $\Delta 12$ elongase (SEQ ID NO:93), derived from <i>Fusarium moniliforme</i> (labeled as "FD12S" in Figure; U.S. Patent 7,504,259); • Aco: Aco terminator sequence from <i>Yarrowia Aco</i> gene (GenBank Accession No. AJ001300)

[0257] The pZKL1-2SR9G85 plasmid was digested with *AscI/SphI*, and then used for transformation of strain Y8154U1 according to the General Methods. The transformant cells were plated onto MM plates and maintained at 30 °C for 3 to 4 days. Single colonies were re-streaked onto MM plates, and then inoculated into liquid MM at 30 °C and shaken at 250 rpm/min for 2 days. The cells were collected by centrifugation, resuspended in HGM and then shaken at 250 rpm/min for 5 days. The cells were subjected to fatty acid analysis, according to the General Methods.

[0258] GC analyses showed that most of the selected 96 strains produced 40-44.5% EPA of total lipids. Five strains (i.e., #44, #46, #47, #66 and #87) that produced about 44.8%, 45.3%, 47%, 44.6% and 44.7% EPA of TFAs were designated as Y8268, Y8269, Y8270, Y8271 and Y8272, respectively. Knockout of the *Lip1* loci (GenBank Accession No. Z50020) was not confirmed in these EPA strains.

[0259] The final genotype of strains Y8268, Y8269, Y8270, Y8271 and Y8272 with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura+*, *Pex3-*, *unknown 1-*, *unknown 2-*, *unknown 3-*, *unknown 4-*, *unknown 5-*, *unknown 6-*, YAT1::ME3S::Pex16, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9eS::Lip2, YAT1::EgD9eS::Lip2, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, YAT1::E389D9eS/EgD8M::Lip1, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, EXP1::EgD5M::Pex16, YAT1::EaD5SM::Oct, FBAIN::EgD5SM::Pex20, GPDIN::EgD5SM::Aco, GPM::EgD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::YCPT::Aco.

Generation Of Strain Y8269U (*Ura3-*)

[0260] In order to disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:130; described in Table 15 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y8269, in a manner similar to that described for pZKUM transformation of strain Y8006 (*supra*). A total of 8 transformants were grown and identified to possess a *Ura-* phenotype.

[0261] GC analyses showed that there were 23.0%, 23.1% and 24.2% EPA of TFAs in pZKUM-transformant strains #2, #3 and #5, respectively. These strains were designated as strains Y8269U1, Y8269U2 and Y8269U3, respectively (collectively, Y8269U).

Generation Of Strain Y8406 And Strain Y8412 To Produce About 51.2% EPA And 55.8% EPA Of TFAs

[0262] Construct pZSCP-Ma83 (FIG. 6B; SEQ ID NO:133) was generated to integrate one $\Delta 8$ desaturase gene, one $C_{16/18}$ elongase gene and one malonyl-CoA synthetase gene into the SCP2 loci (GenBank Accession No. XM_503410) of strain Y8269U1 to thereby enable higher level production of EPA. The pZSCP-Ma83 plasmid contained the following components:

Table 9: Description of Plasmid pZSCP-Ma83 (SEQ ID NO:133)

RE Sites And Nucleotides Within SEQ ID NO:133	Description Of Fragment And Chimeric Gene Components
<i>BsNI/AscI</i> (1-1328)	1327 bp 3' portion of <i>Yarrowia SCP2</i> gene (labeled as "SCP2-3'" in Figure; GenBank Accession No. XM_503410)
<i>SphI/PacI</i> (4036-5816)	1780 bp 5' portion of <i>Yarrowia SCP2</i> gene (labeled as "SCP2-5'" in Figure; GenBank Accession No. XM_503410)
<i>SwaI/BsNI</i> (12994-1)	GPD::ME3S::Pex20, comprising: • GPD: <i>Yarrowia lipolytica</i> GPD promoter (U.S. Patent 7,259,255); • ME3S: codon-optimized $C_{16/18}$ elongase gene (SEQ ID NO:97), derived from <i>M. alpina</i> (U.S. Patent 7,470,532); • Pex20: Pex20 terminator sequence from <i>Yarrowia Pex20</i> gene (GenBank Accession No. AF054613)
<i>PmeI/SwaI</i> (10409-12994)	YAT1::MCS::Lip1 comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006/0094102-A1);

RE Sites And Nucleotides Within SEQ ID NO:133	Description Of Fragment And Chimeric Gene Components
	• MCS: codon-optimized malonyl-CoA synthetase gene (SEQ ID NO:41), derived from <i>Rhizobium leguminosarum</i> bv. viciae 3841 (U.S. Patent Application No. 12/637877); • Lip1: Lip1 terminator sequence from <i>Yarrowia Lip1</i> gene (GenBank Accession No. Z50020)
<i>ClaI/PmeI</i> (7917-10409)	GPD::EaD8S::Pex16 comprising: • GPD: <i>Yarrowia lipolytica</i> GPD promoter (U.S. Patent 7,259,255); • EaD8S: codon-optimized $\Delta 8$ desaturase gene (SEQ ID NO:63), derived from <i>Euglena anabaena</i> (U.S. Pat. Appl. Pub. No. 2008-0254521-A1); • Pex16: Pex16 terminator sequence from <i>Yarrowia Pex16</i> gene (GenBank Accession No. U75433)
<i>SaI/EcoRI</i> (7467-5848)	<i>Yarrowia Ura3</i> gene (GenBank Accession No. AJ306421)

[0263] The pZSCP-Ma83 plasmid was digested with *AscI/SphI*, and then used for transformation of strains Y8269U1, Y8269U2 and Y8269U3, separately, according to the General Methods. The transformant cells were plated onto MM plates and maintained at 30 °C for 3 to 4 days. Single colonies were re-streaked onto MM plates, and then inoculated into liquid MM at 30 °C and shaken at 250 rpm/min for 2 days. The cells were collected by centrifugation, resuspended in HGM and then shaken at 250 rpm/min for 5 days. The cells were subjected to fatty acid analysis, according to the General Methods.

[0264] A total of 96 strains resulting from each pZSCP-Ma83 transformation (i.e., into Y8269U1, Y8269U2 and Y8269U3) were analyzed by GC. Most of the selected 288 strains produced 43-47% EPA of TFAs. Seven strains of Y8269U1 transformed with pZSCP-Ma83 (i.e., #59, #61, #65, #67, #70, #81 and #94) that produced about 51.3%, 47.9%, 50.8%, 48%, 47.8%, 47.8% and 47.8% EPA of TFAs were designated as strains Y8404, Y8405, Y8406, Y8407, Y8408, Y8409 and Y8410, respectively. Three strains of Y8269U2 transformed with pZSCP-Ma83 (i.e., #4, #13 and #17) that produced about 48.8%, 50.8%, and 49.3% EPA of TFAs were designated as Y8411, Y8412 and Y8413, respectively. And, two strains of Y8269U3 transformed with pZSCP-Ma83 (i.e., #2 and #16) that produced about 49.3% and 53.5% EPA of TFAs were designated as Y8414 and Y8415, respectively.

[0265] Knockout of the *SCP2* loci (GenBank Accession No. XM_503410) in strains Y8404, Y8405, Y8406, Y8407, Y8408, Y8409, Y8410, Y8411, Y8412, Y8413, Y8414 and Y8415 was not confirmed in any of these EPA strains, produced by transformation with pZSCP-Ma83.

[0266] The final genotype of strains Y8404, Y8405, Y8406, Y8407, Y8408, Y8409, Y8410, Y8411, Y8412, Y8413, Y8414 and Y8415 with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura+*, *Pex3-*, *unknown 1-*, *unknown 2-*, *unknown 3-*, *unknown 4-*, *unknown 5-*, *unknown 6-*, *unknown 7-*, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9e::Lip2, YAT1::EgD9eS::Lip2, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, GPD::EaD8S::Pex16, YAT1::E389D9eS/EgD8M::Lip1, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, EXP1::EgD5M::Pex16, YAT1::EaD5SM::Oct, FBAIN::EgD5SM::Pex20, GPDIN::EgD5SM::Aco, GPM::EgD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::YICPT::Aco, YAT1::MCS::Lip1.

[0267] *Yarrowia lipolytica* strain Y8406 was deposited with the American Type Culture Collection on May 14, 2009 and bears the designation ATCC PTA-10025. *Yarrowia lipolytica* strain Y8412 was deposited with the American Type Culture Collection on May 14, 2009 and bears the designation ATCC PTA-10026.

Analysis Of Total Lipid Content And Composition By Flask Assay

[0268] Cells from YPD plates of strains Y8404, Y8405, Y8406, Y8407, Y8408, Y8409, Y8410, Y8411, Y8412, Y8413, Y8414 and Y8415 were grown and analyzed for total lipid content and composition, according to the General Methods. Table 10 summarizes the total dry cell weight of the cells ["DCW"], the total lipid content of cells ["TFAs % DCW"], the concentration of each fatty acid as a weight percent of TFAs ["% TFAs"] and the EPA content as a percent of the dry cell weight ["EPA % DCW"]. More specifically, fatty acids are identified as 16:0 (palmitate), 16:1 (palmitoleic acid), 18:0 (stearic acid), 18:1 (oleic acid), 18:2 (LA), ALA, EDA, DGLA, ARA, ETrA, ETA, EPA and other.

Table 10. Total Lipid Content And Composition In *Yarrowia* Strains Y8404, Y8405, Y8406, Y8407, Y8408, Y8409, Y8410, Y8411, Y8412, Y8413, Y8414 And Y8415 By Flask Assay

Strain	DCW (g/L)	TFAs % DCW	% TFAs													EPA % DCW
			16:0	16:1	18:0	18:1	18:2	ALA	EDA	DGLA	ARA	ETrA	ETA	EPA	other	
Y8404	4.1	27.3	2.8	0.8	1.8	5.1	20.4	2.1	2.9	2.5	0.6	0.8	2.4	51.1	6.3	14.0
Y8405	3.9	29.6	2.7	0.5	2.9	5.7	20.5	2.8	2.7	2.1	0.5	0.7	2.0	51.4	5.1	15.2
Y8406	4.0	30.7	2.6	0.5	2.9	5.7	20.3	2.8	2.8	2.1	0.5	0.8	2.1	51.2	5.4	15.7
Y8407	4.0	29.4	2.6	0.5	3.0	5.6	20.5	2.8	2.7	2.1	0.4	0.7	2.1	51.5	5.1	15.2
Y8408	4.1	29.8	2.9	0.6	2.7	5.7	20.2	2.8	2.6	2.1	0.5	0.9	2.1	51.2	5.5	15.3
Y8409	3.9	30.8	2.8	0.5	2.9	5.7	20.6	2.7	2.7	2.1	0.5	0.8	2.1	51.0	5.2	15.7
Y8410	4.0	31.8	2.7	0.5	3.0	5.7	20.5	2.9	2.7	2.1	0.5	0.7	2.1	50.9	5.3	16.2
Y8411	3.6	30.5	2.7	0.3	3.3	5.1	19.9	2.6	2.4	2.0	0.5	0.6	1.8	52.9	5.7	16.1
Y8412	3.2	27.0	2.5	0.4	2.6	4.3	19.0	2.4	2.2	2.0	0.5	0.6	1.9	55.8	5.6	15.1
Y8413	2.9	27.2	3.1	0.4	2.6	5.4	19.9	2.2	2.8	2.0	0.5	0.7	1.8	52.4	5.9	14.2
Y8414	3.7	27.1	2.5	0.7	2.3	6.0	19.9	1.6	3.4	3.4	0.6	0.6	3.1	49.4	6.1	13.4
Y8415	3.6	25.9	1.4	0.3	1.9	4.5	16.0	1.3	2.7	2.9	0.5	0.6	2.5	59.0	6.1	15.3

EXAMPLE 4

Generation Of *Yarrowia lipolytica* Strain Y8647 To Produce About 53.6% EPA Of Total Fatty Acids ["TFAs"] With 37.6% Total Lipid Content

[0269] The present Example describes the construction of strain Y8647, derived from *Yarrowia lipolytica* ATCC #20362, capable of producing about 53.6% EPA relative to the total lipids with 37.6% total lipid content ["TFAs % DCW"] via expression of a $\Delta 9$ elongase/ $\Delta 8$ desaturase pathway. The development of strain Y8647 (FIG. 2) required the construction of strains Y2224, Y4001, Y4001 U, Y4036, Y4036U and L135 (described in Example 1), strains L135U9 and Y8002 (described in Example 2), strains Y8006U6, Y8069, Y8069U, Y8154, Y8154U, Y8269 and Y8269U (described in Example 3) and strain Y8412U6.

Generation Of Strain Y8412U6 (*Ura3-*)

[0270] In order to disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:130; described in Table 15 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y8412 (Example 3) in a manner similar to that described for pZKUM transformation of strain Y8006 (Example 3). A total of 8 transformants were grown and identified to possess a *Ura*-phenotype.

[0271] GC analyses showed that there were 25.9% and 26.9% EPA of TFAs in pZKUM-transformant strains #4 and #6, respectively. These two strains were designated as strains Y8412U6 and Y8412U8, respectively (collectively, Y8412U).

Generation Of Strain Y8647

[0272] Construct pZKL4-398F2 (FIG. 7A; SEQ ID NO:134) was generated to integrate one $C_{16/18}$ elongase gene, one DGLA synthase, and one $\Delta 12$ desaturase gene into the *Yarrowia* lipase-like locus (designated as *Lip4*, GenBank Accession No. XM_503825) of strain Y8412U6 to thereby enable higher level production of EPA. The pZKL4-398F2 plasmid contained the following components:

Table 11. Description Of Plasmid pZKL4-398F2 (SEQ ID NO:134)

RE Sites And Nucleotides Within SEQ ID NO:134	Description Of Fragment And Chimeric Gene Components
<i>AscI</i> / <i>BsNI</i> (11164-10412)	745 bp 5' portion of <i>Yarrowia Lipase 4</i> locus (labeled as "Lip4" in Figure; GenBank Accession No. XM_503825)
<i>PacI</i> / <i>SphI</i> (1-13872)	782 bp 3' portion of <i>Yarrowia Lipase 4</i> locus (labeled as "Lip4-3" in Figure; GenBank Accession No. XM_503825)
<i>EcoRI</i> / <i>PacI</i> (2877-1)	GPDIN::FmD12::Pex16, comprising: • GPDIN: <i>Yarrowia lipolytica</i> GPDIN promoter (U.S. Patent 7,459,546); • FmD12: <i>Fusarium moniliforme</i> Δ 12 desaturase (SEQ ID NO:91) (labeled as "F.D12" in Figure; U.S. Patent 7,504,259); • Pex16: Pex16 terminator sequence from <i>Yarrowia Pex16</i> gene (GenBank Accession No. U75433)
<i>PmeI</i> / <i>SwaI</i> (8361-10256)	YAT1::ME3S::Lip1 comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • ME3S: codon-optimized <i>C_{16:1n}</i> elongase gene (SEQ ID NO:97), derived from <i>M. alpina</i> (U.S. Patent 7,470,532); • Lip1: Lip1 terminator sequence from <i>Yarrowia Lip1</i> gene (GenBank Accession No. Z50020)
<i>SwaI</i> / <i>ClaI</i> (8325-4946)	FBAINm::EaD9eS/EaD8S::Lip2 comprising: • FBAINm: <i>Yarrowia lipolytica</i> FBAINm promoter (U.S. Patent 7,202,356); • EaD9eS/EaD8S: gene fusion comprising a codon-optimized Δ 9 elongase derived from <i>Euglena anabaena</i> ("EaD9eS"), a linker, and a codon-optimized Δ 8 desaturase derived from <i>Euglena anabaena</i> ("EaD8S") (SEQ ID NO:63) (labeled as individually as "EaD9eS", "Linker" and "EaD8S" in Figure; U.S. Pat. Appl. Pub. No. 2008-0254191-A1); • Lip2: Lip2 terminator sequence from <i>Yarrowia Lip2</i> gene (GenBank Accession No. AJ012632)
<i>ClaI</i> / <i>EcoRI</i> (49146-2877)	<i>Yarrowia</i> Ura3 gene (GenBank Accession No. AJ306421)

[0273] The pZKL4-398F2 plasmid was digested with *AscI*/*SphI*, and then used for transformation of strain Y8412U6, according to the General Methods. The transformant cells were plated onto MM plates and maintained at 30 °C for 3 to 4 days. Single colonies were re-streaked onto MM plates, and then inoculated into liquid MM at 30 °C and shaken at 250 rpm/min for 2 days. The cells were collected by centrifugation, resuspended in HGM and then shaken at 250 rpm/min for 5 days. The cells were subjected to fatty acid analysis, according to the General Methods.

[0274] GC analyses showed that most of the selected 96 transformant strains produced 50-52.7% EPA of TFAs. Seven strains (i.e., #31, #35, #38, #41, #60, #61 and #95) that produced about 52.8%, 53.1%, 52.8%, 53.2%, 53.1%, 52.8%, and 52.9% EPA of TFAs were designated as Y8646, Y8647, Y8648, Y8649, Y8650, Y8651 and Y8652, respectively.

[0275] Knockout of the *Lip4* locus (GenBank Accession No. XM_503825) in these EPA strains was not confirmed.

[0276] The final genotype of strains Y8646, Y8647, Y8648, Y8649, Y8650, Y8651 and Y8652 with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura*⁺, *Pex*⁻, *unknown 1*⁻, *unknown 2*⁻, *unknown 3*⁻, *unknown 4*⁻, *unknown 5*⁻, *unknown 6*⁻, *unknown 7*⁻, *unknown 8*⁻, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9eS::Lip2, YAT1::EgD9eS::Lip2, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAINm::EgD8M::Lip1, GPD::EaD8S::Pex16, YAT1::E389D9eS/EgD8M::Lip1, FBAINm::EaD9eS/EaD8S::Lip2, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, GPDIN::FmD12::Pex16, EXP1::EgD5M::Pex16, FBAINm::EgD5SM::Pex20, GPDIN::EgD5SM::Aco, GPM::EgD5SM::Oct, YAT1::EaD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::YCPT::Aco, YAT1::MCS::Lip1.

Analysis Of Total Lipid Content And Composition By Flask Assay

[0277] Cells from YPD plates of strains Y8647, Y8648, Y8649 and Y8650 were grown and analyzed for total lipid content and composition, according to the General Methods. Table 12 summarizes the total dry cell weight of the cells ["DCW"], the total lipid content of cells ["TFAs % DCW"], the concentration of each fatty acid as a weight percent of TFAs ["% TFAs"] and the EPA content as a percent of the dry cell weight ["EPA % DCW"]. Fatty acids are identified as 16:0 (palmitate), 16:1 (palmitoleic acid), 18:0 (stearic acid), 18:1 (oleic acid), 18:2 (LA), ALA, EDA, DGLA, ARA, ETrA, ETA, EPA and other.

Table 12: Total Lipid Content And Composition In *Yarrowia* Strains Y8647, Y8648, Y8649 And Y8650 By Flask Assay

Strain	DCW (g/L)	TFAs % DCW	% TFAs													EPA % DCW
			16:0	16:1	18:0	18:1	18:2	ALA	EDA	DGLA	ARA	ETrA	ETA	EPA	other	
Y8647	3.8	37.6	1.3	0.2	2.1	4.7	20.3	1.7	3.3	3.6	0.7	0.6	3.0	53.6	4.5	20.1
Y8648	3.5	27.8	2.3	0.3	2.7	4.3	18.6	2.3	2.1	2.2	0.6	0.6	1.9	56.7	4.9	15.7
Y8649	3.6	27.9	2.4	0.3	2.9	3.7	18.8	2.2	2.1	2.4	0.6	0.8	2.1	55.8	5.5	15.6
Y8650	3.5	28.2	2.2	0.3	2.9	3.8	18.8	2.4	2.1	2.4	0.6	0.6	2.1	56.1	5.3	15.8

EXAMPLE 5

Generation Of *Yarrowia lipolytica* Strain Y9028 To Produce About 54.5% EPA Of Total Fatty Acids ["TFAs"] With 39.6% Total Lipid Content

[0278] The present Example describes the construction of strain Y9028, derived from *Yarrowia lipolytica* ATCC #20362, capable of producing about 54.5% EPA relative to the total lipids with 39.6% total lipid content ["TFAs % DCW"] via expression of a Δ 9 elongase/ Δ 8 desaturase pathway. The development of strain Y9028 (FIG. 2) required the construction of strains Y2224, Y4001, Y4001 U, Y4036, Y4036U and L135 (described in Example 1), strains L135U8 and Y8002 (described in Example 2), strains Y8006U6, Y8069, Y8069U, Y8154, Y8154U, Y8269 and Y8269U (described in Example 3), strains Y8412U6 and Y8647 (described in Example 4) and strain Y8647U.

Generation Of Strain Y8647U (*Ura*⁻)

[0279] In order to disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:130; described in Table 15 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y8647 (Example 4) in a manner similar to that described for pZKUM transformation of strain Y8006 (Example 3). A total of 12 transformants were grown and identified to possess a *Ura*⁻ phenotype.

[0280] GC analyses showed that there were 30.2%, 29.2%, 28.1% and 29.9% EPA of TFAs in pZKUM-transformant strains #1, #3, #4 and #12, respectively. These four strains were designated as strains Y8647U1, Y8647U2, Y8647U3, and Y8647U6, respectively (collectively, Y8647U).

Generation Of Strain Y9028

[0281] Construct pZP2-85m98F (FIG. 7B; SEQ ID NO:135) was generated to integrate one Δ 8 desaturase gene, one DGLA synthase and one Δ 5 desaturase gene into the *Yarrowia* *Pox2* locus (GenBank Accession No. AJ001300) of strain Y8647U3 to thereby enable higher level production of EPA. The pZP2-85m98F plasmid contained the following components:

Table 13: Description Of Plasmid pZP2-85m98F (SEQ ID NO:135)

RE Sites And Nucleotides Within SEQ ID NO:135	Description Of Fragment And Chimeric Gene Components
<i>AscI</i> / <i>BsNI</i> (5986-5176)	810 bp 5' portion of <i>Yarrowia Pox2</i> gene (GenBank Accession No. AJ001300)
<i>PacI</i> / <i>SphI</i> (9349-8694)	655 bp 3' portion of <i>Yarrowia Pox2</i> gene (GenBank Accession No. AJ001300)

RE Sites And Nucleotides Within SEQ ID NO.135	Description Of Fragment And Chimeric Gene Components
<i>PmeI/SwaI</i> (2493-5020)	EXP1::EgD5SM::Lip1, comprising: - EXP1: <i>Yarrowia lipolytica</i> export protein (EXP1) promoter (labeled as "EXP" in Figure; Intl. App. Pub. No. WO 2006/052870); - EgD5SM: Synthetic mutant $\Delta 5$ desaturase (SEQ ID NO.71; U.S. Pat. Pub. No. 2010-0075386-A1), derived from <i>Euglena gracilis</i> (U.S. Patent 7,678,560); - Lip1: Lip1 terminator sequence from <i>Yarrowia Lip1</i> gene (GenBank Accession No. Z50020)
<i>ClaI/PmeI</i> (1-2493)	GPD::EaD8S::Pex16, comprising: - GPD: <i>Yarrowia lipolytica</i> GPD promoter (U.S. Patent 7,259,255); - EaD8S: codon-optimized $\Delta 8$ desaturase gene (SEQ ID NO.63), derived from <i>Euglena anabaena</i> (U.S. Pat. Appl. Pub. No. 2008-0254521-A1); - Pex16: Pex16 terminator sequence from <i>Yarrowia Pex16</i> gene (GenBank Accession No. U75433)
<i>SalI/EcoRI</i> (14170-12551)	<i>Yarrowia Ura3</i> gene (GenBank Accession No. AJ306421)
<i>EcoRI/PacI</i> (12551-9349)	YAT1::EgD9eS/EgD8M::Aco, comprising: - YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006/0094102-A1); - EgD9eS/EgD8M: gene fusion comprising a codon-optimized $\Delta 9$ elongase derived from <i>Euglena gracilis</i> ("EgD9eS"), a linker, and the synthetic mutant $\Delta 8$ desaturase derived from <i>Euglena gracilis</i> ("EgD8M") (SEQ ID NO.8) (labeled as individually as "EgD9eS", "Linker" and "EgD8M" in Figure; U.S. Pat. Appl. Pub. No. 2008-0254191-A1); - Aco: Aco terminator sequence from <i>Yarrowia Aco</i> gene (GenBank Accession No. AJ001300)

[0282] The pZP2-85m98F plasmid was digested with *AsdI/SphI*, and then used for transformation of strains of Y8647U1, Y8647U2, Y8647U3 and Y8647U6, individually, according to the General Methods. The transformant cells were plated onto MM plates and maintained at 30 °C for 3 to 4 days. Single colonies were re-streaked onto MM plates, and then inoculated into liquid MM at 30 °C and shaken at 250 rpm/min for 2 days. The cells were collected by centrifugation, resuspended in HGM and then shaken at 250 rpm/min for 5 days. The cells were subjected to fatty acid analysis, according to the General Methods.

[0283] GC analyses showed that most of the selected 48 strains of Y8647U1 transformed with pZP2-85m98F produced 49-52% EPA of TFAs. Two strains (i.e., #30 and #31) that produced about 52.6% and 52.1% EPA of TFAs were designated as Y9024 and Y9025, respectively.

[0284] Most of the selected 60 strains of Y8647U2 transformed with pZP2-85m98F produced 49-51.9% EPA of TFAs. Strain #6 produced about 52% EPA of TFAs and was designated as Y9026.

[0285] Most of the selected 60 strains of Y8647U3 transformed with pZP2-85m98F produced 50-52.2% EPA of TFAs. Six strains (i.e., #5, #6, #14, #15, #20 and #34) that produced about 53.2%, 53.7%, 54.0%, 52.9%, 53.4% and 52.3% EPA of TFAs were designated as Y9027, Y9028, Y9029, Y9030, Y9031 and Y9032, respectively.

[0286] Similarly, GC analyses showed that most of the selected 48 strains of Y8647U6 transformed with pZP2-85m98F produced 50-52.1% EPA of TFAs. Two strains (i.e., #27 and #44) that produced about 52.2% and 52.8% EPA of TFAs were designated as Y9033 and Y9034, respectively.

[0287] Knockout of the *Pox2* locus (GenBank Accession No. AJ001300) in strains Y9024, Y9025, Y9026, Y9027, Y9028, Y9029, Y9030, Y9031, Y9032, Y9033 and Y9034 was not confirmed in any of these EPA strains, produced by transformation with pZP2-85m98F.

[0288] The final genotype of these strains with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura+*, *Pex3-*, *unknown 1-*, *unknown 2-*, *unknown 3-*, *unknown 4-*, *unknown 5-*, *unknown 6-*, *unknown 7-*, *unknown 8-*, *unknown 9-*, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9eS::Lip2, YAT1::EgD9eS::Lip2, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, GPD::EaD8S::Pex16 (2 copies), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EaD9eS/EaD8S::Lip2, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, GPDIN::FmD12S::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, GPDIN::EgD5SM::Aco, GPM::EgD5SM::Oct, EXP1::EgD5SM::Lip1, YAT1::EaD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::YICPT::Aco, YAT1::MCS::Lip1.

Analysis Of Total Lipid Content And Composition By Flask Assay

[0289] Cells from YPD plates of strains Y9028, Y9029 and Y9031 were grown and analyzed for total lipid content and composition, according to the General Methods.

[0290] Table 14 below summarizes the total dry cell weight of the cells ["DCW"], the total lipid content of cells ["TFAs % DCW"], the concentration of each fatty acid as a weight percent of TFAs ["% TFAs"] and the EPA content as a percent of the dry cell weight ["EPA % DCW"]. More specifically, fatty acids are identified as 16:0 (palmitate), 16:1 (palmitoleic acid), 18:0 (stearic acid), 18:1 (oleic acid), 18:2 (LA), ALA, EDA, DGLA, ARA, ETrA, ETA, EPA and other.

Table 14: Total Lipid Content And Composition In *Yarrowia* Strains Y9028, Y9029 and Y9031 By Flask Assay

Strain	DCW (g/L)	TFAs % DCW	% TFAs													EPA % DCW
			16:0	16:1	18:0	18:1	18:2	ALA	EDA	DGLA	ARA	ETrA	ETA	EPA	other	
Y9028	3.3	39.6	1.3	0.2	2.1	4.4	19.8	1.7	3.2	2.5	0.8	0.7	1.9	54.5	6.1	21.6
Y9029	3.2	38.4	1.3	0.3	1.7	4.4	19.8	1.5	3.2	3.3	0.9	0.7	2.4	53.8	6.0	20.7
Y9031	3.3	38.6	1.3	0.3	1.8	4.7	20.1	1.7	3.2	3.2	0.9	0.8	2.6	52.3	6.3	20.2

EXAMPLE 6

Generation Of *Yarrowia lipolytica* Strains Y9481 And Y9502, Producing At Least About 57% EPA Of Total Fatty Acids ["TFAs"] With At Least About 35% Total Lipid Content

[0291] The present Example describes the construction of strains Y9481 and Y9502, derived from *Yarrowia lipolytica* ATCC #20362 and expressing a $\Delta 9$ elongase/ $\Delta 8$ desaturase pathway. Strain Y9481 is capable of producing about 60.9% EPA relative to the total lipids with 35% total lipid content ["TFAs % DCW"], while strain Y9502 is capable of producing about 57% EPA relative to the total lipids with 37.1% TFAs % DCW.

[0292] The development of strains Y9481 and Y9502 (FIG. 2) required the construction of strains Y2224, Y4001 U, Y4036, Y4036U and L135 (described in Example 1), strains L135U9 and Y8002 (described in Example 2), strains Y8006U6, Y8069, Y8069U, Y8154, Y8154U, Y8269 and Y8269U (described in Example 3), strains Y8412U6 and Y8647 (described in Example 4), strains Y8467U and Y9028 (described in Example 5) and strain Y9028U.

Generation Of Strain Y9028U (*Ura3-*)

[0293] In order to disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO.130; described in Table 15 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y9028 (Example 5) in a manner similar to that described for pZKUM transformation of strain Y8006 (Example 3). A total of 8 transformants were grown and identified to possess a *Ura*-phenotype.

[0294] GC analyses showed that there were 24.1%, 24.9%, 24.5% and 24.5% EPA of TFAs in pZKUM-transformant strains #1, #3, #4, and #5, respectively. These four strains were designated as strains Y9028U1, Y9028U2, Y9028U3, and Y9028U4, respectively (collectively, Y9028U).

Components Of Integration Vector pZK16-ML8N

[0295] Construct pZK16-ML8N (FIG. 8A; SEQ ID NO:136) was generated to integrate one $\Delta 8$ desaturase gene, one malonyl-CoA synthetase gene, and one lysophosphatidic acid acyltransferase gene ["LPAAT"] into the *Yarrowia* YAL10B14795p locus (GenBank Accession No. XM_500900) of strain Y9028U2. The pZK16-ML8N plasmid contained the following components:

Table 15. Description of Plasmid pZK16-ML8N (SEQ ID NO:136)

RE Sites And Nucleotides Within SEQ ID NO:136	Description Of Fragment And Chimeric Gene Components
<i>AscI/Bst</i> WI (1905-1)	1904 bp 5' portion of YAL10B14795p locus (GenBank Accession No. XM_500900, labeled as "Y8716-5'" in Figure)
<i>PacI/SphI</i> (6414-4613)	1801 bp 3' portion of YAL10B14795p locus (GenBank Accession No. XM_500900, labeled as "Y8716-3'" in Figure)
<i>SwaI/Bst</i> WI (12920 - 1)	YAT1::EgD8M::Pex20, comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • EgD8M: Synthetic mutant $\Delta 8$ desaturase (SEQ ID NO:59; U.S. Patent 7,709,239), derived from <i>Euglena gracilis</i> ("EgD8S"; U.S. Patent 7,256,033); • Pex20: Pex20 terminator sequence from <i>Yarrowia Pex20</i> gene (GenBank Accession No. AF054613)
<i>PmeI/SwaI</i> (10534-12920)	FBA::MCS::Lip1, comprising: • FBA: <i>Yarrowia lipolytica</i> FBA promoter (U.S. Patent 7,202,356); • MCS: codon-optimized malonyl-CoA synthetase gene (SEQ ID NO:41), derived from <i>Rhizobium leguminosarum</i> bv. viciae 3841 (U.S. Patent Application No. 12/637877); • Lip1: Lip1 terminator sequence from <i>Yarrowia Lip1</i> gene (GenBank Accession No. Z50020)
<i>ClaI/PmeI</i> (8515-10534)	YAT1::MaLPAAT1S::Pex16, comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • MaLPAAT1S: codon-optimized lysophosphatidic acid acyltransferase gene (SEQ ID NO:35), derived from <i>Mortierella alpina</i> (U.S. Pat. Appl. Pub. No. 2006-0115881-A1; U.S. Pat. Appl. Pub. No. 2009-0325265-A1); • Pex16: Pex16 terminator sequence from <i>Yarrowia Pex16</i> gene (GenBank Accession No. U75433)
<i>SalI/EcoRI</i> (8065 - 6446)	<i>Yarrowia</i> Ura3 gene (GenBank Accession No. AJ306421)

Components Of Integration Vector pZK16-ML

[0296] Construct pZK16-ML (FIG. 8B; SEQ ID NO:137) was generated to integrate one malonyl-CoA synthetase gene and one lysophosphatidic acid acyltransferase gene ["LPAAT"] into the *Yarrowia* YAL10B14795p locus (GenBank Accession No. XM_500900) of strain Y9028U2. The components of the pZK16-ML plasmid are identical to those of pZK16-ML8N (*supra*); however, the chimeric YAT1::EgD8M::Pex20 gene of pZK16-ML8N is lacking.

Generation Of Strains Y9481 And Y9502

[0297] The pZK16-ML8N plasmid and pZK16-ML plasmid were each individually digested with *AscI/SphI*, and then used separately for transformation of strain Y9028U2, according to the General Methods. The transformant cells were plated onto MM plates and maintained at 30 °C for 3 to 4 days. Single colonies were re-streaked onto MM plates, and then inoculated into liquid MM at 30 °C and shaken at 250 rpm/min for 2 days. The cells were collected by centrifugation, resuspended in HGM and then shaken at 250 rpm/min for 5 days. The cells were subjected to fatty acid analysis, according to the General Methods.

[0298] GC analyses showed that most of the selected 96 strains of Y9028U2 with pZK16-ML8N produced 50-55.4% EPA of TFAs. Fifteen strains (i.e., #8, #18, #21, #24, #29, #48, #60, #66, #68, #75, #76, #78, #90, #95 and #96) that produced about 58.1%, 61.4%, 56.2%, 58.1%, 57.5%, 57.0%, 55.9%, 57.6%, 57.8%, 55.5%, 57.6%, 58.1%, 57.1%, 56.2% and 58.6% EPA of TFAs were designated as Y9472, Y9473, Y9474, Y9475, Y9476, Y9477, Y9478, Y9479, Y9480, Y9481, Y9482, Y9483, Y9484, Y9485 and Y9486, respectively.

[0299] The final genotype of these pZK16-ML8N transformant strains with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura+*, *Pex3-*, *unknown 1-*, *unknown 2-*, *unknown 3-*, *unknown 4-*, *unknown 5-*, *unknown 6-*, *unknown 7-*, *unknown 8-*, *unknown 9-*, *unknown 10-*, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9eS::Lip2, YAT1::EgD9eS::Lip2, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, YAT1::EgD8M::Pex20, GPD::EaD8S::Pex16 (2 copies), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EaD9eS/EaD8S::Lip2, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, GPDIN::FmD12::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, GPDIN::EgD5SM::Aco, GPM::EgD5SM::Oct, EXP1::EgD5SM::Lip1, YAT1::EaD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::YCPT::Aco, YAT1::MCS::Lip1, FBA::MCS::Lip1, YAT1::MaLPAAT1S::Pex16.

[0300] Similarly, GC analyses showed that most of the selected 96 strains of Y9028U2 with pZK16-ML produced 51-55.5% EPA of TFAs. Sixteen strains (i.e., #4, #8, #15, #16, #39, #44, #46, #63, #66, #80, #85, #86, #88, #89, #90 and #96) that produced about 56.5%, 57.4%, 56.8%, 57.0%, 56.4%, 57.3%, 58.2%, 55.6%, 57.8%, 55.6%, 57.6%, 56.8%, 55.8%, 56.4%, 56.1% and 57% EPA of TFAs were designated as Y9496, Y9497, Y9498, Y9499, Y9500, Y9501, Y9502, Y9503, Y9504, Y9505, Y9506, Y9507, Y9508, Y9509, Y9510 and Y9511, respectively.

[0301] The final genotype of these pZK16-ML transformant strains with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura+*, *Pex3-*, *unknown 1-*, *unknown 2-*, *unknown 3-*, *unknown 4-*, *unknown 5-*, *unknown 6-*, *unknown 7-*, *unknown 8-*, *unknown 9-*, *unknown 10-*, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9eS::Lip2, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, GPD::EaD8S::Pex16 (2 copies), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EaD9eS/EaD8S::Lip2, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12S::Aco, PDIN::FmD12::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, GPDIN::EgD5SM::Aco, GPM::EgD5SM::Oct, EXP1::EgD5SM::Lip1, YAT1::EaD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::YCPT::Aco, YAT1::MCS::Lip1, FBA::MCS::Lip1, YAT1::MaLPAAT1S::Pex16.

[0302] Knockout of the YAL10B14795p locus (GenBank Accession No. XM_500900) in strains Y9472, Y9473, Y9474, Y9475, Y9476, Y9477, Y9478, Y9479, Y9480, Y9481, Y94782, Y9483, Y9484, Y9485, Y9486, Y9496, Y9497, Y9498, Y9499, Y9500, Y9501, Y9502, Y9503, Y9504, Y9505, Y9506, Y9507, Y9508, Y9509, Y9510 and Y9511 was not confirmed in any of these EPA strains, produced by transformation with pZK16-ML8N or pZK16-ML.

Analysis Of Total Lipid Content And Composition By Flask Assay

[0303] Cells from YPD plates of strains Y9477, Y9481, Y9486, Y9497, Y9502, Y9504, Y9508 and Y9510 were grown and analyzed for total lipid content and composition, according to the General Methods.

[0304] Table 16 below summarizes the total dry cell weight of the cells ["DCW"], the total lipid content of cells ["TFAs % DCW"], the concentration of each fatty acid as a weight percent of TFAs ["% TFAs"] and the EPA content as a percent of the dry cell weight ["EPA % DCW"]. More specifically, fatty acids are identified as 16:0 (palmitate), 16:1 (palmitoleic acid), 18:0 (stearic acid), 18:1 (oleic acid), 18:2 (LA), ALA, EDA, DGLA, ARA, ETrA, ETA, EPA and other.

Table 16. Total Lipid Content And Composition In *Yarrowia* Strains Y9477, Y9481, Y9486, Y9497, Y9502, Y9504, Y9508 and Y9510 By Flask Assay

Strain	DCW (g/L)	TFAs % DCW	% TFAs													EPA % DCW
			16:0	16:1	18:0	18:1	18:2	ALA	EDA	DGLA	ARA	EtrA	ETA	EPA	other	
Y9477	3.2	32.6	2.6	0.5	3.4	4.8	10.0	0.5	2.5	3.7	1.0	0.5	2.1	61.4	6.9	20.0
Y9481	3.1	35.0	2.5	0.5	3.1	4.7	11.0	0.6	2.6	3.6	0.9	0.5	2.1	60.9	6.8	21.3
Y9486	3.1	32.2	2.1	0.7	1.8	4.2	11.9	0.6	2.9	4.2	1.2	0.7	2.4	60.3	6.7	19.4

Strain	DCW (g/L)	TFAs % DCW	% TFAs													EPA % DCW
			16:0	16:1	18:0	18:1	18:2	ALA	EDA	DGLA	ARA	EtrA	ETA	EPA	other	
Y9497	3.7	33.7	2.4	0.5	3.2	4.6	11.3	0.8	3.1	3.6	0.9	0.7	2.3	58.7	7.1	19.8
Y9502	3.8	37.1	2.5	0.5	2.9	5.0	12.7	0.9	3.5	3.3	0.8	0.7	2.4	57.0	7.5	21.3
Y9504	3.7	33.7	2.2	0.5	3.0	4.5	11.3	0.7	2.9	3.5	0.9	0.7	2.3	59.9	7.1	20.1
Y9508	3.7	34.9	2.3	0.5	2.7	4.4	13.1	0.9	2.9	3.3	0.9	0.7	2.3	58.7	7.3	20.5
Y9510	3.6	35.1	2.5	0.5	2.7	4.4	11.7	0.7	2.9	3.7	0.9	0.7	2.3	58.9	7.8	20.7

EXAMPLE 7**Generation Of *Yarrowia lipolytica* Strain Y8672 To Produce About 61.8% EPA Of Total Fatty Acids ["TFAs"] With 26.5% Total Lipid Content**

[0305] The present Example describes the construction of strain Y8672, derived from *Yarrowia lipolytica* ATCC #20362, capable of producing about 61.8% EPA relative to the total lipids with 26.5% total lipid content ["TFAs % DCW"] via expression of a $\Delta 9$ elongase/ $\Delta 8$ desaturase pathway. The development of strain Y8672 (FIG. 9) required the construction of strains Y2224, Y4001, Y4001 U, Y4036, Y4036U and L135 (described in Example 1), strains L135U9 and Y8002 (described in Example 2), strains Y8006U6, Y8069, Y8069U (described in Example 3) and strains Y8145, Y8145U, Y8259, Y8259U, Y8367 and Y8367U.

Generation Of Strain Y8145 To Produce About 48.5% EPA Of TFAs

[0306] Construct pZKL2-5m89C (FIG. 10; SEQ ID NO:138) was generated to integrate one $\Delta 5$ desaturase gene, one $\Delta 9$ elongase gene, one $\Delta 8$ desaturase gene, and one *Y. lipolytica* diacylglycerol cholinephosphotransferase gene (CPT1) into the *Lip2* loci (GenBank Accession No. AJ012632) of strain Y8069U3 (Example 3) to thereby enable higher level production of EPA. The pZKL2-5m89C plasmid contained the following components:

Table 17: Description of Plasmid pZKL2-5m89C (SEQ ID NO:138)

RE Sites And Nucleotides Within SEQ ID NO:138	Description Of Fragment And Chimeric Gene Components
<i>AscI/Bst</i> WI (730-1)	722 bp 5' portion of <i>Yarrowia Lip2</i> gene (labeled as "Lip2.5N" in Figure; GenBank Accession No. AJ012632)
<i>PacI/SphI</i> (4141-3438)	697 bp 3' portion of <i>Yarrowia Lip2</i> gene (labeled as "Lip2.3N" in Figure; GenBank Accession No. AJ012632)
<i>SwaI/Bst</i> WI (13143-1)	GPD::YICPT1::Aco, comprising: • GPD: <i>Yarrowia lipolytica</i> GPD promoter (U.S. Patent 7,259,255); • YICPT1: <i>Yarrowia lipolytica</i> diacylglycerol cholinephosphotransferase gene (SEQ ID NO:37) (Intl. App. Pub. No. WO 2006/052870); • Aco: Aco terminator sequence from <i>Yarrowia Aco</i> gene (GenBank Accession No. AJ001300)
<i>PmeI/SwaI</i> (10506-13143)	FBAIN::EgD8M::Lip1 comprising: • FBAIN: <i>Yarrowia lipolytica</i> FBAIN promoter (U.S. Patent 7,202,356); • EgD8M: Synthetic mutant $\Delta 8$ desaturase (SEQ ID NO:59; U.S. Patent 7,709,239), derived from <i>Euglena gracilis</i> ("EgD8S"; U.S. Patent 7,256,033); • Lip1: Lip1 terminator sequence from <i>Yarrowia Lip1</i> gene (GenBank Accession No. Z50020)
<i>PmeI/ClaI</i> (10506-8650)	YAT1::EgD9eS::Lip2, comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • EgD9eS: codon-optimized $\Delta 9$ elongase gene (SEQ ID NO:45), derived from <i>Euglena gracilis</i> (U.S. Patent 7,645,604); • Lip2: Lip2 terminator sequence from <i>Yarrowia Lip2</i> gene (GenBank Accession No. AJ012632)
<i>ClaI/EcoRI</i> (8650-6581)	<i>Yarrowia Ura3</i> gene (GenBank Accession No. AJ306421)
<i>EcoRI/PacI</i> (6581-4141)	YAT1::EgD5SM::ACO, comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • EgD5SM: Synthetic mutant $\Delta 5$ desaturase (SEQ ID NO:71; U.S. Pat. Pub. No. 2010-0075386-A1), derived from <i>Euglena gracilis</i> (U.S. Patent 7,678,560); • Aco: Aco terminator sequence from <i>Yarrowia Aco</i> gene (GenBank Accession No. AJ001300)

[0307] The pZKL2-5m89C plasmid was digested with *AscI/SphI*, and then used for transformation of strain Y8069U3 according to the General Methods. The transformant cells were plated onto MM plates and maintained at 30 °C for 3 to 4 days. Single colonies were re-streaked onto MM plates, and then inoculated into liquid MM at 30 °C and shaken at 250 rpm/min for 2 days. The cells were collected by centrifugation, resuspended in HGM and then shaken at 250 rpm/min for 5 days. The cells were subjected to fatty acid analysis, according to the General Methods.

[0308] GC analyses showed that most of the selected 96 strains produced 38-44.5% EPA of TFAs. Four strains (i.e., #10, #50, #70 and #89) that produced about 45.1%, 45.6%, 45.0% and 45.6% EPA of TFAs were designated as Y8143, Y8144, Y8145 and Y8146, respectively. Knockout of the *Lip2* loci (GenBank Accession No. AJ012632) was not confirmed in these EPA strains.

[0309] The final genotype of strains Y8143, Y8144, Y8145 and Y8146 with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura+*, *Pex3-*, *unknown 1-*, *unknown 2-*, *unknown 3-*, *unknown 4-*, *unknown 5-*, *Leu+*, *Lys+*, YAT1::ME3S::Pex16, GPD::FmD12::Pex20, YAT1::FmD12::Oct, GPAT::EgD9e::Lip2, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, YAT1::EgD9eS::Lip2, FBAINm::EgD8M::Pex20, FBAIN::EgD8M::Lip1, EXP1::EgD8M::Pex16, FBAIN::EgD5SM::Pex20, YAT1::EgD5SM::Aco, EXP1::EgD5M::Pex16, YAT1::EaD5SM::Oct, YAT1::PaD17S::Lip1, EXP1::PaD17::Pex16, FBAINm::PaD17::Aco, GPD::YICPT1::Aco. **Analysis Of Total Lipid Content And Composition By Flask Assay**

[0310] Cells from YPD plates of strains Y8143, Y8144, Y8145 and Y8146 were grown and analyzed for total lipid content and composition, according to the General Methods.

Table 18: Total Lipid Content And Composition In *Yarrowia* Strains Y8143, Y8144, Y8145 and Y8146 By Flask Assay

Strain	DCW (g/L)	TFAs % DCW	% TFAs														EPA % DCW
			16:0	16:1	18:0	18:1	18:2	ALA	EDA	DGLA	ARA	EtrA	ETA	EPA	other		
Y8143	4.6	22.3	4.2	1.5	1.4	3.6	18.1	2.6	1.7	1.6	0.6	2.2	1.6	50.3	11.6	11.2	
Y8144	4.3	23	4.0	1.5	1.4	3.3	18.0	2.6	1.8	1.7	0.7	2.3	1.6	50.6	11.5	11.6	
Y8145	4.6	23.1	4.3	1.7	1.4	4.8	18.6	2.8	2.2	1.5	0.6	2.2	1.5	48.5	9.9	11.2	
Y8146	4.5	23.8	4.3	1.7	1.4	4.8	18.7	2.8	2.0	1.5	0.6	2.2	1.5	48.3	11.2	11.5	

Generation Of Strain Y8145U (*Ura3-*)

[0311] In order to disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:130; described in Table 15 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y8145 in a manner similar to that described for pZKUM transformation of strain Y8006 (Example 3). A total of 8 transformants were grown and identified to possess a *Ura-* phenotype.

[0312] GC analyses showed that there were 22.5%, 22.6% and 23.4% EPA of TFAs in pZKUM-transformant strains #5, #6 and #7, respectively. These three strains were designated as strains Y8145U1,

Y8145U2 and Y8145U3, respectively (collectively, Y8145U).

Generation Of Y8259 Strain To Produce About 53.9% EPA Of TFAs

[0313] Construct pZKL1-2SR9G85 (Example 3, FIG. 6A; SEQ ID NO:132) was generated to integrate one DGLA synthase gene, one $\Delta 12$ desaturase gene and one $\Delta 5$ desaturase gene into the *Lip1* loci (GenBank Accession No. Z50020) of strain Y8145U to thereby enable higher level production of EPA.

[0314] The pZKL1-2SR9G85 plasmid was digested with *AscI/SphI*, and then used for transformation of strain Y8145U1, in a manner similar to that described for pZKL1-2SR9G85 transformation of strain Y8154U1 (Example 3). The cells were subjected to fatty acid analysis, according to the General Methods.

[0315] GC analyses showed that most of the selected 96 strains produced 40-44.0% EPA of total lipids. Five strains (i.e., #7, #14, #48, #56 and #60) that produced about 45.2%, 47%, 44.4%, 44.3% and 45.2% EPA of TFAs were designated as Y8255, Y8256, Y8257, Y8258 and Y8259, respectively. Knockout of the *Lip1* loci (GenBank Accession No. Z50020) was not confirmed in these EPA strains.

[0316] The final genotype of these strains with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura+*, *Pex3-*, *unknown 1-*, *unknown 2-*, *unknown 3-*, *unknown 4-*, *unknown 5-*, *unknown 6-*, *Leu+*, *Lys+*, YAT1::ME3S::Pex16, GPD::Fmd12::Pex20, YAT1::Fmd12::Oct, EXP1::Fmd12S::ACO, GPAT::EgD9e::Lip2, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, YAT1::EgD9eS::Lip2, FBAINm::EgD8M::Pex20, FBAIN::EgD8M::Lip1, EXP1::EgD8M::Pex16, YAT1::E389S/EgD8M::Lip1, FBAIN::EgD5SM::Pex20, YAT1::EgD5SM::Aco, GPM::EgD5SM::Oct, EXP1::EgD5M::Pex16, YAT1::EaD5SM::Oct, YAT1::PaD17S::Lip1, EXP1::PaD17::Pex16, FBAINm::PaD17::Aco, GPD::YICPT1::Aco.

[0317] *Yarrowia lipolytica* strain Y8259 was deposited with the American Type Culture Collection on May 14, 2009 and bears the designation ATCC PTA-10027.

Analysis Of Total Lipid Content And Composition By Flask Assay

[0318] Cells from YPD plates of strains Y8256 and Y8259 were grown and analyzed for total lipid content and composition, according to the General Methods.

Table 19: Total Lipid Content And Composition In *Yarrowia* Strains Y8256 and Y8259 By Flask Assay

Strain	DCW (g/L)	TFAs % DCW	% TFAs													EPA % DCW
			16:0	16:1	18:0	18:1	18:2	ALA	EDA	DGLA	ARA	EtrA	ETA	EPA	other	
Y8256	4.0	20.1	3.5	1.4	1.3	3.8	18.8	2.0	2.1	1.6	0.8	2.1	1.7	49.9	11.0	10.0
Y8259	4.7	20.5	3.5	1.3	1.3	4.8	16.9	2.3	1.9	1.7	0.6	1.8	1.6	53.9	8.4	11.0

Generation Of Strain Y8259U (*Ura3-*)

[0319] In order to disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:130; described in Table 15 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y8259 in a manner similar to that described for pZKUM transformation of strain Y8006 (Example 3). A total of 8 transformants were grown and identified to possess a *Ura-* phenotype.

[0320] GC analyses showed that there was 26.6% EPA of TFAs in pZKUM-transformant strain #3. This strain was designated as strain Y8259U.

Generation Of Y8367 Strain To Produce about 58.3% EPA Of TFAs

[0321] Construct pZP2-85m98F (Example 5, FIG. 7B; SEQ ID NO:135) was generated to integrate one $\Delta 8$ desaturase gene, one DGLA synthase, and one $\Delta 5$ desaturase gene into the *Yarrowia* Pox2 locus (GenBank Accession No. AJ001300) of strain Y8259U to thereby enable higher level production of EPA.

[0322] The pZP2-85m98F plasmid was digested with *AscI/SphI*, and then used for transformation of strain Y8259U, in a manner similar to that described for pZP2-85m98F transformation of strain Y8647U3 (Example 5). The cells were subjected to fatty acid analysis, according to the General Methods.

[0323] GC analyses showed that most of the selected 96 strains of Y8259U with pZP2-85m98F produced 41-46% EPA of TFAs. Four strains (i.e., #26, #33, #77 and #81) that produced about 46.7%, 46.5%, 47.4% and 46.9% EPA of TFAs were designated as Y8367, Y8368, Y8369 and Y8370, respectively. Knock out of the Pox2 locus (GenBank Accession No. AJ001300) was not confirmed in these EPA strains.

[0324] The final genotype of strains Y8367, Y8368, Y8369 and Y8370 with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura+*, *Pex3-*, *unknown 1-*, *unknown 2-*, *unknown 3-*, *unknown 4-*, *unknown 5-*, *unknown 6-*, *unknown 7-*, *Leu+*, *Lys+*, YAT1::ME3S::Pex16, GPD::Fmd12::Pex20, YAT1::Fmd12::Oct, EXP1::Fmd12S::ACO, GPAT::EgD9e::Lip2, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, YAT1::EgD9eS::Lip2, FBAINm::EgD8M::Pex20, FBAIN::EgD8M::Lip1, EXP1::EgD8M::Pex16, GPD::EaD8S::Pex16, YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAIN::EgD5SM::Pex20, YAT1::EgD5SM::Aco, GPM::EgD5SM::Oct, EXP1::EgD5M::Pex16, EXP1::EgD5SM::Lip1, YAT1::EaD5SM::Oct, YAT1::PaD17S::Lip1, EXP1::PaD17::Pex16, FBAINm::PaD17::Aco, GPD::YICPT1::Aco.

Analysis Of Total Lipid Content And Composition By Flask Assay

[0325] Cells from YPD plates of strains Y8367, Y8368, Y8369 and Y8370 were grown and analyzed for total lipid content and composition, according to the General Methods.

Table 20: Total Lipid Content And Composition In *Yarrowia* Strains Y8367, Y8368, Y8369 and Y8370 By Flask Assay

Strain	DCW (g/L)	TFAs % DCW	% TFAs													EPA % DCW
			16:0	16:1	18:0	18:1	18:2	ALA	EDA	DGLA	ARA	EtrA	ETA	EPA	other	
Y8367	3.6	18.4	3.7	1.2	1.1	3.4	14.2	1.1	1.5	1.7	0.8	2.1	1.0	58.3	9.9	10.7
Y8368	4.7	19.2	3.0	1.4	1.3	4.3	17.9	1.3	2.4	2.8	1.0	1.8	1.9	52.5	8.4	10.1
Y8369	3.5	19.7	3.7	1.2	1.6	4.2	15.6	1.8	1.7	1.9	0.6	1.7	1.6	55.8	8.6	11.0
Y8370	4.0	23.3	3.4	1.1	1.4	4.0	15.7	1.9	1.7	1.9	0.6	1.8	1.5	56.4	8.6	13.1

Generation Of Strain Y8367U (*Ura3-*)

[0326] In order to disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:130; described in Table 15 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y8367 in a manner similar to that described for pZKUM transformation of strain Y8006 (Example 3). A total of 8 transformants were grown and identified to possess a *Ura-* phenotype.

[0327] GC analyses showed that there were 25.6%, 25.5% and 25.4% EPA of TFAs in pZKUM-transformant strains #2, #3 and #6, respectively. These three strains were designated as strains Y8367U1, Y8367U2 and Y8367U3, respectively (collectively, Y8367U).

Generation Of Y8672 strain To Produce about 61.8% EPA Of TFAs

[0328] Construct pZSCP-Ma83 (Example 3, FIG. 6B; SEQ ID NO:133) was generated to integrate one $\Delta 9$ desaturase gene, one $C_{16/18}$ elongase gene and one malonyl-CoA synthetase gene into the *SCP2* loci (GenBank Accession No. XM_503410) of strain Y8637U to thereby enable higher level production of EPA.

[0329] The pZSCP-Ma83 plasmid was digested with *AscI/SphI*, and then used for transformation of strain Y8367U1, in a manner similar to that described for pZSCP-Ma83 transformation of strain Y8269U1 (Example 3). The cells were subjected to fatty acid analysis, according to the General Methods.

[0330] GC analyses showed that most of the selected 96 strains of Y8367U1 with pZSCP-Ma83 produced 46-52.5% EPA of TFAs. Eight strains (i.e., #8, #40, #43, #44, #61, #63, #68 and #70) that produced about 53.2%, 52.8%, 52.7%, 52.9%, 53.0%, 52.6%, 53.1% and 52.7% EPA of TFAs were designated as Y8666, Y8667, Y8668, Y8669, Y8670, Y8671, Y8672 and Y8673, respectively. Knockout of the *SCP2* loci (GenBank Accession No. XM_503410) was not confirmed in these EPA strains.

[0331] The final genotype of strains Y8666, Y8667, Y8668, Y8669, Y8670, Y8671, Y8672 and Y8673 with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura⁺, Pex3⁻, unknown 1⁻, unknown 2⁻, unknown 3⁻, unknown 4⁻, unknown 5⁻, unknown 6⁻, unknown 7⁻, unknown 8⁻, Leu⁺, Lys⁺, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, GPD::Fmd12::Pex20, YAT1::Fmd12::Oct, EXP1::Fmd12S::ACO, GPAT::EgD9e::Lip2, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip2, YAT1::EgD9eS::Lip2, FBAINm::EgD8M::Pex20, FBAINm::EgD8M::Lip1, EXP1::EgD8M::Pex16, GPD::EaD8S::Pex16 (2 copies), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EgD5SM::Pex20, YAT1::EgD5SM::Aco, GPM::EgD5SM::Oct, EXP1::EgD5M::Pex16, EXP1::EgD5SM::Lip1, YAT1::EaD5SM::Oct, YAT1::PaD17S::Lip1, EXP1::PaD17::Pex16, FBAINm::PaD17::Aco, GPD::YICPT1::Aco, YAT1::MCS::Lip1.*

Analysis Of Total Lipid Content And Composition By Flask Assay

[0332] Cells from YPD plates of strains Y8666, Y8669, Y8679 and Y8672 were grown and analyzed for total lipid content and composition, according to the General Methods.

Table 21: Total Lipid Content And Composition In *Yarrowia* Strains Y8666, Y8669, Y8670 And Y8672 By Flask Assay

Strain	DCW (g/L)	TFAs % DCW	% TFAs													EPA % DCW
			16:0	16:1	18:0	18:1	18:2	ALA	EDA	DGLA	ARA	EtrA	ETA	EPA	other	
Y8666	3.2	25.2	2.3	0.3	2.3	4.1	15.1	1.3	1.7	1.4	0.7	0.6	1.3	62.2	6.7	15.6
Y8669	3.2	26.4	2.3	0.3	2.3	4.1	15.7	1.4	1.8	1.6	0.7	0.5	1.1	61.5	6.7	16.3
Y8670	3.2	27.3	1.9	0.4	3.4	4.3	17.0	1.5	2.2	1.7	0.6	0.5	1.1	60.9	4.5	16.6
Y8672	3.3	26.5	2.3	0.4	2.0	4.0	16.1	1.4	1.8	1.6	0.7	0.4	1.1	61.8	6.4	16.4

EXAMPLE 8

Construction Of Various Expression Vectors Comprising Different LPLAT ORFs

[0333] The present example describes the construction of a series of vectors, each comprising a LPLAT ORF, suitable for expression in *Yarrowia lipolytica*. LPLAT ORFs included the *Saccharomyces cerevisiae* Ale1, *Yarrowia lipolytica* Ale1, *Mortierella alpina* LPAAT1, *Yarrowia lipolytica* LPAAT1 and *Caenorhabditis elegans* LPCAT. Example 9 describes the results obtained following transformation of these vectors into *Yarrowia lipolytica* strain Y8406U.

Origin Of LPLATs

[0334] A variety of LPLATs have been identified in the patent and open literature, but the functionality of these genes has not been previously directly compared. Table 22 summarizes publicly available LPLATs (i.e., ScAle1, ScLPAAT, MaLPAAT1 and CeLPCAT) and LPLAT orthologs identified herein (i.e., YIAle1 and YILPAAT1) that are utilized in the present Example, following codon-optimization of heterologous genes for expression in *Yarrowia lipolytica* (*infra*).

Table 22: LPLATs Functionally Characterized

LPLAT	Organism	ORF Designation	References	SEQ ID NO
Ale1	<i>Saccharomyces cerevisiae</i> *	ORF "YOR175C" or "ScAle1"	GenBank Accession No. NP_014818; U.S. Pat. Appl. Pub. No. 20080145867 (and corresponding to Intl. App. Pub. No. WO 2008/076377); Intl. App. Pub. No. WO 2009/001315	14, 15
	<i>Yarrowia lipolytica</i>	"YALIOF19514p" or "YIAle1"	GenBank Accession No. XP_505624; Intl. App. Pub. No. WO 2009/001315	16, 17
LPAAT	<i>Saccharomyces cerevisiae</i>	ORF "YDL052C" or "ScLPAAT"	GenBank Accession No. NP_010231	32
	<i>Mortierella alpina</i>	"MaLPAAT1"	U.S. Pat. Appl. Pub. No. 2006-0115881-A1; U.S. Pat. Appl. Pub. No. 2009-0325265-A1	28, 29
	<i>Yarrowia lipolytica</i>	"YALIOE18964g" or "YILPAAT1"	GenBank Accession No. XP_504127; U.S. Patent 7,189,559	30, 31
LPCAT	<i>Caenorhabditis elegans</i> *	"clone T06E8.1" or "CeLPCAT"	GenBank Accession No. CAA98276; Intl. App. Pub. No. WO 2004/076617 (corresponding to U.S. Pat. Appl. Pub. No. 2006-0168687-A1)	24, 25

*The *Saccharomyces cerevisiae* Ale1 and *Caenorhabditis elegans* LPCAT were used as comparative Examples.

[0335] More specifically, the ScLPAAT (SEQ ID NO:32) and ScAle1 (SEQ ID NO:15) protein sequences were used as queries to identify orthologs from the public *Y. lipolytica* protein database of the "Yeast project Genolevures" (Center for Bioinformatics, LaBRI, Talence Cedex, France) (see also Dujon, B. et al., Nature, 430(6995):35-44 (2004)). Based on analysis of the best hits, the Ale1 and LPAAT orthologs from *Yarrowia lipolytica* are identified herein as YIAle1 (SEQ ID NO:17) and YILPAAT1 (SEQ ID NO:31), respectively. The identity of YIAle1 and YILPAAT1 as orthologs of ScAle1 and ScLPAAT, respectively, was further confirmed by doing a reciprocal BLAST, i.e., using SEQ ID NOs:17 and 31 as a query against the *Saccharomyces cerevisiae* public protein database to find ScAle1 and ScLPAAT, respectively, as the best hits.

[0336] The LPLAT proteins identified above as ScAle1 (SEQ ID NO:15), YIAle1 (SEQ ID NO:17), ScLPAAT (SEQ ID NO:32), MaLPAAT1 (SEQ ID NO:29), YILPAAT1 (SEQ ID NO:31) and CeLPCAT (SEQ ID NO:25) were aligned using the method of Clustal W (slow, accurate, Gonnert option; Thompson et al., Nucleic Acids Res., 22:4673-4680 (1994)) of the MegAlign™ program (version 8.0.2) of the LASERGENE bioinformatics computing suite (DNASTAR, Inc., Madison, WI). This resulted in creation of Table 23, where percent similarity is shown in the upper triangle of the Table while percent divergence is shown in the lower triangle.

Table 23: Percent Identity And Percent Divergence Among Various LPLATs

YILPAAT1	CeLPCAT	MaLPAAT1	ScAle1	ScLPAAT	YIAle1	
--	26.6	34.0	9.6	43.9	11.7	YILPAAT1
184.3	--	36.4	11.3	32.4	14.5	CeLPCAT
137.5	126.4	--	11.1	34.6	15.0	MaLPAAT1
545.0	442.0	456.0	--	13.5	45.0	ScAle1
97.9	145.7	134.5	365.0	--	15.6	ScLPAAT
426.0	339.0	330.0	94.3	317.0	--	YIAle1

[0337] The percent identities revealed by this method allowed determination of the minimum percent identity between each of the LPAAT polypeptides and the minimum percent identity between each of the Ale1 polypeptides. The range of identity between LPAAT polypeptides was 34.0% identity (MaLPAAT1 and YILPAAT1) to 43.9% identity (ScLPAAT and YILPAAT1), while identity between the ScAle1 and YIAle1 polypeptides was 45%.

[0338] Membrane Bound O-Acyltransferase ("MBOAT") Family Motifs: Orthologs of the ScAle1 protein sequence (SEQ ID NO:15) were identified by conducting a National Center for Biotechnology Information

["NCBI"] BLASTP 2.2.20 (protein-protein Basic Local Alignment Search Tool; Altschul et al., Nucleic Acids Res., 25:3389-3402 (1997); and Altschul et al., FEBS J., 272:5101-5109 (2005)) search using ScAle1 (SEQ ID NO:15) as the query sequence against fungal proteins in the "nr" protein database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure from Brookhaven Protein Data Bank ["PDB"], sequences included in the last major release of the SWISS-PROT protein sequence database, PIR and PRF excluding those environmental samples from WGS projects) using default parameters (expect threshold = 10; word size = 3; scoring parameters matrix = BLOSUM62; gap costs: existence = 11, extension = 1). The following hits were obtained:

Table 24: Fungal Orthologs Of ScAle1 (SEQ ID NO:15) Based On BLAST Analysis

Gen Bank Accession No.	Species
NP_014818.1	<i>Saccharomyces cerevisiae</i>
XP_001643411.1	<i>Vanderwaltozyma polyspora</i> DSM 70294
XP_448977.1	<i>Candida glabrata</i>
XP_455985.1	<i>Kluyveromyces lactis</i>
NP_986937.1	<i>Ashbya gossypii</i> ATCC 10895
XP_001385654.2	<i>Pichia stipitis</i> CBS 6054
XP_001487052.1	<i>Pichia guilliermondii</i> ATCC 6260
EDK36331.2	<i>Pichia guilliermondii</i> ATCC 6260
XP_001525914.1	<i>Lodderomyces elongisporus</i> NRRL YB-4239
XP_461358.1	<i>Debaryomyces hansenii</i> CBS767
XP_713184.1	<i>Candida albicans</i> SC5314
XP_001645053.1	<i>Vanderwaltozyma polyspora</i> DSM 70294
XP_505624.1	<i>Yarrowia lipolytica</i>
XP_001805526.1	<i>Phaeosphaeria nodorum</i> SN15
XP_001598340.1	<i>Sclerotinia sclerotiorum</i> 1980
XP_001907785.1	<i>Podospora anserina</i>
XP_001931658.1	<i>Pyrenophora tritici-repentis</i> Pt-1C-BFP
XP_001560657.1	<i>Botryotinia fuckeliana</i> B05.10
XP_963006.1	<i>Neurospora crassa</i> OR74A
XP_364011.2	<i>Magnaporthe grisea</i> 70-15
XP_001209647.1	<i>Aspergillus terreus</i> NIH2624
XP_001822945.1	<i>Aspergillus oryzae</i> RIB40
XP_001257694.1	<i>Neosartorya fischeri</i> NRRL 181
XP_747591.2	<i>Aspergillus fumigatus</i> Af293
XP_001270060.1	<i>Aspergillus clavatus</i> NRRL 1
NP_596779.1	<i>Schizosaccharomyces pombe</i>
XP_001396584.1	<i>Aspergillus niger</i>
XP_001229385.1	<i>Chaetomium globosum</i> CBS 148.51
XP_001248887.1	<i>Coccidioides immitis</i> RS
XP_664134.1	<i>Aspergillus nidulans</i> FGSC A4
XP_566668.1	<i>Cryptococcus neoformans</i> var. <i>neoformans</i> JEC21
XP_001839338.1	<i>Coprinopsis cinerea</i> okayama T#130
XP_757554.1	<i>Ustilago maydis</i> 521

The yeast and fungal protein sequences of Table 24 were aligned using DNASTAR. Multiple sequence alignments and percent identity calculations were performed using the Clustal W method of alignment (*supra*).

[0339] More specifically, default parameters for multiple protein alignment using the Clustal W method of alignment correspond to: GAP PENALTY=10, GAP LENGTH PENALTY=0.2, Delay Divergent Seqs(%)=30, DNA Transition Weight=0.5, Protein Weight Matrix=Gonnet Series, DNA Weight Matrix=IUB with the 'slow-accurate' option. The resulting alignment was analyzed to determine the presence or absence of the non-plant motifs for Ale1 homologs, as identified in U.S. Pat. Appl. Pub. No. 2008-0145867-A1. Specifically, these include: M-[V/I]-[L/I]-xxK-[L/V/I]-xxxxxDG (SEQ ID NO:102), RxxYYxxWxxx-[E/D]-[A/G]xxxxGxG-[F/Y]-xG (SEQ ID NO:103), EX₁1WNx₂-[T/V]-X₂W (SEQ ID NO:21) and SAxWHGxxPGYxx-[T/F]-F (SEQ ID NO:104), wherein X encodes any amino acid residue. The His residue in SEQ ID NO:104 has been reported to be a likely active site residue within the protein.

[0340] Only one motif, i.e., EX₁1WNx₂-[T/V]-X₂W (SEQ ID NO:21), was completely conserved in all 33 of the organisms aligned. The remaining M-[V/I]-[L/I]-xxK-[L/V/I]-xxxxxDG (SEQ ID NO:102), RxxYYxxWxxx-[E/D]-[A/G]xxxxGxG-[F/Y]-xG (SEQ ID NO:103) and SAxWHGxxPGYxx-[T/F]-F (SEQ ID NO:104) motifs were only partially conserved. Thus, these motifs were appropriately truncated to fit with 0 mismatch (i.e., SAxWHG [SEQ ID NO:20]), 1 mismatch (i.e., RxxYYxxW [SEQ ID NO:18]), or 2 mismatches (i.e., M[V/I](L/I)xxK(LV) [SEQ ID NO:18]) for the purposes of the present methodologies.

[0341] **1-Acyl-sn-Glycerol-3-Phosphate Acyltransferase ("LPAAT") Family Motifs:** Analysis of the protein alignment comprising ScLPAAT (SEQ ID NO:18), MaLPAAT1 (SEQ ID NO:15) and YILPAAT1 (SEQ ID NO:17) revealed that the 1-acyl-sn-glycerol-3-phosphate acyltransferase family motif EGTR (SEQ ID NO:20) was present in each of the LPAAT orthologs. On this basis, MaLPAAT1 was identified as a likely LPAAT, that was clearly distinguishable from the *Mortierella alpina* LPAAT-like proteins disclosed in Intl. App. Pub. No. WO 2004/087902 (i.e., SEQ ID NOs:93 and 95).

[0342] It is noteworthy that the EGTR (SEQ ID NO:20) motif, while lacking in the LPCAT sequences in Intl. App. Pub. No. WO 2004/087902, is present in CeLPCAT (SEQ ID NO:2). It appears that other residues distinguish LPAAT and LPCAT sequences in LPAAT-like proteins. One such residue could be the extension of the EGTR (SEQ ID NO:20) motif. Specifically, whereas the EGTR motif in ScLPAAT (SEQ ID NO:18), MaLPAAT1 (SEQ ID NO:15) and YILPAAT1 (SEQ ID NO:17) is immediately followed by a serine residue, the EGTR motif in CeLPCAT is immediately followed by an asparagine residue. In contrast, the two LPCATs in Intl. App. Pub. No. WO 2004/087902 have a valine substituted for the arginine residue in the EGTR motif and the motif is immediately followed by a valine residue.

Construction Of pY201, Comprising A Codon-Optimized *Saccharomyces cerevisiae* Ale1 Gene

[0343] The *Saccharomyces cerevisiae* ORF designated as "ScAle1" (SEQ ID NO:14) was optimized for expression in *Yarrowia lipolytica*, by DNA 2.0 (Menlo Park, CA). In addition to codon optimization, 5' *Pci*I and 3' *Not*I cloning sites were introduced within the synthetic gene (i.e., ScAle1S; SEQ ID NO:22). None of the modifications in the ScAle1S gene changed the amino acid sequence of the encoded protein (i.e., the protein sequence encoded by the codon-optimized gene [i.e., SEQ ID NO:23] is identical to that of the wildtype protein sequence [i.e., SEQ ID NO:15]). ScAle1S was cloned into pJ201 (DNA 2.0) to result in pJ201:ScAle1S.

[0344] A 1863 bp *Pci*I/*Not*I fragment comprising ScAle1S was excised from pJ201:ScAle1S and used to create pY201 (SEQ ID NO:139; Table 25; FIG. 10A). In addition to comprising a chimeric YAT1::ScAle1S::Lip1 gene, pY201 also contains a *Y. lipolytica* URA3 selection marker flanked by *LoxP* sites for subsequent removal, if needed, by Cre recombinase-mediated recombination. Both the YAT1::ScAle1S::Lip1 chimeric gene and the URA3 gene were flanked by fragments having homology to 5' and 3' regions of the *Y. lipolytica* Pox3 gene to facilitate integration by double homologous recombination, although integration into *Y. lipolytica* is known to usually occur without homologous recombination. Thus, construct pY201 thereby contained the following components:

Table 25: Description of Plasmid pY201 (SEQ ID NO:139)

RE Sites And Nucleotides Within SEQ ID NO:139	Description Of Fragment And Chimeric Gene Components
<i>Bst</i> WI/ <i>Sbf</i> I (1-1706 bp)	LoxP::Ura3::LoxP, comprising:
	• LoxP sequence (SEQ ID NO:140)
	• <i>Yarrowia lipolytica</i> Ura3 gene (GenBank Accession No. AJ306421);

RE Sites And Nucleotides Within SEQ ID NO:139	Description Of Fragment And Chimeric Gene Components
	• LoxP sequence (SEQ ID NO:140)
<i>SbfI/SphI</i> (1706-3043 bp)	3' portion of <i>Yarrowia lipolytica</i> POX3 Acyl-CoA oxidase 3 (GenBank Accession No. YALI024750g) (i.e., bp 2215-3038 in pY201)
<i>SphI/AscI</i> (3043-5743 bp)	• <i>ColE1</i> plasmid origin of replication; • Ampicillin-resistance gene (Amp^R) for selection in <i>E. coli</i> (i.e., bp 3598-4758 [complementary] in pY201); • <i>E. coli</i> f1 origin of replication
<i>AscI/BstWI</i> (5743-6513 bp)	5' portion of <i>Yarrowia lipolytica</i> POX3 Acyl-CoA oxidase 3 (GenBank Accession No. YALI024750g) (i.e., bp 5743-6512 in pY201)
<i>BstWI/BstWI</i> (6514-1 bp)	• YAT1::ScA1e1S::Lip1, comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (U.S. Pat. Appl. Pub. No. 2006/0094102-A1) (i.e., bp 6514-7291 in pY201)
[a <i>NotI</i> site, located between ScA1e1S and Lip1 is present at bp 9154 bp]	• ScA1e1S: codon-optimized Ale1 (SEQ ID NO:22) derived from <i>Saccharomyces cerevisiae</i> YOR175C (i.e., bp 7292-9151 in pY201; labeled as "ScLPCATs ORF" in Figure); • Lip1: Lip1 terminator sequence from <i>Yarrowia Lip1</i> gene (GenBank Accession No. Z50020) (i.e., bp 9160-9481 pY201; labeled as "Lip1-3'" in Figure)

Construction Of pY168, Comprising A *Yarrowia lipolytica* Ale1 Gene

[0345] The *Yarrowia lipolytica* ORF designated as "YIAle1" (GenBank Accession No. XP_505624; SEQ ID NO:16) was amplified by PCR from *Yarrowia lipolytica* ATCC #20362 cDNA library using PCR primers 798 and 799 (SEQ ID NOs:141 and 142, respectively). Additionally, the YAT promoter was amplified by PCR primers 800 and 801 (SEQ ID NOs:143 and 144, respectively) from pY201 (SEQ ID NO:139). Since the primer pairs were designed to create two PCR products having some overlap with one another, a YAT1::YIAle1 fusion fragment was then amplified by overlapping PCR using primers 798 and 801 (SEQ ID NOs:141 and 144, respectively) and the two PCR fragments as template. The PCR was carried out in a RoboCycler Gradient 40 PCR machine (Stratagene) using the manufacturer's recommendations and Pfu Ultra™ High-Fidelity DNA Polymerase (Stratagene, Cat. No. 600380). Amplification was carried out as follows: initial denaturation at 95 °C for 4 min, followed by 30 cycles of denaturation at 95 °C for 30 sec, annealing at 55 °C for 1 min, and elongation at 72 °C for 1 min. A final elongation cycle at 72 °C for 10 min was carried out, followed by reaction termination at 4 °C.

[0346] The PCR product comprising the YAT1::YIAle1 fusion fragment was gel purified and digested with *ClaI/NotI*. This *ClaI-NotI* fragment was ligated into pY201 that had been similarly digested (thereby removing the YAT1::ScA1e1S fragment) to create pY168 (SEQ ID NO:145), comprising a chimeric YAT1::YIAle1::Lip1 gene. The DNA sequence of the *Yarrowia* Ale1 ORF was confirmed by DNA sequencing. The components present in pY168 (FIG. 10B; SEQ ID NO:145) are identical to those present in pY201, with the exception of the YAT1::YIAle1::Lip1 gene in pY168, instead of the YAT1::ScA1e1 S::Lip1 gene in pY201 (FIG. 10A). Note that YIAle1 is labeled as "YI LPCAT" in FIG. 10B.

Construction Of pY208, Comprising A *Mortierella alpina* LPAAT1 Gene

[0347] The *Mortierella alpina* ORF designated as "MaLPAAT1" (SEQ ID NO:28) was optimized for expression in *Yarrowia lipolytica*, by DNA 2.0 (Menlo Park, CA). In addition to codon optimization, 5' *PciI* and 3' *NotI* cloning sites were introduced within the synthetic gene (i.e., MaLPAAT1S; SEQ ID NO:35). None of the modifications in the MaLPAAT1S gene changed the amino acid sequence of the encoded protein (i.e., the protein sequence encoded by the codon-optimized gene [i.e., SEQ ID NO:36] is identical to that of the wildtype protein sequence [i.e., SEQ ID NO:29]). MaLPAAT1S was cloned into pJ201 (DNA 2.0) to result in pJ201:MaLPAAT1 S.

[0348] A 945 bp *PciI/NotI* fragment comprising MaLPAAT1S was excised from pJ201 :MaLPAAT1S and used to create pY208 (SEQ ID NO:146), in a 3-way ligation with two fragments of pY201 (SEQ ID NO:139). Specifically, the MaLPAAT1 fragment was ligated with a 3530 bp *Sph-NotI* pY201 fragment and a 4248 bp *NcoI-SphI* pY201 fragment to result in pY208. The components present in pY208 (FIG. 11A; SEQ ID NO:146) are identical to those present in pY201, with the exception of the YAT1::MaLPAAT1S::Lip1 gene in pY208, instead of the YAT1::ScA1e1S::Lip1 gene in pY201 (FIG. 10A).

Construction Of pY207, Comprising A *Yarrowia lipolytica* LPAAT1 Gene

[0349] A putative LPAAT1 from *Yarrowia lipolytica* (designated herein as "YILPAAT1"; SEQ ID NOs:30 and 31) was described in U.S. Patent 7,189,559 and GenBank Accession No. XP_504127. The protein is annotated as "similar to uniprot|P33333 *Saccharomyces cerevisiae* YDL052c SLC1 fatty acyltransferase".

[0350] The YILPAAT1 ORF (SEQ ID NO:30) was amplified by PCR using a *Yarrowia lipolytica* ATCC #20362 cDNA library as a template and PCR primers 856 and 857 (SEQ ID NOs:147 and 148, respectively). The PCR was conducted using the same components and conditions as described above for amplification of the YAT1::YIAle1 fusion fragment, prior to synthesis of pY168.

[0351] The PCR product comprising the YILPAAT1 ORF was digested with *PciI* and *NotI* and then utilized in a 3-way ligation with two fragments from pY168. Specifically, the YILPAAT1 fragment was ligated with a 3530 bp *Sph-NotI* pY168 fragment and a 4248 bp *NcoI-SphI* pY168 fragment, to produce pY207, comprising a chimeric YAT1::YILPAAT1::Lip1 gene. The *Y. lipolytica* LPAAT1 ORF was confirmed by DNA sequencing. The components present in pY207 (FIG. 11B; SEQ ID NO:149) are identical to those present in pY201, with the exception of the chimeric YAT1::YILPAAT1::Lip1 gene in pY207, instead of the YAT1::ScA1e1S::Lip1 gene in pY201 (FIG. 10A). Note that YILPAAT1 is labeled as "YI LPAT1 ORF" in FIG. 11 B.

Construction Of pY175, Comprising A *Caenorhabditis elegans* LPCAT Gene

[0352] The *Caenorhabditis elegans* ORF designated as "CeLPCAT" (SEQ ID NO:24) was optimized for expression in *Yarrowia lipolytica*, by GenScript Corporation (Piscataway, NJ). In addition to codon optimization, 5' *NcoI* and 3' *NotI* cloning sites were introduced within the synthetic gene (i.e., CeLPCATs; SEQ ID NO:26). None of the modifications in the CeLPCATs gene changed the amino acid sequence of the encoded protein (i.e., the protein sequence encoded by the codon-optimized gene [i.e., SEQ ID NO:27] is identical to that of the wildtype protein sequence [i.e., SEQ ID NO:25]).

[0353] A *NcoI-NotI* fragment comprising CeLPCATs was used to create pY175 (SEQ ID NO:150), in a 3-way ligation with two fragments from pY168 (SEQ ID NO:145). Specifically, the *NcoI-NotI* fragment comprising CeLPCATs was ligated with a 3530 bp *Sph-NotI* pY168 fragment and a 4248 bp *NcoI-SphI* pY168 fragment to result in pY175. The components present in pY175 (FIG. 12A; SEQ ID NO:150) are identical to those present in pY201, with the exception of the YAT1::CeLPCATs::Lip1 gene in pY175, instead of the YAT1::ScA1e1S::Lip1 gene in pY201 (FIG. 10A). Note that CeLPCATs is labeled as "CeLPCATs^{syn}" in FIG. 12A.

EXAMPLE 9

Functional Characterization Of Different LPLATs In EPA-Producing *Yarrowia lipolytica* Strain Y8406

[0354] *Yarrowia lipolytica* strain Y8406U, producing EPA, was used to functionally characterize the effects of overexpression of the *Saccharomyces cerevisiae* Ale1, *Yarrowia lipolytica* Ale1, *Mortierella alpina* LPAAT1, *Yarrowia lipolytica* LPAAT1 and *Caenorhabditis elegans* LPCAT, following their stable integration into the *Yarrowia* host chromosome. This was in spite of the host containing its native LPLATs, i.e., Ale1 and LPAAT1.

Transformation And Growth

[0355] In order to disrupt the *Ura3* gene (Example 3), construct pZKUM (FIG. 5A; SEQ ID NO:130; described in Table 15 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y8406 in a manner similar to that described for pZKUM transformation of strain Y8006 (Example 3). Several transformants were grown and identified to possess a *Ura*-

phenotype.

[0356] GC analyses showed that there were 26.1% EPA of FAMES in pZKUM-transformant strains #4 and #5. These two strains were designated as strains Y8406U1 and Y8406U2, respectively (collectively, Y8406U).

[0357] *Yarrowia lipolytica* strain Y8406U was then individually transformed with linear *SphI*-*AscI* fragments of the integrating vectors described in Example 8, wherein each LPLAT was under the control of the *Yarrowia* YAT1 promoter. Specifically, vectors pY201 (YAT1::ScAle1S::Lip1), pY168 (YAT1::YIAle1::Lip1), pY208 (YAT1::MaLPAAT1S::Lip1), pY207 (YAT1::YILPAAT1::Lip1) and pY175 (YAT1::CeLPCATS::Lip1) were transformed according to the General Methods.

[0358] Each transformation mix was plated on MM agar plates. Several resultant URA⁺ transformants were picked and inoculated into 3 mL FM medium (Biomyx Cat. No. CM-6681, Biomyx Technology, San Diego, CA) containing per L: 6.7 g Difco Yeast Nitrogen Base without amino acids, 5 g Yeast Extract, 6 g KH₂PO₄, 2 g K₂HPO₄, 1.5 g MgSO₄·7H₂O, 1.5 mg thiamine-HCl, and 20 g glucose. After 2 days growth on a shaker at 200 rpm and 30 °C, the cultures were harvested by centrifugation and resuspended in 3 mL HGM medium (Cat. No. 2G2080, Teknova Inc., Hollister, CA) containing 0.63% monopotassium phosphate, 2.7% dipotassium phosphate, 8.0% glucose, adjusted to pH 7.5. After 5 days growth on a shaker at 200 rpm and at 30 °C, 1 mL aliquots of the cultures were harvested by centrifugation and analyzed by GC. Specifically, the cultured cells were collected by centrifugation for 1 min at 13,000 rpm, total lipids were extracted, and fatty acid methyl esters ["FAMES"] were prepared by trans-esterification, and subsequently analyzed with a Hewlett-Packard 6890 GC (General Methods).

[0359] Based on the fatty acid composition of the 3 mL cultures, selected transformants were further characterized. Specifically, clones #5 and #11 of strain Y8406U transformed with expression vector pY201 (comprising ScAle1S) were selected and designated as "Y8406U::ScAle1S-5" and "Y8406U::ScAle1S-11", respectively; clone #16 of strain Y8406U transformed with expression vector pY168 (comprising YIAle1) was selected and designated as "Y8406U::YIAle1"; clone #8 of strain Y8406U transformed with expression vector pY208 (comprising MaLPAAT1S) was selected and designated as "Y8406U::MaLPAAT1S"; clone #21 of strain Y8406U transformed with expression vector pY207 (comprising YILPAAT1) was selected and designated as "Y8406U::YILPAAT1"; and clone #23 of strain Y8406U transformed with expression vector pY175 (comprising CeLPCATS) was selected and designated as "Y8406U::CeLPCATS". Additionally, strain Y8406 (a Ura⁺ strain that was parent to strain Y8406U (Ura⁻)) was used as a control.

[0360] Each selected transformant and the control was streaked onto MM agar plates. Then, one loop of freshly streaked cells was inoculated into 3 mL FM medium and grown overnight at 250 rpm and 30 °C. The OD_{600nm} was measured and an aliquot of the cells were added to a final OD_{600nm} of 0.3 in 25 mL FM medium in a 125 mL flask. After 2 days in a shaker incubator at 250 rpm and at 30 °C, 6 mL of the culture was harvested by centrifugation and resuspended in 25 mL HGM in a 125 mL flask. After 5 days in a shaker incubator at 250 rpm and at 30 °C, a 1 mL aliquot was used for GC analysis (*supra*) and 10 mL dried for dry cell weight ["DCW"] determination.

[0361] For DCW determination, 10 mL culture was harvested by centrifugation for 5 min at 4000 rpm in a Beckman GH-3.8 rotor in a Beckman GS-6R centrifuge. The pellet was resuspended in 25 mL of water and re-harvested as above. The washed pellet was re-suspended in 20 mL of water and transferred to a pre-weighed aluminum pan. The cell suspension was dried overnight in a vacuum oven at 80 °C. The weight of the cells was determined.

Lipid Content, Fatty Acid Composition And Conversion Efficiencies

[0362] A total of four separate experiments were conducted under identical conditions. Experiment 1 compared control strain Y8406 versus strain Y8406U::ScAle1S-5. Experiment 2 compared control strain Y8406 versus strain Y8406U::YIAle1. Experiment 3 compared control strain Y8406 versus strain Y8406U::YIAle1, strain Y8406U::ScAle1 S-11, and strain Y8406U::MaLPAAT1S. Experiment 4 compared control strain Y8406 versus strain Y8406U::MaLPAAT1S, strain Y8406U::YILPAAT1 and strain Y8406U::CeLPCATS.

[0363] In each experiment, the lipid content, fatty acid composition and EPA as a percent of the DCW are quantified for 1, 2 or 3 replicate cultures ["Replicates"] of the control Y8406 strain and the transformant Y8406U strain(s). Additionally, data for each Y8406U transformant is presented as a % of the Y8406 control. Table 26 below summarizes the total lipid content of cells ["TFAs % DCW"], the concentration of each fatty acid as a weight percent of TFAs ["% TFAs"] and the EPA content as a percent of the dry cell weight ["EPA % DCW"]. More specifically, fatty acids are identified as 16:0 (palmitate), 16:1 (palmitoleic acid), 18:0 (stearic acid), 18:1 (oleic acid), 18:2 (LA), ALA, EDA, DGLA, ARA, ETRA, ETA and EPA.

[0364] Table 27 summarizes the conversion efficiency of each desaturase and the Δ9 elongase functioning in the PUFA biosynthetic pathway and which are required for EPA production. Specifically, the Δ12 desaturase conversion efficiency ["Δ12 CE"], Δ8 desaturase conversion efficiency ["Δ8 CE"], Δ5 desaturase conversion efficiency ["Δ5 CE"], Δ17 desaturase conversion efficiency ["Δ17 CE"] and Δ9 elongation conversion efficiency ["Δ9e CE"] are provided for each control Y8406 strain and the transformant Y8406U strain(s); data for each Y8406U transformant is presented as a % of the Y8406 control. Conversion efficiency was calculated according to the formula:

(product)/([substrate+product])*100, where product includes both product and product derivatives.

Table 26: Lipid Content And Composition In LPCAT Transformant Strains Of *Yarrowia lipolytica* Y8406

Expt.	Strain	Repli cates	TFA % DCW	% TFAs												EPA % DCW
				16: 0	16: 1	18: 0	18: 1	18:2	ALA	EDA	DGLA	ARA	ERA	ETA	EPA	
1	Y8406	AVG.3	17.6	3.8	0.7	3.3	6.4	22.6	2.5	2.8	2.2	0.5	1.9	2.0	48.9	8.6
	Y8406U::	AVG.3	18.3	4.2	0.7	3.5	5.7	15.1	0.6	3.3	3.7	0.8	1.8	2.3	56.9	10.4
	ScAle1S-5	% Ctrl	104	111	100	106	89	67	24	118	168	160	95	115	116	121
	Y8406	AVG.3	23.2	3.5	0.6	3.3	6.4	22.3	2.7	2.6	2.1	0.5	1.6	2.0	49.9	11.6
2	Y8406U::	AVG.3	22.3	3.8	0.7	2.9	3.9	12.7	0.4	3.0	3.8	0.8	1.6	2.4	60.9	13.6
	YIAle1	% Ctrl	96	109	117	88	61	57	15	115	181	160	100	120	122	117
	Y8406	1	26.1	2.7	0.7	2.8	6.5	20.5	2.5	3.2	2.3	0.7	0.8	0.0	50.8	13.3
	Y8406U::	AVG.2	23.3	3.3	0.7	2.4	3.6	12.1	0.5	3.2	3.5	0.9	0.0	2.3	62.2	14.5
3	YIAle1	% Ctrl	89	122	100	86	55	59	20	100	152	129	0	na	122	109
	Y8406U::	AVG.2	28.0	3.0	0.7	3.0	5.5	13.1	0.6	3.5	3.8	0.9	0.0	2.4	58.5	16.4
	ScAle1S-11	% Ctrl	107	111	100	107	85	64	24	109	165	129	0	na	115	123
	Y8406U::	AVG.2	23.7	4.4	0.8	4.2	6.6	11.2	0.7	2.7	3.7	0.9	0.0	2.5	57.0	13.5
4	MaLPAAT1S	% Ctrl	91	163	114	150	102	55	28	84	161	129	0	na	112	102
	Y8406	AVG.2	27.9	2.8	0.6	3.1	6.2	20.6	2.9	2.9	2.0	0.6	0.7	2.0	49.4	13.8
	Y8406U::	AVG.2	25.2	4.8	0.8	4.8	6.9	11.6	0.8	2.5	3.0	0.7	0.0	2.3	55.3	14.0
	MaLPAAT1S	% Ctrl	90	171	133	155	111	56	28	86	150	117	0	115	112	101
	Y8406U::	AVG.2	25.2	3.7	0.7	4.2	6.2	13.0	1.2	2.3	2.6	0.6	0.0	2.2	56.7	14.3
	YILPAAT1	% Ctrl	90	132	117	135	100	63	41	79	130	100	0	110	115	104
	Y8406U::	AVG.2	24.7	3.8	0.6	4.6	7.1	13.9	1.6	2.3	2.6	0.6	0.4	2.2	53.6	13.2
	CeLPCATS	% Ctrl	89	136	100	148	115	67	55	79	130	100	57	110	109	96

Table 27: Desaturase And Elongate Conversion Efficiency In LPCAT Transformant Strains Of *Yarrowia lipolytica* Y8406

Expt.	Strain	Replicates	Δ12 CE	Δ9e CE	Δ8 CE	Δ5 CE	Δ17 CE
1	Y8406	AVG.3	93	70	92	92	90
	Y8406U::ScAle1S-5	AVG.3	94	81	93	91	89
		% Ctrl	101	116	101	98	98
2	Y8406	AVG.3	93	70	93	93	91
	Y8406U::YIAle1	AVG.3	96	85	94	91	90
		% Ctrl	103	121	101	98	98
	Y8406	1	93	72	93	96	89
		AVG.2	96	85	96	92	89

Expt.	Strain	Replicates	$\Delta 12$ CE	$\Delta 9e$ CE	$\Delta 8$ CE	$\Delta 5$ CE	$\Delta 17$ CE
3	Y8406U::ScA1e1S-11	% Ctrl	104	119	103	96	100
		AVG.2	94	83	95	91	88
		% Ctrl	101	117	102	95	99
	Y8406U::MaLPAAT1S	AVG.2	92	85	96	90	89
		% Ctrl	100	119	103	94	100
		AVG.2	93	71	94	93	91
4	Y8406U::MaLPAAT1S	AVG.2	92	84	96	91	90
		% Ctrl	99	118	102	99	100
		AVG.2	93	82	96	92	92
	Y8406U::YiLPAAT1	% Ctrl	100	115	103	100	101
		AVG.2	92	80	96	92	91
		% Ctrl	99	113	102	99	100

[0365] Based on the data concerning Experiments 1, 2 and 3 in Table 26 and Table 27, overexpression of LPLAT in EPA strains Y8406U::ScA1e1S-5, Y8406U::ScA1e1S-11, Y8406U::YiA1e1 and Y8406U::MaLPAAT1 S results in significant reduction (to 67% or below of the control) of the concentration of LA (18:2) as a weight % of TFAs ["LA % TFAs"], an increase (to at least 12% of the control) in the concentration of EPA as a weight % of TFAs ["EPA % TFAs"], and an increase (to at least 16% of the control) in the conversion efficiency of the $\Delta 9$ elongase. Compared to Y8406U::ScA1e1S-5 and Y8406U::ScA1e1S-11, Y8406U::YiA1e1 has lower LA % TFAs, higher EPA % TFAs, better $\Delta 9$ elongation conversion efficiency, and slightly lower TFAs % DCW and EPA % DCW. Y8406U::YiA1e1 and Y8406U::MaLPAAT1S are similar except overexpression of MaLPAAT1S resulted in lower LA % TFAs, EPA % TFAs, and EPA % DCW.

[0366] Experiment 4 shows that overexpression of LPLAT in EPA strains Y8406U::YiA1e1, Y8406U::MaLPAAT1 S and Y8406U::CeLPCATS results in significant reduction (to 67% or below of the control) of LA % TFAs, an increase (to at least 9% of the control) in EPA % TFAs, and an increase (to at least 13% of the control) in the conversion efficiency of the $\Delta 9$ elongase. Compared to Y8406U::CeLPCATS, Y8406U::YiLPAAT1 and Y8406U::MaLPAAT1S both have lower LA % TFAs, higher EPA % TFAs, higher EPA % DCW, and slightly better TFAs % DCW. Y8406U::YiLPAAT1 and Y8406U::MaLPAAT1S are similar except overexpression of MaLPAAT1S results in lower LA % TFAs, slightly lower EPA % TFAs and EPA % DCW, and slightly better $\Delta 9$ elongase conversion efficiency.

[0367] It is well known in the art that most desaturations occur at the sn-2 position of phospholipids, while fatty acid elongations occur on acyl-CoAs. Furthermore, ScA1e1S, YiA1e1, MaLPAAT1S and YiLPAAT1 were expected to only incorporate acyl groups from the acyl-CoA pool into the sn-2 position of lysophospholipids, such as lysophosphatidic acid ["LPA"] and lysophosphatidylcholine ["LPC"]. Thus, it was expected that expression of ScA1e1S, YiA1e1, MaLPAAT1S, and YiLPAAT1 would result in improved desaturations due to improved substrate availability in phospholipids, and not result in improved elongations that require improved substrate availability in the CoA pool. Our data (*supra*) shows that unexpectedly, expression of ScA1e1S, YiA1e1, MaLPAAT1S and YiLPAAT1 significantly improved the $\Delta 9$ elongase conversion efficiency in strains of *Yarrowia* producing EPA but did not improve the desaturations (measured as $\Delta 12$ desaturase conversion efficiency, $\Delta 8$ desaturase conversion efficiency, $\Delta 5$ desaturase conversion efficiency or $\Delta 17$ desaturase conversion efficiency).

[0368] CeLPCAT was previously shown to improve $\Delta 6$ elongation conversion efficiency in *Saccharomyces cerevisiae* fed LA or GLA (Int. App. Pub. No. WO 2004/076617). This was attributed to its reversible LPCAT activity that released fatty acids from phospholipids into the CoA pool. An improvement in $\Delta 9$ elongation conversion efficiency in an oleaginous microbe, such as *Yarrowia lipolytica*, engineered for high level LC-PUFA production in the absence of feeding fatty acids was not contemplated in Int. App. Pub. No. WO 2004/076617.

[0369] Furthermore, expression of ScA1e1S, YiA1e1, MaLPAAT1S, YiLPAAT1 and CeLPCATS did not significantly alter either the level of PUFAs accumulated or the total lipid content in strains of *Yarrowia* producing EPA.

[0370] Previous studies have shown that both $\Delta 6$ elongation and $\Delta 9$ elongation are bottlenecks in long chain PUFA biosynthesis due to poor transfer of acyl groups between phospholipid and acyl-CoA pools. Based on the improved $\Delta 9$ elongase conversion efficiency resulting from overexpression of LPLATs, demonstrated above, it is anticipated that the LPLATs described herein and their orthologs, such as ScLPAAT, will also improve $\Delta 6$ elongation conversion efficiency.

EXAMPLE 10

Construction Of Expression Vectors Comprising LPAAT ORFs And An Autonomously Replicating Sequence

[0371] The present example describes the construction of vectors comprising autonomously replicating sequences ["ARS"] and LPAAT ORFs suitable for LPAAT gene expression without integration in *Yarrowia lipolytica*. ORFs included the *Saccharomyces cerevisiae* LPAAT encoding SEQ ID NO:32 and the *Yarrowia lipolytica* LPAAT1 encoding SEQ ID NO:31. Example 11 describes the results obtained following transformation of these vectors into *Y. lipolytica*.

Construction Of pY222, Comprising A Codon-Optimized *Saccharomyces cerevisiae* LPAAT Gene

[0372] The *Saccharomyces cerevisiae* ORF designated as "ScLPAAT" (SEQ ID NO:32) was optimized for expression in *Yarrowia lipolytica*, by DNA 2.0 (Menlo Park, CA). In addition to codon optimization, 5' *Pci*I and 3' *Not*I cloning sites were introduced within the synthetic gene (i.e., ScLPAATS; SEQ ID NO:151). None of the modifications in the ScLPAATS gene changed the amino acid sequence of the encoded protein (i.e., the protein sequence encoded by the codon-optimized gene [i.e., SEQ ID NO:152] is identical to that of the wildtype protein sequence [i.e., SEQ ID NO:32]). ScLPAATS was cloned into pJ201 (DNA 2.0) to result in pJ201 :ScLPAATS.

[0373] A 926 bp *Pci*I/*Not*I fragment comprising ScLPAATS was excised from pJ201 :ScLPAATS and cloned into *Nco*I/*Not*I cut pYAT-DG2-1 to create pY222 (SEQ ID NO:153; Table 28; FIG. 14A). Thus, pY222 contained the following components:

Table 28-Description of Plasmid pY222 (SEQ ID NO:153)

RE Sites And Nucleotides Within SEQ ID NO:153	Description Of Fragment And Chimeric Gene Components
<i>Sa</i> I/ <i>S</i> waI (1-2032)	YAT1::ScLPAATS::Lip1, comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (U.S. Pat. Appl. Pub. No. 2006/0094102-A1); • ScLPAATS: codon-optimized ScLPAATS (SEQ ID NO:151) (labeled as "Sc LPAATs ORF" in Figure); • Lip1: Lip1 terminator sequence from <i>Yarrowia Lip1</i> gene (GenBank Accession No. Z50020) (labeled as "Lip1-3'" in Figure)
<i>S</i> waI/ <i>A</i> vaI (2032-4946)	• <i>ColE1</i> plasmid origin of replication; • Ampicillin-resistance gene (Amp ^R) for selection in <i>E. coli</i> ; • <i>E. coli</i> f1 origin of replication
<i>A</i> vaI- <i>S</i> phI (4946-6330)	<i>Yarrowia lipolytica</i> centromere and autonomously replicating sequence ["ARS"] 18 locus
<i>S</i> phI- <i>S</i> all (6330-1)	<i>Yarrowia lipolytica</i> URA3 gene (GenBank Accession No. AJ306421)

Construction Of pY177, Comprising A *Yarrowia lipolytica* LPAAT1 Gene

[0374] The *Yarrowia lipolytica* centromere and autonomously replicating sequence ["ARS"] was amplified by standard PCR using primer 869 (SEQ ID NO:154) and primer 870 (SEQ ID NO:155), with plasmid pYAT-DG2-1 as template. The PCR product was digested with *Asc*I/*A*vrII and cloned into *Asc*I/*A*vrII digested pY207 (SEQ ID NO:149; Example 8) to create pY177 (SEQ ID NO:156; Table 29; FIG. 14B). Thus, the components present in pY177 are identical to those in pY207 (FIG. 12A), except for the replacement of the 373 bp pY207 sequence between *Asc*I and *A*vrII with the 1341 bp sequence containing ARS. More specifically, pY177 contained the following components:

Table 29: Description of Plasmid pY177 (SEQ ID NO:156)

RE Sites And Nucleotides Within SEQ ID NO:156	Description Of Fragment And Chimeric Gene Components
<i>Bs</i> NI/ <i>Sbf</i> I (1-1706 bp)	LoxP::Ura3::LoxP, comprising: • LoxP sequence (SEQ ID NO:140) • <i>Yarrowia lipolytica</i> Ura3 gene (GenBank Accession No. AJ306421); • LoxP sequence (SEQ ID NO:140)
<i>Sbf</i> I/ <i>Sph</i> I (1706-3043 bp)	3' portion of <i>Yarrowia lipolytica</i> POX3 Acyl-CoA oxidase 3 (GenBank Accession No. YAL10D24750g)
<i>Sph</i> I/ <i>Asc</i> I (3043-5743 bp)	• <i>ColE1</i> plasmid origin of replication; • Ampicillin-resistance gene (Amp ^R) for selection in <i>E. coli</i> ; • <i>E. coli</i> f1 origin of replication
<i>Asc</i> I/ <i>Bs</i> NI (5743-6513 bp)	5' portion of <i>Yarrowia lipolytica</i> POX3 Acyl-CoA oxidase 3 (GenBank Accession No. YAL10D24750g)
<i>Asc</i> I/ <i>Avr</i> II (5743-7084 bp)	<i>Yarrowia lipolytica</i> centromere and autonomously replicating sequence ["ARS"] 18 locus
<i>Avr</i> II/ <i>Bs</i> NI (7084-7481 bp)	5' portion of <i>Yarrowia lipolytica</i> POX3 Acyl-CoA oxidase 3 (GenBank Accession No. YAL10D24750g)
<i>Bs</i> NI/ <i>Bs</i> NI (7481-1 bp)	YAT1::YILPAAT1::Lip1, comprising: • YAT1: <i>Yarrowia lipolytica</i> YAT1 promoter (U.S. Pat. Appl. Pub. No. 2006/0094102-A1); • YILPAAT1: <i>Yarrowia lipolytica</i> LPAAT1 ("YAL10E18964g"; GenBank Accession No. XP_504127) (SEQ ID NO:30) (labeled as "Y1 LPAT1 ORF" in Figure); • Lip1: Lip1 terminator sequence from <i>Yarrowia Lip1</i> gene (GenBank Accession No. Z50020) (labeled as "Lip1-3'" in Figure)

EXAMPLE 11**Functional Characterization Of Different LPAATs In EPA-Producing *Yarrowia lipolytica* Strain Y8406**

[0375] *Yarrowia lipolytica* strain Y8406U, producing EPA, was used to functionally characterize the effects of expression of the *Saccharomyces cerevisiae* LPAATs (SEQ ID NO:151) and *Yarrowia lipolytica* LPAAT1 (SEQ ID NO:30) without integration on self-replicating plasmids. This was in spite of the host containing its native LPAATs.

Transformation And Growth

[0376] *Yarrowia lipolytica* strain Y8406U (Example 9) was individually transformed with uncut plasmids from Example 10. Specifically, vectors pY177 (YAT1::YILPAAT1::Lip1) [SEQ ID NO:156] and pY222 (YAT1::ScLPAATs::Lip1) [SEQ ID NO:153] were transformed according to the General Methods.

[0377] Each transformation mix was plated on MM agar plates. Several resultant URA⁺ transformants were picked and inoculated into 3 mL CSM-U medium (Teknova Cat. No. C8140, Teknova Inc., Hollister, CA), wherein CSM-U medium refers to CM Broth with glucose minus uracil containing 0.13% amino acid dropout powder minus uracil, 0.17% yeast nitrogen base, 0.5% (NH₄)₂SO₄, and 2.0% glucose. After 2 days growth on a shaker at 200 rpm and 30 °C, the cultures were harvested by centrifugation and resuspended in 3 mL HGM medium (Cat. No. 2G2080, Teknova Inc.). After 5 days growth on a shaker, 1 mL aliquots of the cultures were harvested and analyzed by GC, as described in Example 9.

[0378] Based on the fatty acid composition of the 3 mL cultures, selected transformants were further characterized by flask assay. Specifically, clones #5 and #6 of strain Y8406U transformed with expression vector pY222 (comprising ScLPAATs) were selected and designated as "Y8406U::ScLPAATs-5" and "Y8406U::ScLPAATs-6", respectively; clone #1 of strain Y8406U transformed with expression vector pY177 (comprising YILPAAT1) was selected and designated as "Y8406U::YILPAAT1". Additionally, strain Y8406 (a Ura⁺ strain that was parent to strain Y8406U (Ura⁻)) was used as a control.

[0379] Each selected transformant and the control was streaked onto MM agar plates. Then, one loop of freshly streaked cells was inoculated into 3 mL CSM-U medium and grown overnight at 250 rpm and 30 °C. The OD_{600nm} was measured and an aliquot of the cells were added to a final OD_{600nm} of 0.3 in 25 mL CSM-U medium in a 125 mL flask. After 2 days in a shaker incubator at 250 rpm and at 30 °C, 6 mL of the culture was harvested by centrifugation and resuspended in 25 mL HGM in a 125 mL flask. After 5 days in a shaker incubator at 250 rpm and at 30 °C, a 1 mL aliquot was used for GC analysis and 10 mL dried for dry cell weight ("DCW") determination, as described in Example 9.

Lipid Content, Fatty Acid Composition And Conversion Efficiencies

[0380] The lipid content, fatty acid composition and EPA as a percent of the DCW are quantified for 2 replicate cultures ["Replicates"] of the control Y8406 strain and the transformant Y8406U strain(s). Additionally, data for each Y8406U transformant is presented as a % of the Y8406 control. Table 30 below summarizes the total lipid content of cells ["TFAs % DCW"], the concentration of each fatty acid as a weight percent of TFAs ["% TFAs"] and the EPA content as a percent of the dry cell weight ["EPA % DCW"]. More specifically, fatty acids are identified as 16:0 (palmitate), 16:1 (palmitoleic acid), 18:0 (stearic acid), 18:1 (oleic acid), 18:2 (LA), ALA, EDA, DGLA, ARA, ETrA, ETA and EPA.

[0381] Table 31 summarizes the conversion efficiency of each desaturase and the Δ9 elongase functioning in the PUFA biosynthetic pathway and which are required for EPA production, in a manner identical to that described in Example 9.

Table 30: Lipid Content And Composition In ScLPAATs And YILPAAT1 Transformant Strains Of *Yarrowia lipolytica* Y8406

Strain	Replicates	TFA % DCW	% TFAs												EPA % DCW
			16: 0	16: 1	18: 0	18: 1	18:2	ALA	EDA	DGLA	ARA	ERA	ETA	EPA	
Y8406	AVG.2	22.0	2	0	2	4	19	2	3	4	1	2	3	55	12
Y8406U:: YILPAAT1	AVG.2	24.6	2	1	2	6	14	1	3	5	1	2	3	55	14
	% Ctrl	112	98	153	102	148	76	50	120	144	101	109	123	101	113
Y8406U:: ScLPAATs-5	AVG.2	21.6	3	1	3	6	14	1	3	4	1	2	3	57	12
	% Ctrl	98	131	137	125	131	74	56	100	117	86	101	108	104	102
Y8406U:: ScLPAATs-6	AVG.2	21.4	3	1	3	5	14	1	3	4	1	2	3	58	12
	% Ctrl	97	125	133	121	124	72	52	97	119	88	102	111	106	103

Table 31: Desaturase And Elongate Conversion Efficiency In ScLPAATs And YILPAAT1 Transformant Strains Of *Yarrowia lipolytica* Y8406

Strain	Replicates	Δ12 CE	Δ9e CE	Δ8 CE	Δ5 CE	Δ17 CE
Y8406	AVG.2	95	77	92	90	92
Y8406U::YILPAAT1	AVG.2	93	82	92	87	90
	% Ctrl	98	107	99	97	98
Y8406U::ScLPAATs-5	AVG.2	94	83	93	89	92
	% Ctrl	98	108	100	99	100
Y8406U::ScLPAATs-6	AVG.2	94	83	93	89	92
	% Ctrl	99	109	101	99	100

[0382] Based on the data in Table 30 and Table 31 above, overexpression of both ScLPAATS and YILPAAT1 in EPA strains Y8406U::YILPAAT1, Y8406U::ScLPAATS-5 and Y8406U::ScLPAATS-6 resulted in reduction (to 76% or below of the control) of the concentration of LA (18:2) as a weight % of TFAs ["LA % TFAs"], and an increase (to at least 7% of the control) in the conversion efficiency of the $\Delta 9$ elongase. ScLPAATS and YILPAAT1 have a similar effect on lipid profile.

[0383] The results obtained above were then compared to those obtained in Example 9, although different means were utilized to characterize the LPLATs. Specifically, in Example 9, linearized DNA carrying the LPLATs were transformed by chromosomal integration, since the vectors lacked ARS sequences. This resulted in stable integrations and the strains were grown in the relatively rich, non-selective FM growth medium during both preculture and 2 days growth prior to being transferred to HGM.

[0384] In Example 11, the functional characterization of YILPAAT1 and ScLPAATS was done on a replicating plasmid. Thus, *Yarrowia lipolytica* strain Y8406 was transformed with circular DNA carrying each LPAAT and ARS sequence. To maintain these plasmids and assay gene expression without integration, it was necessary to grow the transformants on selective medium (i.e., CSM-U medium) during both preculture and 2 days growth prior to being transferred to HGM.

[0385] These differences described above can contribute to differences in lipid profile and content, as illustrated by the expression of YILPAAT1 in Examples 9 and 11. The change over control in LA % TFAs, EPA % TFAs, and $\Delta 9$ elongase conversion efficiency were 63%, 115%, and 115%, respectively, upon expression of YILPAAT in Example 9, whereas the change over control in LA % TFAs, EPA % TFAs, and $\Delta 9$ elongase conversion efficiency were 76%, 101%, and 107%, respectively, upon expression of YILPAAT in Example 11. Thus, the improvements in $\Delta 9$ elongation and LC-PUFA biosynthesis in Example 11 are minimized when compared to those observed in Example 9. These differences can be attributed to the "position effects" of chromosomal integration and/or different growth conditions.

[0386] Since the improvements in LC-PUFA biosynthesis (measured as reduction in LA % TFAs, increase in EPA % TFAs and increase in $\Delta 9$ elongase conversion efficiency) are similar for both ScLPAATS and YILPAAT when transformed in *Yarrowia lipolytica* strain Y8406 on a replicating plasmid, it is anticipated that both LPLAATs will also function similarly when stably integrated into the host chromosome. Thus, ScLPAATS will likely improve the lipid profile in a manner similar to that observed in Example 9, when YILPAAT1 was stably integrated into the host chromosome.

SEQUENCE LISTING

[0387]

<110> E.I. duPont de Nemours & Company, Inc. Zhu, Quinn

<120> IMPROVED OPTIMIZED STRAINS OF YARROWIA LIPOLYTICA FOR HIGH EICOSAPENTAENOIC ACID PRODUCTION

<130> CL4674

<150> US 61/187366

<151> 2009-06-16

<150> US 61/187368

<151> 2009-06-16

<150> US 61/187359

<151> 2009-06-16

<160> 156

<170> PatentIn version 3.5

<210> 1

<211> 24

<212> PRT

<213> Artificial Sequence

<220>

<223> Multizyme linker

<300>

<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(24)

<300>

<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(24)

<400> 1

Gly Ala Gly Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr Tyr Tyr
1 5 10 15

Asp Ser Leu Ala Val Met Gly Ser
20

<210> 2

<211> 19

<212> PRT

<213> Artificial Sequence

<220>

<223> Multizyme linker

<300>

<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(19)

<300>

<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048

<311> 2008-04-03

<312> 2008-10-16

<313> (1)..(19)

<400> 2

Gly Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr Tyr Tyr Asp Ser
 1 5 10 15
 Leu Ala Val

<210> 3
 <211> 18
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Multizyme linker

<300>
 <302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048
 <311> 2008-04-03
 <312> 2008-10-16
 <313> (1)..(18)

<300>
 <302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1
 <311> 2008-04-03
 <312> 2008-10-16
 <313> (1)..(18)

<400> 3
 Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr Tyr Tyr Asp Ser Leu
 1 5 10 15
 Ala Val

<210> 4
 <211> 18
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Multizyme linker

<300>
 <302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048
 <311> 2008-04-03
 <312> 2008-10-16
 <313> (1)..(18)

<300>
 <302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1
 <311> 2008-04-03
 <312> 2008-10-16
 <313> (1)..(18)

<400> 4
 Pro Thr Arg Pro Ala Gly Pro Pro Ala Thr Tyr Tyr Asp Ser Leu
 1 5 10 15
 Ala Val

<210> 5
 <211> 33
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Multizyme linker

<300>
 <302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048
 <311> 2008-04-03
 <312> 2008-10-16
 <313> (1)..(33)

<300>
 <302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1
 <311> 2008-04-03
 <312> 2008-10-16
 <313> (1)..(33)

<400> 5
 Pro Gly Gly Pro Gly Lys Pro Ser Glu Ile Ala Ser Leu Pro Pro Pro
 1 5 10 15
 Ile Arg Pro Val Gly Asn Pro Pro Ala Ala Tyr Tyr Asp Ala Leu Ala
 20 25 30
 Thr

<210> 6
 <211> 22
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Multizyme linker

<300>
 <302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

```

<310> WO 2008/124048
<311> 2008-04-03
<312> 2008-10-16
<313> (1)..(22)

<300>
<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1
<311> 2008-04-03
<312> 2008-10-16
<313> (1)..(22)

<400> 6
Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr Tyr Tyr Asp Ser Leu
1          5          10          15

Ala Val Ser Gly Arg Thr
20

<210> 7
<211> 36
<212> PRT
<213> Artificial Sequence

<220>
<223> Multizyme linker

<300>
<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048
<311> 2008-04-03
<312> 2008-10-16
<313> (1)..(36)

<300>
<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1
<311> 2008-04-03
<312> 2008-10-16
<313> (1)..(36)

<400> 7
Pro Gly Gly Pro Gly Lys Pro Ser Glu Ile Ala Ser Leu Pro Pro Pro
1          5          10          15

Ile Arg Pro Val Gly Asn Pro Pro Ala Ala Tyr Tyr Asp Ala Leu Ala
20          25          30

Thr Gly Arg Thr
35

<210> 8
<211> 2112
<212> DNA
<213> Artificial Sequence

<220>
<223> Multizyme: EgD9ES/EgD8M gene fusion

<220>
<221> CDS
<222> (1)..(2112)
<223> Multizyme: EgD9ES/EgD8M gene fusion

<300>
<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1
<311> 2008-04-03
<312> 2008-10-16
<313> (1)..(2112)

<300>
<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048
<311> 2008-04-03
<312> 2008-10-16
<313> (1)..(2112)

<400> 8

```

atg gaa gtc gtc aac gaa atc gtc tcc att ggc caa gaa gtt ctt ccc Met Glu Val Val Asn Glu Ile Val Ser Ile Gly Gln Glu Val Leu Pro 1 5 10 15	48
aag gtc gac tat gct caa ctc tgg tct gat gcc tgc cac tgc gaa gtc Lys Val Asp Tyr Ala Gln Leu Trp Ser Asp Ala Ser His Cys Glu Val 20 25 30	96
ctg tac ctc tcc atc gcc ttc gtc atc ctg aag ttc acc ctt ggt cct Leu Tyr Leu Ser Ile Ala Phe Val Ile Leu Lys Phe Thr Leu Gly Pro 35 40 45	144
ctc gga ccc aag ggt caa tct cga atg aag ttt gtc ttc acc aac tac Leu Gly Pro Lys Gly Gln Ser Arg Met Lys Phe Val Phe Thr Asn Tyr 50 55 60	192
aac ctg ctc atg tcc atc tac tgc ctg ggc tcc ttc ctc tot atg gcc Asn Leu Leu Met Ser Ile Tyr Ser Leu Gly Ser Phe Leu Ser Met Ala 65 70 75 80	240
tac gcc atg tac acc att ggt gtc atg tcc gcc aac tgc gaa aag gct Tyr Ala Met Tyr Thr Ile Gly Val Met Ser Asp Asn Cys Glu Lys Ala 85 90 95	288
ttc gac aac aat gtc ttc cga atc acc act cag ctg ttc tac ctc aac Phe Asp Asn Asn Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser 100 105 110	336
aag ttc ctc gaa tac att gac tcc ttc tat ctg ccc ctc atg ggc aag Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Gly Lys 115 120 125	384
cct ctg acc tgg ttg caa ttc ttt cac cat ctc gga gct cct atg aac Pro Leu Thr Trp Leu Gln Phe Phe His His Leu Gly Ala Pro Met Asp 130 135 140	432
atg tgg ctg ttc tac aac tac cga aac gaa gcc gtt tgg atc ttt gtc Met Trp Leu Phe Tyr Asn Tyr Arg Asn Glu Ala Val Trp Ile Phe Val 145 150 155 160	480
ctg ctc aac gcc ttc att cac tgg atc atg tac ggc tac tat tgg acc Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Trp Thr 165 170 175	528
gaa ctg atc aag ctc aag ttc cct atg ccc aag tcc ctg att act tct Arg Leu Ile Lys Leu Lys Phe Pro Met Pro Lys Ser Leu Ile Thr Ser 180 185 190	576
atg caa atc att caa ttc aac gct gcc ttc tac atc gtc tgg aag tac Met Gln Ile Ile Gln Phe Asn Val Gly Phe Tyr Ile Val Trp Lys Tyr 195 200 205	624
agg aac att ccc tgc tac cga cca gat gga atg aga ttt ggc tgg Arg Asn Ile Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Gly Trp 210 215 220	672
ttt ttc aac tac ttc tac gtt ggt act gtc ctg tgt ctg ttc ctc aac Phe Phe Asn Tyr Phe Tyr Val Gly Thr Val Leu Cys Leu Phe Leu Asn 225 230 235 240	720
ttc tac gtc caa acc tac atc gtc cga aag cac aag gga gcc aac aag Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys 245 250 255	768
att caa ggc gcc ggt ccc gct cga cct gcc gga ctt cct ccc gct acc Ile Gln Gly Ala Gly Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr 260 265 270	816
tac tac gac tct ctg gcc gtc atg gga tcc ctg aag gct tat cga caa Tyr Tyr Asp Ser Leu Ala Val Met Gly Ser Val Lys Ala Ser Arg Gln 275 280 285	864
gct ctg ccc ctc gtc atc gaa gga aag gtc tac gac gtc tcc gct tgg Ala Leu Pro Leu Val Ile Asp Gly Lys Val Tyr Asp Val Ser Ala Trp 290 295 300	912
gtg aac ttc cac cct ggt gga gct gaa atc att gag aac tac caa gga Val Asn Phe His Pro Gly Gly Ala Glu Ile Ile Glu Asn Tyr Gln Gly 305 310 315 320	960
cga gat gct act gac gcc ttc atg gtt atg cac tct cag gaa gcc ttc Arg Asp Ala Thr Asp Ala Phe Met Val Met His Ser Gln Glu Ala Phe 325 330 335	1008
gac aac ctc aag cga atc ccc aag atc aac caa gct tcc caa ctg cct Asp Lys Leu Lys Arg Met Pro Lys Ile Asn Gln Ala Ser Glu Leu Pro 340 345 350	1056
ccc caa gct gcc gtc aac gaa gct caa gaa gat ttc cga aag ctc cga Pro Gln Ala Ala Val Asn Glu Ala Gln Glu Asp Phe Arg Lys Leu Arg 355 360 365	1104
gaa gag ctg atc gcc act ggc atg ttt gac gcc tct ccc ctc tgg tac Glu Glu Leu Ile Ala Thr Gly Met Phe Asp Ala Ser Pro Leu Trp Tyr 370 375 380	1152
tgc tac aag atc ttg acc acc ctg ggt ctt ggc gtc ctt gcc ttc ttc Ser Tyr Lys Ile Leu Thr Thr Leu Gly Leu Gly Val Leu Ala Phe Phe 385 390 395 400	1200
atg ctg gtc caa tac cac ctg tac ttc att ggt ctc ctc gtc ctc ggt Met Leu Val Gln Tyr His Leu Tyr Phe Ile Gly Ala Leu Val Leu Gly 405 410 415	1248
atg cac tac caa aag atg gga tgg ctg tct cat gac atc tgc cac cac Met His Tyr Gln Gln Met Gly Trp Leu Ser His Asp Ile Cys His His 420 425 430	1296


```

420          425          430
cag acc ttc aag aac cga aac tgg aat aac gtc ctg ggt ctg gtc ttt 1344
Gln Thr Phe Lys Asn Arg Asn Trp Asn Asn Val Leu Gly Leu Val Phe
435          440          445
ggc aac cga ctc cag ggc ttc tcc ctg acc tgg tgg aag gac aga cag 1392
Gly Asn Gly Leu Gln Gly Phe Ser Val Thr Trp Trp Lys Asp Arg His
450          455          460
aac gcc cat cat tct gct acc aac gtt cag ggt cac ggt ccc gac att 1440
Asn Ala His His Ser Ala Thr Asn Val Gln Gly His Asp Pro Asp Ile
465          470          475
gat aac ctg cct ctg ctc gcc tgg tcc gag gac gat gtc aot cga gct 1488
Asp Asn Leu Pro Leu Leu Ala Trp Ser Gln Asp Asp Val Thr Arg Ala
485          490          495
tot ccc atc tcc cga aag ctc att cag ttc caa cag tac tat ttc ctg 1536
Ser Pro Ile Ser Arg Lys Leu Ile Gln Phe Gln Gln Tyr Tyr Phe Leu
500          505          510
gtc atc tgt att ctc ctg cga ttc atc tgg tgt ttc cag tot gtg ctg 1584
Val Ile Cys Ile Leu Leu Arg Phe Ile Trp Cys Phe Gln Ser Val Leu
515          520          525
acc gtt cga tcc ctc aag gac cga gac aac cag ttc tac cga tot cag 1632
Thr Val Arg Ser Leu Lys Asp Arg Asp Asn Gln Phe Tyr Arg Ser Gln
530          535          540
tac aag aaa gag gcc att gga ctc gct ctg cac tgg aat ctc aag acc 1680
Tyr Lys Lys Gln Ala Ile Gly Leu Ala Leu His Trp Thr Leu Lys Thr
545          550          560
ctg ttc caa ctc ttc ttt atg ccc tcc atc ctg acc tgg atg ctg gtg 1728
Leu Phe His Leu Phe Phe Met Pro Ser Ile Leu Thr Ser Met Leu Val
565          570          575
ttc ttt gtt tcc gag ctc gtc ggt gcc ttc gga att gcc atc gtg gtc 1776
Phe Phe Val Ser Gln Leu Val Gly Gly Phe Gly Ile Ala Ile Val Val
580          585          590
ttc atg aac cac tac cct ctg gag aag atc ggt gat tcc gtc tgg gac 1824
Phe Met Asn His Tyr Pro Leu Gln Lys Ile Gly Asp Ser Val Trp Asp
595          600          605
gga cat ggc ttc tct ctg ggt cag atc cat gag acc atg aac att cga 1872
Gly His Gly Phe Ser Val Gly Gln Ile His Gln Thr Met Asn Ile Arg
610          615          620
cga gcc atc att act gac tgg ttc ttt gga gcc ctg aac tac cag atc 1920
Arg Gly Ile Ile Thr Asp Trp Phe Phe Gly Gly Leu Asn Tyr Gln Ile
625          630          635
gag gac cat ctc tgg ccc aac ctg cct cga cac aac ctc aot gcc gtt 1968
Glu His His Leu Trp Pro Thr Leu Pro Arg His Asn Leu Thr Ala Val
645          650          655
tcc tac cag gtg gaa cag ctg tgc cag aag cac aac ctc ccc tac cga 2016
Ser Tyr Gln Val Glu Gln Leu Cys Gln Lys His Asn Leu Pro Tyr Arg
660          665          670
aac cct ctg ccc cat gaa ggt ctc gtc atc ctg ctc cga tac ctg tcc 2064
Asn Pro Leu Pro His Gln Gly Leu Val Ile Leu Leu Arg Tyr Leu Ser
675          680          685
cag ttc gct cga atg gcc gag aag cag ccc ggt gcc aag gct cag taa 2112
Gln Phe Ala Arg Met Ala Glu Lys Gln Pro Gly Ala Lys Ala Gln
690          695          700
<210> 9
<211> 703
<212> PRT
<213> Artificial Sequence
<220>
<223> Synthetic Construct
<400> 9
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          240
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
245          250          255
Ile Gln Gly Ala Gly Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr
260          265          270
Tyr Tyr Asp Ser Leu Ala Val Met Gly Ser Val Lys Ala Ser Arg Gln
275          280          285
Ala Leu Pro Leu Val Ile Asp Gly Lys Val Tyr Asp Val Ser Ala Trp
290          295          300
Val Asn Phe His Pro Gly Gly Ala Glu Ile Ile Glu Asn Tyr Gln Gly
305          310          315          320
Arg Asp Ala Thr Asp Ala Phe Met Val Met His Ser Gln Glu Ala Phe
325          330          335
Asp Lys Leu Lys Arg Met Pro Lys Ile Asn Gln Ala Ser Glu Leu Pro
340          345          350
Pro Gln Ala Ala Val Asn Glu Ala Gln Glu Asp Phe Arg Lys Leu Arg
355          360          365
Glu Glu Leu Ile Ala Thr Gly Met Phe Asp Ala Ser Pro Leu Trp Tyr
370          375          380
Ser Tyr Lys Ile Leu Thr Thr Leu Gly Leu Gly Val Leu Ala Phe Phe
385          390          395          400
Met Leu Val Gln Tyr His Leu Tyr Phe Ile Gly Ala Leu Val Leu Gly
405          410          415
Met His Tyr Gln Gln Met Gly Trp Leu Ser His Asp Ile Cys His His
420          425          430
Gln Thr Phe Lys Asn Arg Asn Trp Asn Asn Val Leu Gly Leu Val Phe
435          440          445
Gly Asn Gly Leu Gln Gly Phe Ser Val Thr Trp Trp Lys Asp Arg His
450          455          460
Asn Ala His His Ser Ala Thr Asn Val Gln Gly His Asp Pro Asp Ile
465          470          475          480
Asp Asn Leu Pro Leu Leu Ala Trp Ser Glu Asp Asp Val Thr Arg Ala
485          490          495
Ser Pro Ile Ser Arg Lys Leu Ile Gln Phe Gln Gln Tyr Tyr Phe Leu
500          505          510
Val Ile Cys Ile Leu Leu Arg Phe Ile Trp Cys Phe Gln Ser Val Leu
515          520          525
Thr Val Arg Ser Leu Lys Asp Arg Asp Asn Gln Phe Tyr Arg Ser Gln
530          535          540
Tyr Lys Lys Glu Ala Ile Gly Leu Ala Leu His Trp Thr Leu Lys Thr
545          550          555          560
Leu Phe His Leu Phe Phe Met Pro Ser Ile Leu Thr Ser Met Leu Val
565          570          575
Phe Phe Val Ser Glu Leu Val Gly Gly Phe Gly Ile Ala Ile Val Val
580          585          590
Phe Met Asn His Tyr Pro Leu Glu Lys Ile Gly Asp Ser Val Trp Asp
595          600          605
Gly His Gly Phe Ser Val Gly Gln Ile His Glu Thr Met Asn Ile Arg
610          615          620
Arg Gly Ile Ile Thr Asp Trp Phe Phe Gly Gly Lys Asn Tyr Gln Ile
625          630          635          640
Glu His His Leu Trp Pro Thr Leu Pro Arg His Asn Leu Thr Ala Val
645          650          655
Ser Tyr Gln Val Glu Gln Leu Cys Gln Lys His Asn Leu Pro Tyr Arg
660          665          670
Asn Pro Leu Pro His Glu Gly Leu Val Ile Leu Leu Arg Tyr Leu Ser
675          680          685
Gln Phe Ala Arg Met Ala Glu Lys Gln Pro Gly Ala Lys Ala Gln
690          695          700
<210> 10
<211> 2109
<212> DNA
<213> Artificial Sequence
<220>
<223> Multizyme: EaD9ES/EaD8S gene fusion
<220>
<221> CDS
<222> (1) .. (2109)
<223> Multizyme: EaD9ES/EaD8S gene fusion
<300>
<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS
<310> WO 2008/124048
<311> 2008-04-03
<312> 2008-10-16
<313> (1) .. (2109)

```

<300>

<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1

<311> 2008-04-03

<312> 2008-10-16

<313> (1) .. (2109)

```

<400> 10
atg gag gct gcc aag gag ctg gtc ttc atc gtc cag gag gaa ctt ccc 48
Met Glu Ala Ala Lys Glu Leu Val Ser Ile Val Gln Glu Glu Leu Pro
1 5 10 15

aag gtg gac tac gcc cag ctg tgg cag gac gcc tcc tct tgc gag gtt 96
Lys Val Asp Tyr Ala Gln Leu Trp Gln Asp Ala Ser Ser Cys Glu Val
20 25 30

ctg tac ctc tog gtc gct ttc gtg gcc atc aag ttc atg ctt cga cct 144
Leu Tyr Leu Ser Val Ala Phe Val Ala Ile Lys Phe Met Leu Arg Pro
35 40 45

ctg gac ctc aag cga caa gcc acc ctc aaa aag ctg ttc acc gca tac 192
Leu Asp Leu Lys Arg Gln Ala Thr Leu Lys Lys Leu Phe Thr Ala Tyr
50 55 60

aac ttt ctc atg tcc atc tac tgg ttc ggc tcc ttc ctg gcg atg gcc 240
Asn Phe Leu Met Ser Ile Tyr Ser Phe Gly Ser Phe Leu Ala Met Ala
65 70 75 80

tac gct ctc tct gtc act ggt att ctt tcc ggc gat tgt gag act gcc 288
Tyr Ala Leu Ser Val Thr Gly Ile Leu Ser Gly Asp Cys Glu Thr Ala
85 90 95

ttc aac aat gac gtg ttc cga atc acc act cag ctg ttc tac ctc agc 336
Phe Asn Asn Asp Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser
100 105 110

aag ttc gtc gag tac atc gac tcc ttc tac ctt ccc ctc atg gac aag 384
Lys Phe Val Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Asp Lys
115 120 125

ccc ttg tog ttt ctg cag ttc ttt cac cat ctc gga gct ccc atc gac 432
Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly Ala Pro Ile Asp
130 135 140

atg tgg ctg ttc tac aag tat cga aac gaa ggc gtc tgg atc ttt gtt 480
Met Trp Leu Phe Tyr Lys Tyr Arg Asn Glu Gly Val Trp Ile Phe Val
145 150 155 160

ctg ctc aac ggc ttc att cac tgg atc atg tac ggt tac tat tgg acg 528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
165 170 175

cga ctc atc aag ctg aac ttc cct atg ccc aag aac ctc att acc tcc 576
Arg Leu Ile Lys Leu Asn Phe Pro Met Pro Lys Asn Leu Ile Thr Ser
180 185 190

atg caa att atc cag ttc aac gtc gga 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

cga aac gtg ccc tgc tac cgg cag gac ggt atg cga atg ttt gcc tgg 672
Arg Asn Val Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Ala Trp
210 215 220

atc ttc aac tac tgg tat gtc ggc aag gtg ctg ctt ctg ttc ctc aac 720
Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu Phe Leu Asn
225 230 235 240

ttc tac gtc cag acc tac att cgg aag cct cga aag aac cga ggc aac 768
Phe Tyr Val Gln Thr Tyr Ile Arg Lys Pro Arg Lys Asn Arg Gly Lys
245 250 255

aag gag gcc gcc ggt ccc gct cga cct gcc gga ctt cct ccc gct acc 816
Lys Glu Gly Ala Gly Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr
260 265 270

tac tac gac tct ctg gcc gtc atg cga tcc atg gtc aag cga ccc gct 864
Tyr Tyr Asp Ser Leu Ala Val Met Gly Ser Met Val Lys Arg Pro Ala
275 280 285

ctg cct ctc acc gtg gac ggt gtc aec tac gac gtt tct gcc tgg ctc 912
Leu Pro Leu Thr Val Asp Gly Val Thr Tyr Asp Val Ser Ala Trp Leu
290 295 300

aac cac cat ccc gga ggt gcc gac att atc gag aac tac cga ggt cgg 960
Asn His His Pro Gly Gly Ala Asp Ile Ile Glu Asn Tyr Arg Gly Arg
305 310 315 320

gat gct acc gac gtc ttc atg gtt atg cac tcc gag aac gcc gtg tcc 1008
Asp Ala Thr Asp Val Phe Met Val Met His Ser Glu Asn Ala Val Ser
325 330 335

aaa ctc aga cga atg ccc atc atg gaa cct tcc tct ccc ctg act cca 1056
Lys Leu Arg Arg Met Pro Ile Met Glu Pro Ser Ser Pro Leu Thr Pro
340 345 350

aca cct ccc aag cca aac tcc gac gaa cct cag gag gat ttc cga aag 1104
Thr Pro Pro Lys Pro Asn Ser Asp Glu Pro Gln Glu Asp Phe Arg Lys
355 360 365

ctg cga gac gag ctc att gct cca gcc atg ttc gat gcc tct ccc atg 1152
Leu Arg Asp Glu Leu Ile Ala Ala Gly Met Phe Asp Ala Ser Pro Met
370 375 380

tgg tac gct tac aag acc ctg tog act ctc gga ctg ggt gtc ctt gcc 1200
Trp Tyr Ala Tyr Lys Thr Leu Ser Thr Leu Gly Leu Gly Val Leu Ala
385 390 395 400

gtg ctg ttg atg acc cag tgg cac tgg tac ctg gtt ggt gct atc gtc 1248
Val Leu Leu Met Thr Gln Trp His Trp Tyr Leu Val Gly Ala Ile Val
405 410 415

```

```

ctc gcc att cac ttt caa cag atg gga tgg ctc tcg cac gac att tgc      1296
Leu Gly Ile His Phe Gln Gln Met Gly Trp Leu Ser His Asp Ile Cys
420 425 430

cat cac cag ctg ttc aag gac cga tcc atc aac aat gcc att gcc ctg      1344
His His Gln Leu Phe Lys Asp Arg Ser Ile Asn Asn Ala Ile Gly Leu
435 440 445

ctc ttc gga aac gtc ctt cag gcc ttt tct gtc act tgg tgg aag gac      1392
Leu Phe Gly Asn Val Leu Gln Gly Phe Ser Val Thr Trp Trp Lys Asp
450 455 460

cga cac aac gct cat cac tcc gcc aac aac ctg cag ggt cgc gct ccc      1440
Arg His Asn Ala His His Ser Ala Thr Asn Val Gln Gly His Asp Pro
465 470 475 480

gac atc gac aac ctg cct ctc ctg ccg tgg tcc aag gac gac gtc gac      1488
Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser Lys Glu Asp Val Glu
485 490 495

cga gct gcc ccg ttt tct cga ccg atg atc aag tac caa cag tat tac      1536
Arg Ala Gly Pro Phe Ser Arg Arg Met Ile Lys Tyr Gln Gln Tyr Tyr
500 505 510

ttc ttt ttc atc tgt gcc ctt ctg cga ttc atc tgg tgc ttt cag tcc      1584
Phe Phe Phe Ile Cys Ala Leu Leu Arg Phe Ile Trp Cys Phe Gln Ser
515 520 525

att cat act gcc aag ggt ctc aag gct cga agc aat cag tcc tat cga      1632
Ile His Thr Ala Thr Gly Leu Lys Asp Arg Ser Asn Gln Tyr Tyr Arg
530 535 540

aga cag tac gag aag gag tcc gtc ggt ctg gca ctc caa tgg ggt ctc      1680
Arg Gln Tyr Glu Lys Glu Ser Val Gly Leu Ala Leu His Trp Gly Leu
545 550 555 560

aag gcc ttg ttc tac tat ttc tac atg ccc tgg ttt ctc aac gga ctc      1728
Lys Ala Leu Phe Tyr Tyr Phe Tyr Met Pro Ser Phe Leu Thr Gly Leu
565 570 575

atg gtc ttc ttt gtc tcc gac ctg ctt ggt gcc ttc gga att gcc atc      1776
Met Val Phe Phe Val Ser Glu Leu Leu Gly Gly Phe Gly Ile Ala Ile
580 585 590

gtt gtc ttc atg aac cac tcc cct ctg gag aag att cag gac tcc gtg      1824
Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile Gln Asp Ser Val
595 600 605

tgg gct ggt cat gcc ttc tgt gct gga cag att cac gag acc atg aac      1872
Trp Asp Gly His Gly Phe Cys Ala Gly Gln Ile His Gly Thr Met Asn
610 615 620

gtt cag cga gcc ctc gtc aca gac tgg ttt ttc ggt gcc ctc aac tac      1920
Val Gln Arg Gly Leu Val Thr Asp Trp Phe Phe Gly Gly Leu Asn Tyr
625 630 635 640

cag atc gaa cat cac ctg tgg cct act ctt ccc aga aac aac ctc aac      1968
Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg His Asn Leu Thr
645 650 655

cct gcc tcc atc aaa gtg gag cag ctg tgc aag aag aac aac ctg ccc      2016
Ala Ala Ser Ile Lys Val Glu Gln Leu Cys Lys Lys His Asn Leu Pro
660 665 670

tac cga tct cct ccc atg ctc gaa ggt gtc gcc att ctt atc tcc tac      2064

Tyr Arg Ser Pro Pro Met Leu Glu Gly Val Gly Ile Leu Ile Ser Tyr      2109
675 680 685

ctg gcc acc ttc gct cga atg gtt gcc aag gca gac aag gcc taa
Leu Gly Thr Phe Ala Arg Met Val Ala Lys Ala Asp Lys Ala
690 695 700

<210> 11
<211> 702
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 11
Met Glu Ala Ala Lys Glu Leu Val Ser Ile Val Gln Glu Glu Leu Pro
1 5 10 15

Lys Val Asp Tyr Ala Gln Leu Trp Gln Asp Ala Ser Ser Cys Glu Val
20 25 30

Leu Tyr Leu Ser Val Ala Phe Val Ala Ile Lys Phe Met Leu Arg Pro
35 40 45

Leu Asp Leu Lys Arg Gln Ala Thr Leu Lys Lys Leu Phe Thr Ala Tyr
50 55 60

Asn Phe Leu Met Ser Ile Tyr Ser Phe Gly Ser Phe Leu Ala Met Ala
65 70 75 80

Tyr Ala Leu Ser Val Thr Gly Ile Leu Ser Gly Asp Cys Glu Thr Ala
85 90 95

Phe Asn Asn Asp Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser
100 105 110

Lys Phe Val Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Asp Lys
115 120 125

Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly Ala Pro Ile Asp
130 135 140

Met Trp Leu Phe Tyr Lys Tyr Arg Asn Glu Gly 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 Asn Phe Pro Met Pro Lys Asn 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 Val Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Ala Trp
210 215 220

Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu Phe Leu Asn
225 230 235 240

Phe Tyr Val Gln Thr Tyr Ile Arg Lys Pro Arg Lys Asn Arg Gly Lys
245 250 255

Lys Glu Gly Ala Gly Pro Ala Arg Pro Ala Gly Leu Pro Pro Ala Thr
260 265 270

Tyr Tyr Asp Ser Leu Ala Val Met Met Gly Ser Met Val Lys Arg Pro Ala
275 280 285

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

Asn His His Pro Gly Gly Ala Asp Ile Ile Glu Asn Tyr Arg Gly Arg
305 310 315 320

Asp Ala Thr Asp Val Phe Met Val Met His Ser Glu Asn Ala Val Ser
325 330 335

Lys Leu Arg Arg Met Pro Ile Met Glu Pro Ser Ser Pro Leu Thr Pro
340 345 350

Thr Pro Pro Lys Pro Asn Ser Asp Glu Pro Gln Glu Asp Phe Arg Lys
355 360 365

Leu Arg Asp Glu Leu Ile Ala Ala Gly Met Phe Asp Ala Ser Pro Met
370 375 380

Trp Tyr Ala Tyr Lys Thr Leu Ser Thr Leu Gly Leu Gly Val Leu Ala
385 390 395 400

Val Leu Leu Met Thr Gln Trp His Trp Tyr Leu Val Gly Ala Ile Val
405 410 415

Leu Gly Ile His Phe Gln Gln Met Gly Trp Leu Ser His Asp Ile Cys
420 425 430

His His Gln Leu Phe Lys Asp Arg Ser Ile Asn Asn Ala Ile Gly Leu
435 440 445

Leu Phe Gly Asn Val Leu Gln Gly Phe Ser Val Thr Trp Trp Lys Asp
450 455 460

Arg His Asn Ala His His Ser Ala Thr Asn Val Gln Gly His Asp Pro
465 470 475 480

Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser Lys Glu Asp Val Glu
485 490 495

Arg Ala Gly Pro Phe Ser Arg Arg Met Ile Lys Tyr Gln Gln Tyr Tyr
500 505 510

Phe Phe Phe Ile Cys Ala Leu Leu Arg Phe Ile Trp Cys Phe Gln Ser
515 520 525

Ile His Thr Ala Thr Gly Leu Lys Asp Arg Ser Asn Gln Tyr Tyr Arg
530 535 540

Arg Gln Tyr Glu Lys Glu Ser Val Gly Leu Ala Leu His Trp Gly Leu
545 550 555 560

Lys Ala Leu Phe Tyr Tyr Phe Tyr Met Pro Ser Phe Leu Thr Gly Leu
565 570 575

Met Val Phe Phe Val Ser Glu Leu Leu Gly Gly Phe Gly Ile Ala Ile
580 585 590

Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile Gln Asp Ser Val
595 600 605

Trp Asp Gly His Gly Phe Cys Ala Gly Gln Ile His Glu Thr Met Asn
610 615 620

Val Gln Arg Gly Leu Val Thr Asp Trp Phe Phe Gly Gly Leu Asn Tyr
625 630 635 640

Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg His Asn Leu Thr
645 650 655

Ala Ala Ser Ile Lys Val Glu Gln Leu Cys Lys Lys His Asn Leu Pro
660 665 670

Tyr Arg Ser Pro Pro Met Leu Glu Gly Val Gly Ile Leu Ile Ser Tyr
675 680 685

Leu Gly Thr Phe Ala Arg Met Val Ala Lys Ala Asp Lys Ala
690 695 700

<210> 12
<211> 2127
<212> DNA
<213> Artificial Sequence

<220>
<223> Multizyme: E389D9eS/EgD8M gene fusion

<220>
<221> CDS
<222> (1)..(2127)
<223> Multizyme: E389D9eS/EgD8M gene fusion

<300>
<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2008-0254191-A1
<311> 2008-04-03
<312> 2008-10-16
<313> (1)..(2127)

<300>

<302> MULTIZYMES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124048

<311> 2008-04-03

<312> 2008-10-16

<313> (1).(2127)

<400> 12

```

atg gct gcc atc atc gag gtg gcc aac gag ttc gtc gct atc act gcc      48
Met Ala Ala Val Ile Glu Val Ala Asn Glu Phe Val Ala Ile Thr Ala
1      5      10      15

gag acc ctt ccc aag gtg gac tat cag cga ctc tgg cga gac atc tac      96
Glu Thr Leu Pro Lys Val Asp Tyr Gln Arg Leu Trp Arg Asp Ile Tyr
20     25     30

tcc tgc gac ctc ctg tcc ttc tcc att gct ttc gtc atc ctc aag ttt      144
Ser Cys Glu Leu Leu Tyr Phe Ser Ile Ala Phe Val Ile Leu Lys Phe
35     40     45

acc ctt gcc gag ctc tgg gat tct gcc aaa aag att ctg cga gtg ctg      192
Thr Leu Gly Glu Leu Ser Asp Ser Gly Lys Lys Ile Leu Arg Val Leu
50     55     60

ttc aag tgg tac aac ctc ttc atg tcc gtc ttt tgg ctg gtg tcc ttc      240
Phe Lys Trp Tyr Asn Leu Phe Met Ser Val Phe Ser Leu Val Ser Phe
65     70     75     80

ctc tgt atg ggt tac gcc atc tac aac gtt gga ctg tac tcc aac gaa      288
Leu Cys Met Gly Tyr Ala Ile Tyr Thr Val Gly Leu Tyr Ser Asn Glu
85     90     95

tgc gac aga gct ttc gac aac agc tgg ttc cga ttt gcc acc aag gtc      336
Cys Asp Arg Ala Phe Asp Asn Ser Leu Phe Arg Phe Ala Thr Lys Val
100    105    110

ttc tac tat tcc aag ttt ctg gag tac atc gac tct ttc tac ctt ccc      384
Phe Tyr Tyr Ser Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro
115    120    125

ctc atg gcc aag cct ctg tcc ttt ctg cag ttc ttt cat cac ttg cga      432
Leu Met Ala Lys Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly
130    135    140

gct cct atg gac atg tgg ctc ttc gtg cag tac tct gcc gaa tcc att      480

Ala Pro Met Asp Met Trp Leu Phe Val Gln Tyr Ser Gly Glu Ser Ile
145    150    155    160

tgg atc ttt gtg ttc ctg aac gga ttc att cac ttt gtc atg tac gcc      528
Trp Ile Phe Val Phe Leu Asn Gly Phe Ile His Phe Val Met Tyr Gly
165    170    175

tac tat tgg aca cgg ctg atg aag ttc aac ttt ccc atg ccc aag cag      576
Tyr Tyr Trp Thr Arg Leu Met Lys Phe Asn Phe Pro Met Pro Lys Gln
180    185    190

ctc att acc gca atg cag atc acc cag ttc aac gtt gcc ttc tac ctc      624
Leu Ile Thr Ala Met Gln Ile Thr Gln Phe Asn Val Gly Phe Tyr Leu
195    200    205

gtg tgg tgg tac aag gac att ccc tgt tac cga aag gat ccc atg cga      672
Val Trp Trp Tyr Lys Asp Ile Pro Cys Tyr Arg Lys Asp Pro Met Arg
210    215    220

atg ctg gcc tgg atc ttc aac tac tgg tac gtc ggt acc gtt ctt ctg      720
Met Leu Ala Trp Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu
225    230    235    240

ctc ttc atc aac ttc ttt gtc aag tcc tac gtg ttt ccc aag cct aag      768
Leu Phe Ile Asn Phe Phe Val Lys Ser Tyr Val Phe Pro Lys Pro Lys
245    250    255

act gcc gac aaa aag gtc cag gcc gcc ggt ccc gct cga cct gcc gga      816
Thr Ala Asp Lys Lys Val Gln Gly Ala Gly Pro Ala Arg Pro Ala Gly
260    265    270

ctt cct ccc gct acc tac tac gac tct ctg gcc gtc atg gga tcc gtg      864
Leu Pro Pro Ala Thr Tyr Tyr Asp Ser Leu Ala Val Met Gly Ser Val
275    280    285

aag gct tct cga cag gct ctg ccc ctc gtc atc gac gga aag gtg tac      912
Lys Ala Ser Arg Gln Ala Leu Pro Leu Val Ile Asp Gly Lys Val Tyr
290    295    300

gac gtc tcc gct tgg gtg aac ttc cac cct ggt gga gct gaa atc att      960
Asp Val Ser Ala Trp Val Asn Phe His Pro Gly Gly Ala Glu Ile Ile
305    310    315    320

gag aac tac cag gga cga gat gct aat gac gcc ttc atg gtt atg cac      1008
Glu Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val Met His
325    330    335

tct cag gaa gcc ttc gac aag ctc aag cga atg ccc aag atc aac cag      1056
Ser Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile Asn Gln
340    345    350

gct tcc gag ctg cct ccc cag gct gcc gtc aac gaa gct cag gag gat      1104
Ala Ser Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln Glu Asp
355    360    365

ttc cga aag ctc cga gaa gag ctg atc gcc act gcc atg ttt gac gcc      1152
Phe Arg Lys Leu Arg Glu Glu Leu Ile Ala Thr Gly Met Phe Asp Ala
370    375    380

tct ccc ctc tgg tac tgg tac aag atc ttg acc acc ctg ggt ctt gcc      1200
Ser Pro Leu Trp Tyr Ser Tyr Lys Ile Leu Thr Thr Leu Gly Leu Gly
385    390    395    400

gtg ctt gcc ttc ttc atg ctg gtc cag tac cac ctg tac ttc att ggt      1248
Val Leu Ala Phe Phe Met Leu Val Gln Tyr His Leu Tyr Phe Ile Gly

```

405	410	415	
gct ctc gtc ctc ggt atg cac tac cag caa atg gga tgg ctg tct cat Ala Leu Val Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu Ser His 420 425 430			1296
gac atc tgc cac cac cag acc ttc aag aac cga aac tgg aat aac gtc Asp Ile Cys His His Gln Thr Phe Lys Asn Arg Asn Trp Asn Asn Val 435 440 445			1344
ctg ggt ctg gtc ttt ggc aac gga ctc cag ggc ttc tcc gtc acc tgg Leu Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val Thr Trp 450 455 460			1392
tgg aag gac aga cac aac gcc cat cat tct gct acc aac gtt cag ggt Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val Gln Gly 465 470 475			1440
cac gat ccc gac att gat aac ctg cct ctg ctc gcc tgg tcc gac gac His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser Glu Asp 485 490 495			1488
gat gtc act cga gct tct ccc atc tcc cga aag ctc att cag ttc caa Asp Val Thr Arg Ala Ser Pro Ile Ser Arg Lys Leu Ile Gln Phe Gln 500 505 510			1536
cag tac tat ttc ctg gtc atc tgt att ctc ctg cga ttc atc tgg tgt Gln Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile Trp Cys 515 520 525			1584
ttc cag tct gtc ctg acc gtt cga tcc ctc aag gac cga gac aac cag Phe Gln Ser Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp Asn Gln 530 535 540			1632
ttc tac cga tct cag tac aag aaa gag gcc att gga ctc gct ctg cac Phe Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala Leu His 545 550 555			1680
tgg act ctc aag aac ctg ttc csc ctc ttc ttt atg ccc tcc atc ctg Trp Thr Leu Lys Thr Leu Phe His Leu Phe Phe Met Pro Ser Ile Leu 565 570 575			1728
acc tgg atg ctg gtc ttc ttt gtt tcc gag ctc gtc ggt gcc ttc gga Thr Ser Met Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly Phe Gly 580 585 590			1776
att gcc atc gtc gtc ttc atg aac cac tac cct ctg gag aag atc ggt Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile Gly 595 600 605			1824
gat tcc gtc tgg gac gga cat gcc ttc tct gtc ggt cag atc cat gag Asp Ser Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile His Glu 610 615 620			1872
acc atg aac att cga cga gcc atc att act gac tgg ttc ttt gga gcc Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe Gly Gly 625 630 635			1920
ctg aac tac cag atc gag cac cat ctc tgg ccc acc ctg cct cga cac Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg His 645 650 655			1968
aac ctc act gcc gtt tcc tac cag gtc gaa cag ctg tgc cag aag cac Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln Lys His 660 665 670			2016
aac ctc ccc tac cga aac cct ctg ccc cat gaa ggt ctc gtc atc ctg Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val Ile Leu 675 680 685			2064
ctc cga tac ctg tcc cag ttc gct cga atg gcc gag aag cag ccc ggt Leu Arg Tyr Leu Ser Gln Phe Ala Arg Met Ala Glu Lys Gln Pro Gly 690 695 700			2112
gcc aag gct cag taa Ala Lys Ala Gln 705			2127
<210> 13			
<211> 708			
<212> PRT			
<213> Artificial Sequence			
<220>			
<223> Synthetic Construct			
<400> 13			
Met Ala Ala Val Ile Glu Val Ala Asn Glu Phe Val Ala Ile Thr Ala 1 5 10 15			
Glu Thr Leu Pro Lys Val Asp Tyr Gln Arg Leu Trp Arg Asp Ile Tyr 20 25 30			
Ser Cys Glu Leu Leu Tyr Phe Ser Ile Ala Phe Val Ile Leu Lys Phe 35 40 45			
Thr Leu Gly Glu Leu Ser Asp Ser Gly Lys Lys Ile Leu Arg Val Leu 50 55 60			
Phe Lys Trp Tyr Asn Leu Phe Met Ser Val Phe Ser Leu Val Ser Phe 65 70 75 80			
Leu Cys Met Gly Tyr Ala Ile Tyr Thr Val Gly Leu Tyr Ser Asn Glu 85 90 95			
Cys Asp Arg Ala Phe Asp Asn Ser Leu Phe Arg Phe Ala Thr Lys Val 100 105 110			
Phe Tyr Tyr Ser Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro 115 120 125			
Leu Met Ala Lys Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly 130 135 140			
Ala Pro Met Asp Met Trp Leu Phe Val Gln Tyr Ser Gly Glu Ser Ile 145 150 155 160			
Trp Ile Phe Val Phe Leu Asn Gly Phe Ile His Phe Val Met Tyr Gly			

165 170 175

Tyr Tyr Trp Thr Arg Leu Met Lys Phe Asn Phe Pro Met Pro Lys Gln
180 185 190

Leu Ile Thr Ala Met Gln Ile Thr Gln Phe Asn Val Gly Phe Tyr Leu
195 200 205

Val Trp Trp Tyr Lys Asp Ile Pro Cys Tyr Arg Lys Asp Pro Met Arg
210 215 220

Met Leu Ala Trp Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu
225 230 235 240

Leu Phe Ile Asn Phe Phe Val Lys Ser Tyr Val Phe Pro Lys Pro Lys
245 250 255

Thr Ala Asp Lys Lys Val Gln Gly Ala Gly Pro Ala Arg Pro Ala Gly
260 265 270

Leu Pro Pro Ala Thr Tyr Tyr Asp Ser Leu Ala Val Met Gly Ser Val
275 280 285

Lys Ala Ser Arg Gln Ala Leu Pro Leu Val Ile Asp Gly Lys Val Tyr
290 295 300

Asp Val Ser Ala Trp Val Asn Phe His Pro Gly Gly Ala Glu Ile Ile
305 310 315 320

Glu Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val Met His
325 330 335

Ser Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile Asn Gln
340 345 350

Ala Ser Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln Glu Asp
355 360 365

Phe Arg Lys Leu Arg Glu Glu Leu Ile Ala Thr Gly Met Phe Asp Ala
370 375 380

Ser Pro Leu Trp Tyr Ser Tyr Lys Ile Leu Thr Thr Leu Gly Leu Gly
385 390 395 400

Val Leu Ala Phe Phe Met Leu Val Gln Tyr His Leu Tyr Phe Ile Gly
405 410 415

Ala Leu Val Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu Ser His
420 425 430

Asp Ile Cys His His Gln Thr Phe Lys Asn Arg Asn Trp Asn Asn Val
435 440 445

Leu Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val Thr Trp
450 455 460

Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val Gln Gly
465 470 475 480

His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser Glu Asp
485 490 495

Asp Val Thr Arg Ala Ser Pro Ile Ser Arg Lys Leu Ile Gln Phe Gln
500 505 510

Gln Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile Trp Cys
515 520 525

Phe Gln Ser Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp Asn Gln
530 535 540

Phe Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala Leu His
545 550 555 560

Trp Thr Leu Lys Thr Leu Phe His Leu Phe Phe Met Pro Ser Ile Leu
565 570 575

Thr Ser Met Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly Phe Gly
580 585 590

Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile Gly
595 600 605

Asp Ser Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile His Glu
610 615 620

Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe Gly Gly
625 630 635 640

Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg His
645 650 655

Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln Lys His
660 665 670

Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val Ile Leu
675 680 685

Leu Arg Tyr Leu Ser Gln Phe Ala Arg Met Ala Glu Lys Gln Pro Gly
690 695 700

Ala Lys Ala Gln
705

<210> 14
<211> 1860
<212> DNA
<213> Saccharomyces cerevisiae

<220>
<221> CDS
<222> (1)..(1860)
<223> GenBank Accession No. NP_014818; "YOR175C"

<300>
<302> Genes encoding a novel type of lysophosphatidylcholine acyltransferases and their use to increase triacylglycerol production and/or modify fatty acid composition

<310> US-2008-0145867-A1
<311> 2007-06-15

<312> 2008-06-19

<313> (1).(1860)

<300>

<302> Genes encoding a novel type of lysophosphatidylcholine acyltransferases and their use to increase triacylglycerol production and/or modify fatty acid composition

<310> WO 2008/076377

<311> 2007-12-13

<312> 2008-06-26

<313> (1).(1860)

<300>

<302> USE OF A CLASS OF GENES ENCODING LYSOPHOSPHOLIPID ACYL TRANSFERASES FOR APPLICATION IN AGRICULTURE, BIOTECHNOLOGY AND MEDICINE

<310> WO 2009/001315

<311> 2008-06-25

<312> 2008-12-31

<313> (1).(1860)

<400> 14

atg tac aat cct gtg gao got gtt tta aca aag ata att acc aac tat 48
Met Tyr Asn Pro Val Asp Ala Val Leu Thr Lys Ile Ile Thr Asn Tyr
1 5 10 15

ggg att gat agt ttt aca ctg cga tat gct atc tgc tta ttg gga tog 96
Gly Ile Asp Ser Phe Thr Leu Arg Tyr Ala Ile Cys Leu Leu Gly Ser
20 25 30

ttc coa ctg aat got att ttg aag aga att ccc gag aag cgt ata ggt 144
Phe Pro Leu Asn Ala Ile Leu Lys Arg Ile Pro Glu Lys Arg Ile Gly
35 40 45

tta aaa tgt tgt ttt atc att tct atg tog atg ttt tac tta tto ggt 192
Leu Lys Cys Cys Phe Ile Ile Ser Met Ser Met Phe Tyr Leu Phe Gly
50 55 60

gtg ctg aat cta gta agt gga ttc agg acc ctg ttt att agt acc atg 240

Val Leu Asn Leu Val Ser Gly Phe Arg Thr Leu Phe Ile Ser Thr Met 65
70 75 80

ttt act tac ttg atc tca aga ttt taa cgt tcc aag ttt atg coa oac 288
Phe Thr Tyr Leu Ile Ser Arg Phe Tyr Arg Ser Lys Phe Met Pro His
85 90 95

ttg aat ttc atg ttt gtt atg ggt cat ttg gca ata aat cat ata cac 336
Leu Asn Phe Met Phe Val Met Gly His Leu Ala Ile Asn His Ile His
100 105 110

gac coa ttc ctt aac gaa cag act coa act acc gtt gac att aca agt 384
Ala Gln Phe Leu Asn Glu Gln Thr Gln Thr Thr Val Asp Ile Thr Ser
115 120 125

tca coa atg gtt tta gcc atg aaa cta act tct ttt gca tgg tog tac 432
Ser Gln Met Val Leu Ala Met Lys Leu Thr Ser Phe Ala Trp Ser Tyr
130 135 140

tat gat ggt tca tgc act agc gaa agc gat ttc aaa gat ttg act gag 480
Tyr Asp Gly Ser Cys Thr Ser Glu Ser Asp Phe Lys Asp Leu Thr Glu
145 150 155 160

cat coa aaa tct cgt gct gtc aga ggt cat coa ccc tta tta aag ttc 528
His Gln Lys Ser Arg Ala Val Arg Gly His Pro Pro Leu Leu Lys Phe
165 170 175

ctg coa tat gca ttt ttc tat tca acg ttg cta act ggc coa agt ttc 576
Leu Ala Tyr Ala Phe Phe Tyr Ser Thr Leu Leu Thr Gly Pro Ser Phe
180 185 190

gat tat gac gat ttt gac agc tgg ttg aat tgt gag atg ttc cgt gac 624
Asp Tyr Ala Asp Phe Asp Ser Trp Leu Asn Cys Glu Met Phe Arg Asp
195 200 205

ttg cct gaa agc aaa aag cct atg aga aga cac ccc cct ggt gaa aga 672
Leu Pro Glu Ser Lys Lys Pro Met Arg Arg His His Pro Gly Glu Arg
210 215 220

aga cag att coa aag aat ggt aaa ctt gca tta tgg aaa gtt gtt coa 720
Arg Gln Ile Pro Lys Asn Gly Lys Leu Ala Leu Trp Lys Val Val Gln
225 230 235 240

ggt ctt gct tgg atg att tta agt aca cta gga atg aag cac ttc ccc 768
Gly Leu Ala Trp Met Ile Leu Ser Thr Leu Gly Met Lys His Phe Pro
245 250 255

gta aaa tac gtt ttg gac aaa gat ggc ttc coa acg aga tot ttt ata 816
Val Lys Tyr Val Leu Asp Lys Asp Gly Phe Pro Thr Arg Ser Phe Ile
260 265 270

ttc aga atc cat tac tta ttc ttg ctt ggt ttc atc cat aga ttc aag 864
Phe Arg Ile His Tyr Leu Phe Leu Leu Gly Phe Ile His Arg Phe Lys
275 280 285

tac tac cct gcc tgg act att tgg gaa gga tct tgt att ttg tgc ggt 912
Tyr Tyr Ala Ala Trp Thr Ile Ser Glu Gly Ser Cys Ile Leu Cys Gly
290 295 300

ttg ggt tat aat ggt tat gat tca aag aca coa aag atc aga tgg gat 960
Leu Gly Tyr Asn Gly Tyr Asp Ser Lys Thr Gln Lys Ile Arg Trp Asp
305 310 315 320

cgt gtc aga aat att gac att tgg acc gta gaa acg ggc cag aat acg 1008
Arg Val Arg Asn Ile Asp Ile Trp Thr Val Glu Thr Ala Gln Asn Thr

```

325          330          335
cgt gaa atg ttg gaa gca tgg aat atg aat act aac aag tgg cta aaa 1056
Arg Glu Met Leu Glu Ala Trp Asn Met Asn Thr Asn Lys Trp Leu Lys
340          345          350

tac tct gtt tat tta cgt gtc aca aag aag gtc aaa aaa cct ggt ttc 1104
Tyr Ser Val Tyr Leu Arg Val Thr Lys Lys Gly Lys Lys Pro Gly Phe
355          360          365

cgc tca act ttg ttt act ttc cta act tcc gca ttt tgg cat ggt acc 1152
Arg Ser Thr Leu Phe Thr Phe Leu Thr Ser Ala Phe Trp His Gly Thr
370          375          380

aga cct ggg tac tat ctg act ttt gcg aca ggg gct tlg tac cca aca 1200
Arg Pro Gly Tyr Tyr Leu Thr Phe Ala Thr Gly Ala Leu Tyr Gln Thr
385          390          400

tgt ggt aaa atc tac aga cgc aat ttt aga cca att ttc ttg cca gaa 1248
Cys Gly Lys Ile Tyr Arg Arg Asn Phe Arg Pro Ile Phe Leu Arg Glu
405          410          415

gat ggt gtc act cct ttg cct tct aaa aaa atc tac gat tta gtt ggc 1296
Asp Gly Val Thr Pro Leu Pro Ser Lys Lys Ile Tyr Asp Leu Val Gly
420          425          430

ata tat gca att aaa cta gca ttt ggt tac atg gtg cca cca ttt att 1344
Ile Tyr Ala Ile Lys Leu Ala Phe Gly Tyr Met Val Gln Pro Phe Ile
435          440          445

atc ctt gat ttg aag cca tct tta atg gta tgg ggc tct gtt tat ttc 1392
Ile Leu Asp Leu Lys Pro Ser Leu Met Val Trp Gly Ser Val Tyr Phe
450          455          460

tat gtt cat att att gtt gct ttc tca ttt ttc cta ttc aga gga cca 1440
Tyr Val His Ile Ile Val Ala Phe Ser Phe Phe Leu Phe Arg Gly Pro
465          470          475

tat gct aaa cca gtt act gaa ttt ttt ana tcc aaa cca cct aaa gaa 1488
Tyr Ala Lys Gln Val Thr Glu Phe Phe Lys Ser Lys Gln Pro Lys Glu
485          490          495

ata ttc att aga aaa cca aag aag ttg gaa aaa gat att tct gca agc 1536
Ile Phe Ile Arg Lys Gln Lys Lys Leu Glu Lys Asp Ile Ser Ala Ser
500          505          510

tct cca aac ttg ggt ggt ata ttg aag gca aag att gaa cat gaa aag 1584
Ser Pro Asn Leu Gly Gly Ile Leu Lys Ala Lys Ile Glu His Glu Lys
515          520          525

gga aag aca gca gaa gaa gaa gaa atg aac tta ggt att cca cca att 1632
Gly Lys Thr Ala Glu Glu Glu Glu Met Asn Leu Gly Ile Pro Pro Ile
530          535          540

gag tta gaa aag tgg gac aat gct aag gaa gat tgg gaa gat ttc tgc 1680
Glu Leu Glu Lys Trp Asp Asn Ala Lys Glu Asp Trp Glu Asp Phe Cys
545          550          555          560

aaa gat tac aaa gaa tgg aga aat aaa aat ggt ctt gaa ala gaa gaq 1728
Lys Asp Tyr Lys Lys Glu Trp Arg Asn Lys Asn Gly Leu Glu Ile Glu Glu
565          570          575

gaa aac ctt tct aaa gct ttt gaa aga ttc aag cag gaa ttt tct aac 1776
Glu Asn Leu Ser Lys Ala Phe Glu Arg Phe Lys Gln Glu Phe Ser Asn
580          585          590

gct gca agt gga tca ggt gaa cgt gtg aga aaa atg agt ttt agt ggt 1824
Ala Ala Ser Gly Ser Gly Glu Arg Val Arg Lys Met Ser Phe Ser Gly
595          600          605

tac tca cca aag cct att tca aaa aag gaa gag tag 1860
Tyr Ser Pro Lys Pro Ile Ser Lys Lys Glu Glu
610          615

<210> 15
<211> 619
<212> PRT
<213> Saccharomyces cerevisiae

<400> 15
Met Tyr Asn Pro Val Asp Ala Val Leu Thr Lys Ile Ile Thr Asn Tyr
1          5          10          15

Gly Ile Asp Ser Phe Thr Leu Arg Tyr Ala Ile Cys Leu Leu Gly Ser
20          25          30

Phe Pro Leu Asn Ala Ile Leu Lys Arg Ile Pro Glu Lys Arg Ile Gly
35          40          45

Leu Lys Cys Cys Phe Ile Ile Ser Met Ser Met Phe Tyr Leu Phe Gly
50          55          60

Val Leu Asn Leu Val Ser Gly Phe Arg Thr Leu Phe Ile Ser Thr Met
65          70          75          80

Phe Thr Tyr Leu Ile Ser Arg Phe Tyr Arg Ser Lys Phe Met Pro His
85          90          95

Leu Asn Phe Met Phe Val Met Gly His Leu Ala Ile Asn His Ile His
100         105         110

Ala Gln Phe Leu Asn Glu Gln Thr Gln Thr Thr Val Asp Ile Thr Ser
115         120         125

Ser Gln Met Val Leu Ala Met Lys Leu Thr Ser Phe Ala Trp Ser Tyr
130         135         140

Tyr Asp Gly Ser Cys Thr Ser Glu Ser Asp Phe Lys Asp Leu Thr Glu
145         150         155         160

His Gln Lys Ser Arg Ala Val Arg Gly His Pro Pro Leu Leu Lys Phe
165         170         175

Leu Ala Tyr Ala Phe Phe Tyr Ser Thr Leu Leu Thr Gly Pro Ser Phe
180         185         190

```

```

Asp Tyr Ala Asp Phe Asp Ser Trp Leu Asn Cys Glu Met Phe Arg Asp
195          200          205

Leu Pro Glu Ser Lys Lys Pro Met Arg Arg His His Pro Gly Glu Arg
210          215          220

Arg Gln Ile Pro Lys Asn Gly Lys Leu Ala Leu Trp Lys Val Val Gln
225          230          235          240

Gly Leu Ala Trp Met Ile Leu Ser Thr Leu Gly Met Lys His Phe Pro
245          250          255

Val Lys Tyr Val Leu Asp Lys Asp Gly Phe Pro Thr Arg Ser Phe Ile
260          265          270

Phe Arg Ile His Tyr Leu Phe Leu Leu Gly Phe Ile His Arg Phe Lys
275          280          285

Tyr Tyr Ala Ala Trp Thr Ile Ser Glu Gly Ser Cys Ile Leu Cys Gly
290          295          300

Leu Gly Tyr Asn Gly Tyr Asp Ser Lys Thr Gln Lys Ile Arg Trp Asp
305          310          315          320

Arg Val Arg Asn Ile Asp Ile Trp Thr Val Glu Thr Ala Gln Asn Thr
325          330          335

Arg Glu Met Leu Glu Ala Trp Asn Met Asn Thr Asn Lys Trp Leu Lys
340          345          350

Tyr Ser Val Tyr Leu Arg Val Thr Lys Lys Gly Lys Lys Pro Gly Phe
355          360          365

Arg Ser Thr Leu Phe Thr Phe Leu Thr Ser Ala Phe Trp His Gly Thr
370          375          380

Arg Pro Gly Tyr Tyr Leu Thr Phe Ala Thr Gly Ala Leu Tyr Gln Thr
385          390          395          400

Cys Gly Lys Ile Tyr Arg Arg Asn Phe Arg Pro Ile Phe Leu Arg Glu
405          410          415

Asp Gly Val Thr Pro Leu Pro Ser Lys Lys Ile Tyr Asp Leu Val Gly
420          425          430

Ile Tyr Ala Ile Lys Leu Ala Phe Gly Tyr Met Val Gln Pro Phe Ile
435          440          445

Ile Leu Asp Leu Lys Pro Ser Leu Met Val Trp Gly Ser Val Tyr Phe
450          455          460

Tyr Val His Ile Ile Val Ala Phe Ser Phe Phe Leu Phe Arg Gly Pro
465          470          475          480

Tyr Ala Lys Gln Val Thr Glu Phe Phe Lys Ser Lys Gln Pro Lys Glu
485          490          495

Ile Phe Ile Arg Lys Gln Lys Lys Leu Glu Lys Asp Ile Ser Ala Ser
500          505          510

Ser Pro Asn Leu Gly Gly Ile Leu Lys Ala Lys Ile Glu His Glu Lys
515          520          525

Gly Lys Thr Ala Glu Glu Glu Met Asn Leu Gly Ile Pro Pro Ile
530          535          540

Glu Leu Glu Lys Trp Asp Asn Ala Lys Glu Asp Trp Glu Asp Phe Cys
545          550          555          560

Lys Asp Tyr Lys Glu Trp Arg Asn Lys Asn Gly Leu Glu Ile Glu Glu
565          570          575

Glu Asn Leu Ser Lys Ala Phe Glu Arg Phe Lys Gln Glu Phe Ser Asn
580          585          590

Ala Ala Ser Gly Ser Gly Glu Arg Val Arg Lys Met Ser Phe Ser Gly
595          600          605

Tyr Ser Pro Lys Pro Ile Ser Lys Lys Glu Glu
610          615

<210> 16
<211> 1539
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> CDS
<222> (1)..(1539)
<223> GenBank Accession No. XP_505624; "YALI0F19514p"

<400> 16
atg gcc ttt oca tgg gca gat aag tgg gca gcc gat gcg tot gca tct      48
Met Ala Phe Pro Trp Ala Asp Lys Trp Ala Ala Asp Ala Ser Ala Ser
1          5          10          15

aca ggg ctg cct ccg gac ctc ctc aag att gca ttc act ctg gtc atg      96
Thr Gly Leu Pro Pro Asp Leu Leu Lys Ile Ala Phe Thr Leu Val Met
20          25          30

tct tat ccg ctg agt tct ctc atg aaa cgg ctg cca gat gac gcc aaa      144
Ser Tyr Pro Leu Ser Ser Leu Met Lys Arg Leu Pro Asp Asp Ala Lys

```

35	40	45	
aac ctc aag atc atc tat atc atc tcc gtg tcc atc ttc tac atg gtg			192
Asn Leu Lys Ile Ile Tyr Ile Ile Ser Val Ser Ile Phe Tyr Met Val			
50 55 60			
ggt gtc ttc tcc ctc tat gcc gga gct gcc act ctg ctc ttc tcc tca			240
Gly Val Phe Ser Leu Tyr Gly Gly Ala Ala Thr Leu Leu Phe Ser Ser			
65 70 75 80			
atg ggt acc ttc ttc atc acc cca tgg aag agc cct tac atg ccc tgg			288
Met Gly Thr Phe Phe Ile Thr Gln Trp Lys Ser Pro Tyr Met Pro Trp			
85 90 95			
gtc aat ttt ggt ttt gtc atg acc cat ctc ttc gtc aat cnc ctg cgt			336
Val Asn Phe Gly Phe Val Met Thr His Leu Phe Val Asn His Leu Arg			
100 105 110			
tcc caq ttt ttc ccc gaa aca tac gac ccc aat gtc att gac atc acc			384
Ser Gln Phe Phe Pro Glu Thr Tyr Asp Pro Asn Val Ile Asp Ile Thr			
115 120 125			
gga gca cag atg gtt ctg tgt atg aag cta tcc tot ttt gga tgg aac			432
Gly Ala Gln Met Val Leu Cys Met Lys Leu Ser Ser Phe Gly Trp Asn			
130 135 140			
gtc tac gat gga tgg cag att gag aag ggt gag cag ctc agc gag ttc			480
Val Tyr Asp Gly Trp Gln Ile Glu Lys Gly Glu Gln Leu Ser Glu Phe			
145 150 155 160			
cag act aaa agy gct gtt ctc aag cnc ccc agt ctt atg gac ttc cta			528
Gln Thr Lys Arg Ala Val Leu Lys His Pro Ser Leu Met Asp Phe Leu			
165 170 175			
gct ttt gtg ttc tcc ttc cct tcc att ctg aca ggt cct tot tac gac			576
Ala Phe Val Phe Tyr Phe Pro Ser Ile Leu Thr Gly Pro Ser Tyr Asp			
180 185 190			
tat atg cag ttc cat aac tgg ctc gat ctc agc ctg ttc aag gag ctg			624
Tyr Met Glu Phe His Asn Trp Leu Asp Leu Ser Leu Phe Lys Glu Leu			
195 200 205			
gag aaa gat aag gac ccc aag cga gct gct cga cga aag cga cnc aag			672
Glu Lys Asp Lys Asp Pro Lys Arg Ala Ala Arg Lys Arg His Lys			
210 215 220			
atc ccc cga tct gga atc gct gct tcc aag aaa ctc gcc gct ggt atc			720
Ile Pro Arg Ser Gly Ile Ala Ala Ser Lys Lys Leu Ala Ala Gly Ile			
225 230 235 240			
ttc tgg atc gtt ctg tgg acc cag gtg gac tot cga atc tcc acc gcc			768
Phe Trp Ile Val Leu Trp Thr Gln Val Asp Ser Arg Ile Ser Thr Ala			
245 250 255			
tac gct tac tca gac gca ttc acc aag gag cnc aac atc ttt gga cga			816
Tyr Ala Tyr Ser Asp Ala Phe Thr Lys Glu His Asn Ile Phe Gly Arg			
260 265 270			
att gtg tac ctc tac atg ctc gct ttc atg tac cga ctc aag tac tac			864
Ile Val Tyr Leu Tyr Met Leu Gly Phe Met Tyr Arg Leu Lys Tyr Tyr			
275 280 285			
gga gcc tgg tcc att tcc gag gga gcc tgc atc ttg tct gcc ctc gga			912
Gly Ala Trp Ser Ile Ser Glu Gly Ala Cys Ile Leu Ser Gly Leu Gly			
290 295 300			
ttc cat gcc gtg gac ccc aaa act gcc aag tac aag tgg gac cgt gtc			960
Phe His Gly Val Asp Pro Lys Thr Gly Lys Tyr Lys Trp Asp Arg Val			
305 310 315 320			
cag aac gtg gac ccg tgg gga ttc gaa act ggt cna aac cna aag gct			1008
Gln Asn Val Asp Pro Trp Gly Phe Glu Thr Gly Gln Asn Thr Lys Ala			
325 330 335			
ctg ctg gag gcc tgg aac cag aac act aac aag tgg cta cga aac tat			1056
Leu Leu Glu Ala Trp Asn Gln Asn Thr Asn Lys Trp Leu Arg Asn Tyr			
340 345 350			
gtg tac ctc cga gtg gtg ccc aaa gcc cna aag cct gga ttc cga gcc			1104
Val Tyr Leu Arg Val Val Pro Lys Gly Gln Lys Pro Gly Phe Arg Ala			
355 360 365			
act atc ttc aca ttt gtg gtt tcc gcc ttc tgg cat gga act cga cct			1152
Thr Ile Phe Thr Phe Val Val Ser Ala Phe Trp His Gly Thr Arg Pro			
370 375 380			
gcc tac tat ctc acc ttt gtg acc gct gcc atg tac cag tot gtt ggt			1200
Gly Tyr Tyr Leu Thr Phe Val Thr Ala Ala Met Tyr Gln Ser Val Gly			
385 390 395 400			
aag ttc ttc cga cga tac ctg cga ccc ttc ttc atg gag tot gat gga			1248
Lys Phe Phe Arg Arg Tyr Leu Arg Pro Phe Phe Met Glu Ser Asp Gly			
405 410 415			
aag act gcc ggt ccc tat aag atc tac tac gac att gtg tgt tgg atc			1296
Lys Thr Ala Gly Pro Tyr Lys Ile Tyr Tyr Asp Ile Val Cys Trp Ile			
420 425 430			
ggt gtc cna acc gca ttt gga tac gct acc cag tcc ttt atg att cta			1344
Val Val Gln Thr Ala Phe Gly Tyr Ala Thr Gln Ser Phe Met Ile Leu			
435 440 445			
gac ttc tgg ctg tcc ctc aag tgt tgg aag aac tcc tgg ttc ctg tac			1392
Asp Phe Trp Leu Ser Leu Lys Cys Trp Lys Asn Ser Trp Phe Leu Tyr			
450 455 460			
cac att gct ctg gcc gca atc ttt gca att tot agc ccc tac aag gca			1440
His Ile Ala Leu Gly Ala Ile Phe Ala Ile Ser Ser Pro Tyr Lys Ala			
465 470 475 480			
tgg gcy att ccc aag atc aag aaa aag cag gct gga gcc gtc act gac			1488
Trp Ala Ile Pro Lys Ile Lys Lys Lys Gln Ala Gly Ala Val Thr Asp			
485 490 495			
aag aag gac gcc aag gag gag gtg aag aag gac acc atc aag acc aag			1536
Lys Lys Asp Ala Lys Glu Glu Val Lys Lys Asp Thr Ile Lys Thr Lys			
500 505 510			
taa			1539
<210> 17			
<211> 512			
<212> PRT			
<213> Yarrowia lipolytica			
<400> 17			
Met Ala Phe Pro Trp Ala Asp Lys Trp Ala Ala Asp Ala Ser Ala Ser			
1 5 10 15			

Thr Gly Leu Pro Pro Asp Leu Leu Lys Ile Ala Phe Thr Leu Val Met
 20 25 30
 Ser Tyr Pro Leu Ser Ser Leu Met Lys Arg Leu Pro Asp Asp Ala Lys
 35 40 45
 Asn Leu Lys Ile Ile Tyr Ile Ile Ser Val Ser Ile Phe Tyr Met Val
 50 55 60
 Gly Val Phe Ser Leu Tyr Gly Gly Ala Ala Thr Leu Leu Phe Ser Ser
 65 70 75 80
 Met Gly Thr Phe Phe Ile Thr Gln Trp Lys Ser Pro Tyr Met Pro Trp
 85 90 95
 Val Asn Phe Gly Phe Val Met Thr His Leu Phe Val Asn His Leu Arg
 100 105 110
 Ser Gln Phe Phe Pro Glu Thr Tyr Asp Pro Asn Val Ile Asp Ile Thr
 115 120 125
 Gly Ala Gln Met Val Leu Cys Met Lys Leu Ser Ser Phe Gly Trp Asn
 130 135 140
 Val Tyr Asp Gly Trp Gln Ile Glu Lys Gly Glu Gln Leu Ser Glu Phe
 145 150 155 160
 Gln Thr Lys Arg Ala Val Leu Lys His Pro Ser Leu Met Asp Phe Leu
 165 170 175
 Ala Phe Val Phe Tyr Phe Pro Ser Ile Leu Thr Gly Pro Ser Tyr Asp
 180 185 190
 Tyr Met Glu Phe His Asn Trp Leu Asp Leu Ser Leu Phe Lys Glu Leu
 195 200 205
 Glu Lys Asp Lys Asp Pro Lys Arg Ala Ala Arg Arg Lys Arg His Lys
 210 215 220
 Ile Pro Arg Ser Gly Ile Ala Ala Ser Lys Lys Leu Ala Ala Gly Ile
 225 230 235 240
 Phe Trp Ile Val Leu Trp Thr Gln Val Asp Ser Arg Ile Ser Thr Ala
 245 250 255
 Tyr Ala Tyr Ser Asp Ala Phe Thr Lys Glu His Asn Ile Phe Gly Arg
 260 265 270
 Ile Val Tyr Leu Tyr Met Leu Gly Phe Met Tyr Arg Leu Lys Tyr Tyr
 275 280 285
 Gly Ala Trp Ser Ile Ser Glu Gly Ala Cys Ile Leu Ser Gly Leu Gly
 290 295 300
 Phe His Gly Val Asp Pro Lys Thr Gly Lys Tyr Lys Trp Asp Arg Val
 305 310 315 320
 Gln Asn Val Asp Pro Trp Gly Phe Glu Thr Gly Gln Asn Thr Lys Ala
 325 330 335
 Leu Leu Glu Ala Trp Asn Gln Asn Thr Asn Lys Trp Leu Arg Asn Tyr
 340 345 350
 Val Tyr Leu Arg Val Val Pro Lys Gly Gln Lys Pro Gly Phe Arg Ala
 355 360 365
 Thr Ile Phe Thr Phe Val Val Ser Ala Phe Trp His Gly Thr Arg Pro
 370 375 380
 Gly Tyr Tyr Leu Thr Phe Val Thr Ala Ala Met Tyr Gln Ser Val Gly
 385 390 395 400
 Lys Phe Phe Arg Arg Tyr Leu Arg Pro Phe Phe Met Glu Ser Asp Gly
 405 410 415
 Lys Thr Ala Gly Pro Tyr Lys Ile Tyr Tyr Asp Ile Val Cys Trp Ile
 420 425 430
 Val Val Gln Thr Ala Phe Gly Tyr Ala Thr Gln Ser Phe Met Ile Leu
 435 440 445
 Asp Phe Trp Leu Ser Leu Lys Cys Trp Lys Asn Ser Trp Phe Leu Tyr
 450 455 460
 His Ile Ala Leu Gly Ala Ile Phe Ala Ile Ser Ser Pro Tyr Lys Ala
 465 470 475 480
 Trp Ala Ile Pro Lys Ile Lys Lys Lys Gln Ala Gly Ala Val Thr Asp
 485 490 495
 Lys Lys Asp Ala Lys Glu Glu Val Lys Lys Asp Thr Ile Lys Thr Lys
 500 505 510
 <210> 18
 <211> 7
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> membrane bound O-acyltransferase motif
 <220>
 <221> MISC_FEATURE
 <222> (2)..(2)
 <223> Xaa = Val [V] or Ile [I]
 <220>
 <221> misc_feature
 <222> (4)..(5)
 <223> Xaa can be any naturally occurring amino acid
 <400> 18
 Met Xaa Leu Xaa Xaa Lys Leu
 1 5

<210> 19
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> membrane bound O-acyltransferase motif

<220>
<221> misc_feature
<222> (2)..(2)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (6)..(7)
<223> Xaa can be any naturally occurring amino acid

<400> 19
Arg Xaa Lys Tyr Tyr Xaa Xaa Trp
1 5

<210> 20
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> membrane bound O-acyltransferase motif

<220>
<221> misc_feature
<222> (3)..(3)
<223> Xaa can be any naturally occurring amino acid

<400> 20
Ser Ala Xaa Trp His Gly
1 5

<210> 21
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> motif

<220>
<221> misc_feature
<222> (2)..(12)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (15)..(16)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> MISC_FEATURE
<222> (17)..(17)
<223> Xaa = Thr [T] or Val [V]

<220>
<221> misc_feature
<222> (18)..(19)
<223> Xaa can be any naturally occurring amino acid

<300>
<302> GENES ENCODING A NOVEL TYPE OF LYSOPHOPHATIDYLCHOLINE ACYLTRANSFERASES AND THEIR USE TO INCREASE TRIACYLGLYCEROL PRODUCTION AND/OR MODIFY FATTY ACID COMPOSITION

<310> U.S. Patent Publication No. 2008-0145867-A1
<311> 2007-06-15
<312> 2008-06-19
<313> (1)..(15)

<400> 21
Glu Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Trp Asn Xaa Xaa
1 5 10 15

Xaa Xaa Xaa Trp
20

<210> 22
<211> 1870
<212> DNA
<213> Saccharomyces cerevisiae

<220>
<221> CDS
<222> (3)..(1862)
<223> synthetic Ale1 (codon-optimized for Yarrowia lipolytica)

<400> 22
ac atg tac aac ooo gtg gac gca gtg ttg act aag atb atb aca aac
Met Tyr Asn Pro Val Asp Ala Val Leu Thr Lys Ile Ile Thr Asn
1 5 10 15
47

tac gga att gat tct ttt acc ctg cga tat gcc att tgt ctg ttg gga 95
Tyr Gly Ile Asp Ser Phe Thr Leu Arg Tyr Ala Ile Cys Leu Leu Gly
20 25 30

tct ttt out ctt aac gct att ctg aag cgg att cct gaa aag cga atc 143
Ser Phe Pro Leu Asn Ala Ile Leu Lys Arg Ile Pro Glu Lys Arg Ile
35 40 45

ggc ctg aag tgt tgt ttt atc att tct atg tcc atg ttt tat ctc ttc 191
Gly Leu Lys Cys Cys Phe Ile Ile Sar Met Sar Met Thr Tyr Leu Phe
50 55 60

ggc gtt ctg aat ctc gtc agc gga ttt cga acc ctc ttc att tcc aca 239
Gly Val Leu Asn Leu Val Ser Gly Phe Arg Thr Leu Phe Ile Ser Thr
65 70 75

atg ttc aca tac ctt atc tct cgg ttc tac cga tcc aag ttt atg ccc 287
Met Phe Thr Tyr Leu Ile Ser Arg Phe Tyr Arg Ser Lys Phe Met Pro
80 85 90 95

cat ctc aac ttc atg ttc gtc atg ggc cac ttg gct atc aac cac att 335
His Leu Asn Phe Met Phe Val Met Gly His Leu Ala Ile Asn His Ile
100 105 110

cat gct cag ttc ctg aac gaa aca act cca acg acc gtc gat att aca 383
His Ala Gln Phe Leu Asn Glu Gln Thr Gln Thr Thr Val Asp Ile Thr
115 120 125

tcc tgg cag atg gtc ctg gct atg aag ctg aca agc ttt gcc tgg tct 431
Ser Ser Gln Met Val Leu Ala Met Lys Leu Thr Ser Phe Ala Trp Ser
130 135 140

tac tat gac ggt tgg tgt agc ago gag tcc gac ttc aag gac ctt acc 479
Tyr Tyr Asp Gly Ser Cys Thr Ser Glu Ser Asp Phe Lys Asp Leu Thr
145 150 155

gaa cac cag aag tcc cga gcc gtc cga ggc cat cct ccc ctt ctg aaa 527
Glu His Gln Lys Ser Arg Ala Val Arg Gly His Pro Pro Leu Leu Lys
160 165 170 175

ttt ttg gct tac gcc ttt ttc tac tct acc ctt ctc aac ggt ccc tcc 575
Phe Leu Ala Tyr Ala Phe Phe Tyr Ser Thr Leu Leu Phe Gly Pro Ser
180 185 190

ttc gat tac gct gat ttc gac tct tgg ctg aac tgc gaa atg ttc cgg 623
Phe Asp Tyr Ala Asp Phe Asp Ser Trp Leu Asn Cys Glu Met Phe Arg
195 200 205

gac ctt ccc gag tcc aag aca ccc atg cga cga cat cat cct ggt gag 671
Asp Leu Pro Glu Ser Lys Lys Pro Met Arg Arg His His Pro Gly Glu
210 215 220

cgg cgt cag att ccc aag aac ggc aag ctc gcc ctg tgg aag gtt gtc 719
Arg Arg Gln Ile Pro Lys Asn Gly Lys Leu Ala Leu Trp Lys Val Val
225 230 235

cag ggc ctc gcc tgg atg att ctg agc acc ttg ggt atg aag cac ttc 767
Gln Gly Leu Ala Trp Met Ile Leu Ser Thr Leu Glu Met Lys His Phe
240 245 250 255

ccc gtc aag tac atg ctg gac aag gac gaa ttt cct acc cgt tcc ttt 815
Pro Val Lys Tyr Val Leu Asp Lys Asp Gly Phe Pro Thr Arg Ser Phe
260 265 270

atc ttc cgt att cat tat ctg ttt ctg ctg gga ttc atc cac cga ttt 863

Ile Phe Arg Ile His Tyr Leu Phe Leu Leu Gly Phe Ile His Arg Phe 911
275 280 285

aag tat tac gct ggc tgg acc att ago gaa ggt tgg tgc att ctc tgt 959
Lys Tyr Tyr Ala Ala Trp Thr Ile Ser Glu Gly Ser Cys Ile Leu Cys
290 295 300

ggt ctt ggt tat aat gga tac gat tct aag acc cag aag atc cgg tgg 959
Gly Leu Gly Tyr Asn Gly Tyr Asp Ser Lys Thr Gln Lys Ile Arg Trp
305 310 315

gat cga gtc cgg aat att gat att tgg aca gtc gag aat cga aac 1007
Asp Arg Val Arg Asn Ile asp ile Trp Phe Val Glu Thr Ala Gln Asn
320 325 330 335

acc cga gag atg ctg gaa gcc tgg aac atg act act aac aca tgg ctg 1055
Thr Arg Glu Met Leu Glu Ala Trp Asn Met Asn Thr Asn Lys Trp Leu
340 345 350

aag tat agc gtc tat ctt aga gtc act aag aag ggt aag aag cca ggt 1103
Lys Tyr Ser Val Tyr Leu Arg Val Thr Lys Lys Gly Lys Pro Gly
355 360 365

ttt cga tct acc ctg ttt acc ttc ctg acc tcc gcc ttt tgg csc ggt 1151
Phe Arg Ser Thr Leu Phe Thr Phe Leu Thr Ser Ala Thr Trp His Gly
370 375 380

acc cgt cct gga tac tac cct acc ttc gaa act ggt gcc ctg tac aca 1199
Thr Arg Pro Gly Tyr Tyr Leu Thr Phe Ala Thr Gly Ala Leu Tyr Gln
385 390 395

acc tgt cga aag atc tat aga cga aac ttt cgt ccc atc ttt ctg aga 1247
Thr Cys Gly Lys Ile Tyr Arg Arg Asn Phe Arg Pro Ile Phe Leu Arg
400 405 410 415

gaa gat ggc gtc aca cct ctc cgg tcc aag aag att tac gac ctg gtc 1295
Glu Asp Gly Val Thr Pro Leu Pro Ser Lys Lys Ile Tyr Asp Leu Val
420 425 430

ggc att tac gct att aag ctg gcc ttt ggt tac atg gtt caa ccc ttc 1343
Gly Ile Tyr Ala Ile Lys Leu Ala Phe Gly Tyr Met Val Gln Pro Phe
435 440 445

att atc ctt gac ctg aag ccc tat ctt atg gtt tgg gaa tcc gtc tat 1391
Ile Ile Leu Asp Leu Lys Pro Ser Leu Met Val Trp Gly Ser Val Tyr
450 455 460

ttc tac gtc gat att att ggc gcc ttc tgg ttc ttt ctg ttc cga gga 1439
Phe Tyr Val His Ile Ile Val Ala Phe Ser Phe Phe Leu Phe Arg Gly
465 470 475

cca tac gtt aag cag gtt act gaa ttt ttc aca agc aag cca cag aag 1487
Pro Tyr Ala Lys Gln Val Thr Glu Phe Phe Lys Ser Lys Gln Pro Lys
480 485 490 495

gag atc ttc atc cga aag cag aag aag tgg gaa aca gac atc tct gcc 1535
Glu Ile Phe Ile Arg Lys Gln Lys Lys Leu Glu Lys Asp Ile Ser Ala
500 505 510

tct tcc ccc aac ctc gga ggt att ctt aag cga aca atc gaa cat gag 1583
Ser Ser Pro Asn Leu Gly Gly Ile Leu Lys Ala Lys Ile Glu His Glu
515 520 525

aag gaa aag acg gaa gag gaa gaa gag atg aac ttg gcc att cca ccc 1631
Lys Gly Lys Thr Ala Glu Glu Glu Glu Met Asn Leu Gly Ile Pro Pro
530 535 540

<210> 23

<211> 619

<212> PRT

<213> Saccharomyces cerevisiae

<400> 23

```

Met Tyr Asn Pro Val Asp Ala Val Leu Thr Lys Ile Ile Thr Asn Tyr
1      5      10      15

Gly Ile Asp Ser Phe Thr Leu Arg Tyr Ala Ile Cys Leu Leu Gly Ser
20     25     30

Phe Pro Leu Asn Ala Ile Leu Lys Arg Ile Pro Glu Lys Arg Ile Gly
35     40     45

Leu Lys Cys Cys Phe Ile Ile Ser Met Ser Met Phe Tyr Leu Phe Gly
50     55     60

Val Leu Asn Leu Val Ser Gly Phe Arg Thr Leu Phe Ile Ser Thr Met
65     70     75     80

Phe Thr Tyr Leu Ile Ser Arg Phe Tyr Arg Ser Lys Phe Met Pro His
85     90     95

Leu Asn Phe Met Phe Val Met Gly His Leu Ala Ile Asn His Ile His
100    105    110

Ala Gln Phe Leu Asn Glu Gln Thr Gln Thr Thr Val Asp Ile Thr Ser
115    120    125

Ser Gln Met Val Leu Ala Met Lys Leu Thr Ser Phe Ala Trp Ser Tyr
130    135    140

Tyr Asp Gly Ser Cys Thr Ser Glu Ser Asp Phe Lys Asp Leu Thr Glu
145    150    155    160

His Gln Lys Ser Arg Ala Val Arg Gly His Pro Pro Leu Leu Lys Phe
165    170    175

Leu Ala Tyr Ala Phe Phe Tyr Ser Thr Leu Leu Thr Gly Pro Ser Phe
180    185    190

Asp Tyr Ala Asp Phe Asp Ser Trp Leu Asn Cys Glu Met Phe Arg Asp
195    200    205

Leu Pro Glu Ser Lys Lys Pro Met Arg Arg His His Pro Gly Glu Arg
210    215    220

Arg Gln Ile Pro Lys Asn Gly Lys Leu Ala Leu Trp Lys Val Val Gln
225    230    235    240

Gly Leu Ala Trp Met Ile Leu Ser Thr Leu Gly Met Lys His Phe Pro
245    250    255

Val Lys Tyr Val Leu Asp Lys Asp Gly Phe Pro Thr Arg Ser Phe Ile
260    265    270

Phe Arg Ile His Tyr Leu Phe Leu Leu Gly Phe Ile His Arg Phe Lys
275    280    285

Tyr Tyr Ala Ala Trp Thr Ile Ser Glu Gly Ser Cys Ile Leu Cys Gly
290    295    300

Leu Gly Tyr Asn Gly Tyr Asp Ser Lys Thr Gln Lys Ile Arg Trp Asp
305    310    315    320

Arg Val Arg Asn Ile Asp Ile Trp Thr Val Glu Thr Ala Gln Asn Thr
325    330    335

Arg Glu Met Leu Glu Ala Trp Asn Met Asn Thr Asn Lys Trp Leu Lys
340    345    350

Tyr Ser Val Tyr Leu Arg Val Thr Lys Lys Gly Lys Lys Pro Gly Phe
355    360    365

Arg Ser Thr Leu Phe Thr Phe Leu Thr Ser Ala Phe Trp His Gly Thr
370    375    380

Arg Pro Gly Tyr Tyr Leu Thr Phe Ala Thr Gly Ala Leu Tyr Gln Thr
385    390    395    400

Cys Gly Lys Ile Tyr Arg Arg Asn Phe Arg Pro Ile Phe Leu Arg Glu
405    410    415

Asp Gly Val Thr Pro Leu Pro Ser Lys Lys Ile Tyr Asp Leu Val Gly
420    425    430

Ile Tyr Ala Ile Lys Leu Ala Phe Gly Tyr Met Val Gln Pro Phe Ile
435    440    445

Ile Leu Asp Leu Lys Pro Ser Leu Met Val Trp Gly Ser Val Tyr Phe
450    455    460

Tyr Val His Ile Ile Val Ala Phe Ser Phe Phe Leu Phe Arg Gly Pro
465    470    475    480

Tyr Ala Lys Gln Val Thr Glu Phe Phe Lys Ser Lys Gln Pro Lys Glu
485    490    495

Ile Phe Ile Arg Lys Gln Lys Lys Leu Glu Lys Asp Ile Ser Ala Ser
500    505    510

Ser Pro Asn Leu Gly Gly Ile Leu Lys Ala Lys Ile Glu His Glu Lys
515    520    525

Gly Lys Thr Ala Glu Glu Glu Met Asn Leu Gly Ile Pro Pro Ile
530    535    540

Glu Leu Glu Lys Trp Asp Asn Ala Lys Glu Asp Trp Glu Asp Phe Cys
545    550    555    560

Lys Asp Tyr Lys Glu Trp Arg Asn Lys Asn Gly Leu Glu Ile Glu Glu
565    570    575

Glu Asn Leu Ser Lys Ala Phe Glu Arg Phe Lys Gln Glu Phe Ser Asn
580    585    590

Ala Ala Ser Gly Ser Gly Glu Arg Val Arg Lys Met Ser Phe Ser Gly
595    600    605

Tyr Ser Pro Lys Pro Ile Ser Lys Lys Glu Glu
610    615

```


<210> 24
 <211> 849
 <212> DNA
 <213> *Caenorhabditis elegans*

 <220>
 <221> CDS
 <222> (1)..(849)
 <223> GenBank Accession No. CAA98276; "clone T06E8.1"

 <300>
 <302> METHOD FOR THE PRODUCTION OF POLYUNSATURATED FATTY ACIDS

 <310> US 2006-0168687-A1
 <311> 2004-01-29
 <312> 2006-07-27
 <313> (1)..(849)

 <300>
 <302> METHOD FOR THE PRODUCTION OF POLYUNSATURATED FATTY ACIDS

 <310> WO 2004/076617
 <311> 2004-01-29
 <312> 2004-09-10
 <313> (1)..(849)

 <400> 24
 atg gag aac ttc tgg tcg atc gtc gtg ttt ttt cta ctc tca att ctc 48
 Met Glu Asn Phe Trp Ser Ile Val Val Phe Phe Leu Leu Ser Ile Leu
 1 5 10 15
 ttc att tta tat aac ata tcg aca gta tgc cac tac tat atg cgg att 96
 Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile
 20 25 30
 tcg ttt tat tac ttc aca att tta ttg cat gga atg gaa gtt tgt gtt 144
 Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val
 35 40 45
 aca atg atc cct tct tgg cta aat ggg aag ggt gct gat tac gtg ttt 192
 Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe
 50 55 60
 cac tcg ttt ttc tat tgg tgt aaa tcg act ggt gtt cat aca aca ctc 240
 His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val
 65 70 75 80
 tat gga tat gaa aaa aca caa gtt gaa ggt ccg gct gta gtt att tgt 288
 Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys
 85 90 95
 aat cat cag agt tct ctc gac att cta tcg atg gca tca atc tgg ccg 336
 Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro
 100 105 110
 aag aat tgt gtt gta atg atg aaa cga att ctt gcc tat gtt cca ttc 384
 Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe
 115 120 125
 ttc aat ctc gga gcc tac ttt tcc aac aca atc ttc atc gat cga tat 432
 Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr
 130 135 140
 aac cgt gaa cgt gcg atg gct tca gct gat tat tgt goa tct gaa atg 480
 Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met
 145 150 155 160
 aag aac aga aat ctt aaa ctt tgg gta ttt ccg gaa gga aca aga aat 528
 Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn
 165 170 175
 cgt gaa gga ggg ttc att cca ttc aag aaa gga gca ttc aat att goa 576
 Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala
 180 185 190
 gtt cgt gcg cag att ccc att att cca gtt gta ttc tca gac tat ccg 624
 Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg
 195 200 205
 gat ttc tac tca aag cca gcc cga tat ttc aag aat gat gga gaa gtt 672
 Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val
 210 215 220
 gtt att cga gtt ctg gat gcg att cca aca aaa ggg ctc act ctt gat 720
 Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp
 225 230 235 240
 gac gtc agc gag ttg tct gat atg tgt ccg gac gtt atg ttg gca gcc 768
 Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala
 245 250 255
 tat aag gaa gtt act cta gaa gct cag caa cga aat gcg aca cgg cgt 816
 Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg
 260 265 270
 gga gaa aca aaa gac ggg aag aaa tct gag taa 849
 Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu
 275 280

 <210> 25
 <211> 282
 <212> PRT
 <213> *Caenorhabditis elegans*

 <400> 25

Met	Glu	Asn	Phe	Trp	Ser	Ile	Val	Val	Phe	Phe	Leu	Leu	Ser	Ile	Leu	
1				5				10					15			
Phe	Ile	Leu	Tyr	Asn	Ile	Ser	Thr	Val	Cys	His	Tyr	Tyr	Met	Arg	Ile	
	20					25						30				
Ser	Phe	Tyr	Tyr	Phe	Thr	Ile	Leu	Leu	His	Gly	Met	Glu	Val	Cys	Val	
	35					40					45					
Thr	Met	Ile	Pro	Ser	Trp	Leu	Asn	Gly	Lys	Gly	Ala	Asp	Tyr	Val	Phe	
	50					55					60					
His	Ser	Phe	Phe	Tyr	Trp	Cys	Lys	Trp	Thr	Gly	Val	His	Thr	Thr	Val	
65				70					75				80			
Tyr	Gly	Tyr	Glu	Lys	Thr	Gln	Val	Glu	Gly	Pro	Ala	Val	Val	Ile	Cys	
	85							90					95			
Asn	His	Gln	Ser	Ser	Leu	Asp	Ile	Leu	Ser	Met	Ala	Ser	Ile	Trp	Pro	
	100						105						110			
Lys	Asn	Cys	Val	Val	Met	Met	Lys	Arg	Ile	Leu	Ala	Tyr	Val	Pro	Phe	
	115						120					125				
Phe	Asn	Leu	Gly	Ala	Tyr	Phe	Ser	Asn	Thr	Ile	Phe	Ile	Asp	Arg	Tyr	
	130						135					140				
Asn	Arg	Glu	Arg	Ala	Met	Ala	Ser	Val	Asp	Tyr	Cys	Ala	Ser	Glu	Met	
145				150					155					160		
Lys	Asn	Arg	Asn	Leu	Lys	Leu	Trp	Val	Phe	Pro	Glu	Gly	Thr	Arg	Asn	
	165							170					175			
Arg	Glu	Gly	Gly	Phe	Ile	Pro	Phe	Lys	Lys	Gly	Ala	Phe	Asn	Ile	Ala	
	180						185						190			
Val	Arg	Ala	Gln	Ile	Pro	Ile	Ile	Pro	Val	Val	Phe	Ser	Asp	Tyr	Arg	
	195						200					205				
Asp	Phe	Tyr	Ser	Ser	Lys	Pro	Gly	Arg	Tyr	Phe	Lys	Asn	Asp	Gly	Glu	Val
	210					215						220				
Val	Ile	Arg	Val	Leu	Asp	Ala	Ile	Pro	Thr	Lys	Gly	Leu	Thr	Leu	Asp	
225					230				235					240		
Asp	Val	Ser	Glu	Leu	Ser	Asp	Met	Cys	Arg	Asp	Val	Met	Leu	Ala	Ala	
	245							250					255			
Tyr	Lys	Glu	Val	Thr	Leu	Glu	Ala	Gln	Gln	Arg	Asn	Ala	Thr	Arg	Arg	
	260						265						270			
Gly	Glu	Thr	Lys	Asp	Gly	Lys	Lys	Ser	Glu							
	275						280									
<210> 26																
<211> 859																
<212> DNA																
<213> Caenorhabditis elegans																
<220>																
<221> CDS																
<222> (3)..(851)																
<223> synthetic LPCAT (codon-optimized for Yarrowia lipolytica)																
<400> 26																
cc	atg	gag	aac	ttc	tgg	tcg	atc	gtc	gtg	ttc	ttt	ctg	ctc	tcg	att	47
	Met	Glu	Asn	Phe	Trp	Ser	Ile	Val	Val	Phe	Leu	Leu	Ser	Ile		
1				5				10					15			
ctg	ttc	atc	ctc	tac	aac	att	tgg	aca	gtc	tcg	cac	tac	tac	atg	cga	95
Leu	Phe	Ile	Leu	Tyr	Asn	Ile	Ser	Thr	Val	Cys	His	Tyr	Tyr	Met	Arg	
	20							25					30			
atc	tcg	ttc	tac	tac	ttt	acc	atc	ctg	ctt	cac	ggc	atg	gag	gtg	tcg	143
Ile	Ser	Phe	Tyr	Tyr	Phe	Thr	Ile	Leu	Leu	His	Gly	Met	Glu	Val	Cys	
	35						40					45				
gtt	acc	atg	att	ccc	tct	tgg	ctc	aac	ggc	aag	ggt	gcc	gac	tac	gtg	191
Val	Thr	Met	Ile	Pro	Ser	Trp	Leu	Asn	Gly	Lys	Gly	Ala	Asp	Tyr	Val	

50	55	60	
ttt cac tcg ttc ttc tac tgg tgc aag tgg act gga gtc cac acc act			239
Phe His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr			
65 70 75			
gtg tat ggc tac gag aag acc cag gtc gaa ggt cct gcc gtg gtc atc			287
Val Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile			
80 85 90 95			
tgc aac cat cag tcc tcg ctc gac att ctg tct atg gct tcc atc tgg			335
Cys Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp			
100 105 110			
ccc aag aac tgt gtt gtc atg atg aag cgg att ctt gcc tac gtt ccc			383
Pro Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro			
115 120 125			
ttc ttc aac ctg gga gcc tac ttt tcc aac acc atc ttc atc gac cga			431
Phe Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg			
130 135 140			
tac aac cga gag cga gct atg gct tct gtc gac tac tgt gcc tcc gag			479
Tyr Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu			
145 150 155			
atg aag aac cga aac ctg aag ctc tgg gtg ttt ccc gaa gcc act cgg			527
Met Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg			
160 165 170 175			
aat cga gag ggt gga ttc att ccc ttc aag aaa ggt gcc ttc aac atc			575
Asn Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile			
180 185 190			
gct gtt cga gcc cag att ccc atc att cct gtc gtg ttc tct gac tat			623
Ala Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr			
195 200 205			
cga gac ttc tac tcc aag cct gcc cga tac ttc aag aac gat gga gag			671
Arg Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu			
210 215 220			
gtc gtg atc cga gtc ctg gat gcc att ccc acc aag ggt ctg acc ctc			719
Val Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu			
225 230 235			
gat gac gtc tct gag ctt tgg gac atg tgt cga gac gtc atg ctg gct			767
Asp Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala			
240 245 250 255			
gcc tac aag gaa gtt acc ctc gag gct cag caa cga aac gcc act cga			815
Ala Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg			
260 265 270			
aga gga gag acc aag gac gcc aag aaa tcc gag taa gggggcgc			859
Arg Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu			
275 280			
<210> 27			
<211> 282			
<212> PRT			
<213> Caenorhabditis elegans			
<400> 27			
Met Glu Asn Phe Trp Ser Ile Val Val Phe Leu Leu Ser Ile Leu			
1 5 10 15			
Phe Ile Leu Tyr Asn Ile Ser Thr Val Cys His Tyr Tyr Met Arg Ile			
20 25 30			
Ser Phe Tyr Tyr Phe Thr Ile Leu Leu His Gly Met Glu Val Cys Val			
35 40 45			
Thr Met Ile Pro Ser Trp Leu Asn Gly Lys Gly Ala Asp Tyr Val Phe			
50 55 60			
His Ser Phe Phe Tyr Trp Cys Lys Trp Thr Gly Val His Thr Thr Val			
65 70 75 80			
Tyr Gly Tyr Glu Lys Thr Gln Val Glu Gly Pro Ala Val Val Ile Cys			
85 90 95			
Asn His Gln Ser Ser Leu Asp Ile Leu Ser Met Ala Ser Ile Trp Pro			
100 105 110			
Lys Asn Cys Val Val Met Met Lys Arg Ile Leu Ala Tyr Val Pro Phe			
115 120 125			
Phe Asn Leu Gly Ala Tyr Phe Ser Asn Thr Ile Phe Ile Asp Arg Tyr			
130 135 140			
Asn Arg Glu Arg Ala Met Ala Ser Val Asp Tyr Cys Ala Ser Glu Met			
145 150 155 160			
Lys Asn Arg Asn Leu Lys Leu Trp Val Phe Pro Glu Gly Thr Arg Asn			
165 170 175			
Arg Glu Gly Gly Phe Ile Pro Phe Lys Lys Gly Ala Phe Asn Ile Ala			
180 185 190			
Val Arg Ala Gln Ile Pro Ile Ile Pro Val Val Phe Ser Asp Tyr Arg			
195 200 205			
Asp Phe Tyr Ser Lys Pro Gly Arg Tyr Phe Lys Asn Asp Gly Glu Val			
210 215 220			
Val Ile Arg Val Leu Asp Ala Ile Pro Thr Lys Gly Leu Thr Leu Asp			
225 230 235 240			
Asp Val Ser Glu Leu Ser Asp Met Cys Arg Asp Val Met Leu Ala Ala			
245 250 255			
Tyr Lys Glu Val Thr Leu Glu Ala Gln Gln Arg Asn Ala Thr Arg Arg			
260 265 270			
Gly Glu Thr Lys Asp Gly Lys Lys Ser Glu			
275 280			
<210> 28			
<211> 945			
<212> DNA			
<213> Mortierella alpina			
<220>			
<221> CDS			
<222> (1)..(945)			
<223> LPAAT1			
<300>			

<302> High eicosapentaenoic acid producing strains of *Yarrowia lipolytica*

<310> US-2006-0115881-A1

<311> 2005-11-02

<312> 2006-06-01

<313> (1)..(945)

<300>

<302> High eicosapentaenoic acid producing strains of *Yarrowia lipolytica*

<310> WO 2006/052870

<311> 2005-11-03

<312> 2006-05-18

<313> (1)..(945)

<400> 28

```

atg tcc ata ggt tct tcc aat cct gtc ctg ctg gca ggc atc ccc ttc      48
Met Ser Ile Gly Ser Ser Asn Pro Val Leu Leu Ala Ala Ile Pro Phe
1      5      10      15

gtc tac ctc ttc gtc ctc cct cgt gtc ctc gcc ttc ctc cct caa aaq      96
Val Tyr Leu Phe Val Leu Pro Arg Val Leu Ala Phe Leu Pro Gln Lys
20      25      30

gcc cag ttc ctc gca aaa tgc atc gtc gtc ttg atc gcc acc ctt atc      144
Ala Gln Phe Leu Ala Lys Cys Ile Val Val Leu Ile Ala Thr Leu Ile
35      40      45

atg tcc gtc gca ggc tgc ttc att tcc atc gtc tgt ggc ctc ctc gat      192
Met Ser Val Ala Gly Cys Phe Ile Ser Ile Val Cys Ala Leu Leu Asp
50      55      60

aaa cgc tat gtc atc aac tac gtc tgc tca aga ctc ttc tca ttc ctc      240
Lys Arg Tyr Val Ile Asn Tyr Val Val Ser Arg Leu Phe Ser Phe Leu
65      70      75      80

gct gca aga ccc tgc ggt gtc acc tac aag atc gtc gcc gag gaa cat      288
Ala Ala Arg Pro Cys Gly Val Thr Tyr Lys Ile Val Gly Glu Glu His
85      90      95

ctg gac aag tac ccc gcc att gtc gtc tgc aac cac cag agc tcc atg      336
Leu Asp Lys Tyr Pro Ala Ile Val Val Cys Asn His Gln Ser Ser Met
100     105     110

gac atg atg gtc ctg gga cgc gtc ttc cca aag cac tgt gtc gtc atg      384

Asp Met Met Val Leu Gly Arg Val Phe Pro Lys His Cys Val Val Met
115     120     125

gca aag aag gaa ctt ctt tac ttc ccc ttc ctg gcc atg ttt atg aag      432
Ala Lys Lys Glu Leu Leu Tyr Phe Pro Phe Leu Gly Met Phe Met Lys
130     135     140

ctg agt aac gcc atc ttc att gac cgc aag aac cac aag aag gcg atc      480
Leu Ser Asn Ala Ile Phe Ile Asp Arg Lys Asn His Lys Lys Ala Ile
145     150     155     160

gag tcc acc acc caa gct gtc gcc gac atg aag aag cac aac tct gga      528
Glu Ser Thr Thr Gln Ala Val Ala Asp Met Lys Lys His Asn Ser Gly
165     170     175

ate tgg att ttc ccc gaa gga aca cgt tcc cgc ttg gac aag gcc gat      576
Ile Trp Ile Phe Pro Glu Gly Thr Arg Ser Arg Leu Asp Lys Ala Asp
180     185     190

ctc ttg ccc ttc aag aag gga gcc ttc cac ctc gcc att caa gcc caa      624
Leu Leu Pro Phe Lys Lys Gly Ala Phe His Leu Ala Ile Gln Ala Gln
195     200     205

ctc ccc atc ctc ccc atc atc tgc caa gga tac tca cac atc tac gat      672
Leu Pro Ile Leu Pro Ile Ile Ser Gln Gly Tyr Ser His Ile Tyr Asp
210     215     220

tcg tca aaa cgc tac ttc ccc ggt gga gag ctc gag atc aga gtc ctg      720
Ser Ser Lys Arg Tyr Phe Pro Gly Gly Glu Leu Glu Ile Arg Val Leu
225     230     235     240

gaa cct atc ccc acc acg gga ttg acc aca gac gat gtc aac gac ctg      768
Glu Pro Ile Pro Thr Thr Gly Leu Thr Thr Asp Asp Val Asn Asp Leu
245     250     255

atg gac aag act cgc aac ctg atg ctg aag cac ctc aag gag atg gat      816
Met Asp Lys Thr Arg Asn Leu Met Leu Lys His Leu Lys Glu Met Asp
260     265     270

tct caa tac tcc tcc tcc acc gct gaa aac gga tcg acc cat att gac      864
Ser Gln Tyr Ser Ser Ser Thr Ala Glu Asn Gly Ser Thr His Ile Asp
275     280     285

gcc gat atc gca aag tca act gcc aca tcg atc gga aac acg gac gat      912
Ala Asp Ile Ala Lys Ser Thr Ala Thr Ser Ile Gly Asn Thr Asp Asp
290     295     300

gct atc aca aag agg agg aca cca aaa gag tag      945
Ala Ile Thr Lys Arg Arg Thr Pro Lys Glu
305     310

```

<210> 29

<211> 314

<212> PRT

<213> *Mortierella alpina*

<400> 29

```

Met Ser Ile Gly Ser Ser Asn Pro Val Leu Leu Ala Ala Ile Pro Phe
1      5      10      15

Val Tyr Leu Phe Val Leu Pro Arg Val Leu Ala Phe Leu Pro Gln Lys
20      25      30

```

Ala Gln Phe Leu Ala Lys Cys Ile Val Val Leu Ile Ala Thr Leu Ile
35 40 45

Met Ser Val Ala Gly Cys Phe Ile Ser Ile Val Cys Ala Leu Leu Asp
50 55 60

Lys Arg Tyr Val Ile Asn Tyr Val Val Ser Arg Leu Phe Ser Phe Leu
65 70 75 80

Ala Ala Arg Pro Cys Gly Val Thr Tyr Lys Ile Val Gly Glu Glu His
85 90 95

Leu Asp Lys Tyr Pro Ala Ile Val Val Cys Asn His Gln Ser Ser Met
100 105 110

Asp Met Met Val Leu Gly Arg Val Phe Pro Lys His Cys Val Val Met
115 120 125

Ala Lys Lys Glu Leu Leu Tyr Phe Pro Phe Leu Gly Met Phe Met Lys
130 135 140

Leu Ser Asn Ala Ile Phe Ile Asp Arg Lys Asn His Lys Lys Ala Ile
145 150 155 160

Glu Ser Thr Thr Gln Ala Val Ala Asp Met Lys Lys His Asn Ser Gly
165 170 175

Ile Trp Ile Phe Pro Glu Gly Thr Arg Ser Arg Leu Asp Lys Ala Asp
180 185 190

Leu Leu Pro Phe Lys Lys Gly Ala Phe His Leu Ala Ile Gln Ala Gln
195 200 205

Leu Pro Ile Leu Pro Ile Ile Ser Gln Gly Tyr Ser His Ile Tyr Asp
210 215 220

Ser Ser Lys Arg Tyr Phe Pro Gly Gly Glu Leu Glu Ile Arg Val Leu
225 230 235 240

Glu Pro Ile Pro Thr Thr Gly Leu Thr Thr Asp Asp Val Asn Asp Leu
245 250 255

Met Asp Lys Thr Arg Asn Leu Met Leu Lys His Leu Lys Glu Met Asp
260 265 270

Ser Gln Tyr Ser Ser Ser Thr Ala Glu Asn Gly Ser Thr His Ile Asp
275 280 285

Ala Asp Ile Ala Lys Ser Thr Ala Thr Ser Ile Gly Asn Thr Asp Asp
290 295 300

Ala Ile Thr Lys Arg Arg Thr Pro Lys Glu
305 310

<210> 30

<211> 1549

<212> DNA

<213> Yarrowia lipolytica

<220>

<221> CDS

<222> (501)..(1349)

<223> LPAAT1

<300>

<302> High eicosapentaenoic acid producing strains of Yarrowia lipolytica

<310> US-2006-0115881-A1

<311> 2005-11-02

<312> 2006-06-01

<313> (1)..(1549)

<300>

<302> High eicosapentaenoic acid producing strains of Yarrowia lipolytica

<310> WO 2006/052870

<311> 2005-11-03

<312> 2006-05-18

<313> (1)..(1549)

<400> 30

cacacgataa taccacggca tgaccccgct gactcgaacc ttcatcttggc cacatgtagg	60
tgacaaaggg acttcaagag gggccaattt catggcgaca catggcgcaa aaaaaccccg	120
actttgatta cacagacacg taataacgac gaagccgaga tgagccacag tggccaagtc	180
tgccaatggc cccctggacc cccctgacaa agtttcccaa caagcccgag cgtgcattgt	240
gtgtttttgt ggggagacac acgccaatta ggotcatttg aggttatgca gcgaaaaaaa	300
attagtgtgg gtagtgtgtt tgcaggaaac aagtgggtgg ttgaaaaaca agaaaagagc	360
acgacaagag agagagaaaa agagagagag actccataaa gcgtgcataca aaattaaagt	420
gtgtgactat ccgaaaaacca aacatgaaca gtggatata tgcgtctgtg attgcagttg	480
ctgcggttct cattgcccga atg tcc gtt gca tcc aag etc gtc ttc tac gtc	533
Met Ser Val Ala Ser Lys Leu Val Phe Tyr Val	
1 5 10	
oqc gcc gcc atc gcc gtg gtc atc ttt gcc gcc tgt gcc acc tac gcc	581
Arg Ala Ala Ile Ala Val Val Ile Phe Ala Ala Cys Ala Thr Tyr Gly	
15 20 25	
gtg ctg gcc tcc acc att etc acc gcc atc gcc aag cag gcc ctg gcc	629
Val Leu Ala Ser Thr Ile Leu Thr Ala Ile Gly Lys Gln Gly Leu Ala	
30 35 40	

```

caa tgg acc gtt gcc aga gcc ttc tac tac tgg gtc gcc atc ttc ctg      677
Gln Trp Thr Val Ala Arg Ala Phe Tyr Tyr Ser Val Arg Ile Phe Leu
45          50          55

ggt atc agc atc aag ctg cgt agc cgg cag gtg acc gga acc gcc ggt      725
Gly Ile Ser Ile Lys Leu Arg Ser Arg Gln Val Thr Gly Thr Ala Gly
60          65          70          75

ctg gat gcc tcc aag atc cag gtc gcc aac acc acc aag ccc att gcc      773
Leu Asp Ala Ser Lys Ile Gln Val Ala Asn Thr Thr Lys Pro Ile Asp
80          85          90          95

gac atc acc aaa ccc ctg ccc cga cca tgc att ctg att tcc aac ccc      821
Asp Ile Thr Lys His Leu Pro Arg Pro Cys Ile Leu Ile Ser Asn His
95          100          105          110

cag aac gaa atg gac att ctg gtg ctc ggt gcc atc ttc ccc cag tac      869
Gln Asn Glu Met Asp Ile Leu Val Leu Gly Arg Ile Phe Pro Gln Tyr
110          115          120          125

tgc tcc gtc acc gcc aaa aag gcc ctc aag tgg tac ccc ctg ctg gcc      917
Cys Ser Val Thr Ala Lys Lys Ala Leu Lys Trp Tyr Pro Leu Leu Gly
125          130          135          140

cag ttc atg gcg ctg tcc ggc acc atc ttc ctg gac cga aag gac cga      965
Gln Phe Met Ala Leu Ser Gly Thr Ile Phe Leu Asp Arg Lys Asp Arg
140          145          150          155

acc aag tcc gtg cag acc ctc ggc gcc gcc gtc aag acc atc cag agc      1013
Thr Lys Ser Val Gln Thr Leu Gly Gly Ala Val Lys Thr Ile Gln Ser
160          165          170          175

ggc aac gga gcc aag gcc cag ago gtc ttc atg ttc ccc gag gga acc      1061
Gly Asn Gly Gly Lys Gly Gln Ser Val Phe Met Phe Pro Glu Gly Thr
175          180          185          190

cga tcc tac tcc aag gac gtc gcc atc atg ccc ttc aag aag gcc tgt      1109
Arg Ser Tyr Ser Lys Asp Val Gly Ile Met Pro Phe Lys Lys Gly Cys
190          195          200          205

ttc ccc ctg gcg gtc cag tgg gcc gct ccc att gtc ccc gtg gtg gtc      1157
Phe His Leu Ala Val Gln Ser Gly Ala Pro Ile Val Pro Val Val Val
205          210          215          220

cag aac acc tcc cga atg ttt ttt ttc ggc cga gcc aag ctg gcc gcc      1205
Gln Asn Thr Ser Arg Met Phe Ser Phe Gly Arg Gly Lys Leu Asp Ala
220          225          230          235

gga gag atc ctt gtc gac gtc ctg agc ccc att gag acc aag ggt ctg      1253
Gly Glu Ile Leu Val Asp Val Leu Ser Pro Ile Glu Thr Lys Gly Leu
240          245          250          255

gac gcc agc aac gtc gac gct ctc atg gcc acc act tat aag gcc atg      1301
Asp Ala Ser Asn Val Asp Ala Leu Met Ala Thr Thr Tyr Lys Ala Met
255          260          265          270

tgc gag act gcc gac cag att gcc tac gct gcc cag aag act cag tag      1349
Cys Glu Thr Ala Asp Gln Ile Gly Tyr Ala Gly Gln Lys Thr Gln
270          275          280          285

agactgcagc acagaagctg ctgttgctta ctttagggaga gagtaggta atatgaaca      1409
tttttcagat cgaacccac gccgaacact tggctgtgga gctatgggtg aatggattaa      1469
tatagcaacg aaatctacct cgaattacaa cgcgaacaga gcccaatttc tctgtactgt      1529

gctatatcgt gtatacccca      1549

<210> 31
<211> 282
<212> PRT
<213> Yarrowia lipolytica

<400> 31
Met Ser Val Ala Ser Lys Leu Val Phe Tyr Val Arg Ala Ala Ile Ala
1          5          10          15

Val Val Ile Phe Ala Ala Cys Ala Thr Tyr Gly Val Leu Ala Ser Thr
20          25          30          35

Ile Leu Thr Ala Ile Gly Lys Gln Gly Leu Ala Gln Trp Thr Val Ala
35          40          45          50

Arg Ala Phe Tyr Tyr Ser Val Arg Ile Phe Leu Gly Ile Ser Ile Lys
50          55          60          65

Leu Arg Ser Arg Gln Val Thr Gly Thr Ala Gly Leu Asp Ala Ser Lys
65          70          75          80

Ile Gln Val Ala Asn Thr Thr Lys Pro Ile Asp Asp Ile Thr Lys His
85          90          95          100

Leu Pro Arg Pro Cys Ile Leu Ile Ser Asn His Gln Asn Glu Met Asp
100          105          110          115

Ile Leu Val Leu Gly Arg Ile Phe Pro Gln Tyr Cys Ser Val Thr Ala
115          120          125          130

Lys Lys Ala Leu Lys Trp Tyr Pro Leu Leu Gly Gln Phe Met Ala Leu
130          135          140          145

Ser Gly Thr Ile Phe Leu Asp Arg Lys Asp Arg Thr Lys Ser Val Gln
145          150          155          160

Thr Leu Gly Gly Ala Val Lys Thr Ile Gln Ser Gly Asn Gly Gly Lys
165          170          175          180

Gly Gln Ser Val Phe Met Phe Pro Glu Gly Thr Arg Ser Tyr Ser Lys
180          185          190          195

Asp Val Gly Ile Met Pro Phe Lys Lys Gly Cys Phe His Leu Ala Val
195          200          205          210

Gln Ser Gly Ala Pro Ile Val Pro Val Val Val Gln Asn Thr Ser Arg
210          215          220          225

Met Phe Ser Phe Gly Arg Gly Lys Leu Asp Ala Gly Glu Ile Leu Val
225          230          235          240

Asp Val Leu Ser Pro Ile Glu Thr Lys Gly Leu Asp Ala Ser Asn Val
245          250          255          260

Asp Ala Leu Met Ala Thr Thr Tyr Lys Ala Met Cys Glu Thr Ala Asp
260          265          270          275

Gln Ile Gly Tyr Ala Gly Gln Lys Thr Gln
275          280          285          290

```

```

<210> 32
<211> 303
<212> PRT
<213> Saccharomyces cerevisiae

<220>

```

<221> MISC_FEATURE
 <222> (1).(303)
 <223> Slc1p; GenBank Accession No. NP_010231

 <400> 32
 Met Ser Val Ile Gly Arg Phe Leu Tyr Tyr Leu Arg Ser Val Leu Val
 1 5 10 15

 Val Leu Ala Leu Ala Gly Cys Gly Phe Tyr Gly Val Ile Ala Ser Ile
 20 25 30

 Leu Cys Thr Leu Ile Gly Lys Gln His Leu Ala Gln Trp Ile Thr Ala
 35 40 45

 Arg Cys Phe Tyr His Val Met Lys Val Met Leu Gly Leu Asp Val Lys
 50 55 60

 Val Val Gly Glu Glu Asn Leu Ala Lys Lys Pro Tyr Ile Met Ile Ala
 65 70 75 80

 Asn His Gln Ser Thr Leu Asp Ile Phe Met Leu Gly Arg Ile Phe Pro
 85 90 95

 Pro Gly Cys Thr Val Thr Ala Lys Lys Ser Leu Lys Tyr Val Pro Phe
 100 105 110

 Leu Gly Trp Phe Met Ala Leu Ser Gly Thr Tyr Phe Leu Asp Arg Ser
 115 120 125

 Lys Arg Gln Glu Ala Ile Asp Thr Leu Asn Lys Gly Leu Glu Asn Val
 130 135 140

 Lys Lys Asn Lys Arg Ala Leu Trp Val Phe Pro Glu Gly Thr Arg Ser
 145 150 155 160

 Tyr Thr Ser Glu Leu Thr Met Leu Pro Phe Lys Lys Gly Ala Phe His
 165 170 175

 Leu Ala Gln Gln Gly Lys Ile Pro Ile Val Pro Val Val Val Ser Asn
 180 185 190

 Thr Ser Thr Leu Val Ser Pro Lys Tyr Gly Val Phe Asn Arg Gly Cys
 195 200 205

 Met Ile Val Arg Ile Leu Lys Pro Ile Ser Thr Glu Asn Leu Thr Lys
 210 215 220

 Asp Lys Ile Gly Glu Phe Ala Glu Lys Val Arg Asp Gln Met Val Asp
 225 230 235 240

 Thr Leu Lys Glu Ile Gly Tyr Ser Pro Ala Ile Asn Asp Thr Thr Leu
 245 250 255

 Pro Pro Gln Ala Ile Glu Tyr Ala Ala Leu Gln His Asp Lys Lys Val
 260 265 270

 Asn Lys Lys Ile Lys Asn Glu Pro Val Pro Ser Val Ser Ile Ser Asn
 275 280 285

 Asp Val Asn Thr His Asn Glu Gly Ser Ser Val Lys Lys Met His
 290 295 300

 <210> 33
 <211> 7
 <212> PRT
 <213> Artificial Sequence

 <220>
 <223> 1-acyl-sn-glycerol-3-phosphate acyltransferase motif

 <220>
 <221> misc_feature
 <222> (3).(6)
 <223> Xaa can be any naturally occurring amino acid

 <400> 33
 Asn His Xaa Xaa Xaa Xaa Asp
 1 5

 <210> 34
 <211> 4
 <212> PRT
 <213> Artificial Sequence

 <220>
 <223> 1-acyl-sn-glycerol-3-phosphate acyltransferase motif

 <400> 34
 Glu Gly Thr Arg
 1

 <210> 35
 <211> 955
 <212> DNA
 <213> Mortierella alpina

 <220>
 <221> CDS
 <222> (3).(947)
 <223> synthetic LPAAT1 (codon-optimized for Yarrowia lipolytica)

 <400> 35

```

ac atg tct att ggt tgg tcc aac ccc gtg ctc ttg gct gcg att ccc      47
Met Ser Ile Gly Ser Ser Asn Pro Val Leu Leu Ala Ala Ile Pro
1      5      10      15

ttc gtc tac ctg ttt gtc ctc cca cga gtc ctg gct ttc ctg cct cag      95
Phe Val Tyr Leu Phe Val Leu Pro Arg Val Leu Ala Phe Leu Pro Gln
20     25     30

aag gct cag ttc ctg gcc aaa tgt att gtg gtc ctg att gcc acg ctt      143
Lys Ala Gln Phe Leu Ala Lys Cys Ile Val Val Leu Ile Ala Thr Leu
35     40     45

atc atg tcc gtt gca gcc tgc ttc atc tgg atc gtg tgc gct ctt ctg      191
Ile Met Ser Val Ala Gly Cys Phe Ile Ser Ile Val Cys Ala Leu Leu
50     55     60

gac aag aga tac gtc atc aat tac gtt gtg tgg cga ttg ttc tcc ttc      239
Asp Lys Arg Tyr Val Ile Asn Tyr Val Val Ser Arg Leu Phe Ser Phe
65     70     75

ctt gcc gct cga cgg tgt ggt gtg acc tat aag att gtt ggt gag gaa      287
Leu Ala Ala Arg Pro Cys Gly Val Thr Tyr Lys Ile Val Gly Glu Glu
80     85     90     95

cac ctc gat aag tac cct gct atc gtc gtc tgt aac cat caa tcc tct      335
His Leu Asp Lys Tyr Pro Ala Ile Val Val Cys Asn His Gln Ser Ser
100    105    110

atg gat atg atg gtt ttg gga cga gtt ttt cca aag cac tgc gtt gtc      383
Met Asp Met Met Val Leu Gly Arg Val Phe Pro Lys His Cys Val Val
115    120    125

atg gcg aag aag gaa ctc ctg tac ttt ccc ttt ttg gga atg ttt atg      431
Met Ala Lys Lys Glu Leu Leu Tyr Phe Pro Phe Leu Gly Met Phe Met
130    135    140

aaa ctg agc aac gct atc ttc atc gac cgg aag aac cac aag aaa gcc      479
Lys Leu Ser Asn Ala Ile Phe Ile Asp Arg Lys Asn His Lys Lys Ala
145    150    155

atc gag tct acc acc caa gcc gtg gcg gac atg aag aag cac aac tct      527
Ile Glu Ser Thr Thr Gln Ala Val Ala Asp Met Lys Lys His Asn Ser
160    165    170    175

gga atc tgg att ttc cca gag gcc acc cgg tct aga ctg gac aag gca      575
Gly Ile Trp Ile Phe Pro Glu Gly Thr Arg Ser Arg Leu Asp Lys Ala
180    185    190

gac ctg ctg ccc ttc aag aaa ggt gcc ttt cat ctt gca att cag gcc      623
Asp Leu Leu Pro Phe Lys Lys Gly Ala Phe His Leu Ala Ile Gln Ala
195    200    205

cag ctc cct att ctc ccc att atc tgg cag gcc tat tcc cat atc tac      671
Gln Leu Pro Ile Leu Pro Ile Ile Ser Gln Gly Tyr Ser His Ile Tyr
210    215    220

gac tct tgg aag cgg tac ttc ccc ggt gga gag ctc gag atc aga gtc      719
Asp Ser Ser Lys Arg Tyr Phe Pro Gly Gly Glu Leu Glu Ile Arg Val
225    230    235

ctg gag ccc att cct aca act ggc ctc act act gat gat gtg aac gac      767
Leu Glu Pro Ile Pro Thr Thr Gly Leu Thr Thr Asp Asp Val Asn Asp
240    245    250    255

ctg atg gac aag aca cga aac ctt atg ctc aag cac ttg aag gag atg      815
Leu Met Asp Lys Thr Arg Asn Leu Met Leu Lys His Leu Lys Glu Met
260    265    270

gat tcc cag tat tgg tgg agc act gct gaa aat gga tcc acg cac atc      863
Asp Ser Gln Tyr Ser Ser Ser Thr Ala Glu Asn Gly Ser Thr His Ile
275    280    285

gac gcc gat att gcc aag tct aca gcc acc agc att ggc aac act gac      911
Asp Ala Asp Ile Ala Lys Ser Thr Ala Thr Ser Ile Gly Asn Thr Asp
290    295    300

gac goa att aca aaa cgt cgt acc cct aag gaa taa ggggcggc      955
Asp Ala Ile Thr Lys Arg Arg Thr Pro Lys Glu
305    310

<210> 36
<211> 314
<212> PRT
<213> Mortierella alpina

<400> 36
Met Ser Ile Gly Ser Ser Asn Pro Val Leu Leu Ala Ala Ile Pro Phe
1      5      10      15

Val Tyr Leu Phe Val Leu Pro Arg Val Leu Ala Phe Leu Pro Gln Lys
20     25     30

Ala Gln Phe Leu Ala Lys Cys Ile Val Val Leu Ile Ala Thr Leu Ile
35     40     45

Met Ser Val Ala Gly Cys Phe Ile Ser Ile Val Cys Ala Leu Leu Asp
50     55     60

Lys Arg Tyr Val Ile Asn Tyr Val Val Ser Arg Leu Phe Ser Phe Leu
65     70     75     80

```


Ala Ala Arg Pro Cys Gly Val Thr Tyr Lys Ile Val Gly Glu Glu His
85 90 95

Leu Asp Lys Tyr Pro Ala Ile Val Val Cys Asn His Gln Ser Ser Met
100 105 110

Asp Met Met Val Leu Gly Arg Val Phe Pro Lys His Cys Val Val Met
115 120 125

Ala Lys Lys Glu Leu Leu Tyr Phe Pro Phe Leu Gly Met Phe Met Lys
130 135 140

Leu Ser Asn Ala Ile Phe Ile Asp Arg Lys Asn His Lys Lys Ala Ile
145 150 155 160

Glu Ser Thr Thr Gln Ala Val Ala Asp Met Lys Lys His Asn Ser Gly
165 170 175

Ile Trp Ile Phe Pro Glu Gly Thr Arg Ser Arg Leu Asp Lys Ala Asp
180 185 190

Leu Leu Pro Phe Lys Lys Gly Ala Phe His Leu Ala Ile Gln Ala Gln
195 200 205

Leu Pro Ile Leu Pro Ile Ile Ser Gln Gly Tyr Ser His Ile Tyr Asp
210 215 220

Ser Ser Lys Arg Tyr Phe Pro Gly Gly Glu Leu Glu Ile Arg Val Leu
225 230 235 240

Glu Pro Ile Pro Thr Thr Gly Leu Thr Thr Asp Asp Val Asn Asp Leu
245 250 255

Met Asp Lys Thr Arg Asn Leu Met Leu Lys His Leu Lys Glu Met Asp
260 265 270

Ser Gln Tyr Ser Ser Ser Thr Ala Glu Asn Gly Ser Thr His Ile Asp
275 280 285

Ala Asp Ile Ala Lys Ser Thr Ala Thr Ser Ile Gly Asn Thr Asp Asp
290 295 300

Ala Ile Thr Lys Arg Arg Thr Pro Lys Glu
305 310

<210> 37
<211> 1185
<212> DNA
<213> Yarrowia lipolytica

<220>
<221> CDS
<222> (1)..(1185)
<223> diacylglycerol cholinephosphotransferase (Y1CPT1)

<300>
<302> HIGH EICOSAPENTAENOIC ACID PRODUCING STRAINS OF

<310> WO 2006/052870
<311> 2005-11-03
<312> 2006-05-18
<313> (1)..(1185)

<300>
<302> HIGH EICOSAPENTAENOIC ACID PRODUCING STRAINS OF

<310> US 2006-0115881-A1
<311> 2005-11-02
<312> 2006-06-01
<313> (1)..(1185)

<400> 37

atg ggc gta ttc att asa 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 gac 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 gcc 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 sac 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 tcc 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 tot 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 att tcg ctg ctg cct gcc agc ttc ctg gac gtc acc atc acc 672
 Ser Ile Thr Ser 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 gcc 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
 gcc gcc tgc gga aac gtg atc aag tac tac aac aac aag gcc 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 gcc atc ttc tac gtc 816
 Ala Leu Pro Ala Ile Leu Gly Ile Ala Pro Phe Gly Ile Phe Tyr Val
 260 265 270
 gcc 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
 gcc 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 gcc tgg gac gcc gat atc gtg tgg 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 gcc 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> 38
 <211> 394
 <212> PRT
 <213> Yarrowia lipolytica
 <400> 38
 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 80
 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
 145 150 155 160
 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
 225 230 235 240
 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> 39
<211> 1515
<212> DNA
<213> Rhizobium leguminosarum bv. viciae 3841

<220>
<221> CDS
<222> (1)..(1515)
<223> GenBank Accession No. YP_766603

<400> 39
gtg agc aac cat ctt ttc gac gcc atg cgg gct gcc gcg ccc ggt gac      48
Val Ser Asn His Leu Phe Asp Ala Met Arg Ala Ala Ala Pro Gly Asp
1          5          10          15

gca cgg ttc atc cgg atc gat aac gcg cgc aca tgg acc tat gat gac      96
Ala Pro Phe Ile Arg Ile Asp Asn Ala Arg Thr Trp Thr Tyr Asp Asp
20          25          30

gcg atc gct ctt tcc gcc cgt att gcc gcc gcg atg gac gcg ctc gcc      144
Ala Ile Ala Leu Ser Gly Arg Ile Ala Gly Ala Met Asp Ala Leu Gly
35          40          45

att cgc ccc gcc gac cgg gcg gcg gcg cag gtc gag aaa sgt gcc gag      192
Ile Arg Pro Gly Asp Arg Val Ala Val Gln Val Glu Lys Ser Ala Glu
50          55          60

gcg ttg atc ctc tat ctc gcc tgt ctt cga acc gcc gcc gtc tac ctg      240
Ala Leu Ile Leu Tyr Leu Ala Cys Leu Arg Thr Gly Ala Val Tyr Leu

65          70          75          80

ccg ctc aac acc gcc tat acg ctg gct gag ctc gat tac ttt atc gcc      288
Pro Leu Asn Thr Ala Tyr Thr Leu Ala Glu Leu Asp Tyr Phe Ile Gly
85          90          95

gat gcg gag cgg cgt ttg gcg gtc gtt gcg cgg gcg gct cga gcc gcc      336
Asp Ala Glu Pro Arg Leu Val Val Val Ala Pro Ala Ala Asp Gly Gly
100          105          110

gtg gag aca atc gcc aag cgc cgc gcc gcg atc gtc gaa acg ctc gat      384
Val Glu Thr Ile Ala Lys Arg His Gly Ala Ile Val Glu Thr Leu Asp
115          120          125

gct gat gcc gcc gcc tca ttg ctg gat ctc gca cgc gat gag cgg gcc      432
Ala Asp Gly Arg Gly Ser Leu Leu Asp Leu Ala Arg Asp Glu Pro Ala
130          135          140

gac ttt ctc gat gcc tcg cgt tcc ccg gat gat ctg gcc gcc atc ctc      480
Asp Phe Val Asp Ala Ser Arg Ser Ala Asp Asp Leu Ala Ala Ile Leu
145          150          155          160

tac acg tcg gga acg acc gga cgc tcc aag ggg gcg atg ctc acg cat      528
Tyr Thr Ser Gly Thr Thr Gly Arg Ser Lys Gly Ala Met Leu Thr His
165          170          175

ggg aac ctg ctc tcg aac gcc ctg acc ttg cga gac tat tgg cgc gtc      576
Gly Asn Leu Leu Ser Asn Ala Leu Thr Leu Arg Asp Tyr Trp Arg Val
180          185          190

acc gat gac gat cgg ctg atc cat gcc ttg cgg atc ttc cac aag cat      624
Thr Ala Asp Asp Arg Leu Ile His Ala Leu Pro Ile Phe His Thr His
195          200          205

ggg ctg ttc gtc gcc aag aac gtc aca ctg ctt gcc gcc gcc tcg atg      672
Gly Leu Phe Val Ala Thr Asn Val Thr Leu Leu Ala Gly Ala Ser Met
210          215          220

ttc ctg ctg tcg aaa ttc gac gcc gac gag gtc gtt tca ctg atg cca      720
Phe Leu Leu Ser Lys Phe Asp Ala Asp Glu Val Val Ser Leu Met Pro
225          230          235          240

cag gca act atg ctg atg gcc gcg acc ttc tac ctg cgc ctc ctg      768
Gln Ala Thr Met Leu Met Gly Val Pro Thr Phe Tyr Val Arg Leu Leu
245          250          255

caa agc cgg cgc ctc gaa aaa ggg gcg gtc gcc agc atc cgc ctc ttc      816
Gln Ser Pro Arg Leu Glu Lys Gly Ala Val Ala Ser Ile Arg Leu Phe
260          265          270

att tcc ggt tcg gct ccc ctg ctg gcc gaa acc cat gcc gag ttc cat      864
Ile Ser Gly Ser Ala Pro Leu Leu Ala Glu Thr His Ala Glu Phe His
275          280          285

gcg cgt acc ggt cgc gcc att ctc gag cgc tac gcc atg aag gaa acc      912
Ala Arg Thr Gly His Ala Ile Leu Glu Arg Tyr Gly Met Thr Glu Thr
290          295          300

aat atg aac acg tcc aac ccc tat gaa gcc aac cgg att gcc gaa acc      960
Asn Met Asn Thr Ser Asn Pro Tyr Glu Gly Lys Arg Ile Ala Gly Thr
305          310          315          320

gtt ggt ttc cgg ctg cct gat gtc aag gtc cgc gtc acc gat acc gcc      1008
Val Gly Phe Pro Leu Pro Asp Val Thr Val Arg Val Thr Asp Pro Ala
325          330          335

```

acc ggg ctc gtg ctg cca cct gaa gaq acg ggc atg atc gag atc aag 1056
 Thr Gly Leu Val Leu Pro Pro Glu Glu Thr Gly Met Ile Glu Ile Lys
 340 345 350

gga ccg aac gtc ttc aag ggc tat tgg cgc atg ccg gaa aag acc gcg 1104
 Gly Pro Asn Val Phe Lys Gly Tyr Trp Arg Met Pro Glu Lys Thr Ala
 355 360 365

gac gaa ttc acc gcc gac ggt ttc ttc atc agc ggc gat ctc gcc aag 1152
 Ala Glu Phe Thr Ala Asp Gly Phe Phe Ile Ser Gly Asp Leu Gly Lys
 370 375 380

atc gac cgg gaa ggt tat gtc cac atc gtc ggt cgc ggc aag gat ctg 1200
 Ile Asp Arg Glu Gly Tyr Val His Ile Val Gly Arg Gly Lys Asp Leu
 385 390 395 400

gtg atc tgg ggt gga tac aac atc tat ccg aaa gag gtc gaa gcc gag 1248
 Val Ile Ser Gly Gly Tyr Asn Ile Tyr Pro Lys Glu Val Glu Gly Glu
 405 410 415

atc gac cag atc gag ggt gtg gtt gag agc gct gtg atc ggt gtg ccg 1296
 Ile Asp Gln Ile Glu Gly Val Val Glu Ser Ala Val Ile Gly Val Pro
 420 425 430

cat ccc gat ttc gga gaa gcc gta acc gct gtc gtc gta tgt aag ccc 1344
 His Pro Asp Phe Gly Glu Gly Val Thr Ala Val Val Val Cys Lys Pro
 435 440 445

ggc gcc gtc ctc gat gaa aag acc atc gtc agc gcc ctc cag gac cgt 1392
 Gly Ala Val Leu Asp Glu Lys Thr Ile Val Ser Ala Leu Gln Asp Arg
 450 455 460

ctc gcc cgc tac aaa caa ccc aag cgc atc atc ttc gcc gac gac ctg 1440
 Leu Ala Val Tyr Lys Gln Pro Lys Arg Ile Ile Phe Ala Asp Asp Leu
 465 470 475 480

ccg cgc aac act atg ggt aag gtt cag aag aat atc ctg cgg cag caa 1488
 Pro Arg Asn Thr Met Gly Lys Val Gln Lys Asn Ile Leu Arg Gln Gln
 485 490 495

tac gcc gat ctt tat acc agg agg taa 1515
 Tyr Ala Asp Leu Tyr Thr Arg Arg
 500

<210> 40

<211> 504

<212> PRT

<213> Rhizobium leguminosarum bv. viciae 3841

<400> 40

Val Ser Asn His Leu Phe Asp Ala Met Arg Ala Ala Ala Pro Gly Asp
 1 5 10 15

Ala Pro Phe Ile Arg Ile Asp Asn Ala Arg Thr Trp Thr Tyr Asp Asp
 20 25 30

Ala Ile Ala Leu Ser Gly Arg Ile Ala Gly Ala Met Asp Ala Leu Gly
 35 40 45

Ile Arg Pro Gly Asp Arg Val Ala Val Gln Val Glu Lys Ser Ala Glu
 50 55 60

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

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

Asp Ala Glu Pro Arg Leu Val Val Val Ala Pro Ala Ala Arg Gly Gly
 100 105 110

Val Glu Thr Ile Ala Lys Arg His Gly Ala Ile Val Glu Thr Leu Asp
 115 120 125

Ala Asp Gly Arg Gly Ser Leu Leu Asp Leu Ala Arg Asp Glu Pro Ala
 130 135 140

Asp Phe Val Asp Ala Ser Arg Ser Ala Asp Asp Leu Ala Ala Ile Leu
 145 150 155 160

Tyr Thr Ser Gly Thr Thr Gly Arg Ser Lys Gly Ala Met Leu Thr His
 165 170 175

Gly Asn Leu Leu Ser Asn Ala Leu Thr Leu Arg Asp Tyr Trp Arg Val
 180 185 190

Thr Ala Asp Asp Arg Leu Ile His Ala Leu Pro Ile Phe His Thr His
 195 200 205

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

Phe Leu Leu Ser Lys Phe Asp Ala Asp Glu Val Val Ser Leu Met Pro
 225 230 235 240

Gln Ala Thr Met Leu Met Gly Val Pro Thr Phe Tyr Val Arg Leu Leu
 245 250 255

Gln Ser Pro Arg Leu Glu Lys Gly Ala Val Ala Ser Ile Arg Leu Phe
 260 265 270

Ile Ser Gly Ser Ala Pro Leu Leu Ala Glu Thr His Ala Glu Phe His
 275 280 285

Ala Arg Thr Gly His Ala Ile Leu Glu Arg Tyr Gly Met Thr Glu Thr
 290 295 300

Asn Met Asn Thr Ser Asn Pro Tyr Glu Gly Lys Arg Ile Ala Gly Thr

```

305          310          315          320
Val Gly Phe Pro Leu Pro Asp Val Thr Val Arg Val Thr Asp Pro Ala
325          330          335
Thr Gly Leu Val Leu Pro Pro Glu Glu Thr Gly Met Ile Glu Ile Lys
340          345          350
Gly Pro Asn Val Phe Lys Gly Tyr Trp Arg Met Pro Glu Lys Thr Ala
355          360          365
Ala Glu Phe Thr Ala Asp Gly Phe Phe Ile Ser Gly Asp Leu Gly Lys
370          375          380
Ile Asp Arg Glu Gly Tyr Val His Ile Val Gly Arg Gly Lys Asp Leu
385          390          395          400
Val Ile Ser Gly Gly Tyr Asn Ile Tyr Pro Lys Glu Val Glu Gly Glu
405          410          415
Ile Asp Gln Ile Glu Gly Val Val Glu Ser Ala Val Ile Gly Val Pro
420          425          430
His Pro Asp Phe Gly Glu Gly Val Thr Ala Val Val Val Cys Lys Pro
435          440          445
Gly Ala Val Leu Asp Glu Lys Thr Ile Val Ser Ala Leu Gln Asp Arg
450          455          460
Leu Ala Arg Tyr Lys Gln Pro Lys Arg Ile Ile Phe Ala Asp Asp Leu
465          470          475          480
Pro Arg Asn Thr Met Gly Lys Val Gln Lys Asn Ile Leu Arg Gln Gln
485          490          495
Tyr Ala Asp Leu Tyr Thr Arg Arg
500
<210> 41
<211> 1518
<212> DNA
<213> Rhizobium leguminosarum bv. viciae 3841
<220>
<221> CDS
<222> (1)..(1518)
<223> synthetic malonyl-CoA synthetase (codon-optimized for Yarrowia lipolytica)
<400> 41
atg gtc tcc aac cac ctg ttc gac gcc atg cga gct gcc gct ccc gga          48
Met Val Ser Asn His Leu Phe Asp Ala Met Arg Ala Ala Ala Pro Gly
1      5      10      15
gac gca cct ttc att cga atc gac aac gct cga acc tgg acg tac gat          96
Asp Ala Pro Phe Ile Arg Ile Asp Asn Ala Arg Thr Trp Thr Tyr Asp
20     25     30
gac gcc att gct ctt tcc ggt cga atc gct gga gct atg gac gca ctc          144
Asp Ala Ile Ala Leu Ser Gly Arg Ile Ala Gly Ala Met Asp Ala Leu
35     40     45
ggc att cga ccc gga gac aga gtt gcc gtg cag gtc gag aag tct gcc          192
Gly Ile Arg Pro Gly Asp Arg Val Ala Val Gln Val Glu Lys Ser Ala
50     55     60
gag gag ttg ctt ctc tac ctg gcc tgt ctt cga acc gga gct gtc tac          240
Glu Ala Leu Ile Leu Tyr Leu Ala Cys Leu Arg Thr Gly Ala Val Tyr
65     70     75     80
ctg cct ctc aac act gcc tac acc ctg gcc gag ctc gac tac ttc atc          288
Leu Pro Leu Asn Thr Ala Tyr Thr Leu Ala Glu Leu Asp Tyr Phe Ile
85     90     95
ggc gat gcc gaa ccg cgt ctg gtg gtc gtt gct ccc gca gct cga ggt          336
Gly Asp Ala Glu Pro Arg Leu Val Val Val Ala Pro Ala Ala Arg Gly
100    105    110
ggc gtg gag aca att gcc aag cga aac ggt gct atc gtc gaa acc ctc          384
Gly Val Glu Thr Ile Ala Lys Arg His Gly Ala Ile Val Glu Thr Leu
115    120    125
gac gcc gat gga cga ggc tcc ttg ctg gac ctt gct aga gat gag cct          432
Asp Ala Asp Gly Arg Gly Ser Leu Leu Asp Leu Ala Arg Asp Glu Pro
130    135    140
gcc gac ttt gtc gat gct tgg cga tct gcc gac gat ctg gct gct att          480
Ala Asp Phe Val Asp Ala Ser Arg Ser Ala Asp Asp Leu Ala Ala Ile
145    150    155    160
ctc tac act tcc ggt aca acc gga cga tgg aag ggt gcc atg ctt act          528
Leu Tyr Thr Ser Gly Thr Thr Gly Arg Ser Lys Gly Ala Met Leu Thr
165    170    175
cat gcc aat ctg ctc tcc aac gct ctc acc ttg cga gac tat tgg aga          576
His Gly Asn Leu Leu Ser Asn Ala Leu Thr Leu Arg Asp Tyr Trp Arg
180    185    190
gtt acc gca gac gat cga ctc atc cgt gcc ttg cca atc ttt cgc act          624
Val Thr Ala Asp Asp Arg Leu Ile His Ala Leu Pro Ile Phe His Thr
195    200    205
cat ggt ctg ttc gtt gct acg aac gtc aca ctg ctt gca gga gcc tgg          672
His Gly Leu Phe Val Ala Thr Asn Val Thr Leu Leu Ala Gly Ala Ser
210    215    220
atg ttt ctg ctc tcc aag ttc gat gcc gac gag gtc gtt tct ctc atg          720
Met Phe Leu Leu Ser Lys Phe Asp Ala Asp Glu Val Val Ser Leu Met
225    230    235    240
cca cag gcc acc atg ctt atg gcc gtg ccc aca ttc tac gtt cga ttg          768
Pro Gln Ala Thr Met Leu Met Gly Val Pro Thr Phe Tyr Val Arg Leu
245    250    255
ctg cag agt cct cga ctc gag aag ggt gct gtg gcc agc atc aga ctg          816
Leu Gln Ser Pro Arg Leu Glu Lys Gly Ala Val Ala Ser Ile Arg Leu

```

260	265	270	
ttc att tct gga tca gct ccc ttg ctt gcc gaa acc cac gcc gag ttt			864
Phe Ile Ser Gly Ser Ala Pro	Leu Leu Ala Glu Thr His Ala Glu Phe		
275	280	285	
oat gct cgt act ggt oac gcc att ctc gag cga tac gcc atg acg gaa			912
His Ala Arg Thr Gly His Ala Ile Leu Glu Arg Tyr Gly Met Thr Glu			
290	295	300	
acc aac atg aat act tcc aac ccc tac gag gcc aag cgt att gcc gga			960
Thr Asn Met Asn Thr Ser Asn Pro Tyr Glu Gly Lys Arg Ile Ala Gly			
305	310	315	320
acc gtt ggt ttt cct ctg ccc gac gtc act gtg cga gtc acc gat ccc			1008
Thr Val Gly Phe Pro Leu Pro Asp Val Thr Val Arg Val Thr Asp Pro			
325	330	335	
ccc acc ggt ctc gtt ctt cca cct gaa gag act gcc atg atc gag atc			1056
Ala Thr Gly Leu Val Leu Pro Pro Glu Glu Thr Gly Met Ile Ala Ile			
340	345	350	
aag gga ccc aac gtc ttc aag gcc tat tgg cga atg ccc gaa aag acc			1104
Lys Gly Pro Asn Val Phe Lys Gly Tyr Trp Arg Met Pro Glu Lys Thr			
355	360	365	
cct gcc gag ttt acc gca gac ggt ttc ttt atc tct gga gat ctc ggc			1152
Ala Ala Glu Phe Thr Ala Asp Gly Phe Phe Ile Ser Gly Asp Leu Gly			
370	375	380	
aag atc gac cga gaa ggt tac gtt cac att gtg cga cga gcc aag gac			1200
Lys Ile Asp Arg Glu Gly Tyr Val His Ile Val Gly Arg Gly Lys Asp			
385	390	395	400
ctg gtc att tcc ggt gcc tac aac atc tat ccc aaa gag gtc gaa gcc			1248
Leu Val Ile Ser Gly Gly Tyr Asn Ile Tyr Pro Lys Glu Val Glu Gly			
405	410	415	
gag atc gac cag atc gag ggt gtg gtc gag tot gct gtc att ggt gtt			1296
Glu Ile Asp Glu Ile Glu Gly Val Val Glu Ser Ala Val Ile Gly Val			
420	425	430	
cct cat ccc gat ttc gga gaa ggt gtc acc gct gtt gtc gtg tgc aaa			1344
Pro His Pro Asp Phe Gly Glu Gly Val Thr Ala Val Val Val Cys Lys			
435	440	445	
cct ggt gcc gtt ctc gac gaa aag acc atc gtg tot gct ctg cag gac			1392
Pro Gly Ala Val Leu Asp Glu Lys Thr Ile Val Ser Ala Leu Glu Asp			
450	455	460	
cgt ctt gcc cga tac aag caa ccc aag cgg att atc ttt gcc gac gat			1440
Arg Leu Ala Arg Tyr Lys Glu Pro Lys Arg Ile Ile Phe Ala Asp Asp			
465	470	475	480
ctg cct cga aac act atg gga aag gtt cag aag aac att ctt cga cag			1488
Leu Pro Arg Asn Thr Met Gly Lys Val Gln Lys Asn Ile Leu Arg Gln			
485	490	495	
caa tac gcc gat ctc tac acc aga cga taa			1518
Gln Tyr Ala Asp Leu Tyr Thr Arg Arg			
500	505		
<210> 42			
<211> 505			
<212> PRT			
<213> Rhizobium leguminosarum bv. viciae 3841			
<400> 42			
Met Val Ser Asn His Leu Phe Asp Ala Met Arg Ala Ala Ala Pro Gly			
1	5	10	15
Asp Ala Pro Phe Ile Arg Ile Asp Asn Ala Arg Thr Trp Thr Tyr Asp			
20	25	30	
Asp Ala Ile Ala Leu Ser Gly Arg Ile Ala Gly Ala Met Asp Ala Leu			
35	40	45	
Gly Ile Arg Pro Gly Asp Arg Val Ala Val Gln Val Glu Lys Ser Ala			
50	55	60	
Glu Ala Leu Ile Leu Tyr Leu Ala Cys Leu Arg Thr Gly Ala Val Tyr			
65	70	75	80
Leu Pro Leu Asn Thr Ala Tyr Thr Leu Ala Glu Leu Asp Tyr Phe Ile			
85	90	95	
Gly Asp Ala Glu Pro Arg Leu Val Val Val Ala Pro Ala Ala Arg Gly			
100	105	110	
Gly Val Glu Thr Ile Ala Lys Arg His Gly Ala Ile Val Glu Thr Leu			
115	120	125	
Asp Ala Asp Gly Arg Gly Ser Leu Leu Asp Leu Ala Arg Asp Glu Pro			
130	135	140	
Ala Asp Phe Val Asp Ala Ser Arg Ser Ala Asp Asp Leu Ala Ala Ile			
145	150	155	160
Leu Tyr Thr Ser Gly Thr Thr Gly Arg Ser Lys Gly Ala Met Leu Thr			
165	170	175	
His Gly Asn Leu Leu Ser Asn Ala Leu Thr Leu Arg Asp Tyr Trp Arg			
180	185	190	
Val Thr Ala Asp Asp Arg Leu Ile His Ala Leu Pro Ile Phe His Thr			
195	200	205	
His Gly Leu Phe Val Ala Thr Asn Val Thr Leu Leu Ala Gly Ala Ser			
210	215	220	
Met Phe Leu Leu Ser Lys Phe Asp Ala Asp Glu Val Val Ser Leu Met			
225	230	235	240

```

Pro Gln Ala Thr Met Leu Met Gly Val Pro Thr Phe Tyr Val Arg Leu
245 250 255

Leu Gln Ser Pro Arg Leu Glu Lys Gly Ala Val Ala Ser Ile Arg Leu
260 265 270

Phe Ile Ser Gly Ser Ala Pro Leu Leu Ala Glu Thr His Ala Glu Phe
275 280 285

His Ala Arg Thr Gly His Ala Ile Leu Glu Arg Tyr Gly Met Thr Glu
290 295 300

Thr Asn Met Asn Thr Ser Asn Pro Tyr Glu Gly Lys Arg Ile Ala Gly
305 310 315 320

Thr Val Gly Phe Pro Leu Pro Asp Val Thr Val Arg Val Thr Asp Pro
325 330 335

Ala Thr Gly Leu Val Leu Pro Pro Glu Glu Thr Gly Met Ile Glu Ile
340 345 350

Lys Gly Pro Asn Val Phe Lys Gly Tyr Trp Arg Met Pro Glu Lys Thr
355 360 365

Ala Ala Glu Phe Thr Ala Asp Gly Phe Phe Ile Ser Gly Asp Leu Gly
370 375 380

Lys Ile Asp Arg Glu Gly Tyr Val His Ile Val Gly Arg Gly Lys Asp
385 390 395 400

Leu Val Ile Ser Gly Gly Tyr Asn Ile Tyr Pro Lys Glu Val Glu Gly
405 410 415

Glu Ile Asp Gln Ile Glu Gly Val Val Glu Ser Ala Val Ile Gly Val
420 425 430

Pro His Pro Asp Phe Gly Glu Gly Val Thr Ala Val Val Val Cys Lys
435 440 445

Pro Gly Ala Val Leu Asp Glu Lys Thr Ile Val Ser Ala Leu Gln Asp
450 455 460

Arg Leu Ala Arg Tyr Lys Gln Pro Lys Arg Ile Ile Phe Ala Asp Asp
465 470 475 480

Leu Pro Arg Asn Thr Met Gly Lys Val Gln Lys Asn Ile Leu Arg Gln
485 490 495

Gln Tyr Ala Asp Leu Tyr Thr Arg Arg

500 505

<210> 43
<211> 777
<212> DNA
<213> Euglena gracilis

<220>
<221> CDS
<222> (1)..(777)
<223> delta-9 elongase

<300>
<302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2007/061742
<311> 2006-11-16
<312> 2007-05-31
<313> (1)..(777)

<300>
<302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US 2007-0117190-A1
<311> 2006-11-16
<312> 2007-05-24
<313> (1)..(777)

<400> 43
atg gag gtg atg aat gaa ata gtc tca att ggg cag gaa gtt tta ccc 48
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 ctt 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 ggt 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
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 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
ttc tat gtg caa acg tat atc gtc agg aag csc 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
Ile Gln
777

<210> 44
<211> 258
<212> PRT
<213> Euglena gracilis

<400> 44
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      240
Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
245      250      255

Ile Gln

<210> 45
<211> 777
<212> DNA
<213> Euglena gracilis

<220>
<221> CDS
<222> (1)..(777)
<223> synthetic delta-9 elongase (codon-optimized for Yarrowia lipolytica)

<300>
<302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2007/061742
<311> 2006-11-16
<312> 2007-05-31
<313> (1)..(777)

<300>
<302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US-2007-0117190-A1
<311> 2006-11-16
<312> 2007-05-24
<313> (1)..(777)

<400> 45

```


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 gaa 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 gcc 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 gcc ttc att cac tgg atc atg tac gcc 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 tot 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 gtc gcc 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> 46
 <211> 258
 <212> PRT
 <213> Euglena gracilis
 <400> 46
 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 Lys Cys Leu Phe Leu Asn
 225 230 235 240
 Phe Tyr Val Gln Thr Tyr Ile Val Arg Lys His Lys Gly Ala Lys Lys
 245 250 255
 Ile Gln
 <210> 47
 <211> 792
 <212> DNA
 <213> Eutreptiella sp. CCMP389
 <220>

<221> CDS
 <222> (1)..(792)
 <223> delta-9 elongase

<300>
 <302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2007/061742
 <311> 2006-11-16
 <312> 2007-05-31
 <313> (1)..(792)

<300>
 <302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US 2007-0117190-A1
 <311> 2006-11-16
 <312> 2007-05-24
 <313> (1)..(792)

<400> 47
 atg gct ggc gtg ata gag gtc gcc aac gag ttt gta gcc atc acg gca 48
 Met Ala Ala Val Ile Glu Val Ala Asn Glu Phe Val Ala Ile Thr Ala
 1 5 10 15
 gaa acg ctc ccc aaa gtt gac tat caa cga cta tgg cga gac att tac 96
 Glu Thr Leu Pro Lys Val Asp Tyr Gln Arg Leu Trp Arg Asp Ile Tyr
 20 25 30
 agt tgt gag cta ctg tat ttc tcc att gcc ttc gtg atc ttg aag ttt 144
 Ser Cys Glu Leu Leu Tyr Phe Ser Ile Ala Phe Val Ile Leu Lys Phe
 35 40 45
 acg ttg gcc gag ttg agc gac agc gga aaa aag att ttg aga gtg ttg 192
 Thr Leu Gly Glu Leu Ser Asp Ser Gly Lys Lys Ile Leu Arg Val Leu
 50 55 60
 ttc aag tgg tac aat ctc ttc atg tcc gtg ttc tcc ttg gtg tct ttc 240
 Phe Lys Trp Tyr Asn Leu Phe Met Ser Val Phe Ser Leu Val Ser Phe
 65 70 75 80
 ctt tgc atg ggc tat gcc att tat acc ttg ggc cta tac tct aac gaa 288
 Leu Cys Met Gly Tyr Ala Ile Tyr Thr Val Gly Leu Tyr Ser Asn Glu
 85 90 95
 tgc gac agc gct ttc gac aac tcg ttg ttc cgc ttt gca aca aag gtg 336
 Cys Asp Arg Ala Phe Asp Asn Ser Leu Phe Arg Phe Ala Thr Lys Val
 100 105 110
 ttc tac tac agt aag ttt ttg gag tac atc gac tct ttt tat ctt ccg 384
 Phe Tyr Tyr Ser Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro
 115 120 125
 ctc atg gcc aag ccg ctg tct ttc ctg caa ttc ttc cat cac ttg gga 432
 Leu Met Ala Lys Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly
 130 135 140
 gcc ccc atg gac atg tgg ctc ttt gtc caa tat tct ggc gaa tct att 480
 Ala Pro Met Asp Met Trp Leu Phe Val Gln Tyr Ser Gly Glu Ser Ile
 145 150 155 160
 tgg atc ttt gtg ttt ttg aat ggg ttc att cac ttt gtt atg tac ggg 528
 Trp Ile Phe Val Phe Leu Asn Gly Phe Ile His Phe Val Met Tyr Gly
 165 170 175
 tac tac tgg act ccg ctg atg aag ttc aat ttc cca atg ccc aag cag 576
 Tyr Tyr Trp Thr Arg Leu Met Lys Phe Asn Phe Pro Met Pro Lys Gln
 180 185 190
 ttg att ccc gcg atg cag atc acg cag ttc aac gtt ggt ttc tac ctc 624
 Leu Ile Thr Ala Met Gln Ile Thr Gln Phe Asn Val Gly Phe Tyr Leu
 195 200 205
 gtg tgg tgg tac aaa gat att ccc tgc tac cga aag gat ccc atg cga 672
 Val Trp Trp Tyr Lys Asp Ile Pro Cys Tyr Arg Lys Asp Pro Met Arg
 210 215 220
 atg ttg gcc tgg atc ttc aat tac tgg tat gtt ggc act gtc ttg ctg 720
 Met Leu Ala Trp Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu
 225 230 235 240
 ctg ttc att aat ttc ttc gtc aaa tcc tat gtg ttc cca aag ccg aag 768
 Leu Phe Ile Asn Phe Phe Val Lys Ser Tyr Val Phe Pro Lys Pro Lys
 245 250 255
 act gca gat aaa aag gtc caa tag 792
 Thr Ala Asp Lys Lys Val Gln
 260

<210> 48
 <211> 263
 <212> PRT
 <213> Eutrepiaella sp. CCMP389

<400> 48
 Met Ala Ala Val Ile Glu Val Ala Asn Glu Phe Val Ala Ile Thr Ala
 1 5 10 15

Glu Thr Leu Pro Lys Val Asp Tyr Gln Arg Leu Trp Arg Asp Ile Tyr
 20 25 30
 Ser Cys Glu Leu Leu Tyr Phe Ser Ile Ala Phe Val Ile Leu Lys Phe
 35 40 45
 Thr Leu Gly Glu Leu Ser Asp Ser Gly Lys Lys Ile Leu Arg Val Leu
 50 55 60
 Phe Lys Trp Tyr Asn Leu Phe Met Ser Val Phe Ser Leu Val Ser Phe
 65 70 75 80
 Leu Cys Met Gly Tyr Ala Ile Tyr Thr Val Gly Leu Tyr Ser Asn Glu
 85 90 95
 Cys Asp Arg Ala Phe Asp Asn Ser Leu Phe Arg Phe Ala Thr Lys Val
 100 105 110
 Phe Tyr Tyr Ser Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro
 115 120 125
 Leu Met Ala Lys Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly
 130 135 140
 Ala Pro Met Asp Met Trp Leu Phe Val Gln Tyr Ser Gly Glu Ser Ile
 145 150 155 160
 Trp Ile Phe Val Phe Leu Asn Gly Phe Ile His Phe Val Met Tyr Gly
 165 170 175
 Tyr Tyr Trp Thr Arg Leu Met Lys Phe Asn Phe Pro Met Pro Lys Gln
 180 185 190
 Leu Ile Thr Ala Met Gln Ile Thr Gln Phe Asn Val Gly Phe Tyr Leu
 195 200 205
 Val Trp Trp Tyr Lys Asp Ile Pro Cys Tyr Arg Lys Asp Pro Met Arg
 210 215 220
 Met Leu Ala Trp Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu
 225 230 235 240
 Leu Phe Ile Asn Phe Phe Val Lys Ser Tyr Val Phe Pro Lys Pro Lys
 245 250 255
 Thr Ala Asp Lys Lys Val Gln
 260
 <210> 49
 <211> 792
 <212> DNA
 <213> Eutrepitella sp. CCMP389
 <220>
 <221> CDS
 <222> (1)..(792)
 <223> synthetic delta-9 elongase (codon-optimized for Yarrowia lipolytica)
 <300>
 <302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS
 <310> WO 2007/061742
 <311> 2006-11-16
 <312> 2007-05-31
 <313> (1)..(792)
 <300>
 <302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS
 <310> US 2007-0117190-A1
 <311> 2006-11-16
 <312> 2007-05-24
 <313> (1)..(792)
 <400> 49
 atg gct gcc gtc atc gag gtg gcc aac gag ttc gtc gct atc act gcc 48
 Met Ala Ala Val Ile Glu Val Ala Asn Glu Phe Val Ala Ile Thr Ala
 1 5 10 15
 gag acc ctt ccc aag ctg gac tat caa cga ctc tgg cga gac atc tac 96
 Glu Thr Leu Pro Lys Val Asp Tyr Gln Arg Leu Trp Arg Asp Ile Tyr
 20 25 30
 tcc tgc gag ctc ctg tac ttc tcc att gct ttc gtc atc ctc aag ttt 144
 Ser Cys Glu Leu Leu Tyr Phe Ser Ile Ala Phe Val Ile Leu Lys Phe
 35 40 45
 acc ctt gcc gag ctc tgg gat tct gcc aaa aag att ctg cga gtg ctg 192
 Thr Leu Gly Glu Leu Ser Asp Ser Gly Lys Lys Ile Leu Arg Val Leu
 50 55 60
 ttc aag tgg tac aac ctc ttc atg tcc gcc ttt tcc ctg ctg gtg tcc ttc 240
 Phe Lys Trp Tyr Asn Leu Phe Met Ser Val Phe Ser Leu Val Ser Phe
 65 70 75 80
 ctc tgt atg ggt tac gcc atc tac acc gtt gga ctg tac tcc aac gaa 288
 Leu Cys Met Gly Tyr Ala Ile Tyr Thr Val Gly Leu Tyr Ser Asn Glu
 85 90 95
 tgc gac aga gct ttc gac aac agc ttg ttc cga ttt gcc acc aag gtc 336
 Cys Asp Arg Ala Phe Asp Asn Ser Leu Phe Arg Phe Ala Thr Lys Val
 100 105 110
 ttc tac tat tcc aag ttt ctg gag tac atc gac tct ttc tac ctt ccc 384
 Phe Tyr Tyr Ser Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro
 115 120 125
 ctc atg gcc aag cct ctg tcc ttt ctg cag ttc ttt cat cac ttg gga 432
 Leu Met Ala Lys Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly
 130 135 140

```

gct cct atg gac atg tgg ctg ttc gtc caq tac tct ggc gaa tcc att      480
Ala Pro Met Asp Met Trp Leu Phe Val Gln Tyr Ser Gly Glu Ser Ile
145          150          155          160

tgg atc ttt gtg ttc ctg aac gga ttc att cnc ttt gtc atg tac ggc      528
Trp Ile Phe Val Phe Leu Asn Gly Phe Ile His Phe Val Met Tyr Gly
165          170          175

tac tat tgg acg cgg ctg atg aag ttc aac ttt ccc atg ccc aag caq      576
Tyr Tyr Trp Thr Arg Leu Met Lys Phe Asn Phe Pro Met Pro Lys Gln
180          185          190

ctc att acc gca atg caq atc acc caq ttc aac gtt ggc ttc tac ctc      624
Leu Ile Thr Ala Met Gln Ile Thr Gln Phe Asn Val Gly Phe Tyr Leu
195          200          205

gtg tgg tgg tac aag gac att ccc tgt tac cga aag gat ccc atg cga      672
Val Trp Trp Tyr Lys Asp Ile Pro Cys Tyr Arg Lys Asp Pro Met Arg
210          215          220

atg ctg gcc tgg atc ttc aac tac tgg tac gtc ggt acc gtt ctt ctg      720
Met Leu Ala Trp Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu
225          230          235          240

ctc ttc atc aac ttc ttt gtc aag tcc tac gtc ttt ccc aag cct aag      768
Leu Phe Ile Asn Phe Phe Val Lys Ser Tyr Val Phe Pro Lys Pro Lys
245          250          255

act gcc gac aaa aag gtc cag tag      792
Thr Ala Asp Lys Lys Val Gln
260

```

<210> 50

<211> 263

<212> PRT

<213> Eutrepella sp. CCMP389

<400> 50

```

Met Ala Ala Val Ile Glu Val Ala Asn Glu Phe Val Ala Ile Thr Ala      1
1          5          10          15

Glu Thr Leu Pro Lys Val Asp Tyr Gln Arg Leu Trp Arg Asp Ile Tyr      20
20          25          30

Ser Cys Glu Leu Leu Tyr Phe Ser Ile Ala Phe Val Ile Leu Lys Phe      35
35          40          45

Thr Leu Gly Glu Leu Ser Asp Ser Gly Lys Lys Ile Leu Arg Val Leu      50
50          55          60

Phe Lys Trp Tyr Asn Leu Phe Met Ser Val Phe Ser Leu Val Ser Phe      65
65          70          75          80

Leu Cys Met Gly Tyr Ala Ile Tyr Thr Val Gly Leu Tyr Ser Asn Glu      85
85          90          95

Cys Asp Arg Ala Phe Asp Asn Ser Leu Phe Arg Phe Ala Thr Lys Val      100
100          105          110

Phe Tyr Tyr Ser Lys Phe Leu Glu Tyr Ile Asp Ser Phe Tyr Leu Pro      115
115          120          125

Leu Met Ala Lys Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly      130
130          135          140

Ala Pro Met Asp Met Trp Leu Phe Val Gln Tyr Ser Gly Glu Ser Ile      145
145          150          155          160

Trp Ile Phe Val Phe Leu Asn Gly Phe Ile His Phe Val Met Tyr Gly      165
165          170          175

Tyr Tyr Trp Thr Arg Leu Met Lys Phe Asn Phe Pro Met Pro Lys Gln      180
180          185          190

Leu Ile Thr Ala Met Gln Ile Thr Gln Phe Asn Val Gly Phe Tyr Leu      195
195          200          205

Val Trp Trp Tyr Lys Asp Ile Pro Cys Tyr Arg Lys Asp Pro Met Arg      210
210          215          220

Met Leu Ala Trp Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu      225
225          230          235          240

Leu Phe Ile Asn Phe Phe Val Lys Ser Tyr Val Phe Pro Lys Pro Lys      245
245          250          255

Thr Ala Asp Lys Lys Val Gln      260

```

<210> 51

<211> 774

<212> DNA

<213> Euglena anabaena UTEX373

<220>

<221> CDS

<222> (1)..(774)

<223> delta-9 elongase

<300>

<302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> U.S. 2008-0254522-A1

<311> 2008-04-15

<312> 2008-10-16

<313> (1)..(774)

<300>

<302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/128241

<311> 2008-04-16

<312> 2008-10-23

<313> (1)..(774)

<400> 51

```

atg gaa gca gcc aaa gaa ttg gtt tcc atc gtc caa gaq gaq ctc ccc      48
Met Glu Ala Ala Lys Glu Leu Val Ser Ile Val Gln Glu Glu Leu Pro
1      5      10      15

aag gtg gac tat gcc cag ctt tgg cag gat gcc agc agc tgt gaq gtc      96
Lys Val Asp Tyr Ala Gln Leu Trp Gln Asp Ala Ser Ser Cys Glu Val
20      25      30

ctt tac ctc tcg gtg gca ttc gtg cgc atc aag ttc atg ctg cgc cca      144
Leu Tyr Leu Ser Val Ala Phe Val Ala Ile Lys Phe Met Leu Arg Pro
35      40      45

ctg gac ctg aag cgc cag gcc acc ttg aag aag ctg ttc aca gca tac      192
Leu Asp Leu Lys Arg Gln Ala Thr Leu Lys Lys Leu Phe Thr Ala Tyr
50      55      60

aac ttc ctc atg tcg atc tat tcc ttt ggc tcc ttc ctg gcc atg gcc      240
Asn Phe Leu Met Ser Ile Tyr Ser Phe Gly Ser Phe Leu Ala Met Ala
65      70      75      80

tat gcc cta tca gta act ggc atc ctc tcc ggc gac tgt gaq acg gcg      288
Tyr Ala Leu Ser Val Thr Gly Ile Leu Ser Gly Asp Cys Glu Thr Ala
85      90      95

ttc aac aac gat gtg ttc agg atc aca act cag ctg ttc tac ctc agc      336
Phe Asn Asn Asp Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser
100      105      110

aag ttc gta gag tac atc gac tcc ttc tac ctt ccc ctt atg gac aag      384
Lys Phe Val Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Asp Lys
115      120      125

cca ctg tcg ttc ctt cag ttc ttc cat cat ttg ggg gcc ccc att gac      432
Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly Ala Pro Ile Asp
130      135      140

atg tgg cta ttc tac aaa tac cgc aac gaa gga gtc tgg atc ttt gtc      480
Met Trp Leu Phe Tyr Lys Tyr Arg Asn Glu Gly Val Trp Ile Phe Val
145      150      155      160

ctg ttg aat ggg ttc att cac tgg atc atg tac ggt tac tat tgg acg      528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
165      170      175

cgg ctc atc aag ctg aac ttc ccc atg ccc aag aac ctg atc acc tcc      576
Arg Leu Ile Lys Leu Asn Phe Pro Met Pro Lys Asn Leu Ile Thr Ser
180      185      190

atg cag atc atc cag ttc aat gtc ggg 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

cgc aat gtg cca tgc tao cgc cag gat ggc atg cgc atg ttt gcc tgg      672
Arg Asn Val Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Ala Trp
210      215      220

atc ttc aac tac tgg tat gtc ggg acg gtc ttg ctg ctg ttc ctc aac      720
Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu Leu Phe Leu Asn
225      230      235      240

ttt tac gtg cag acg tac atc cgg aag ccg agg aag aac cga ggg aag      768
Phe Tyr Val Gln Thr Tyr Ile Arg Lys Pro Arg Lys Asn Arg Gly Lys
245      250      255

aag gag
Lys Glu

<210> 52
<211> 258
<212> PRT
<213> Euglena anabaena UTEX373

<400> 52
Met Glu Ala Ala Lys Glu Leu Val Ser Ile Val Gln Glu Glu Leu Pro
1      5      10      15

Lys Val Asp Tyr Ala Gln Leu Trp Gln Asp Ala Ser Ser Cys Glu Val
20      25      30

Leu Tyr Leu Ser Val Ala Phe Val Ala Ile Lys Phe Met Leu Arg Pro
35      40      45

Leu Asp Leu Lys Arg Gln Ala Thr Leu Lys Lys Leu Phe Thr Ala Tyr
50      55      60

Asn Phe Leu Met Ser Ile Tyr Ser Phe Gly Ser Phe Leu Ala Met Ala
65      70      75      80

Tyr Ala Leu Ser Val Thr Gly Ile Leu Ser Gly Asp Cys Gln Thr Ala
85      90      95

Phe Asn Asn Asp Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser
100      105      110

Lys Phe Val Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Asp Lys
115      120      125

Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly Ala Pro Ile Asp
130      135      140

Met Trp Leu Phe Tyr Lys Tyr Arg Asn Glu Gly 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 Asn Phe Pro Met Pro Lys Asn 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 Val Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Ala Trp
210      215      220

Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu Leu Phe Leu Asn
225      230      235      240

Phe Tyr Val Gln Thr Tyr Ile Arg Lys Pro Arg Lys Asn Arg Gly Lys
245      250      255

Lys Glu

<210> 53
<211> 774
<212> DNA
<213> Euglena anabaena UTEX373

<220>
<221> CDS

```

<222> (1)..(774)

<223> synthetic delta-9 elongase (codon-optimized for Yarrowia lipolytica)

<300>

<302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> U.S. 2008-0254522-A1

<311> 2008-04-15

<312> 2008-10-16

<313> (1)..(774)

<300>

<302> DELTA-9 ELONGASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/128241

<311> 2008-04-16

<312> 2008-10-23

<313> (1)..(774)

<400> 53

```

atg gag gct gcc aag gag ctg gtc tcc atc gtc cag gag gaa ctt ccc      48
Met Glu Ala Ala Lys Glu Leu Val Ser Ile Val Gln Glu Glu Leu Pro
1      5      10      15

```

```

aag gtg gac tac gcc cag ctg tgg cag gac gcc tcc tct tgc gag gtt      96
Lys Val Asp Tyr Ala Gln Leu Trp Gln Asp Ala Ser Ser Cys Glu Val
20     25     30

```

```

ctg tac ctc tgc gtc gct ttc gtg gcc atc aag ttc atg ctt cga cct      144
Leu Tyr Leu Ser Val Ala Phe Val Ala Ile Lys Phe Met Leu Arg Pro
35     40     45

```

```

ctg gac ctc aag cga caa gcc acc ctc aaa aag ctg ttc acc gca tac      192
Leu Asp Leu Lys Arg Gln Ala Thr Leu Lys Lys Leu Phe Thr Ala Tyr
50     55     60

```

```

aac ttt ctc atg tcc atc tac tgc ttc gcc tcc ttc ctg gag atg gcc      240
Asn Phe Leu Met Ser Ile Tyr Ser Phe Gly Ser Phe Leu Ala Met Ala
65     70     75     80

```

```

tac gct ctc tct gtc act ggt att ctt tcc ggc gat tgt gag act gcc      288
Tyr Ala Leu Ser Val Thr Gly Ile Leu Ser Gly Asp Cys Glu Thr Ala
85     90     95

```

```

ttc aac aat gac gtg ttc cga atc acc act cag ctg ttc tac ctc agc      336
Phe Asn Asn Asp Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser
100    105    110

```

```

aag ttc gtc gag tac atc gac tcc ttc tac ctt ccc ctc atg gac aag      384
Lys Phe Val Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Asp Lys
115    120    125

```

```

ccc ttg tgc ttt ctg cag ttc ttc cac cat ctc gga gct ccc atc gac      432
Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly Ala Pro Ile Asp
130    135    140

```

```

atg tgg ctg ttc tac aag tat cga aac gaa gcc gtc tgg atc ttt gtt      480
Met Trp Leu Phe Tyr Lys Tyr Arg Asn Glu Gly Val Trp Ile Phe Val
145    150    155    160

```

```

ctg ctc aac gcc ttc att cac tgg atc atg tac ggt tac tat tgg acg      528
Leu Leu Asn Gly Phe Ile His Trp Ile Met Tyr Gly Tyr Tyr Trp Thr
165    170    175

```

```

cga ctc atc aag ctg aac ttc cct atg ccc aag aac ctc att acc tcc      576
Arg Leu Ile Lys Leu Asn Phe Pro Met Pro Lys Asn Leu Ile Thr Ser
180    185    190

```

```

atg caa att atc cag ttc aac gtc gga 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

```

```

cga aac gtg ccc tgc tac cgg cag gac ggt atg cga atg ttt gcc tgg      672
Arg Asn Val Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Ala Trp
210    215    220

```

```

atc ttc aac tac tgg tat gtc gcc acg gtg ctg ctt ctg ttc ctc aac      720
Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu Phe Leu Asn
225    230    235    240

```

```

ttc tac gtc cag acc tcc att cgg aag cct cga aag aac cga gcc aaa      768
Phe Tyr Val Gln Thr Tyr Ile Arg Lys Pro Arg Lys Asn Arg Gly Lys
245    250    255

```

```

aag gag
Lys Glu
774

```

<210> 54

<211> 258

<212> PRT

<213> Euglena anabaena UTEX373

<400> 54

```

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

```

```

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

```

```

20          25          30
Leu Tyr Leu Ser Val Ala Phe Val Ala Ile Lys Phe Met Leu Arg Pro
35          40          45

Leu Asp Leu Lys Arg Gln Ala Thr Leu Lys Lys Leu Phe Thr Ala Tyr
50          55          60

Asn Phe Leu Met Ser Ile Tyr Ser Phe Gly Ser Phe Leu Ala Met Ala
65          70          75          80

Tyr Ala Leu Ser Val Thr Gly Ile Leu Ser Gly Asp Cys Glu Thr Ala
85          90          95

Phe Asn Asn Asp Val Phe Arg Ile Thr Thr Gln Leu Phe Tyr Leu Ser
100         105         110

Lys Phe Val Glu Tyr Ile Asp Ser Phe Tyr Leu Pro Leu Met Asp Lys
115         120         125

Pro Leu Ser Phe Leu Gln Phe Phe His His Leu Gly Ala Pro Ile Asp
130         135         140

Met Trp Leu Phe Tyr Lys Tyr Arg Asn Glu Gly 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 Asn Phe Pro Met Pro Lys Asn 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 Val Pro Cys Tyr Arg Gln Asp Gly Met Arg Met Phe Ala Trp
210         215         220

Ile Phe Asn Tyr Trp Tyr Val Gly Thr Val Leu Leu Leu Phe Leu Asn
225         230         235         240

Phe Tyr Val Gln Thr Tyr Ile Arg Lys Pro Arg Lys Asn Arg Gly Lys
245         250         255

Lys Glu

<210> 55
<211> 1271
<212> DNA
<213> Euglena gracilis

<220>
<221> CDS
<222> (4)..(1269)
<223> delta-8 desaturase ("Eg5" or "EgD8")

<300>
<302> DELTA-8 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2006/012325 and WO 2006/012326
<311> 2005-06-24
<312> 2006-02-02
<313> (1)..(1271)

<300>
<302> DELTA-8 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> U.S. 7,256,033
<311> 2005-06-24
<312> 2007-08-14
<313> (1)..(1271)

<400> 55
gaa atg aag tca aag cgc caa gcg ctt ccc ctt aca att gat gga aca      48
Met Lys Ser Lys Arg Gln Ala Leu Pro Leu Thr Ile Asp Gly Thr
1          5          10          15

aca tat gat gtg tct gcc tgg gtc aat ttc cac cct ggt ggt gcg gaa      96
Thr Tyr Asp Val Ser Ala Trp Val Asn Phe His Pro Gly Gly Ala Glu
20         25         30

att ata gag aat tac caa gga agg gat gcc act gat gcc ttc atg gtt      144
Ile Ile Glu Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val
35         40         45

atg cac tct caa gaa gcc ttc gac aag ctc aag cgc atg ccc aaa atc      192
Met His Ser Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile
50         55         60

aat ccc agt tct gag ttg cca ccc cag gct gca gtg aat gaa gct caa      240
Asn Pro Ser Ser Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln
65         70         75

gag gat ttc cgg aag ctc cga gaa gag ttg atc gca act ggc atg ttt      288
Glu Asp Phe Arg Lys Leu Arg Glu Glu Leu Ile Ala Thr Gly Met Phe
80         85         90         95

gat gcc tcc ccc ctc tgg tac tca tac aaa atc agc acc acc ctg ggc      336
Asp Ala Ser Pro Leu Trp Tyr Ser Tyr Lys Ile Ser Thr Thr Leu Gly
100        105        110

ctt gga gtg ctg ggt tat ttc ctg atg gtt cag tat cag atg tat ttc      384
Leu Gly Val Leu Gly Tyr Phe Leu Met Val Gln Tyr Gln Met Tyr Phe
115        120        125

att ggg gca atg ttg ctt ggg atg cac tat caa cag atg ggc tgg ctt      432
Ile Gly Ala Val Leu Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu
130        135        140

tct cat gac att tgc cao cac cag act ttc aag aac cgg aac tgg aac      480
Ser His Asp Ile Cys His Gln Thr Phe Lys Asn Arg Asn Trp Asn
145        150        155

```

aac ctc gtc gga ctg gta ttt ggc aat ggt ctg caa ggt ttt tcc gtc 528
 Asn Leu Val Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val
 160 165 170 175
 aca tgg tgg aag gac aga cac aat gca cat cat tgg gca acc aat gtt 576
 Thr Trp Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val
 180 185 190
 caa ggg cac gac cct gat att gac aac ctc ccc ctc tta gcc tgg tct 624
 Gln Gly His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser
 195 200 205
 gag gat gac gtc aca cgg ggc tca ccg att tcc cgc aag ctc att cag 672
 Glu Asp Asp Val Thr Arg Ala Ser Pro Ile Ser Arg Lys Leu Ile Gln
 210 215 220
 ttc cag cag tac tat ttc ttg gtc atc tgt atc ttg ttg cgg ttc att 720
 Phe Gln Gln Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile
 225 230 235
 tgg tgt ttc cag agc gtc ttg acc gtc cgc agt ttg aag gac aga gat 768
 Trp Cys Phe Gln Ser Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp
 240 245 250 255
 aac caa ttc tat cgc tct cag tat aag aag gag gcc att ggc ctc gcc 816
 Asn Gln Phe Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala
 260 265 270
 ctg cac tgg acc ttg aag acc ctg ttc caa tta ttc ttt atg ccc agc 864
 Leu His Trp Thr Leu Lys Thr Leu Phe His Leu Phe Phe Met Pro Ser
 275 280 285
 atc ctc aca tgg ctg ttg gtc ttt ttc gtt tgg gag ctg gtt ggc ggc 912
 Ile Leu Thr Ser Leu Leu Val Phe Phe Val Ser Gln Leu Val Gly Gly
 290 295 300
 ttc ggc att ggc atc gtc gtc ttc atg aac cac tac caa ctg gag aag 960
 Phe Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys
 305 310 315
 atc ggg gac tca gtc tgg gat ggc cat cga ttc tgg gtt ggc cag atc 1008
 Ile Gly Asp Ser Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile
 320 325 330 335
 cat gag acc atg aac att cgg cga ggg att atc aca gat tgg ttt ttc 1056
 His Glu Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe
 340 345 350
 gga ggc ttg aat tac cag att gag cac cat ttg tgg ccg acc ctc cct 1104
 Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro
 355 360 365
 cgc cac aac ctg aca cgc gtt agc tac cag ctg gaa cag ctg tgc cag 1152
 Arg His Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln
 370 375 380
 aag cac aac ctg ccg tat cgg aac ccg ctg ccc cac gaa ggg ttg gtc 1200
 Lys His Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val
 385 390 395
 atc ctg ctg cgc tat ctg ggc gtc ttc gcc cgg atg ggc gag aag caa 1248
 Ile Leu Leu Arg Tyr Leu Ala Val Phe Ala Arg Met Ala Glu Lys Gln
 400 405 410 415
 ccc ggc ggg aag gct cta taa gg 1271
 Pro Ala Gly Lys Ala Leu
 420
 <210> 56
 <211> 421
 <212> PRT
 <213> Euglena gracilis
 <400> 56
 Met Lys Ser Lys Arg Gln Ala Leu Pro Leu Thr Ile Asp Gly Thr Thr
 1 5 10 15
 Tyr Asp Val Ser Ala Trp Val Asn Phe His Pro Gly Gly Ala Glu Ile
 20 25 30
 Ile Glu Asn Tyr Gln Gly Arg Asp Ala Thr Asp Ala Phe Met Val Met
 35 40 45
 His Ser Gln Glu Ala Phe Asp Lys Leu Lys Arg Met Pro Lys Ile Asn
 50 55 60
 Pro Ser Ser Glu Leu Pro Pro Gln Ala Ala Val Asn Glu Ala Gln Glu
 65 70 75 80
 Asp Phe Arg Lys Leu Arg Glu Glu Leu Ile Ala Thr Gly Met Phe Asp
 85 90 95
 Ala Ser Pro Leu Trp Tyr Ser Tyr Lys Ile Ser Thr Thr Leu Gly Leu
 100 105 110
 Gly Val Leu Gly Tyr Phe Leu Met Val Gln Tyr Gln Met Tyr Phe Ile
 115 120 125
 Gly Ala Val Leu Leu Gly Met His Tyr Gln Gln Met Gly Trp Leu Ser
 130 135 140
 His Asp Ile Cys His His Gln Thr Phe Lys Asn Arg Asn Trp Asn Asn
 145 150 155 160
 Leu Val Gly Leu Val Phe Gly Asn Gly Leu Gln Gly Phe Ser Val Thr
 165 170 175
 Trp Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn Val Gln
 180 185 190
 Gly His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp Ser Glu
 195 200 205
 Asp Asp Val Thr Arg Ala Ser Pro Ile Ser Arg Lys Leu Ile Gln Phe


```

210          215          220
Gln Gln Tyr Tyr Phe Leu Val Ile Cys Ile Leu Leu Arg Phe Ile Trp
225          230          235          240
Cys Phe Gln Ser Val Leu Thr Val Arg Ser Leu Lys Asp Arg Asp Asn
245          250          255
Gln Phe Tyr Arg Ser Gln Tyr Lys Lys Glu Ala Ile Gly Leu Ala Leu
260          265          270
His Trp Thr Leu Lys Thr Leu Phe His Leu Phe Phe Met Pro Ser Ile
275          280          285
Leu Thr Ser Leu Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly Phe
290          295          300
Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys Ile
305          310          315          320
Gly Asp Ser Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile His
325          330          335
Glu Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe Gly
340          345          350
Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro Arg
355          360          365
His Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln Lys
370          375          380
His Asn Leu Pro Tyr Arg Asn Pro Leu Pro Pro His Glu Gly Leu Val Ile
385          390          395          400
Leu Leu Arg Tyr Leu Ala Val Phe Ala Arg Met Ala Glu Lys Gln Pro
405          410          415
Ala Gly Lys Ala Leu
420
<210> 57
<211> 1272
<212> DNA
<213> Euglena gracilis
<220>
<221> CDS
<222> (2)..(1270)
<223> synthetic delta-8 desaturase (codon-optimized for Yarrowia lipolytica) ("D8SF" or "EgD8S")
<300>
<302> DELTA-8 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY ACIDS
<310> WO 2006/012325 and WO 2006/012326
<311> 2005-06-24
<312> 2006-02-02
<313> (1)..(1272)
<300>
<302> DELTA-8 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY ACIDS
<310> U.S. 7,256,033
<311> 2005-06-24
<312> 2007-08-14
<313> (1)..(1272)
<400> 57
c atg gtg aag tcc aag cga cag gct ctg ccc ctc acc atc gac gga act      49
Met Val Lys Ser Lys Arg Gln Ala Leu Pro Leu Thr Ile Asp Gly Thr
1      5      10      15
acc tac gac gtc tcc gct tgg gtg aac ttc cac cct ggt gga gct gaa      97
Thr 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 ccc tcc tcc gag ctg cct ccc cag gct gcc gtc aac gaa gct cag      241
Asn Pro Ser 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 gcc 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 tgg tac aag atc tcc acc acc ctg ggt      337
Asp Ala Ser Pro Leu Trp Tyr Ser Tyr Lys Ile Ser Thr Thr Leu Gly
100    105    110
ctt gcc gtg ctt gga tac ttc ctg atg gtc cag tac cag atg tac ttc      385
Leu Gly Val Leu Gly Tyr Phe Leu Met Val Gln Tyr Gln Met Tyr Phe
115    120    125
att ggt gct gtg ctg ctc ggt atg cac tac cag caa atg gga tgg ctg      433
Ile Gly Ala Val Leu 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 ctc gtg ggt ctg gtc ttt gcc aac gga ctc cag gcc ttc tcc gtg      529
Asn Leu Val 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

```

160	185	190	
oag 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 ctc cag tct gtc ctc 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 cga 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 tgg ctc ctg gtc ttc ttt gct tcc gag ctc gtc ggt ggc			913
Ile Leu Thr Ser Leu Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly			
290 295 300			
ttc gga att gcc atc ctg 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 gcc ttc tct gtc 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 gcc 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 gcc 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 gcc gtc ttc gct cga atg gcc gag aag cag			1249
Ile Leu Leu Arg Tyr Leu Ala Val Phe Ala Arg Met Ala Glu Lys Gln			
405 410 415			
ccc gct ggc aag gct ctc taa gc			1272
Pro Ala Gly Lys Ala Leu			
420			
<210> 58			
<211> 422			
<212> PRT			
<213> Euglena gracilis			
<400> 58			
Met Val Lys Ser Lys Arg Gln Ala Leu Pro Leu Thr Ile Asp Gly Thr			
1 5 10 15			
Thr 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 Pro Ser Ser Glu Leu Pro Pro Gln Ala Ala Val Asn Gln 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 Ser Thr Thr Leu Gly			
100 105 110			
Leu Gly Val Leu Gly Tyr Phe Leu Met Val Gln Tyr Gln Met Tyr Phe			
115 120 125			
Ile Gly Ala Val Leu 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 Leu Val 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 Leu Leu Val Phe Phe Val Ser Glu Leu Val Gly Gly
 290 295 300
 Phe Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys
 305 310 315 320
 Ile Gly Asp Ser Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile
 325 330 335
 His Glu Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe
 340 345 350
 Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro
 355 360 365
 Arg His Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln
 370 375 380
 Lys His Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val
 385 390 395 400
 Ile Leu Leu Arg Tyr Leu Ala Val Phe Ala Arg Met Ala Glu Lys Gln
 405 410 415
 Pro Ala Gly Lys Ala Leu
 420

<210> 59

<211> 1272

<212> DNA

<213> Artificial Sequence

<220>

<223> mutant EgD8S-23 delta-8 desaturase (also "EgD8M")

<220>

<221> CDS

<222> (2)..(1270)

<300>

<302> MUTANT DELTA-8 DESATURASE GENES ENGINEERED BY TARGETED MUTAGENESIS AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/073271

<311> 2007-12-05

<312> 2008-06-19

<313> (1)..(1272)

<300>

<302> MUTANT DELTA-8 DESATURASE GENES ENGINEERED BY TARGETED MUTAGENESIS AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US 2008-0138868-A1

<311> 2006-12-07

<312> 2008-06-12

<313> (1)..(1272)

<400> 59

o 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 oct 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 tgg 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 ctg 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 ccc 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 ccc 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 tgg atg ctg gtg ttc ttt gtt tcc gag ctc gtc ggt ggc 913
 Ile Leu Thr Ser 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 ccc taa 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 gcc 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 gcc 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 taa 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> 60
 <211> 422
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Synthetic Construct
 <400> 60
 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

Phe Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu Lys
305          310          315          320

Ile Gly Asp Ser Val Trp Asp Gly His Gly Phe Ser Val Gly Gln Ile
325          330          335

His Glu Thr Met Asn Ile Arg Arg Gly Ile Ile Thr Asp Trp Phe Phe
340          345          350

Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu Pro
355          360          365

Arg His Asn Leu Thr Ala Val Ser Tyr Gln Val Glu Gln Leu Cys Gln
370          375          380

Lys His Asn Leu Pro Tyr Arg Asn Pro Leu Pro His Glu Gly Leu Val
385          390          395          400

Ile Leu Leu Arg Tyr Leu Ser Gln Phe Ala Arg Met Ala Glu Lys Gln
405          410          415

Pro Gly Ala Lys Ala Gln
420

<210> 61
<211> 1260
<212> DNA
<213> Euglena anabaena UTEX373

<220>
<221> CDS
<222> (1)..(1260)
<223> delta-8 desaturase

<300>
<302> DELTA-8 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> U.S. 2008-0254521-A1
<311> 2008-04-09
<312> 2008-10-16
<313> (1)..(1260)

<300>
<302> DELTA-8 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124194
<311> 2008-04-10
<312> 2008-10-16
<313> (1)..(1260)
<400> 61
atg gtg aaa agg cca gca ctt ccg ctg acc gtt gat ggt gtc acc tat      48
Met Val Lys Arg Pro Ala Leu Pro Leu Thr Val Asp Gly Val Thr Tyr
1          5          10          15

gat gtg tct gcc tgg ttg aac cat cca ggg ggt gct gac atc att      96
Asp Val Ser Ala Trp Leu Asn His His Pro Gly Gly Ala Asp Ile Ile
20          25          30

gag aac tac cgc ggt cgt gat gcc act gat gtc ttt atg gtt atg cac      144
Glu Asn Tyr Arg Gly Arg Asp Ala Thr Asp Val Phe Met Val Met His
35          40          45

tct gaa aat gct gtg agt aaa cta aga agg atg cct atc atg gaa cca      192
Ser Glu Asn Ala Val Ser Lys Leu Arg Arg Met Pro Ile Met Glu Pro
50          55          60

tca tct cca ctg acg cct acg cca ccg aaa ccc aac tca gac gaa ccg      240
Ser Ser Pro Leu Thr Pro Thr Pro Pro Lys Pro Asn Ser Asp Glu Pro
65          70          75          80

cag gag gat ttc cgc aag ctg cga gat gag ctg atc gca gca gga atg      288
Gln Glu Asp Phe Arg Lys Leu Arg Asp Glu Leu Ile Ala Ala Gly Met
85          90          95

ttc gac gca tca ccg atg tgg tac gca tat aag acg ctg agt acg ctg      336
Phe Asp Ala Ser Pro Met Trp Tyr Ala Tyr Lys Thr Leu Ser Thr Leu
100          105          110

ggc ctg ggg gtc ctg ccg gtg cta ttg atg acc cag tgg cac tgg tac      384
Gly Leu Gly Val Leu Ala Val Leu Leu Met Thr Gln Trp His Trp Tyr
115          120          125

ctc gtc ggg gca atc gtg ttg ggc att cac ttc caa caa atg ggt tgg      432
Leu Val Gly Ala Ile Val Leu Gly Ile His Phe Gln Gln Met Gly Trp
130          135          140

ttg tog cac gat atc tgc cac cat cag ctg ttc aag gac cga tgc atc      480
Leu Ser His Asp Ile Cys His His Gln Leu Phe Lys Asp Arg Ser Ile
145          150          155          160

aac aac gcc atc ggc ttg ctt ttc ggg aac gtc ttg caa ggg ttc tct      528
Asn Asn Ala Ile Gly Leu Leu Phe Gly Asn Val Leu Gln Gly Phe Ser
165          170          175

gtg acc tgg tgg aag gac agg cac aat gca csc cac toc gcc acc aac      576
Val Thr Trp Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn
180          185          190

gtg caa gcc cac gac ccc gac att gac aac ctg ccg ctg gca tgg      624
Val Gln Gly His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp
195          200          205

tcc aag gag gac gtg gag agg gcc ggc ccg ttc tca cgg cgg atg atc      672
Ser Lys Glu Asp Val Glu Arg Ala Gly Pro Phe Ser Arg Arg Met Ile
210          215          220

aag tac cag caa tac tac ttc ttc atc tgt gcc ctg agg ttc      720

```

Lys Tyr Gln Gln Tyr Tyr Phe Phe Phe Ile Cys Ala Leu Leu Arg Phe
 225 230 235 240
 atc tgg tgc ttc cag agc atc cac aca gcc acg ggc ctg aag gat cgc 768
 ile Trp Cys Phe Gln Ser Ile His Thr Ala Thr Gly Leu Lys Asp Arg
 245 250 255
 agc aac cag tac tac cgc agc cag tac gag aca cag agc ctg ggc ctg 816
 Ser Asn Gln Tyr Tyr Arg Arg Gln Tyr Glu Lys Glu Ser Val Gly Leu
 260 265 270
 gcc ctc cac tgg ggc ctg aag cgc ttg ttc tac tac ttt tat atg cca 864
 Ala Leu His Trp Gly Leu Lys Ala Leu Phe Tyr Tyr Phe Tyr Met Pro
 275 280 285
 agc ttc ttg aac gga ctc atg gtg ttt ttc gtg tcc gag ttg ctt ggg 912
 Ser Phe Leu Thr Ser Leu Met Val Phe Val Ser Glu Leu Leu Gly
 290 295 300
 gcc ttc gcc atc gcc atc gtg gtg ttc atg aac cac tac ccc ctg gag 960
 Gly Phe Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu
 305 310 315 320
 aag atc cag gac tgg gtg tgg gag gcc cac gcc ttt tgc gcc ggc cag 1008
 Lys Ile Gln Asp Ser Val Trp Asp Gly His Gly Phe Cys Ala Gly Gln
 325 330 335
 att cac gaa acg atg aac gtc cag cgc gga ctc gtc acg gac tgg ttc 1056
 Ile His Glu Thr Met Asn Val Gln Arg Gly Leu Val Thr Asp Trp Phe
 340 345 350
 ttc ggt ggg ctg aat tac caa atc gag cac cac ctg tgg ccg acg ctg 1104
 Phe Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu
 355 360 365
 ccc cgg cac aac ctg acg ggc gcc agc atc aaa gtg gag cag ttg tgc 1152
 Pro Arg His Asn Leu Thr Ala Ala Ser Ile Lys Val Glu Gln Leu Cys
 370 375 380
 aag aag cac aac ttg ccg tat cgc agc ccc cca atg ctg gag ggg gtg 1200
 Lys Lys His Asn Leu Pro Tyr Arg Ser Pro Met Leu Glu Gly Val
 385 390 395 400
 gcc atc ctg atc agc tac ctg ggc acc ttt gcc cgc atg gtg gca aag 1248
 Gly Ile Leu Ile Ser Tyr Leu Gly Thr Phe Ala Arg Met Val Ala Lys
 405 410 415
 gcc gac aag ggc 1260
 Ala Asp Lys Ala
 420
 <210> 62
 <211> 420
 <212> PRT
 <213> Euglena anabaena UTEX373
 <400> 62
 Met Val Lys Arg Pro Ala Leu Pro Leu Thr Val Asp Gly Val Thr Tyr
 1 5 10 15
 Asp Val Ser Ala Trp Leu Asn His His Pro Gly Gly Ala Asp Ile Ile
 20 25 30
 Glu Asn Tyr Arg Gly Arg Asp Ala Thr Asp Val Phe Met Val Met His
 35 40 45
 Ser Glu Asn Ala Val Ser Lys Leu Arg Arg Met Pro Ile Met Glu Pro
 50 55 60
 Ser Ser Pro Leu Thr Pro Thr Pro Pro Lys Pro Asn Ser Asp Glu Pro
 65 70 75 80
 Gln Glu Asp Phe Arg Lys Leu Arg Asp Glu Leu Ile Ala Ala Gly Met
 85 90 95
 Phe Asp Ala Ser Pro Met Trp Tyr Ala Tyr Lys Thr Leu Ser Thr Leu
 100 105 110
 Gly Leu Gly Val Leu Ala Val Leu Leu Met Thr Gln Trp His Trp Tyr
 115 120 125
 Leu Val Gly Ala Ile Val Leu Gly Ile His Phe Gln Gln Met Gly Trp
 130 135 140
 Leu Ser His Asp Ile Cys His His Gln Leu Phe Lys Asp Arg Ser Ile
 145 150 155 160
 Asn Asn Ala Ile Gly Leu Leu Phe Gly Asn Val Leu Gln Gly Phe Ser
 165 170 175
 Val Thr Trp Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn
 180 185 190
 Val Gln Gly His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp
 195 200 205
 Ser Lys Glu Asp Val Glu Arg Ala Gly Pro Phe Ser Arg Arg Met Ile
 210 215 220
 Lys Tyr Gln Gln Tyr Tyr Phe Phe Phe Ile Cys Ala Leu Leu Arg Phe
 225 230 235 240
 Ile Trp Cys Phe Gln Ser Ile His Thr Ala Thr Gly Leu Lys Asp Arg
 245 250 255
 Ser Asn Gln Tyr Tyr Arg Arg Gln Tyr Glu Lys Glu Ser Val Gly Leu
 260 265 270
 Ala Leu His Trp Gly Leu Lys Ala Leu Phe Tyr Tyr Phe Tyr Met Pro
 275 280 285

```

Ser Phe Leu Thr Gly Leu Met Val Phe Phe Val Ser Glu Leu Leu Gly
290          295          300

Gly Phe Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu
305          310          315          320

Lys Ile Gln Asp Ser Val Trp Asp Gly His Gly Phe Cys Ala Gly Gln
325          330          335

Ile His Glu Thr Met Asn Val Gln Arg Gly Leu Val Thr Asp Trp Phe
340          345          350

Phe Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu
355          360          365

Pro Arg His Asn Leu Thr Ala Ala Ser Ile Lys Val Glu Gln Leu Cys
370          375          380

Lys Lys His Asn Leu Pro Tyr Arg Ser Pro Pro Met Leu Glu Gly Val
385          390          395          400

Gly Ile Leu Ile Ser Tyr Leu Gly Thr Phe Ala Arg Met Val Ala Lys
405          410          415

Ala Asp Lys Ala
420

<210> 63
<211> 1260
<212> DNA
<213> Euglena anabaena UTEX 373

<220>
<221> CDS
<222> (1)..(1260)
<223> synthetic delta-8 desaturase (codon-optimized for Yarrowia lipolytica)

<300>
<302> DELTA-8 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> U.S. 2008-0254521-A1
<311> 2008-04-09
<312> 2008-10-16
<313> (1)..(1260)

<300>
<302> DELTA-8 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/124194
<311> 2008-04-10
<312> 2008-10-16
<313> (1)..(1260)
<400> 63
atg gtc aag cga ccc gct ctg cct cto acc gtg gac ggt gtc aac tac      48
Met Val Lys Arg Pro Ala Leu Pro Leu Thr Val Asp Gly Val Thr Tyr
1          5          10          15

gac gtt tct gcc tgg ctc aac cac cat ccc gga ggt gcc gac att atc      96
Asp Val Ser Ala Trp Leu Asn His His Pro Gly Gly Ala Asp Ile Ile
20          25          30

gag aac tac cga ggt cgg gat gct acc gac gtc ttc atg gtt atg aac      144
Glu Asn Tyr Arg Gly Arg Asp Ala Thr Asp Val Phe Met Val Met His
35          40          45

tcc gag aac gcc gtg tcc aac ctc aga cga atg ccc atc atg gaa cct      192
Ser Glu Asn Ala Val Ser Lys Leu Arg Arg Met Pro Ile Met Glu Pro
50          55          60

tcc tct ccc ctg aat cca aca cct ccc aag cca aac tcc gac gaa cct      240
Ser Ser Pro Leu Thr Pro Thr Pro Pro Lys Pro Asn Ser Asp Glu Pro
65          70          75          80

cag gag gat ttc cga aag ctg cga gac gag ctc att gct gca gcc atg      288
Gln Glu Asp Phe Arg Lys Leu Arg Asp Glu Leu Ile Ala Ala Gly Met
85          90          95

ttc gat gcc tct ccc atg tgg tac gct tac aag acc ctg tcy act ctc      336
Phe Asp Ala Ser Pro Met Trp Tyr Ala Tyr Lys Thr Leu Ser Thr Leu
100          105          110

gga ctg ggt gtc ctt gcc ctg ctg ttg atg acc caa tgg aac tgg tac      384
Gly Leu Gly Val Leu Ala Val Leu Leu Met Thr Gln Trp His Trp Tyr
115          120          125

ctg gtt ggt gct atc gtc ctc gcc att cac ttt caa aag atg gga tgg      432
Leu Val Gly Ala Ile Val Leu Gly Ile His Phe Gln Gln Met Gly Trp
130          135          140

cto tgg ccc gac att tgc cat cac cag ctg ttc aag gac cga too atc      480
Leu Ser His Asp Ile Cys His His Gln Leu Phe Lys Asp Arg Ser Ile
145          150          155          160

aac aat gcc att gcc ctg cto ttc gga aac gtg ctt cag gcc ttt tct      528
Asn Asn Ala Ile Gly Leu Leu Phe Gly Asn Val Leu Gln Gly Phe Ser
165          170          175

gtc act tgg tgg aag gac cga ccc aac gct cat cac tcc gcc acc aac      576
Val Thr Trp Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn
180          185          190

gtg cag ggt cac gat ccc gac atc gac aac ctg cct ctc ctg gcg tgg      624
Val Gln Gly His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp
195          200          205

tcc aag gag gac gtc gag cga gct gcc ccg ttt tct cga cgg atg atc      672
Ser Lys Glu Asp Val Glu Arg Ala Gly Pro Phe Ser Arg Arg Met Ile
210          215          220

aag tac caa cag tat tac ttc ttt ttc atc tgt gcc ctt ctg cga ttc      720
Lys Tyr Gln Gln Tyr Tyr Phe Phe Phe Ile Cys Ala Leu Leu Arg Phe
225          230          235          240

ato tgg tgc ttt cag too att aat aat gcc aag ggt ctc aag gat cga      768
Ile Trp Cys Phe Gln Ser Ile His Thr Ala Thr Gly Leu Lys Asp Arg
245          250          255

```

```

agc aat cag tac tat cga aga cag tac gag aag gaq tcc gtc ggt ctg      816
Ser Asn Gln Tyr Tyr Arg Arg Gln Tyr Glu Lys Glu Ser Val Gly Leu
      260      265      270

gca ctc cac tgg ggt ctc aag gcc ttg ttc tac tat ttc tac atg ccc      864
Ala Leu His Trp Gly Leu Lys Ala Leu Phe Tyr Tyr Phe Tyr Met Pro
      275      280      285

tcg ttt ctc acc gga ctc atg gtg ttc ttt gtc tcc gag ctg ctt ggt      912
Ser Phe Leu Thr Gly Leu Met Val Phe Phe Val Ser Glu Leu Leu Gly
      290      295      300

ggc ttc gga att gcc atc gtt gtc ttc atg aac cac tac cct ctg gag      960
Gly Phe Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu
      305      310      315      320

aag att cag cac tcc gtg tgg gat ggt cat gcc ttc tgt gct gga cag      1008
Lys Ile Gln Asp Ser Val Trp Asp Gly His Gly Phe Cys Ala Gly Gln
      325      330      335

att cac gag acc atg aac gtt cag cga gcc ctc ctc aca gac tgg ttt      1056
Ile His Glu Thr Met Asn Val Gln Arg Gly Leu Val Thr Asp Trp Phe
      340      345      350

ttc ggt gcc ctc aac tac cag atc gaa cat cac ctg tgg cct act ctt      1104
Phe Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu
      355      360      365

ccc aqa cac aac ctc acc gct gcc tcc atc aaa gtg gag cag ctg tgc      1152
Pro Arg His Asn Leu Thr Ala Ala Ser Ile Lys Val Glu Gln Leu Cys
      370      375      380

aag aag cac aac ctg ccc tac cga tct cct ccc atg ctc gaa ggt gtc      1200
Lys Lys His Asn Leu Pro Tyr Arg Ser Pro Pro Met Leu Glu Gly Val
      385      390      395      400

ggc att ctt atc tcc tac ctg gcc acc ttc gct cga atg gtt gcc aag      1248
Gly Ile Leu Ile Ser Tyr Leu Gly Thr Phe Ala Arg Met Val Ala Lys
      405      410      415

gca gac aag gcc      1260
Ala Asp Lys Ala
      420

<210> 64
<211> 420
<212> PRT
<213> Euglena anabaena UTEX373

<400> 64
Met Val Lys Arg Pro Ala Leu Pro Leu Thr Val Asp Gly Val Thr Tyr
1      5      10      15

Asp Val Ser Ala Trp Leu Asn His His Pro Gly Gly Ala Asp Ile Ile
20      25      30

Glu Asn Tyr Arg Gly Arg Asp Ala Thr Asp Val Phe Met Val Met His
35      40      45

Ser Glu Asn Ala Val Ser Lys Leu Arg Arg Met Pro Ile Met Glu Pro
50      55      60

Ser Ser Pro Leu Thr Pro Thr Pro Pro Lys Pro Asn Ser Asp Glu Pro
65      70      75      80

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

Phe Asp Ala Ser Pro Met Trp Tyr Ala Tyr Lys Thr Leu Ser Thr Leu
100      105      110

Gly Leu Gly Val Leu Ala Val Leu Leu Met Thr Gln Trp His Trp Tyr
115      120      125

Leu Val Gly Ala Ile Val Leu Gly Ile His Phe Gln Gln Met Gly Trp
130      135      140

Leu Ser His Asp Ile Cys His His Gln Leu Phe Lys Asp Arg Ser Ile
145      150      155      160

Asn Asn Ala Ile Gly Leu Leu Phe Gly Asn Val Leu Gln Gly Phe Ser
165      170      175

Val Thr Trp Trp Lys Asp Arg His Asn Ala His His Ser Ala Thr Asn
180      185      190

Val Gln Gly His Asp Pro Asp Ile Asp Asn Leu Pro Leu Leu Ala Trp
195      200      205

Ser Lys Glu Asp Val Glu Arg Ala Gly Pro Val Ser Arg Arg Met Ile
210      215      220

Lys Tyr Gln Gln Tyr Tyr Phe Phe Phe Ile Cys Ala Leu Leu Arg Phe
225      230      235

Ile Trp Cys Phe Gln Ser Ile His Thr Ala Thr Gly Leu Lys Asp Arg
245      250      255

Ser Asn Gln Tyr Tyr Arg Arg Gln Tyr Glu Lys Glu Ser Val Gly Leu
260      265      270

Ala Leu His Trp Gly Leu Lys Ala Leu Phe Tyr Tyr Phe Tyr Met Pro
275      280      285

Ser Phe Leu Thr Gly Leu Met Val Phe Phe Val Ser Glu Leu Leu Gly
290      295      300

Gly Phe Gly Ile Ala Ile Val Val Phe Met Asn His Tyr Pro Leu Glu
305      310      315      320

```


Lys Ile Gln Asp Ser Val Trp Asp Gly His Gly Phe Cys Ala Gly Gln
 325 330 335
 Ile His Glu Thr Met Asn Val Gln Arg Gly Leu Val Thr Asp Trp Phe
 340 345 350
 Phe Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu Trp Pro Thr Leu
 355 360 365
 Pro Arg His Asn Leu Thr Ala Ala Ser Ile Lys Val Glu Gln Leu Cys
 370 375 380
 Lys Lys His Asn Leu Pro Tyr Arg Ser Pro Pro Met Leu Glu Gly Val
 385 390 395 400
 Gly Ile Leu Ile Ser Tyr Leu Gly Thr Phe Ala Arg Met Val Ala Lys
 405 410 415
 Ala Asp Lys Ala
 420
 <210> 65
 <211> 1350
 <212> DNA
 <213> Euglena gracilis
 <220>
 <221> CDS
 <222> (1)..(1350)
 <223> delta-5 desaturase
 <300>
 <302> DELTA-5 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY ACIDS
 <310> WO 2007/136671
 <311> 2007-05-16
 <312> 2007-11-29
 <313> (1)..(1350)
 <300>
 <302> DELTA-5 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY ACIDS
 <310> US 2007-0292924-A1
 <311> 2007-05-15
 <312> 2007-12-20
 <313> (1)..(1350)
 <400> 65
 atg got 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 ggc att gat ggc atc ctc tac gac ctt gaa ggc 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 act 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 ggc 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 aag tat gat tct ccc ttc gaa caa 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 ctg atg aaa ggg agg aac tgg tac gaa 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 ggc atc atc gcc ctg 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 aag ggg atg ctg 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 ggc 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 gcc 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 gcc ctc gac ctg gat cgc 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 ctg ttc cac aac tac ccc gcc gaa 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 got tgg tac atg tac ctt gtc 720
 Ala Arg Lys Trp Phe His Arg Phe Gln Ala Trp Tyr Met Tyr Leu Val
 225 230 235 240
 ctg ggc gaa tac ggg gaa tcg ctg ctg 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 gcc ttt ctg cgg cgc tac cgc aca ctt gaa ttc gtc 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 cgg      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

tte 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 Thr Gly Gly Leu Asn Phe
370                375                380

cag atc gag cac cac ctc ttt ccc cgg atg tgg 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 cgc 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 gcc 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> 66
<211> 449
<212> PRT
<213> Euglena gracilis

<400> 66
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                80

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                160

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 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                320

```

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> 67

<211> 1350

<212> DNA

<213> Euglena gracilis

<220>

<221> CDS

<222> (1)..(1350)

<223> synthetic delta-5 desaturase (codon-optimized for Yarrowia lipolytica)

<300>

<302> DELTA-5 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2007/136671

<311> 2007-05-16

<312> 2007-11-29

<313> (1)..(1350)

<300>

<302> DELTA-5 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US 2007-0292924-A1

<311> 2007-05-15

<312> 2007-12-20

<313> (1)..(1350)

<400> 67

atg gct ctc tcc ctt act acc gag cag ctg ctc gag cga cga 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 gga 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 cag cct tac gtc aag ccc gag aac tgg 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 cag 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 aac tac gac tct ooc 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 gcc 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 gac 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 gcc 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 tgg tcc aga tgg atc tgg ctg cag tct cgc atc atg	576
Asp Val Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile Met	
180 185 190	
cga cat cgc acc tac acc aac 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 cgc aac tac cct gct gcc aac act	672
Ser Ala Glu Pro Phe Leu Val Phe His Asn Tyr Pro Ala Ala Asn Thr	
210 215 220	
gct cga aag tgg ttt cgc 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 tgg 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 gcc ttt ctg cga cgg tac cga acc ott 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 gcc tog 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 tog 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 gcc aac atc att tcc acc ctc aag      1296
Tyr Val Phe 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 gcc 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> 68
<211> 449
<212> PRT
<213> Euglena gracilis

<400> 68
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 80

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
145 150 155 160

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
305 310 315 320

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> 69

<211> 1350

<212> DNA

<213> Euglena gracilis

<220>

<221> CDS

<222> (1)..(1350)

<223> mutant delta-5 desaturase

<400> 69

atg get 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 ggc 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 tcc gat ttg att ctc gct tct ggt gcc tat gat gcc tcc 144
His Pro Gly Ser 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 aca ccg ggg aac 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 tog 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 aag tat gat tct ooc 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 ctg atc aaa cgg aag aac tgc tac cca acc cct gcc 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 aag 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 aag ggg atg gtg ctg tgg gcc ctg tgg aat 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 ggc tcc ccc 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 gcc att 528
Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Leu Phe Ala Tyr Gly Ile
165 170 175

gac gtc atc gaa tgg tcc cgg tgg att tgg ctg cag tog ccc 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 gcc ctc gac ctg gat ggc gag 624
Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala Glu
195 200 205

tgg 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 cag 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 gla tgg 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 gac tct gtc aag gcc atg ccg 816
Arg Met Gln His Asn Asp Thr Ile Pro Glu Ser Val Thr Ala Met Arg
260 265 270

gaa 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 cgg      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 acc 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 gcc 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 cga tct gag aag gac gtt      1056
Arg Ile Pro Asp Lys Asn Cys Lys Val Lys Arg Ser Glu Lys Asp Val
340          345          350

gag gct gac caa att gac tgg tat cgg cgg 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 gcc ccc atc gcc atg ttc ttc act gcc 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 tgg 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 cgg gtc cgg gag tgt tgc gaa cga 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 gcc 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 gca 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> 70
<211> 449
<212> PRT
<213> Euglena gracilis

<400> 70
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 Ser 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          80

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          160

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          240

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 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          320
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 Arg 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> 71
<211> 1350
<212> DNA
<213> Euglena gracilis

<220>
<221> CDS
<222> (1)..(1350)
<223> synthetic mutant delta-5 desaturase (codon-optimized for Yarrowia lipolytica)

<400> 71
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 tcc gac ttg atc ctc gct tct ggt gcc tcc gat gct tct      144
His Pro Gly Ser 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 tgg 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 cag 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 acc 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 ggt 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 tgg 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 gcc 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 gcc 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 tgg tcc aga tgg atc tgg ctg cag tct csc 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 gaa gaa cca ttc ctt gtg ttc cac aac tac cct gct gcc aac act      672
Ser 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
ctt gga gcc tac gcc gtt tgg 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 acc 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 gtc ccc act ggt acc ggt gcc      960
Thr Ser Leu Leu Ile Thr Ile Pro Leu Val Pro Thr Ala Thr Gly Ala

```

```

305          310          315          320
tte 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
oga 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
tte gtg cag caa gct gtc oga gag tgt tgc gaa oga cac gga gtt cgg 1248
Phe Val Gln Gln Ala Val Arg Glu Cys Cys Glu Arg His Gly Val Arg
405          410          415
tae gtc ttc tac cct acc att gtg ggc aac atc att tcc acc ctc aag 1296
Tyr Val Phe 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> 72
<211> 449
<212> PRT
<213> Euglena gracilis

<400> 72
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 Ser 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          80
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          160
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          240
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          320
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> 73
<211> 1392
<212> DNA
<213> Peridinium sp. CCMP626

<220>
<221> CDS
<222> (1)..(1392)
<223> delta-5 desaturase

<300>
<302> DELTA-5 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2007/136646
<311> 2007-05-16
<312> 2007-11-29
<313> (1)..(1392)

<300>
<302> DELTA-5 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US 2007-0271632-A1
<311> 2007-05-15
<312> 2007-11-22
<313> (1)..(1392)

<400> 73
atg gct cca gat gcg gac aag ttg aga cag cgc aag gcg caa tcg att      48
Met Ala Pro Asp Ala Asp Lys Leu Arg Gln Arg Lys Ala Gln Ser Ile
1          5          10          15

caa gac acg gct gat tcg caa gct acc gaa ctc aag att gcc acc ctg      96

Gln Asp Thr Ala Asp Ser Gln Ala Thr Glu Leu Lys Ile Gly Thr Leu
20          25          30

aag gcc ttg cag ggg aca gaa atc gtc att gat gga gac att tac gat      144
Lys Gly Leu Gln Gly Thr Glu Ile Val Ile Asp Gly Asp Ile Tyr Asp
35          40          45

ata aaa cac ttt gat cac ccc ggt ggt gaa tcc atc atg act ttt ggg      192
Ile Lys Asp Phe Asp His Pro Gly Gly Glu Ser Ile Met Thr Phe Gly
50          55          60

gga aac gat gtc acc gcc acg tac aag atg atc cac ccc tac cac tct      240
Gly Asn Asp Val Thr Ala Thr Tyr Lys Met Ile His Pro Tyr His Ser
65          70          75          80

aag cac cat ttg gag aag atg aag aaa gtg gga cga gtt ccg gac tac      288
Lys His His Leu Glu Lys Met Lys Lys Val Gly Arg Val Pro Asp Tyr
85          90          95

acc tcg gaa tac aag ttt gat act ccc ttt gag cgt gaa atc aag caa      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 cga gcc cgc gag ttt gga aca cct gga      384
Glu Val Phe Lys Ile Val Arg Arg Gly Arg Glu Phe Gly Thr Pro Gly
115          120          125

tac ttc ttc cgg gct ttc tgc tac att gga ctt ttc ttt tac ttg cag      432
Tyr Phe Phe Arg Ala Phe Cys Tyr Ile Gly Leu Phe Phe Tyr Leu Gln
130          135          140

tat ttg tgg gtc aag act ccc act acc ttt gcc ttg ggg atc ttc tat      480
Tyr Leu Trp Val Thr Thr Pro Thr Thr Phe Ala Leu Ala Ile Phe Tyr
145          150          155          160

ggt gtt tcg caa gct ttc att ggt ttg aac gta caa cat gat 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 aag aag cct tgg atc aat aac ttg cta gga ttg      576
His Gly Ala Ala Ser Lys Lys Pro Trp Ile Asn Asn Leu Leu Gly Leu
180          185          190

cgg gct qac ttt atc gga ggt tcc aaa tgg ttg ttg atg aac caa cac      624
Gly Ala Asp Phe Ile Gly Gly Ser Lys Trp Leu Trp Met Asn Gln His
195          200          205

tgg aag caa cgc aca tac acc aac cgc cat gag aag gat ccc gat gcc      672
Trp Thr His His Thr Tyr Thr Asn His His Glu Lys Asp Pro Asp Ala
210          215          220

ttg gcc gct gaa cca atg ttg ttg ttc aat gat 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

cca aag cgt act ttg att cac cac ttc cag gcc ttc tat tac ctt ttc      768
Pro Lys Arg Thr Leu Ile His His Phe Gln Ala Phe Tyr Tyr Leu Phe
245          250          255

gtc ttg gcc gga tac tgg gtc tct tcg gtc ttc aac cct caa att ttg      816
Val Leu Ala Gly Tyr Trp Val Ser Ser Val Phe Asn Pro Gln Ile Leu
260          265          270

gac ttg caa cac cgc ggt gct caa cgg ggt gga atg aaa atg gag aac      864
Asp Leu Gln His Arg Gly Ala Gln Ala Val Gly Met Lys Met Glu Asn

```

```

275          280          285
gat tac att gcc aaa agc cga aag tat gcc atc ttc ttg cgt ctc ttg      912
asp tyr ile ala lys ser arg lys tyr ala ile phe leu arg leu leu
290          295          300

tat att tac acc aac att gtc gct ccg atc cca aac caa ggc ttc tgg      960
tyr ile tyr thr asn ile val ala pro ile gln asn gln gly phe ser
305          310          315

ttg acc gtg gtc gcc csc att ttg acc atg ggc gtc gct tcc agt ttg      1008
leu thr val val ala his ile leu thr met gly val ala ser ser leu
325          330          335

act ttg gcg act ctt ttt gcc ttg tgg cac aat ttt gaa aac gcg gat      1056
thr leu ala thr leu phe ala leu ser his asn phe glu asn ala asp
340          345          350

cgc gat ccc act tac gag gcc cgc aag gga gga gag cct gtt tgt tgg      1104
arg asp pro thr tyr glu ala arg lys gly gly glu pro val cys trp
355          360          365

ttc aag tgg caa gtc gaa acc tgg tca act tac gga ggt ttc atc tgg      1152
phe lys ser gln val glu thr ser ser thr tyr gly gly phe ile ser
370          375          380

ggt tgc ttg acg ggc gga ctc aac ttc caa gtg gaa cac cac ttg ttc      1200
gly cys leu thr gly gly leu asn phe gln val glu his his leu phe
385          390          395

oct cgt atg agt tgg gcc tgg tac ccc tac att gcc oct act gtt cga      1248
pro arg met ser ser ala trp tyr pro tyr ile ala pro thr val arg
405          410          415

gag gtt tgc aaa aag csc gga gtc aag tac gca tac tat ccc tgg gtc      1296
glu val cys lys lys his gly val lys tyr ala tyr tyr pro trp val
420          425          430

tgg caa aac ttg att tca act gtc aag tat ctg cat caa agc gga act      1344
trp gln asn leu ile ser thr val lys tyr leu his gln ser gly thr
435          440          445

gga tcc aac tgg aag aat ggc gcc aac ccc tac tgg gga aaa ttg taa      1392
gly ser asn trp lys asn gly ala asn pro tyr ser gly lys leu
450          455          460

<210> 74
<211> 463
<212> PRT
<213> Peridinium sp. CCMP626

<400> 74
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 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 Val Cys 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 Val Glu 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 Val Cys Lys Lys His Gly Val Lys 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> 75

<211> 1392

<212> DNA

<213> Peridinium sp. CCMP626

<220>

<221> CDS

<222> (1)..(1392)

<223> synthetic delta-5 desaturase (codon-optimized for Yarrowia lipolytica)

<300>

<302> DELTA-5 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2007/136646

<311> 2007-05-16

<312> 2007-11-29

<313> (1)..(1392)

<300>

<302> DELTA-5 DESATURASE AND ITS USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US 2007-0271632-A1

<311> 2007-05-15

<312> 2007-11-22

<313> (1)..(1392)

<400> 75

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 gcc 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 gcc acc gag atc gtc att gat gcc 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 gcc 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	
aag cat cac ctg gag aag atg aaa aag gtc ggt cga gtc 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 gtc cga aga ggt cga gag ttt gga aca cct gcc	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 gtc tct cag gcc ttc att gcc 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 gcc 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 gcc 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 ccc 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 ccc 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 ccc 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 gcc 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 gcc 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 gcc 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> 76
<211> 463
<212> PRT
<213> Peridinium sp. CCMP626

<400> 76
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 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 Val Cys 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 Val Glu 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 Val Cys Lys Lys His Gly Val Lys 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> 77
<211> 1362
<212> DNA
<213> Euglena anabaena UTEX 373

<220>
<221> CDS
<222> (1)..(1362)
<223> delta-5 desaturase

<300>
<302> DELTA-5 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> U.S. 2008-0274521-A1
<311> 2008-04-29
<312> 2008-11-06
<313> (1)..(1362)

<300>
<302> DELTA-5 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/137532
<311> 2008-05-01
<312> 2008-11-13
<313> (1)..(1362)

<400> 77
atg gcc acc atc tct ttg act act gag caa ctt tta gaa cac cca gaa      48
Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu
1          5          10          15

ctg gtt gca att gat ggg gtg ttg tac gat ctc ttc gga ctg gcg aaa      96
Leu Val Ala Ile Asp Gly Val Leu Tyr Asp Leu Phe Gly Leu Ala Lys
20         25         30

gtg cat cca ggt ggc aac ctg att gaa gcc gcc ggt gcc tcc qac qga     144
Val His Pro Gly Gly Asn Leu Ile Glu Ala Ala Gly Ala Ser Asp Gly
35         40         45

acc gcc ctg ttc tac tcc atg cac cct cct gga gtg aag cca gag aat tog     192
Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser
50         55         60

aag ctg ctg cag caa ttt gcc cga gcc aaa cac gaa cga agc tcg aag     240
Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys
65         70         75         80

gac coa gtg tac acc ttt gac agt ccc ttc gcc cag gat gtc aag cag     288
Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln
85         90         95

agc gtt cgg gag gtc atg aag ggg cgc aac tgg tac gcc acg ccc ggc     336
Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly
100        105        110

ttt tgg ctg cgg acc gcg ctg atc atc gcg tgc act gcc ata gcc gaa     384
Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu
115        120        125

tgg tat tgg atc atc acc ggg gca ctg atg tgg ggc atc ttc acc ggg     432
Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly
130        135        140

tac ttc cac agc cag att ggg ttg gcg att cca cac gat gcc tct cac     480
Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Ser His
145        150        155        160

gga gcc atc agc aaa aag ccc tgg gtg aac gcc ttt ttc gcc tac ggc     528
Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly
165        170        175

atc gac gcc att gga tcc tcc cgc tgg atc tgg ctg cag tcc cac att     576
Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile
180        185        190

atg cgc cac cac acc tcc acc aac cag cat ggc ctg gac ctg gac gct     624
Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala
195        200        205

gcc tog gcg gag ccg ttc att ttg ttc cac tcc tcc ccc gca aca aat     672

```

Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn
210 215 220

cgc tca cga aag tgg tac cat cgg ttc cag gcg tgg tac atg tac atc 720
Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile
225 230 235 240

gtt ttg ggg atg tat ggt gtg tgg atg gtg tac aat ccg atg tac ttg 768
Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu
245 250 255

tte acg atg cag cac aac gac aca atc cca gag gcc acc tct ctt aga 816
Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg
260 265 270

cca gcc agc ttt ttc aac cgg cag cgc gcc ttc gcc gtt tcc ctc cgc 864
Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg
275 280 285

cta ctg ttc atc ttc cgc aac gcc ttc ctc ccc tgg tac atc cgc gcc 912
Leu Leu Phe Ile Phe Arg Asn Ala Phe Leu Pro Trp Tyr Ile Ala Gly
290 295 300

gcc tct ccg ctg ctc acc atc ctg ctg gtg cca acg gtc aca gcc atc 960
Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile
305 310 315 320

tte ttg aca ttt gtt ttt gtg ctg tcc cat aac ttt gaa gcc gct gag 1008
Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu
325 330 335

cgg acc ccc gaa aag aac tgc aag gcc aaa agc gcc aag gag ggg aag 1056
Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys
340 345 350

cag gtc cgc gat gta gag gag gac cgg gtg gac tgg tac cgg gcc cag 1104
Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln
355 360 365

ccc gag acc gcc gcc acc tac cgg gcc agc gtc ggg atg atg ctg acc 1152
Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr
370 375 380

cgc ggt ttg aac ctg cag atc gag cac cac ttg ttc ccc cgc atg tcc 1200
Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser
385 390 395 400

tct tgg cac tac ccc ttc atc caa gat acg gtg cgg gaa tgt tgc aag 1248
Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Cys Lys
405 410 415

cgc cat gcc gtg cgc tac aca tac tac ccg acc atc ctg gag aat ata 1296
Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile
420 425 430

atg tcc acg ctc cgc tac atg cag aag gtg gcc gtg gcc cac aca att 1344
Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile
435 440 445

cag gat gcc cag gaa ttc 1362
Gln Asp Ala Gln Glu Phe
450

<210> 78
<211> 454
<212> PRT
<213> Euglena anabaena UTEX373

<400> 78
Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu
1 5 10 15

Leu Val Ala Ile Asp Gly Val Leu Tyr Asp Leu Phe Gly Leu Ala Lys
20 25 30

Val His Pro Gly Gly Asn Leu Ile Glu Ala Ala Gly Ala Ser Asp Gly
35 40 45

Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser
50 55 60

Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys
65 70 75 80

Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln
85 90 95

Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly
100 105 110

Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu
115 120 125

Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly
130 135 140

Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Ser His
145 150 155 160

Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly
165 170 175

Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile
180 185 190

Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala
195 200 205

Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn
210 215 220

Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile
225 230 235 240

```

Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu
245 250 255

Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg
260 265 270

Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg
275 280 285

Leu Leu Phe Ile Phe Arg Asn Ala Phe Leu Pro Trp Tyr Ile Ala Gly
290 295 300

Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile
305 310 315 320

Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu
325 330 335

Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys
340 345 350

Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln
355 360 365

Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr
370 375 380

Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser
385 390 395 400

Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Lys
405 410 415

Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile
420 425 430

Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile
435 440 445

Gln Asp Ala Gln Glu Phe
450

<210> 79
<211> 1362
<212> DNA
<213> Euglena anabaena UTEX 373

<220>
<221> CDS
<222> (1)..(1362)
<223> synthetic delta-5 desaturase (codon-optimized for Yarrowia lipolytica)

<300>
<302> DELTA-5 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> U.S. 2008-0274521-A1
<311> 2008-04-29
<312> 2008-11-06
<313> (1)..(1362)

<300>
<302> DELTA-5 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/137532
<311> 2008-05-01
<312> 2008-11-13
<313> (1)..(1362)

<400> 79
atg gcc acc atc tcc ctg act acc gaq cag ctc ctg gaa cac ccc gaq 48
Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu
1 5 10 15

ctc gtt gcc atc gac gga gtc ctg tac gat ctc ttc ggt ctg gcc aag 96
Leu Val Ala Ile Asp Gly Val Leu Tyr Asp Leu Phe Gly Leu Ala Lys
20 25 30

gtg cat cca gga ggc aac ctc atc gaa gct gcc ggt gca tcc gac gga 144
Val His Pro Gly Gly Asn Leu Ile Glu Ala Ala Gly Ala Ser Asp Gly
35 40 45

acc gct ctg ttc tac tcc atg atc cct gga gtc aag cca gag aac tcg 192
Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser
50 55 60

aag ctt ctg cag caa ttt gcc cga ggc aag cac gaa cga agc tcc aag 240
Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys
65 70 75 80

gat ccc gtg tac acc ttc gac tct ccc ttt gct cag gac gtc aag cag 288
Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln
85 90 95

tcc gtt cga gag gtc atg aag ggt cga aac tgg tac gcc act cct gcc 336
Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly
100 105 110

ttc tgg ctg aga acc gca ctc atc atc gct tgt act gcc att gcc gag 384
Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu
115 120 125

tgg tac tgg atc acg acc gga gca gtc atg tgg ggt atc ttt act gga 432
Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly
130 135 140

tac ttc cac tcg cag att gcc ttg gcc att cca cac gat gct tct cac 480
Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Ser His
145 150 155 160

gga gcc atc agc aaa aag ccc tgg gtc aac gcc ttt ttc gct tat gcc 528
Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly
165 170 175

```

```

atc gac gcc att ggt tcc tct cgt tgg atc tgg ctg cag tcc cac att      576
ile asp ala ile gly ser ser arg trp ile trp leu gln ser his ile
180      185      190

atg cga cat cac act tac acc aac cag cat ggc ctg gac ctg gat gat      624
met arg his his thr thr thr asn gln his gly leu asp leu asp ala
195      200      205

gcc tgg gca gag cng ttc atc ttg ttc cac tcc tat cct gat acc aac      672
ala ser ala glu pro phe phe leu phe his ser tyr trp ala thr asn
210      215      220

gcc tct cga sag tgg tac cac cga ttt cag gcg tgg tac atg tac atc      720
ala ser arg lys trp tyr his arg phe gln ala trp tyr met tyr ile
225      230      235

gtt ctg cga atg tat ggt ctg tgg atg gtg tac aat ccc atg tac ctg      768
val leu gly met tyr gly val ser met val tyr asn pro met tyr leu
245      250      255

ttc aca atg cag cac aac gac acc att ccc gag gcc act tct ctg aga      816
phe thr met gln his asn asp thr ile pro glu ala thr ser leu arg
260      265      270

cca gcc agc ttt ttc aat cgg cag cga ggt ttc gcc gtt tcc ctt cga      864
pro gly ser phe phe asn arg gln arg ala phe ala val ser leu arg
275      280      285

ctg ctg ttc atc ttc cga aac gcc ttt ctt ccc tgg tac att ggt ggt      912
leu leu phe ile phe arg asn ala phe leu pro trp tyr ile ala gly
290      295      300

gcc tct cct ctg ctg acc att ctt ctg gtg ccc aag gtc aca gcc atc      960
ala ser pro leu leu thr ile leu leu val pro thr val thr gly ile
305      310      315

ttc ctg acc ttt gtg ttc gtt ctg tcc cat aac ttc gag gga gcc gaa      1008
phe leu thr phe phe val phe val leu ser his asn phe glu gly ala glu
325      330      335

ogg acc cca gag aag aac tgc aag gcc aca cga gct aag gaa gcc aag      1056
arg thr pro glu lys asn cys lys ala lys arg ala lys glu gly lys
340      345      350

gag gtc cga gac gtg gaa gag gat cga gtc gac tgg tac cga gaa cag      1104
glu val arg asp val glu glu asp arg val asp trp tyr arg ala gln
355      360      365

gcc gag act gct gcc acc tac ggt gcc agc gtg gaa atg atg ctt aca      1152
ala glu thr ala ala thr tyr gly gly ser val gly met met leu thr
370      375      380

gcc ggt ctg aac ctg cag atc gag cat cac ttg ttt ccc cga atg tcc      1200
gly gly leu asn leu gln ile glu his his leu phe pro arg met ser
385      390      395

tct tgg cac tat ccc ttc att cca gac acc gtt cgg gag tgt tgc aag      1248
ser trp his tyr pro phe ile gln asp thr val arg glu cys cys lys
405      410      415

cga cat gcc gtc cgt tac aca tac tat cct acc att ctg gag aac atc      1296
arg his gly val arg tyr thr tyr tyr pro thr ile leu glu asn ile
420      425      430

atg tcc act ctt cga tac atg cag aag gtg ggt ggt gcc acc acc att      1344

Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile
435      440      445

cag gat gcc cag gag ttc      1362
Gln Asp Ala Gln Glu Phe
450

<210> 80
<211> 454
<212> PRT
<213> Euglena anabaena UTEX373

<400> 80
Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu
1      5      10      15

Leu Val Ala Ile Asp Gly Val Leu Tyr Asp Leu Phe Gly Leu Ala Lys
20      25      30

Val His Pro Gly Gly Asn Leu Ile Glu Ala Ala Gly Ala Ser Asp Gly
35      40      45

Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser
50      55      60

Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys
65      70      75      80

Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln
85      90      95

Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly
100      105      110

Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu
115      120      125

Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly
130      135      140

Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Ser His
145      150      155      160

Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly
165      170      175

Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile
180      185      190

Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala
195      200      205

```


Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn
210 215 220

Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile
225 230 235 240

Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu
245 250 255

Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg
260 265 270

Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg
275 280 285

Leu Leu Phe Ile Phe Arg Asn Ala Phe Leu Pro Trp Tyr Ile Ala Gly
290 295 300

Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile
305 310 315 320

Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu
325 330 335

Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys
340 345 350

Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln
355 360 365

Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr
370 375 380

Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser
385 390 395 400

Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Cys Lys
405 410 415

Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile
420 425 430

Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile
435 440 445

Gln Asp Ala Gln Glu Phe
450

<210> 81
<211> 1365
<212> DNA
<213> Euglena anabaena UTEX373

<220>
<221> CDS
<222> (1)..(1365)
<223> synthetic mutant delta-5 desaturase (codon-optimized for Yarrowia lipolytica)

<400> 81

atg gcc acc atc tcc ctg act acc gaa gac ctc ctg gaa cac ccc gag Met Ala Thr Ile Ser Leu Thr Thr Gln Gln Leu Leu Glu His Pro Glu 1 5 10 15	48
ctc gtt gcc atc gac gga gtc ctg tac gat ctc ttc ggt ctg gcc aag Leu Val Ala Ile Asp Gly Val Leu Tyr Asp Leu Phe Gly Leu Ala Lys 20 25 30	96
gtc cat gcc gga ggc aac ctc atc gaa gct gcc ggt gca tcc gac gga Val His Ala Gly Gly Asn Leu Ile Glu Ala Ala Gly Ala Ser Asp Gly 35 40 45	144
acc gct ctg ttc tac tcc atg cat cct gga gtc aag cca gag aac tgg Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser 50 55 60	192
aaq ctt ctg caa caa ttt gcc cga gcc aag cac gaa cga agc tcc aaq Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys 65 70 75 80	240
gat ccc gtg tac acc ttc gac tct ccc ttt gct caa gac gtc aaq caa Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln 85 90 95	288
tcc gtt cga gag gtc atg aag ggt cga aac tgg tac gcc act cct gcc Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly 100 105 110	336
ttc tgg ctg aga acc gca ctc atc atc gct tgt act gcc att gcc gag Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu 115 120 125	384
tgg tac tgg atc aca acc gga gca gtc atg tgg ggt atc ttt act gga Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly 130 135 140	432
tac ttc cac tog cag att gcc ttg gcc att cca cac gat got tot cac Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Ser His 145 150 155 160	480
gga gcc atc agc aaa aag ccc tgg gtc aac gcc ttt ttc got tat gcc Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly 165 170 175	528
atc gac gcc att ggt tcc tct cgt tgg atc tgg ctg caa tcc cac att Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile 180 185 190	576
atg cga cat cac act tac acc aac cag cat gcc ctc gac ctg gat gct Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala 195 200 205	624

```

goc tgg gca gag cgg ttc atc ttg ttc cac tcc tat cct gct acc aac      672
Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn
210          215          220

goc tct cga aag tgg tac cac oga ttt cag gog tgg tac atg tac atc      720
Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile
225          230          235          240

glt ctg gga atg tat ggt gtc tcg atg gtg tac aat ccc atg tac ctc      768
Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu
245          250          255

tto aca atg cag cac aac gac acc att ooo gag gcc aot tot cto aga      816
Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg
260          265          270

cca gcc agc ttt ttc aat cgg cag oga gct ttc gcc gtt tcc ctt cga      864
Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg
275          280          285

ctg ctc ttc atc ttc oga aac gcc ttt ctt ccc tgg tac att gct ggt      912
Leu Leu Phe Ile Phe Arg Asn Ala Phe Leu Pro Trp Tyr Ile Ala Gly
290          295          300

goc tct cct ctg ctc acc att ctt ctg gtg ccc acg gtc aca ggc atc      960
Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile
305          310          315          320

tto ctc aoc ttt gtg ttc gtt ctg tcc aat aac ttc gag gga gcc gaa      1008
Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu
325          330          335

cgg acc cca gag aag aac tgc aag gcc aaa oga gct aag gaa gcc aag      1056
Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys
340          345          350

gag gtc aga gac gtg gaa gag gat oga gtc gac tgg tac oga gca cag      1104
Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln
355          360          365

gcc gag act gct gcc acc tac ggt gcc agc gtg gga atg atg ctt aca      1152
Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr
370          375          380

ggc ggt ctc aac ctg cag atc gag cat cac ttg ttt ccc oga atg tcc      1200
Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser
385          390          395          400

tot tgg cac tat ccc ttc att caa gac aoc gtt cgg gag tgt tgc aag      1248
Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Cys Lys
405          410          415

oga cat gcc gtc cgt tac aca tac tat cct acc att ctc gag aac atc      1296
Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile
420          425          430

atg tcc act ctt cga tac atg cag aag gtg ggt gct gct cac acc att      1344
Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile
435          440          445

cag gat gcc cag gag ttc taa      1365
Gln Asp Ala Gln Glu Phe
450

<210> 82
<211> 454
<212> PRT
<213> Euglena anabaena UTEX373

<400> 82
Met Ala Thr Ile Ser Leu Thr Thr Glu Gln Leu Leu Glu His Pro Glu
1      5      10      15

Leu Val Ala Ile Asp Gly Val Leu Tyr Asp Leu Phe Gly Leu Ala Lys
20      25      30

Val His Ala Gly Gly Asn Leu Ile Glu Ala Ala Gly Ala Ser Asp Gly
35      40      45

Thr Ala Leu Phe Tyr Ser Met His Pro Gly Val Lys Pro Glu Asn Ser
50      55      60

Lys Leu Leu Gln Gln Phe Ala Arg Gly Lys His Glu Arg Ser Ser Lys
65      70      75      80

Asp Pro Val Tyr Thr Phe Asp Ser Pro Phe Ala Gln Asp Val Lys Gln
85      90      95

Ser Val Arg Glu Val Met Lys Gly Arg Asn Trp Tyr Ala Thr Pro Gly
100     105     110

Phe Trp Leu Arg Thr Ala Leu Ile Ile Ala Cys Thr Ala Ile Gly Glu
115     120     125

Trp Tyr Trp Ile Thr Thr Gly Ala Val Met Trp Gly Ile Phe Thr Gly
130     135     140

Tyr Phe His Ser Gln Ile Gly Leu Ala Ile Gln His Asp Ala Ser His
145     150     155     160

Gly Ala Ile Ser Lys Lys Pro Trp Val Asn Ala Phe Phe Ala Tyr Gly
165     170     175

Ile Asp Ala Ile Gly Ser Ser Arg Trp Ile Trp Leu Gln Ser His Ile
180     185     190

Met Arg His His Thr Tyr Thr Asn Gln His Gly Leu Asp Leu Asp Ala
195     200     205

Ala Ser Ala Glu Pro Phe Ile Leu Phe His Ser Tyr Pro Ala Thr Asn
210     215     220

Ala Ser Arg Lys Trp Tyr His Arg Phe Gln Ala Trp Tyr Met Tyr Ile

```

```

225          230          235          240
Val Leu Gly Met Tyr Gly Val Ser Met Val Tyr Asn Pro Met Tyr Leu
245          250          255

Phe Thr Met Gln His Asn Asp Thr Ile Pro Glu Ala Thr Ser Leu Arg
260          265          270

Pro Gly Ser Phe Phe Asn Arg Gln Arg Ala Phe Ala Val Ser Leu Arg
275          280          285

Leu Leu Phe Ile Phe Arg Asn Ala Phe Leu Pro Trp Tyr Ile Ala Gly
290          295          300

Ala Ser Pro Leu Leu Thr Ile Leu Leu Val Pro Thr Val Thr Gly Ile
305          310          315          320

Phe Leu Thr Phe Val Phe Val Leu Ser His Asn Phe Glu Gly Ala Glu
325          330          335

Arg Thr Pro Glu Lys Asn Cys Lys Ala Lys Arg Ala Lys Glu Gly Lys
340          345          350

Glu Val Arg Asp Val Glu Glu Asp Arg Val Asp Trp Tyr Arg Ala Gln
355          360          365

Ala Glu Thr Ala Ala Thr Tyr Gly Gly Ser Val Gly Met Met Leu Thr
370          375          380

Gly Gly Leu Asn Leu Gln Ile Glu His His Leu Phe Pro Arg Met Ser
385          390          395          400

Ser Trp His Tyr Pro Phe Ile Gln Asp Thr Val Arg Glu Cys Lys
405          410          415

Arg His Gly Val Arg Tyr Thr Tyr Tyr Pro Thr Ile Leu Glu Asn Ile
420          425          430

Met Ser Thr Leu Arg Tyr Met Gln Lys Val Gly Val Ala His Thr Ile
435          440          445

Gln Asp Ala Gln Glu Phe
450

<210> 83
<211> 1086
<212> DNA
<213> Phytophthora ramorum

<220>
<221> CDS
<222> (1)..(1086)
<223> delta-17 desaturase

<300>
<302> DELTA-17 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2007/123999
<311> 2007-04-19
<312> 2007-11-01
<313> (1)..(1086)

<300>
<302> DELTA-17 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US 2007-0249026-A1
<311> 2007-04-18
<312> 2007-10-25
<313> (1)..(1086)

<400> 83
atg gcg act aag cag ccg tac cag ttc ccg acc ctg acg gag atc aag      48
Met Ala Thr Lys Gln Pro Tyr Gln Phe Pro Thr Leu Thr Glu Ile Lys
1      5      10      15

cgg tcg ctg ccc agc gag tgc ttt gaq gcc tcg gtg ccg ctg tcg ctc      96
Arg Ser Leu Pro Ser Glu Cys Phe Glu Ala Ser Val Pro Leu Ser Leu
20     25     30

tac tac acg ctg cgc atc gtg gcc atc gcc gtg ccg ctg gcg ttc gcc      144
Tyr Tyr Thr Val Arg Ile Val Ala Ile Ala Val Ala Leu Ala Phe Gly
35     40     45

ctc aac tac gcg cgc gcg ctg ccc gtg gtc gaq agc ttg tgg gcg ctg      192
Leu Asn Tyr Ala Arg Ala Leu Pro Val Val Glu Ser Leu Trp Ala Leu
50     55     60

gac gct gcg ctc tgc tgc ggt tac gtg ctg ctg cag gcc atc gtg ttc      240
Asp Ala Ala Leu Cys Cys Gly Tyr Val Leu Leu Gln Gly Ile Val Phe
65     70     75     80

tgg gcc ttc ttc acg gtg ggc cat gac gcc gcc cag gcc gcc ttc tgg      288
Trp Gly Phe Phe Thr Val Gly His Asp Ala Gly His Gly Ala Phe Ser
85     90     95

cgt tac cac ctg ctc aac ttc gtg gtg gcc acc ttc atc cac tcg ctc      336
Arg Tyr His Leu Leu Asn Phe Val Val Gly Thr Phe Ile His Ser Leu
100    105    110

atc ctc acg ccc ttc gaq tgg tgg aag ctc acg cac cgc cac cac cac      384
Ile Leu Thr Pro Phe Glu Ser Trp Lys Leu Thr His Arg His His His
115    120    125

aag aac acg gcc aac att gac cgc gac gag atc ttc tac ccg cag ccg      432
Lys Asn Thr Gly Asn Ile Asp Arg Asp Glu Ile Phe Tyr Pro Gln Arg
130    135    140

aag gcc gac gac cac ccg ctg tcg cgc aac ctc gtg ctg gcg ctc gcc      480
Lys Ala Asp Asp His Pro Leu Ser Arg Asn Leu Val Leu Ala Leu Gly
145    150    155    160

gcc gcg tgg ttc gcc tac ctg gtc gac gcc ttc ccg ccc cgc aag gtc      528
Ala Ala Trp Phe Ala Tyr Leu Val Glu Gly Phe Pro Pro Arg Lys Val
165    170    175

```

```

aac cac ttc aac cca ttc gag cag ctg ttt gtg cgc caa gtg gcc gcc      576
asn his phe asn pro phe glu pro leu phe val arg gln val ala ala
180 185 190

gtc gtc atc tgc ctc tcc ggc caa ttc gcc gtg ttg ggc ctg tcc gtg      624
val val ile ser leu ser ala his phe ala val leu ala leu ser val
195 200 205

tat ctg agc ttc cag ttc ggt ctc aag acc atg ggc ctc tac tac tac      672
tyr leu ser phe gln phe gly leu lys thr met ala leu tyr tyr tyr
210 215 220

ggc ccc gtc ttc gtg ttc ggc agc atg att gtg atc acc acc ttc ctg      720
gly pro val phe val phe gly ser met leu val ile thr thr phe leu
225 230 235 240

cat cac aat gac gag gag acc cca tgg tac gga gac tcc gac tgg acc      768
his his asn asp glu glu thr pro trp tyr gly asp ser asp trp thr
245 250 255

tac gtc aag ggc aac ctg tgc tcc gtg gac cgg tcc tac ggc ggc ttc      816
tyr val lys gly asn leu ser ser val asp arg ser tyr gly ala phe
260 265 270

atc gac aac ctg agc ccc aac atc gcc agc ccc cag atc cac ccc ctc      864
ile asp asn leu ser his asn ile gly thr his gln ile his his leu
275 280 285

ttc ccc atc atc ccc ccc tac aag ctc aac cgc gct acg ggc gca ttc      912
phe pro ile ile pro his tyr lys leu asn arg ala thr ala ala phe
290 295 300

cac cag gcc ttc ccc gag ctc gtg cgc aag agc gac gag ccc atc ctc      960
his gln ala phe pro glu leu val arg lys ser asp glu pro ile leu
305 310 315 320

aag gcc ttc tgg cgc gtc ggc cga ctg tac gcc aac tac ggc gtc gtg      1008
lys ala phe trp arg val gly arg leu tyr ala asn tyr gly val val
325 330 335

gac ccc gag gcc aag ctc ttc acg ctc aag gag gcc aag ggc ggc tcc      1056
asp pro asp ala lys leu phe thr leu lys glu ala lys ala ala ser
340 345 350

gag ggc ggc acc aag acc aag gcc acc taa
glu ala ala thr lys thr lys ala thr
355 360

<210> 84
<211> 361
<212> PRT
<213> Phytophthora ramorum

<400> 84
Met Ala Thr Lys Gln Pro Tyr Gln Phe Pro Thr Leu Thr Glu Ile Lys
1 5 10 15

Arg Ser Leu Pro Ser Glu Cys Phe Glu Ala Ser Val Pro Leu Ser Leu
20 25 30

Tyr Tyr Thr Val Arg Ile Val Ala Ile Ala Val Ala Leu Ala Phe Gly
35 40 45

Leu Asn Tyr Ala Arg Ala Leu Pro Val Val Glu Ser Leu Trp Ala Leu
50 55 60

Asp Ala Ala Leu Cys Cys Gly Tyr Val Leu Leu Gln Gly Ile Val Phe
65 70 75 80

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

Arg Tyr His Leu Leu Asn Phe Val Val Gly Thr Phe Ile His Ser Leu
100 105 110

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

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

Lys Ala Asp Asp His Pro Leu Ser Arg Asn Leu Val Leu Ala Leu Gly
145 150 155 160

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

Asn His Phe Asn Pro Phe Glu Pro Leu Phe Val Arg Gln Val Ala Ala
180 185 190

Val Val Ile Ser Leu Ser Ala His Phe Ala Val Leu Ala Leu Ser Val
195 200 205

Tyr Leu Ser Phe Gln Phe Gly Leu Lys Thr Met Ala Leu Tyr Tyr Tyr
210 215 220

Gly Pro Val Phe Val Phe Gly Ser Met Leu Val Ile Thr Thr Phe Leu
225 230 235 240

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

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

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

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

His Gln Ala Phe Pro Glu Leu Val Arg Lys Ser Asp Glu Pro Ile Leu
305 310 315 320

Lys Ala Phe Trp Arg Val Gly Arg Leu Tyr Ala Asn Tyr Gly Val Val
325 330 335

Asp Pro Asp Ala Lys Leu Phe Thr Leu Lys Glu Ala Lys Ala Ala Ser
340 345 350

Glu Ala Ala Thr Lys Thr Lys Ala Thr
355 360

<210> 85
<211> 1086
<212> DNA
<213> Phytophthora ramorum

<220>

```

<221> CDS
 <222> (1)..(1086)
 <223> synthetic delta-17 desaturase (codon-optimized for Yarrowia lipolytica)

<300>
 <302> DELTA-17 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2007/123999
 <311> 2007-04-19
 <312> 2007-11-01
 <313> (1)..(1086)

<300>
 <302> DELTA-17 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US 2007-0249026-A1
 <311> 2007-04-18
 <312> 2007-10-25
 <313> (1)..(1086)

<400> 85

atg gct acc aag cag ccc tac cag ttc cct act ctg acc gag atc aag	48
Met Ala Thr Lys Gln Pro Tyr Gln Phe Pro Thr Leu Thr Glu Ile Lys	
1 5 10 15	
aga tct ctt ccc tcc gag tgc ttt gaa gcc tgg gtc cct ctg tcc ttg	96
Arg Ser Leu Pro Ser Glu Cys Phe Glu Ala Ser Val Pro Leu Ser Leu	
20 25 30	
tac tac acc gtg cga atc gtc gct att gcc gtt gct ctg gcc ttc gga	144
Tyr Tyr Thr Val Arg Ile Val Ala Ile Ala Val Ala Leu Ala Phe Gly	
35 40 45	
ctc aac tac gct cga gcc ctt ccc gtg gtc gag tct ctg tgg gaa ctc	192
Leu Asn Tyr Ala Arg Ala Leu Pro Val Val Gln Ser Leu Trp Ala Leu	
50 55 60	
gac gct gcc ctt tgt tgc ggt tac gtt ctg ctc aac gcc att gtc ttc	240
Asp Ala Ala Leu Cys Cys Gly Tyr Val Leu Leu Gln Gly Ile Val Phe	
65 70 75 80	
tgg gga ttc ttt acc gtg ggt cag gat gct gga cat ggt gcc ttc tct	288
Trp Gly Phe Phe Thr Val Gly His Asp Ala Gly His Gly Ala Phe Ser	
85 90 95	
aga tac cac ctg ctc aac ttt gtc gtt gcc acc ttt atc cac tcc ctc	336
Arg Tyr His Leu Leu Asn Phe Val Val Gly Thr Phe Ile His Ser Leu	
100 105 110	
att ctt act ccc ttc gag tgg tgg aag ctc aca cat cga cac cat cac	384
Ile Leu Thr Pro Phe Glu Ser Trp Lys Leu Thr His Arg His His His	
115 120 125	
aag aac acc gga aac atc gac cga gaa atc ttc tac cct cag cga	432
Lys Asn Thr Gly Asn Ile Asp Arg Asp Glu Ile Phe Tyr Pro Gln Arg	
130 135 140	
aag gcc gac gat cat cct ctg tct cga aac ctc gtc ctg gct ctc ggt	480
Lys Ala Asp Asp His Pro Leu Ser Arg Asn Leu Val Leu Ala Leu Gly	
145 150 155 160	
gcc gct tgg ttt gcc tac att gtc gag gcc ttt cct ccc aga aag gtc	528
Ala Ala Trp Phe Ala Tyr Leu Val Glu Gly Phe Pro Pro Arg Lys Val	
165 170 175	
aac cac ttc aac ccc ttc gaa cct ctg ttt gtg cga cag gtg gct gcc	576
Asn His Phe Asn Trp Phe Glu Pro Leu Phe Val Arg Gln Val Ala Ala	
180 185 190	
gtt gtc att tcc ctc tct gct cag ttc gcc gtc ctg gaa ctg tcc gtg	624
Val Val Ile Ser Leu Ser Ala His Phe Ala Val Leu Ala Leu Ser Val	
195 200 205	
tat ctg agc ttt cag ttc ggt ctc aag aca atg gct ctg tac tac tat	672
Tyr Leu Ser Phe Gln Phe Gly Leu Lys Thr Met Ala Leu Tyr Tyr Tyr	
210 215 220	
gga ccc gtc ttc gtg ttc ggc tcc atg ctc gtc att act acc ttt ctg	720
Gly Pro Val Phe Val Phe Gly Ser Met Leu Val Ile Thr Thr Phe Leu	
225 230 235 240	
cat cac aat gac gag gaa act cct tgg tac gga gat tcc gac tgg acc	768
His His Asn Asp Glu Glu Thr Pro Trp Tyr Gly Asp Ser Asp Trp Thr	
245 250 255	
tac gtc aag gcc aac tbg tct tcc ctg gac cga tct tac ggt gcc ttc	816
Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr Gly Ala Phe	
260 265 270	
atc gac aac ctc tgg cac aac att gcc aca cac cag atc cac cat ctg	864
Ile Asp Asn Leu Ser His Asn Ile Gly Thr His Gln Ile His His Leu	
275 280 285	
ttt ccc atc att cct cac tac aag ctc aac cga gcc acc gct gcc ttc	912
Phe Pro Ile Ile Pro His Tyr Lys Leu Asn Arg Ala Thr Ala Ala Phe	
290 295 300	
cac cag gcc ttt ccc gaa ctt gtc cga aag agc gac gag ccc att ctc	960
His Gln Ala Phe Pro Glu Leu Val Arg Lys Ser Asp Glu Pro Ile Leu	
305 310 315 320	
aag gct ttc tgg aga gtt ggt cga ctt tac gcc aac tac gga gtc gtg	1008
Lys Ala Phe Trp Arg Val Gly Arg Leu Tyr Ala Asn Tyr Gly Val Val	
325 330 335	
gat ccc gac gca aag ctg ttt act ctc aag gag gcc aaa gct gcc tcc	1056
Asp Pro Asp Ala Lys Leu Phe Thr Leu Lys Glu Ala Lys Ala Ala Ser	
340 345 350	
gag gct gcc acc aag acc aag gct act taa	1086
Glu Ala Ala Thr Lys Thr Lys Ala Thr	
355 360	

<210> 86
 <211> 361
 <212> PRT
 <213> Phytophthora ramorum

<400> 86

Met Ala Thr Lys Gln Pro Tyr Gln Phe Pro Thr Leu Thr Glu Ile Lys
1 5 10 15
Arg Ser Leu Pro Ser Glu Cys Phe Glu Ala Ser Val Pro Leu Ser Leu
20 25 30
Tyr Tyr Thr Val Arg Ile Val Ala Ile Ala Val Ala Leu Ala Phe Gly
35 40 45
Leu Asn Tyr Ala Arg Ala Leu Pro Val Val Glu Ser Leu Trp Ala Leu
50 55 60
Asp Ala Ala Leu Cys Cys Gly Tyr Val Leu Leu Gln Gly Ile Val Phe
65 70 75 80
Trp Gly Phe Phe Thr Val Gly His Asp Ala Gly His Gly Ala Phe Ser
85 90 95
Arg Tyr His Leu Leu Asn Phe Val Val Gly Thr Phe Ile His Ser Leu
100 105 110
Ile Leu Thr Pro Phe Glu Ser Trp Lys Leu Thr His Arg His His His
115 120 125
Lys Asn Thr Gly Asn Ile Asp Arg Asp Glu Ile Phe Tyr Pro Gln Arg
130 135 140
Lys Ala Asp Asp His Pro Leu Ser Arg Asn Leu Val Leu Ala Leu Gly
145 150 155 160
Ala Ala Trp Phe Ala Tyr Leu Val Glu Gly Phe Pro Pro Arg Lys Val
165 170 175
Asn His Phe Asn Pro Phe Glu Pro Leu Phe Val Arg Gln Val Ala Ala
180 185 190
Val Val Ile Ser Leu Ser Ala His Phe Ala Val Leu Ala Leu Ser Val
195 200 205
Tyr Leu Ser Phe Gln Phe Gly Leu Lys Thr Met Ala Leu Tyr Tyr Tyr
210 215 220
Gly Pro Val Phe Val Phe Gly Ser Met Leu Val Ile Thr Thr Phe Leu
225 230 235 240
His His Asn Asp Glu Glu Thr Pro Trp Tyr Gly Asp Ser Asp Trp Thr
245 250 255
Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr Gly Ala Phe
260 265 270
Ile Asp Asn Leu Ser His Asn Ile Gly Thr His Gln Ile His His Leu
275 280 285
Phe Pro Ile Ile Pro His Tyr Lys Leu Asn Arg Ala Thr Ala Ala Phe
290 295 300
His Gln Ala Phe Pro Glu Leu Val Arg Lys Ser Asp Glu Pro Ile Leu
305 310 315 320
Lys Ala Phe Trp Arg Val Gly Arg Leu Tyr Ala Asn Tyr Gly Val Val
325 330 335
Asp Pro Asp Ala Lys Leu Phe Thr Leu Lys Glu Ala Lys Ala Ala Ser
340 345 350
Glu Ala Ala Thr Lys Thr Lys Ala Thr
355 360
<210> 87
<211> 1080
<212> DNA
<213> Pythium aphanidermatum
<220>
<221> CDS
<222> (1)..(1080)
<223> delta-17 desaturase
<300>
<302> DELTA-17 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS
<310> WO 2008/054565
<311> 2007-07-19
<312> 2008-05-08
<313> (1)..(1080)
<300>
<302> DELTA-17 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS
<310> US 2008-0125326-A1
<311> 2007-07-19
<312> 2008-05-29
<313> (1)..(1080)
<400> 87

atg gct tct tcc act gtt gct gcg ccg tac gag ttc ccg acg ctg acg Met Ala Ser Ser Thr Val Ala Ala Pro Tyr Glu Phe Pro Thr Leu Thr 1 5 10 15	48
gag atc aag cgc tcg ctg cca gcg cac tgc ttt gag gcc tcg gtc ccg Glu Ile Lys Arg Ser Leu Pro Ala His Cys Phe Glu Ala Ser Val Pro 20 25 30	96
tcg tcg ctc tac tac acc gtg cgc gcg ctg gcc atc gcc gcc tcg ctc Trp Ser Leu Tyr Tyr Thr Val Arg Ala Leu Gly Ile Ala Gly Ser Leu 35 40 45	144
gcg ctc gcc ctc tac tac ccg cgc gcg ctc gcg atc gtg cag gag ttt Ala Leu Gly Leu Tyr Tyr Ala Arg Ala Leu Ala Ile Val Gln Glu Phe 50 55 60	192
gcc ctg ctg gat ccg ctg ctc tgc acg gcc tac att ctg ctg cag gcc Ala Leu Leu Asp Ala Val Leu Cys Thr Gly Tyr Ile Leu Leu Gln Gly 65 70 75 80	240
atc gta ttc tcg gcg ttc ttc acc atc gcc cat gac tgc gcc cac gcc Ile Val Phe Trp Gly Phe Phe Thr Ile Gly His Asp Cys Gly His Gly 85 90 95	288
gcg ttc tcg cgt tcg cac ctg ctc aac ttc agc gtc gcc acg ctc att Ala Phe Ser Arg Ser His Leu Leu Asn Phe Ser Val Gly Thr Leu Ile 100 105 110	336
cac tcg atc atc ctc acg ccg tac gag tca tgg aag atc tcg cac ccg His Ser Ile Ile Leu Thr Pro Tyr Glu Ser Trp Lys Ile Ser His Arg 115 120 125	384
cac cac cac aag aac acg gcc aac atc gac aag gac gag att ttc tac His His His Lys Asn Thr Gly Asn Ile Asp Lys Asp Glu Ile Phe Tyr 130 135 140	432
ccg cag cgc gag gcc gac tcg cac cca ctg tcc cga cac atg gtg atc Pro Gln Arg Glu Ala Asp Ser His Pro Leu Ser Arg His Met Val Ile 145 150 155 160	480
tcg ctc gcc tcg gcc tgg ttc gcg tac ctc gtt ccg gcc ttc cct cct Ser Leu Gly Ser Ala Trp Phe Ala Tyr Leu Val Ala Gly Phe Pro Pro 165 170 175	528
gcg aag gtg aac cac ttc aac cct tcg gaa ccg ttg tac ctg cgc ccg Arg Lys Val Asn His Phe Asn Pro Trp Glu Pro Leu Tyr Leu Arg Arg 180 185 190	576
atg tct gcc gtc atc atc tca ctc gcc tcg ctc gtg gcg ttc gcg gcc Met Ser Ala Val Ile Ile Ser Leu Gly Ser Leu Val Ala Phe Ala Gly 195 200 205	624
ttg tat gcg tat ctc acc tac gtc tat gcc ctt aag acc atg gcg ctg Leu Tyr Ala Tyr Leu Thr Tyr Val Tyr Gly Leu Lys Thr Met Ala Leu 210 215 220	672
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 ccg cca tgg tac gcc gac tcg Thr Phe Leu His His Asn Asp Glu Glu Thr Pro Trp Tyr Ala Asp Ser 245 250 255	768
gag tcg acg tac gtc aag gcc aac ctc tcg tcc gtg gac cgc tcg tac Glu Trp Thr Tyr Val Lys Gly Asn Leu Ser Ser Val Asp Arg Ser Tyr 260 265 270	816
ggc gcg ctc atc gac aac ctg agc cac aac atc gcc acg cac cag atc Gly Ala Leu Ile Asp Asn Leu Ser His Asn Ile Gly Thr His Gln Ile 275 280 285	864
cac cac ctg ttt ccg atc atc ccg cac tac aag ctg aac gag gcg acg His His Leu Phe Pro Ile Ile Pro His Tyr Lys Leu Asn Glu Ala Thr 290 295 300	912
gca gcg ttc gcg cag gcg ttc ccg gag ctc gtg cgc aag agc gcg tcg Ala Ala Phe Ala Gln Ala Phe Pro Glu Leu Val Arg Lys Ser Ala Ser 305 310 315 320	960
ccg atc atc ccg acg ttc atc cgc atc gcg ctc atg tac gcc aag tac Pro Ile Ile Pro Thr Phe Ile Arg Ile Gly Leu Met Tyr Ala Lys Tyr 325 330 335	1008
ggc gtc gtg gac aag gcc gcc aag atg ttt acg ctc aag gag gcc aag Gly Val Val Asp Lys Asp Ala Lys Met Phe Thr Leu Lys Glu Ala Lys 340 345 350	1056
gcc gcc aag acc aag gcc aac tag Ala Ala Lys Thr Lys Ala Asn 355	1080
<210> 88	
<211> 359	
<212> PRT	
<213> Pythium aphanidermatum	
<400> 88	
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 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 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> 89
<211> 1080
<212> DNA
<213> Pythium aphanidermatum

<220>
<221> CDS
<222> (1)..(1080)
<223> synthetic delta-17 desaturase (codon-optimized for Yarrowia lipolytica)

<300>
<302> DELTA-17 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> WO 2008/054565
<311> 2007-07-19
<312> 2008-05-08
<313> (1)..(1080)

<300>
<302> DELTA-17 DESATURASES AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

<310> US 2008-0125326-A1
<311> 2007-07-19
<312> 2008-05-29
<313> (1)..(1080)

<400> 89
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 atc 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 gcc 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 cac 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 tgg cga tcc ccc 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 tgc 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 gcc 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 gcc 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 tgg 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 tgg 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 Phe 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 gag gcc aag 1056
 Gly Val Val 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> 90
 <211> 359
 <212> PRT
 <213> Pythium aphanidermatum
 <400> 90
 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> 91
<211> 1434
<212> DNA
<213> *Fusarium moniliforme*

<220>
<221> CDS
<222> (1)..(1434)
<223> delta-12 desaturase

<300>
<302> DELTA-12 DESATURASES SUITABLE FOR ALTERING LEVELS OF POLYUNSATURATED FATTY ACIDS IN OLEAGINOUS YEAST

<310> WO 2005/047485
<311> 2004-11-12
<312> 2005-05-26
<313> (1)..(1434)

<300>
<302> DELTA-12 DESATURASES SUITABLE FOR ALTERING LEVELS OF POLYUNSATURATED FATTY ACIDS IN OLEAGINOUS YEAST

<310> U.S. 7,504,259
<311> 2004-11-10
<312> 2009-03-17
<313> (1)..(1434)

<400> 91

atg	gcg	tcg	act	tcg	gct	ctg	ccc	aag	cag	aac	ccg	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 gct tct gng 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

ccg tca gac tct ccc cgc cac tgg gcc tct tcc acc tgg ctc tgg tcc 144

Pro Ser Asp Ser Pro Arg His Ser Ala Ser Ser Thr Ser Leu Ser Ser

35 40 45

atg tcc cag 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 cag tgc ttc aag cgc tcc ggt ctc 288

Asp Ile Tyr Asn Ala Ile Pro Lys His Cys Phe Lys Arg Ser Ala Leu

85 90 95

aag gga taa ggt tat atc ctc cgc gac att gtc ctc ctc act acc act 336

Lys Gly Tyr Gly Tyr Ile Leu Arg Asp Ile Val Leu Leu Thr Thr Thr

100 105 110

ttc ago atc tgg taa 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 cag 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 cag 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 taa ttc ago tgg caa atc tcc cag gga aag 576

Ser Ser 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 gat cat gtt act gag gag acc ccc gtt 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 gcc 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 gcc cac aac taa cac gag cgc cgt gag ggt cgc gcc aag gcc 816

Thr Gly His Asn Tyr His Glu Arg Gln Arg Glu Arg Gly Lys Gly

260 265 270

aag cat aac gcc ctc ggc ggt ggt gtt aac cac ttc gat ccc gcc agc 864

Lys His Asn Gly Leu Gly Gly Gly Val Asn His Phe Asp Pro Arg Ser

275 280 285

```

oct ctg tac gaq aac sgt 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 tac ttc 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

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 gcc ttc atc gcc 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 gcc      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 gcc 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 sgt 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 gcc acc ccc ccc gct gtt atc aag ccc gtt got taa      1434
Asn Val Gly Thr Pro Pro Ala Val Ile Lys Pro Val Ala
465                470                475

<210> 92
<211> 477
<212> PRT
<213> Fusarium moniliforme

<400> 92
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 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          400
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> 93
<211> 1434
<212> DNA
<213> Fusarium moniliforme
<220>
<221> CDS
<222> (1)..(1434)
<223> synthetic delta-12 desaturase (codon-optimized for Yarrowia lipolytica)
<300>
<302> DELTA-12 DESATURASES SUITABLE FOR ALTERING LEVELS OF POLYUNSATURATED FATTY ACIDS IN OLEAGINOUS YEAST
<310> WO 2005/047485
<311> 2004-11-12
<312> 2005-05-26
<313> (1)..(1434)
<300>
<302> DELTA-12 DESATURASES SUITABLE FOR ALTERING LEVELS OF POLYUNSATURATED FATTY ACIDS IN OLEAGINOUS YEAST
<310> U.S. 7,504,259
<311> 2004-11-10
<312> 2009-03-17
<313> (1)..(1434)
<400> 93
atg gcc tcc acc tcg gct ctg ccc aag cag aac oct 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 tot 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
oct 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 cag 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 gcc 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

ctg 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 gcc aag gga      816
Thr Gly His Asn Tyr His Glu Arg Gln Arg Glu Gly Arg Gly Lys Gly
260          265          270

aag cao aac ggt ctt ggc ggt gga gtt aac cat ttc gat ccc oga tct      864
Lys His Asn Gly Leu Gly Gly Gly Val Asn His Phe Asp Pro Arg Ser
275          280          285

cct ctg tac gaa aac aac gac gcc aag ctc atc gtg ctc too 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 aac 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 cao tgg ctc gtc gcc att aac 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 cao tac aac 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

cca aac atc gac cga gag atg ggc ttc att gaa cgt cct ctg ctc cao      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 tat 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 ggc gac gaa gct acc gag gcc atc aaa cct atc atg ggc      1248
Phe Tyr Asn Ala Asp Glu Ala Thr Glu Ala Ile Lys Pro Ile Met Gly
405          410          415

aag cao tat oga gct gat gtc cag gac ggt cct oga gga ttc att oga      1296
Lys His Tyr Arg Ala Asp Val Glu Asp Gly Pro Arg Gly Phe Ile Arg
420          425          430

gcc atg tac oga tct gca cga atg tgc cag tgg gtt gaa ccc too gct      1344
Ala Met Tyr Arg Ser Ala Arg Met Cys Gln Trp Val Glu Pro Ser Ala
435          440          445

ggt gcc gag gaa gct gcc aag ggt gtc ctg ttc ttt oga aac oga 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      1424
Asn Val Gly Thr Pro Pro Ala Val Ile Lys Pro Val Ala
465          470          475

<210> 94
<211> 477
<212> PRT
<213> Fusarium moniliforme

<400> 94
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 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      400
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> 95
<211> 828
<212> DNA
<213> Mortierella alpina

<220>
<221> CDS
<222> (1)..(828)
<223> C16/18 elongase

<300>
<302> A MORTIERELLA ALPINA C16/18 FATTY ACID ELONGASE

<310> WO 2007/046817
<311> 2005-11-04
<312> 2007-04-26
<313> (1)..(828)

<300>
<302> A MORTIERELLA ALPINA C16/18 FATTY ACID ELONGASE

<310> US 2007-0087420-A1
<311> 2005-10-19
<312> 2007-04-19
<313> (1)..(828)

<400> 95
atg gag tct gga cct atg cct gcc ggg atc ccc ttc cct gaa tac tat      48
Met Glu Ser Gly Pro Met Pro Ala Gly Ile Pro Phe Pro Glu Tyr Tyr
1      5      10      15
gac ttt ttc atg gac tgg aag aca ccc ctg gca att gct gcc acc tac      96
Asp Phe Phe Met Asp Trp Lys Thr Pro Leu Ala Ile Ala Ala Thr Tyr
20      25      30
acc gcc gct gtt ggg ctc ttc aac ccc aag gtt ggc aaa gtc tgg cgc      144
Thr Ala Ala Val Gly Leu Phe Asn Pro Lys Val Gly Lys Val Ser Arg
35      40      45
gtg gta gcc aag tgg gct aac gcc aag ccg gca gag cgc acg cag tcc      192
Val Val Ala Lys Ser Ala Asn Ala Lys Pro Ala Glu Arg Thr Gln Ser
50      55      60
ggc gcc gcc atg acc gcc ttt gtc ttt gtc cnc aac ctt atc ctc tgc      240
Gly Ala Ala Met Thr Ala Phe Val Phe Val His Asn Leu Ile Leu Cys
65      70      75      80
gtg tac tct gga atc acc ttc tac tac atg ttc cca gcc atg gtc aag      288
Val Tyr Ser Gly Ile Thr Phe Tyr Tyr Met Phe Pro Ala Met Val Lys
85      90      95
aac ttt aga aca cat acc ctc cat gag gcc tac tgc gat acg gat cag      336
Asn Phe Arg Thr His Thr Leu His Glu Ala Tyr Cys Asp Thr Asp Gln
100      105      110
agc ctg tgg aac aac gcc ctt ggc tac tgg ggc tac ctc ttc tac ctt      384
Ser Leu Trp Asn Asn Ala Leu Gly Tyr Trp Gly Tyr Leu Phe Tyr Leu
115      120      125
tca aag ttt tac gag gtc att gac acc atc atc atc atc ttg aag ggg      432
Ser Lys Phe Tyr Glu Val Ile Asp Thr Ile Ile Ile Ile Leu Lys Gly
130      135      140
cgc cgc tgg tcc ctg ctc cag acc tac cnc cnc gcc ggc gct atg atc      480
Arg Arg Ser Ser Leu Leu Gln Thr Tyr His His Ala Gly Ala Met Ile
145      150      155      160

```

```

acc atg tgg tcc ggc atc aac tac caq gca acg ccc att tgg att ttt      528
Thr Met Trp Ser Gly Ile Asn Tyr Gln Ala Thr Pro Ile Trp Ile Phe
      165      170      175

gtc gtc ttc aac tgg ttc atc 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 tca atc ggc ttc cac ccc cca ggc aag aag tac ctc acc tcc      624
Phe Thr Ser Ile Gly Phe His Pro Pro Gly Lys Lys Tyr Leu Thr Ser
      195      200      205

atg caq atc acc caq ttt ttg gtc ggc atc act atc gcc gtc tct tat      672
Met Gln Ile Thr Gln Phe Leu Val Gly Ile Thr Ile Ala Val Ser Tyr
      210      215      220

ctc ttc gtc cct gga tgt atc cgc aca ccc ggt gct caq atg gct gtc      720
Leu Phe Val Pro Gly Cys Ile Arg Thr Pro Gly Ala Gln Met Ala Val
      225      230      235

tgg atc aac gtc gga taa ctc ttt ccc ctc act tat ctc ttt gtg gat      768
Trp Ile Asn Val Gly Tyr Leu Phe Pro Leu Thr Tyr Leu Phe Val Asp
      245      250      255

ttt gcc aag cgt act tac tcc aag cgt agt gcc atc gcc gct caq aag      816
Phe Ala Lys Arg Thr Tyr Ser Lys Arg Ser Ala Ile Ala Ala Gln Lys
      260      265      270

aag gcc caq taa
Lys Ala Gln
      275

```

<210> 96

<211> 275

<212> PRT

<213> Mortierella alpina

<400> 96

```

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      80

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      160

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      240

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> 97

<211> 828

<212> DNA

<213> Mortierella alpina

<220>

<221> CDS

<222> (1).(828)

<223> synthetic C16/18 elongase (codon-optimized for Yarrowia lipolytica)

<300>

<302> A MORTIERELLA ALPINA C16/18 FATTY ACID ELONGASE

<310> U.S. 7,470,532

<311> 2005-10-19

<312> 2008-12-30

<313> (1).(828)

<300>

<302> A MORTIERELLA ALPINA C16/18 FATTY ACID ELONGASE

<310> WO 2007/046817

<311> 2005-11-04

<312> 2007-04-26

<313> (1).(828)

<400> 97
 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
 atg gtt gcc aag tcg gct aac gaa aag cct gcc aag 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 gcc 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
 tit 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> 98

<211> 275

<212> PRT

<213> Mortierella alpina

<400> 98

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 80
 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
 150 155 160
 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
225 230 235 240

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> 99
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> membrane bound O-acyltransferase motif

<220>
<221> misc_feature
<222> (4)..(6)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (9)..(11)
<223> Xaa can be any naturally occurring amino acid

<300>
<301> Hideo Shindou, Miki Etoa, Ryo Morimotoa and Takao Shimizua <302> Identification of membrane O-acyltransferase family motifs <303> Biochemical and Biophysical Research Communications <304>
383 <305> 3 <306> 320-325 <307> 2009-04-08
<313> (1)..(12)

<400> 99
Trp His Gly Xaa Xaa Xaa Gly Tyr Xaa Xaa Xaa Phe
1 5 10

<210> 100
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> membrane bound O-acyltransferase motif

<220>
<221> misc_feature
<222> (2)..(5)
<223> Xaa can be any naturally occurring amino acid

<300>
<301> Hideo Shindou, Miki Etoa, Ryo Morimotoa and Takao Shimizua <302> Identification of membrane O-acyltransferase family motifs <303> Biochemical and Biophysical Research Communications <304>
383 <305> 3 <306> 320-325 <307> 2009-04-08
<313> (1)..(6)

<400> 100
Tyr Xaa Xaa Xaa Xaa Phe
1 5

<210> 101
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> membrane bound O-acyltransferase motif

<220>
<221> misc_feature
<222> (2)..(4)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (7)..(8)
<223> Xaa can be any naturally occurring amino acid

<300>
<301> Hideo Shindou, Miki Etoa, Ryo Morimotoa and Takao Shimizua <302> Identification of membrane O-acyltransferase family motifs <303> Biochemical and Biophysical Research Communications <304>
383 <305> 3 <306> 320-325 <307> 2009-04-08
<313> (1)..(9)

<400> 101
Tyr Xaa Xaa Xaa Tyr Phe Xaa Xaa His
1 5

<210> 102
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> motif

<220>
 <221> MISC_FEATURE
 <222> (2)..(2)
 <223> Xaa = Val [V] or Ile [I]

 <220>
 <221> MISC_FEATURE
 <222> (3)..(3)
 <223> Xaa = Leu [L] or Ile [I]

 <220>
 <221> misc_feature
 <222> (4)..(5)
 <223> Xaa can be any naturally occurring amino acid

 <220>
 <221> MISC_FEATURE
 <222> (7)..(7)
 <223> Xaa = Leu [L] or Ile [I] or Val [V]

 <220>
 <221> misc_feature
 <222> (8)..(13)
 <223> Xaa can be any naturally occurring amino acid

 <300>
 <302> GENES ENCODING A NOVEL TYPE OF LYSOPHOPHATIDYLCHOLINE ACYLTRANSFERASES AND THEIR USE TO INCREASE TRIACYLGLYCEROL PRODUCTION AND/OR MODIFY FATTY ACID COMPOSITION

 <310> U.S. Patent Publication No. 2008-0145867-A1
 <311> 2007-06-15
 <312> 2008-06-19
 <313> (1)..(15)

 <400> 102
 Met Xaa Xaa Xaa Xaa Lys Xaa Xaa Xaa Xaa Xaa Xaa Asp Gly
 1 5 10 15

 <210> 103
 <211> 24
 <212> PRT
 <213> Artificial Sequence

 <220>
 <223> motif

 <220>
 <221> misc_feature
 <222> (2)..(2)
 <223> Xaa can be any naturally occurring amino acid

 <220>
 <221> misc_feature
 <222> (6)..(7)
 <223> Xaa can be any naturally occurring amino acid

 <220>
 <221> misc_feature
 <222> (9)..(11)
 <223> Xaa can be any naturally occurring amino acid

 <220>
 <221> MISC_FEATURE
 <222> (12)..(12)
 <223> Xaa = Glu [E] or Asp [D]

 <220>
 <221> MISC_FEATURE
 <222> (13)..(13)
 <223> Xaa = Ala [A] or Gly [G]

 <220>
 <221> misc_feature
 <222> (14)..(18)
 <223> Xaa can be any naturally occurring amino acid

 <220>
 <221> misc_feature
 <222> (20)..(20)
 <223> Xaa can be any naturally occurring amino acid

 <220>
 <221> MISC_FEATURE
 <222> (22)..(22)
 <223> Xaa = Phe [F] or Tyr [Y]

 <220>
 <221> misc_feature
 <222> (23)..(23)
 <223> Xaa can be any naturally occurring amino acid

 <300>
 <302> GENES ENCODING A NOVEL TYPE OF LYSOPHOPHATIDYLCHOLINE ACYLTRANSFERASES AND THEIR USE TO INCREASE TRIACYLGLYCEROL PRODUCTION AND/OR MODIFY FATTY ACID COMPOSITION

 <310> U.S. Patent Publication No. 2008-0145867-A1
 <311> 2007-06-15
 <312> 2008-06-19
 <313> (1)..(15)

<400> 103
Arg Xaa Lys Tyr Tyr Xaa Xaa Trp Xaa Xaa Xaa Xaa Xaa Xaa Xaa
1 5 10 15

Xaa Xaa Gly Xaa Gly Xaa Xaa Gly
20

<210> 104
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> motif

<220>
<221> misc_feature
<222> (3)..(3)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (7)..(8)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (12)..(13)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> MISC_FEATURE
<222> (14)..(14)
<223> Xaa = Thr [T] or Phe [F]

<300>
<302> GENES ENCODING A NOVEL TYPE OF LYSOPHOPHATIDYLCHOLINE ACYLTRANSFERASES AND THEIR USE TO INCREASE TRIACYLGLYCEROL PRODUCTION AND/OR MODIFY FATTY ACID COMPOSITION

<310> U.S. Patent Publication No. 2008-0145867-A1
<311> 2007-06-15
<312> 2008-06-19
<313> (1)..(15)

<400> 104
Ser Ala Xaa Trp His Gly Xaa Xaa Pro Gly Tyr Xaa Xaa Xaa Phe
1 5 10 15

<210> 105
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> 1-acyl-sn-glycerol-3-phosphate acyltransferase motif

<220>
<221> misc_feature
<222> (2)..(3)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> Xaa = Asp [D] or Arg [R]

<300>
<301> Tal M. Lewin, Ping Wang, and Rosalind A. Coleman <302> Analysis of Amino Acid Motifs Diagnostic for the sn-Glycerol-3-phosphate Acyltransferase Reaction <303> Biochemistry <304> 38 <305> 18
<306> 57645771 <307> 1999-04-15
<313> (1)..(7)

<300>
<301> Atsushi Yamashita, Hiroki Nakanishia, Hiroshi Suzukia, Ryo Kamataa, Ken Tanakaa, Keizo Wakua and Takayuki Sugiura <302> Topology of acyltransferase motifs and substrate specificity and accessibility in 1-acyl-sn-glycerol-3-phosphate acyltransferase 1 <303> Biochimica et Biophysica Acta <304> 1771 <305> 9 <306> 1202-1215 <307> 2007-07-17
<313> (1)..(7)

<400> 105
Gly Xaa Xaa Phe Ile Xaa Arg
1 5

<210> 106
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> 1-acyl-sn-glycerol-3-phosphate acyltransferase motif

<220>
<221> MISC_FEATURE
<222> (1)..(1)
<223> Xaa = Val [V] or Ile [I]

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> Xaa = Pro [P] or Xaa

<220>
<221> MISC_FEATURE
<222> (3)..(3)
<223> Xaa = Ile [I] or Val [V] or Leu [L]

```

<220>
<221> MISC_FEATURE
<222> (4)..(4)
<223> Xaa = Ile [I] or Val [V]

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> Xaa = Val [V] or Ile [I]

<300>
<301> Atsushi Yamashita, Hiroki Nakanishia, Hiroshi Suzukia, Ryo Kamataa, Ken Tanakaa, Keizo Wakua and Takayuki Sugiura <302> Topology of acyltransferase motifs and substrate specificity and
accessibility in 1-acyl-sn-glycero-3-phosphate acyltransferase 1 <303> Biochimica et Biophysica Acta <304> 1771 <305> 9 <306> 1202-1215 <307> 2007-07-17
<313> (1)..(6)

<400> 106
Xaa Xaa Xaa Xaa Pro Xaa
1 5

<210> 107
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> 1-acyl-sn-glycerol-3-phosphate acyltransferase motif

<300>
<301> Atsushi Yamashita, Hiroki Nakanishia, Hiroshi Suzukia, Ryo Kamataa, Ken Tanakaa, Keizo Wakua and Takayuki Sugiura <302> Topology of acyltransferase motifs and substrate specificity and
accessibility in 1-acyl-sn-glycero-3-phosphate acyltransferase 1 <303> Biochimica et Biophysica Acta <304> 1771 <305> 9 <306> 1202-1215 <307> 2007-07-17
<313> (1)..(6)

<400> 107
Ile Val Pro Ile Val Met
1 5

<210> 108
<211> 1024
<212> PRT
<213> Yarrowia lipolytica

<220>
<221> MISC_FEATURE
<222> (1)..(1024)
<223> YIPex1p; GenBank Accession No. CAG82178

<400> 108
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
225 230 235 240

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
305 310 315 320

Asp Gly Val Ser Ala Asp Asp Gln Phe Ser Leu Leu Glu Glu Leu Lys
325 330 335

```

Lys Leu Gln Met 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 Len 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
 385 390 395 400
 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 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 Val Gly Gly Arg Ala Thr Leu
 1010 1015 1020
 Ala
 <210> 109
 <211> 381
 <212> PRT
 <213> Yarrowia lipolytica
 <220>
 <221> MISC_FEATURE
 <222> (1)..(381)
 <223> YIPex2p; GenBank Accession No. CAG77647
 <400> 109
 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 Tyr 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 Phe Asp
 355 360 365
 Asp Asp Asp Glu Asp Ser Asn Phe Gln Leu Met Lys Asp
 370 375 380
 <210> 110

<211> 431
 <212> PRT
 <213> Yarrowia lipolytica

 <220>
 <221> MISC_FEATURE
 <222> (1)..(431)
 <223> YIPex3p; GenBank Accession No. CAG78565

 <400> 110
 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 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 Tyr Trp Val Ser Leu
 225 230 235 240

 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 Tyr Val Leu
 290 295 300

 Ala Gln Asn Pro Ser Asp Thr Ala Ala Pro Ile Val Gly Pro Thr Leu
 305 310 315 320

 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
 370 375 380

 Leu Ala Asn Met Ala Arg Gln Ala Gly Gly Met Ala Gln Gly Ser Pro
 385 390 395 400

 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> 111
 <211> 395
 <212> PRT
 <213> Yarrowia lipolytica

 <220>
 <221> MISC_FEATURE
 <222> (1)..(395)
 <223> YIPex3Bp; GenBank Accession No. CAG83356

 <400> 111

```

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 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    105    110

Gly Thr Glu Thr Thr Ala Ser Met Ala Asp Gly Gln Lys Met Ser Lys
115    120    125

Ile Gln Leu Trp Asp Glu Leu Lys Ile Glu Ser Leu Thr Arg Ile Val
130    135    140

Thr Leu Ile Tyr Cys Val Ser Leu Leu Asn Tyr Leu Ile Arg Leu Gln
145    150    155    160

Thr Asn Ile Val Gly Arg Lys Arg Tyr Gln Asn Glu Ala Gly Pro Ala
165    170    175

Gly Ala Thr Tyr Asp Met Ser Leu Glu Gln Cys Tyr Thr Trp Leu Leu
180    185    190

Thr Arg Gly Trp Lys Ser Val Val Asp Asn Val Arg Arg Ser Val Gln
195    200    205

Gln Val Phe Thr Gly Val Asn Pro Arg Gln Asn Leu Ser Leu Asp Glu
210    215    220

Phe Ala Thr Leu Leu Lys Arg Val Gln Thr Leu Val Asn Ser Pro Pro
225    230    235    240

Tyr Ser Thr Thr Pro Asn Thr Phe Leu Thr Ser Leu Leu Pro Pro Arg
245    250    255

Glu Leu Glu Gln Leu Arg Leu Glu Lys Glu Lys Gln Ser Leu Ser Pro
260    265    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    320

Val Asn Pro Pro Ser Thr 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> 112
<211> 153
<212> PRT
<213> Yarrowia lipolytica

<220>
<221> MISC_FEATURE
<222> (1)..(153)
<223> YIPex4p; GenBank Accession No. CAG79130

<400> 112
Met Ala Ser Gln Lys Arg Leu Ile Lys Glu Leu Ala Ala Tyr Lys Lys
1      5      10      15

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     80

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> 113
<211> 598
<212> PRT
<213> Yarrowia lipolytica

<220>
<221> MISC_FEATURE
<222> (1)..(598)
<223> YIPexSp; GenBank Accession No. CAG78803

<400> 113
Met Ser Phe Met Arg Gly Gly Ser Glu Cys Ser Thr Gly Arg Asn Pro
1      5      10      15

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     80

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    160

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 Glu Thr Thr Thr Glu Asp Lys Pro
225    230    235    240

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    320

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    400

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    480

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 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
545          550          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> 114
<211> 1024
<212> PRT
<213> Yarrowia lipolytica

<220>
<221> MISC_FEATURE
<222> (1)..(1024)
<223> YIPex&p; GenBank Accession No. CAG82306

<400> 114
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
 465 470 475 480
 Val Gly Lys Ser Thr Val Leu Arg Ser Val Ala Ala Gln Cys Gly Ile
 485 490 495
 Ser Val Phe Glu Ile Ser Cys Phe Gly Leu Ile Gly Asp Asn Glu Ala
 500 505 510
 Gln Thr Leu Gly Thr Leu Arg Ala Lys Leu Asp Arg Ala Tyr Gly Cys
 515 520 525
 Ser Pro Cys Val Val Val Leu Gln His Leu Glu Ser Ile Ala Lys Lys
 530 535 540
 Ser Asp Gln Asp Gly Lys Asp Glu Gly Ile Val Ser Lys Leu Val Asp
 545 550 555 560
 Val Leu Ala Asp Tyr Ser Gly His Gly Val Leu Leu Ala Ala Thr Ser
 565 570 575
 Asn Asp Pro Asp Lys Ile Ser Glu Ala Ile Arg Ser Arg Phe Gln Phe
 580 585 590
 Glu Ile Glu Ile Gly Val Pro Ser Glu Pro Gln Arg Arg Gln Ile Phe
 595 600 605
 Ser His Leu Thr Lys Ser Gly Pro Gly Gly Asp Ser Ile Arg Asn Ala
 610 615 620
 Pro Ile Ser Leu Arg Ser Asp Val Ser Val Glu Asn Leu Ala Leu Gln
 625 630 635 640
 Ser Ala Gly Leu Thr Pro Pro Asp Leu Thr Ala Ile Val Gln Thr Thr
 645 650 655
 Arg Leu Arg Ala Ile Asp Arg Leu Asn Lys Leu Thr Lys Asp Ser Asp
 660 665 670
 Thr Thr Leu Asp Asp Leu Leu Thr Leu Ser His Gly Thr Leu Gln Leu
 675 680 685
 Thr Pro Ser Asp Phe Asp Asp Ala Ile Ala Asp Ala Arg Gln Lys Tyr
 690 695 700
 Ser Asp Ser Ile Gly Ala Pro Arg Ile Pro Asn Val Gly Trp Asp Asp
 705 710 715 720
 Val Gly Gly Met Glu Gly Val Lys Lys Asp Ile Leu Asp Thr Ile Glu
 725 730 735
 Thr Pro Leu Lys Tyr Pro His Trp Phe Ser Asp Gly Val Lys Lys Arg
 740 745 750
 Ser Gly Ile Leu Phe Tyr Gly Pro Pro Gly Thr Gly Lys Thr Leu Leu
 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> 115

<211> 356

<212> PRT

<213> Yarrowia lipolytica

<220>

<221> MISC_FEATURE

<222> (1)..(356)

<223> YIPex7p; GenBank Accession No. CAG78389

<400> 115

```

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> 116

<211> 671

<212> PRT

<213> Yarrowia lipolytica

<220>

<221> MISC_FEATURE

<222> (1)..(671)

<223> YIPex8p; GenBank Accession No. CAG80447

<400> 116

```

Met Asn Lys Tyr Leu Val Pro Pro Gln Ala Asn Arg Thr Val Thr
1      5      10      15

Asn Leu Asp Leu Thr 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    240

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    320

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    400

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 Val Phe 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> 117
<211> 377
<212> PRT
<213> Yarrowia lipolytica

<220>
<221> MISC_FEATURE
<222> (1)..(377)
<223> YIPex10p; GenBank Accession No. CAG81606

<400> 117
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> 118
<211> 408
<212> PRT
<213> Yarrowia lipolytica

<220>
<221> MISC_FEATURE
<222> (1).(408)
<223> YIPex12p; GenBank Accession No. CAG81532

<400> 118
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 Met 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 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> 119
<211> 412
<212> PRT
<213> Yarrowia lipolytica
<220>
<221> MISC_FEATURE
<222> (1)..(412)
<223> YIPex13p; GenBank Accession No. CAG81789
<400> 119
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          80
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          160
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          320
Ala Glu Ile Asp Pro Ser Lys Leu Glu Phe Cys Arg Ala Asn Phe Asp
325          330          335
Phe Val Pro Gln Asn Pro Gln Leu Glu Leu Glu Leu Arg Lys Gly Asp
340          345          350
Leu Val Ala Val Leu Ala Lys Thr Asp Pro Met Gly Asn Pro Ser Gln
355          360          365
Trp Trp Arg Val Arg Thr Arg Asp Gly Arg Ser Gly Tyr Val Pro Ala
370          375          380
Asn Tyr Leu Glu Val Ile Pro Arg Pro Ala Val Glu Ala Pro Lys Lys
385          390          395          400
Val Glu Glu Ile Gly Ala Ser Ala Val Pro Val Asn
405          410
<210> 120
<211> 380
<212> PRT
<213> Yarrowia lipolytica
<220>

```


<221> MISC_FEATURE

<222> (1)..(380)

<223> YIPex14p; GenBank Accession No. CAG79323

<400> 120

```
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
```

<210> 121

<211> 391

<212> PRT

<213> Yarrowia lipolytica

<220>

<221> MISC_FEATURE

<222> (1)..(391)

<223> YIPex16p; GenBank Accession No. CAG79622

<400> 121

```

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     80

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    160

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    240

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    320

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> 122
<211> 225
<212> PRT
<213> Yarrowia lipolytica

<220>
<221> MISC_FEATURE
<222> (1)..(225)
<223> YIPex17p; GenBank Accession No. CAG84025

<400> 122
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> 123
<211> 324
<212> PRT
<213> Yarrowia lipolytica

<220>
<221> MISC_FEATURE
<222> (1).(324)
<223> YIPex19p; GenBank Accession No. AAK84827

<400> 123
Met Ser His Glu Glu Asp Leu Asp 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> 124
<211> 417
<212> PRT
<213> Yarrowia lipolytica

<220>

```

<221> MISC_FEATURE
 <222> (1)..(417)
 <223> YIPex20p; GenBank Accession No. CAG79226
 <400> 124
 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 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> 125

<211> 195

<212> PRT

<213> Yarrowia lipolytica

<220>

<221> MISC_FEATURE

<222> (1)..(195)

<223> YIPex22p; GenBank Accession No. CAG77876

<400> 125

```

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 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> 126
<211> 386
<212> PRT
<213> Yarrowia lipolytica

<220>
<221> MISC_FEATURE
<222> (1)..(386)
<223> YIPex26p; GenBank Accession No. NC_006072, antisense translation of nucleotides 117230-118387

<400> 126
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 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> 127

<211> 8739

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pY116

<400> 127

ggccggccacc gggggccgag attccggcct ctccggccgc caagcgaccc ggttggaagt 60
 ctgagagttac ctagcaatta acagatagtt tgcgggtgat aattctctta aactccacaa 120
 ctccctttgac ataacgattt atgtaacgaa actgaattt gaccagatat tgtgtccgcg 180
 gtggagctcc agcttttgtt ccccttagtg aggtttaa ccagcttgcc gtaatcatgg 240
 tcatagctgt ttcctgtgtg aaattgttat ccgtccaaa tccacacaa cgtacgagcc 300
 ggaagcataa sgtgtaagc ctgggtgccc taatgagtga gctaatccaa attaatgccc 360
 ttgogctaac tgcgcgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc 420
 ggccaaacgg cggggagagc cggtttgagt attggcgagt ctcccgcttc ctgctcaact 480
 gactcgtccg gctcgtcgtt tccgtccgcg ccagcgggat cagctcaact aaagcccgta 540
 atccggttat ccacagatcc aggggataac gcaggaagaa acatgtgagc aaagggccag 600
 caaaaggcca ggaacgttaa aaagccgcgc ttgtcgcggt ttctccatag gtcgcgcgcc 660
 cctgcagagc atccaaaaa tcgacgctca agtcagaggt ggcgaaaccc gacaggacta 720
 taagaatacc aggcgtttcc cccctggaag tccctcgtgc gctctcctgt tccgacccgt 780
 ccgttacccg gataccctgc ccgctttctc ccttcgggaa gactggcgct ttctcataga 840
 tcccgctgta gctatctcag ttccgtgtag gtccctcgtt ccaagctggg ctgtgtgcac 900
 gaaccccccg ttacgcccga ccgtctgcgc ttatccgcta aetatcgtct tgagtccaac 960
 ccggttaagac ccgacttacc gccactggca gccgcacatg gtaacaggat tagcagagcg 1020
 aggtatgtag ccggtgctac agagttcttg aagtgctggc ctaactacgg ctacactaga 1080
 agaacagtat ttggtatctg ccgtctgctg aaqcacgta ccttcgggaa aagagttggt 1140
 agctcttgat ccggcaaaaa aaccacgcct gctagccgtg gttttttgt ttgcagagca 1200
 ccgattacgc gcagaaaaaa aggtctccta gaagatcctt tgatcttttc taagggtctc 1260
 gacgtccagt ggaacgaaaa ctcccgctaa gggattttgg taatgagatt atcaaaaagg 1320
 atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcatctta aagtatatat 1380
 gagttaacct ggtctgacag ttaaccaatgc ttaatcagtg aggcacctat ctacgcagtc 1440
 tgtctatttc gtctatccat agttgcctga ctcccgctcg tctagataac taagatacgg 1500
 gaggccttac catctggccc ccgtctctga atgataccgc gacacccacc ctacccggct 1560
 ccagatttat cagcaataaa ccagccagcc ggaaggcccg agcgagaaag tgtctcctga 1620
 accttaacg cctccatcca gctatataat tgttgccggg aagctagagt aagtgttcg 1680
 ccagtaata gtttgccaaa ccgtcttgcc attgctaacg ccactcgtgt gtcacgctcg 1740
 tctgttgcta tggcttcatt cagctccggt tcccaacgat caaggcgagt taactgatac 1800
 cccatgttgt gcaaaaaagc ggttagctcc ttccgtctcc ccagctgtgt cagaagtaag 1860
 ttggccgagc tgttatcact catgttatg gcagcaatgc aaattctct tactgtcctg 1920
 cctcccgtaa gatcttttc tgtgactggt gactactcaa ccagtcattc ctgagatag 1980

tgtatgoggo gacogagttg etottgocog gegtcaatac gggataatac ogogocacat 2040
 agcagaactt taaaagtgt catcatgga aaacgttctt oggggogaaa actctcaagg 2100
 atcttacocg tgttagagatc agttcagtg taacccactc gtgcacccaa ctgatcttca 2160
 gcatctttta ctttaacog ogttttctgg tgagcaaaaa caggaaggaa aaatgocga 2220
 aaaaaggaaa taaggggagac aoggaatgt tgaatactca tactctctct ttttaaatat 2280
 tattgaagca ttatacagg ttattgtctc atgagcggat acatattga atgtatttag 2340
 aaaaataaac aaatagggtt tocgogacac ttctccgaaa aagtgcacac tgaogogoc 2400
 tglagogggc cattaaogc goggggtgt gtgttacgc gogcgtgac ogtaaacctt 2460
 gogagogoc tagogocogc tctttcgtt ttcttctctt cctttctcgc caogttgoc 2520
 ggcttccccc gtcaagctct aaatcggggg ctccctttag ggttccgatt tagtgcttta 2580
 cggcactcgc accocaaaaa ccttgattag ggtgaggtt cagtgatgg gocatogoc 2640
 tgatagaggg tttttcgcct tttagcgttg gagtccactt totttaaatg tggactcttg 2700
 ttccaaactg gaacaaactc caacctctc tcggtctatt ctlttgattt ataaaggatt 2760
 ttgocgattt cggcctattg gttaaaaaat gagctgattt aacaaaaatt taacoggaat 2820
 tttacaaaaa tattaagct tacaatttcc attcgcattt caggctgogo aactgttggg 2880
 aaggcgatc ggtgcgggoc tcttgcctat tacgcagctt ggcgaaggg ggtgtgtgt 2940
 caagcgattt asgttgggta acgcaggggt ttccocagtc acgacgttgt aaaaagcgg 3000
 ccagtgaatt gtaatacagc tcaactaggg gcaattggg tacggggoccc cccctogagg 3060
 togatgtgtt cgaataagct gatacgaat tcatgtcaca caaacgacgc ttocctcaa 3120
 ggaacactaa ttctacatcc gagagactgc cgaactcag tctacactga ttaattttcg 3180
 ggccaatat ttaaaaaaat oggtttatat aatattatat gtattatata tatatacat 3240
 gatgatactg acagtctgtt cccattgcta aatagacaga ctccactcgc cgcctcaac 3300
 tgaigtcttc aatatttaag ggtctctcgc gcatgttta ataataaaca gactcactc 3360
 accgcctcca aatgatgttc tcaaatata ttgtatgaac ttatttttat tacttagtat 3420
 tattagacaa cttaacttgt ttatgaanaa cactctcat ttaggaaaca attataatg 3480
 gcagtctgtt catttaacaa tttatgtaga ataatgtta taagtogta tgggaactct 3540
 taatatgga tagcataaat gatactgca ttgcctaatt cgaacacac agcaacgaaa 3600
 aaaaacccct gtacacata aatagtcac gagaaatato aactacaaa gaacagctat 3660
 tcaacagctta ctattgagat tattattgga cgaataatca acactcaact gtcttctct 3720
 ctcttagaaa taacgggaca agtatgtact attctcattg tctactcttc tagtacttc 3780
 atccacacata ttcttggat ttctctccaa tgaatgaact tctactcttc aaattcaaca 3840
 attataataa gatatacaaa agtagcggtg tagtggcaat caaaaagctt ctctgtgtg 3900

ctctcgtas ttatttttat totaatgac cattaaaggt atatatattt ttctgttat 3960
 ataactcttt tgtttattac atgggtgga taataaagg tatattgatt taatttttg 4020
 cttaattcca atcccccct gttaagtgc aactgaatg taggaattt acatacttt 4080
 tgaagaagca aaaaaaatga aagaaaaaaa aatcgtatt tccagtttag acgttcgca 4140
 gaactagaa tgcgtatgc ggtacattgt tcttgcagag taagaattgc gctccctgag 4200
 atattgaca ttttgottt taacagtaaa agtactatgt acaactatgt actactgtt 4260
 atgcactcac aacagtgtt ttgtttttt ttgtttttt tttttctaa tgattcaata 4320
 cgcctatga tactactgt tactctagt aacccgggtt attgoggtto aattaatcat 4380
 agactatga atctgcaggg tgtgcctgc gactacttt tagctatgc atgcactgt 4440
 gggttaatat tgggtatgt tcggaactca accgtctgc aacgcatttc gacagtaatt 4500
 aattaatttg aatcgaatgc gagctaaaa tgaacccag tatatctcat aaattctcgc 4560
 gtgagaggtc tgtgactgc agtacaaggt gcttcaata tgcctcaac cttaacatac 4620
 ctactgaat gtatgtacc tctaaaaatg aaatacagtg ccaaaagcca aggcactgag 4680
 ctgcctcaa ggaactgata taacacaaat taacacaaat gaagaagaa acagtcttt 4740
 gtactattg taacaaatc actgtacaaa ctgaagttat gaactccac aatattccca 4800
 aagtccaccc ctttcaaat tgtcatgctt acaactcata taacaaagca taactacaa 4860
 aacacacata aaaccccaaa aaatatatct taacgaatat aagtaacaaa gatacaacaa 4920
 cactcgttg gtgcagtcgc cagcttaag atactatcc aactacgcca caactccctt 4980
 ccttaataa accgactaca cccctggcta ttgaagttat gactgaatat actgtagaca 5040
 agacacttct aagaagactg ttccaaaaa gtacactgt cctcactaac aaacacaccc 5100
 aatctgcttc tcttagtcaa ggttgtaca ccggttaatt ataactatc atttcttag 5160
 caggcgagg ccccttttat agagtcttat acactagcgg aacgtgcgg tagacacac 5220
 ogcagcgcg toagtttgt ccttccatca atgcctgta gaaagactt actctctt 5280
 gacagctac ttgacttgt tggcaacaa tctcgaact cggaggtgga ggaagagct 5340
 ccgatatcg cgttagtcat accgcctcg acgaactct tgaaggcag ctcaacagcg 5400
 tcaccggcg gcttcatgtt aagcagaaac ttgaacatca tggcgccaga cagaatgtg 5460
 gcaatgggtt gacactctg cttgcogaga tgggggcaag atccgtgaac gggctgtac 5520
 ngacogaaog cctcgttgt gtccggcaga gaagcagag aggcggagg cagcagaccc 5580
 agagaacgg ggaatgagga ggcctcgtcg gagatgatat cgcacacat gtcgtgtgtg 5640
 atgatgata cattcatctt ggaagctgc ttgatgaga tcatggcgc cagctcagtc 5700
 agctgttgt tgaactcag ctgagggaa ttgtcttga ggaactcagat gacagtctt 5760
 cgcacagtc gagaggagc cagcagcttg gcttgtcaa gagacacac ggaagaggg 5820
 ggttgttgt gaaggcgag gaaggcgcc attcggcaca ttgcctcaac ctacggaacg 5880

gagtaggtct cgggtgtaga agagacgcca gatcogtcat cctcctttag ctctccaaag 5940
 tagatacctc cgaacagctc toggacaatg atgaagtogg tgcctccaac gtttcggatg 6000
 gggagagat cggcagactt cggcagacag agctggcagg gtccagactt ggcctacagg 6060
 ttcaaggtct ttccagactt gaggagaccc tgcctgggtc gacgtcgggt tgcctcgtcg 6120
 ggaatggctc atcaggtggt ggaagcgcct ccgacagcac cgaagataat agagtccagc 6180
 ttccggcaga gtccagagat agcgtcgggt atgggtcgc cctccttctc aatggcagct 6240
 cctccaatga gtccgtctcc aaacacaaac tgggtgcggg aggcctccag aacagacttg 6300
 agccacttga cggcctcggc aatccctcgc gggccacaga agtcgcgcgc gagaagaaca 6360
 atctctctgg agtcagctct ggtctctctt gtttcgggtt ccattgtgga tgtgtgtggt 6420
 tgtatgtgtg atgtcgtgtg tggagtgaac atcgtgtgct ggcacacgct cttgtatata 6480
 taagcaactt tgcctcgtct atgtggaaga ctaaacctcc gaagattgtg actcaggtag 6540
 tcgggtatcg gctaggagcc caaaccttgt cgtatccgat agcgtatcgc aacgtacccc 6600
 agccggccgc gaatgtctgc gaggcagcat acgagatcgt caaggttttg tggcuaactg 6660
 gtatttaaat gtatgttaac gtacagggcg aactactggt acatacctcc ccgggatat 6720
 gtacaggaat aatcgtatc tgtggacat gtggtcgttg cgcactatg taagcagcgt 6780
 gtactcctct gactgtccat atggtttgtt ccactccacc ctcactgttt tcatgttca 6840
 caggcgccca caaaaaaact gtctctctc cttctctctt cgccttagtc tactcggacc 6900
 agtttttgtt tagcttggcg ccaatggata aatgagaccc caggccttgt gatgaggagg 6960
 tcactatga agcagttag gagggtcttg tatgtataga gaagacacca aaataaag 7020
 aataataata aaacaggggg cgttgtcatt tcatctcgtg ttttccact caataaccc 7080
 caaaacatg cccctcatgt cggcagcccc aatattgtcc tgtagtccaa ctctatgca 7140
 ctgctatctt attgagcaag taaaactctg tcagccgata ttgcgcgacc cgcgacaagg 7200
 gtccacaaag tgggttaagg ccttcgcaga agtcaaaact gtgcacaaac aacatctaga 7260
 gctcctttgg tgtctctcgc atatatctwa tggcgtctct taactatttg cgcctcggta 7320
 ccggaataat ttccgtatcc ccccaatcgc cttttctctc gacgtctcca acagtgtcca 7380
 tgatctatcc acataaatg gtccatgatg gcgtataatt tgcgtgtgct gataataatt 7440
 cccatattt tgacacaaaa ctcccccacc tagacataca tctccaaatc tcaactctg 7500
 tgcctctgtc acacatctcc tccagctgac ttaactccac accctcgcgc cagtgtgtct 7560
 acaggggat aaggtttctc cgcataagag tgccacactc ctcccgatac ttgttttgt 7620
 gacttgtgg tcacgacata tatatctaca ccacttgccc caccctttgg ttcttcagc 7680
 acaacaaaaa ccagcacgcg taacactgga caatttactg accgtacacc aaatttgc 7740
 tgcatlaacg gtcatgcaa cagatgata ggttcgcaag aacgtgatgg acatgttca 7800

ggaicgcacg cgtttttctg agcataactg gaaaatgctt ctgtccggtt gcgggtcgtg 7860
 gggcgcatgg tgcaggttga ataacccgaa atggtttccc gcagacactg aagaigtctg 7920
 cgaattatct ctatatcttc aggcgcggcg tctggcagta aaactatccc agcaacattt 7980
 gggccagata aacatgcttc atcgtcggtc cgggctgcca cgaaccaagt acagcaatgc 8040
 tgcctcactg gttatcgccc ggcctcgaaa agaaaaagtt gatgcggctg aacgtgcaaa 8100
 acaggctcta gcgttcgaac gcaactgatt cgaaccaagt cgttccacta tggaaaatag 8160
 cgcctcctgc caggatatac gtaactctgc atttctggg atgtctata acacctggtt 8220
 acgtatagcc gaatttgcca ggcacagggt taagatatac tcaactctg acggtgggag 8280
 aatgttaact catattggca gaacgaaaac gctgcctaga accgcaggtg tagagaaagc 8340
 acttagcctg gggglaacta aactggtoga gcgtggatt tccgtctctg gtgtaactga 8400
 tgatcgaat aactactgt tttgcgggt cagaaaaaat ggtgttgcgc cgcactctgc 8460
 caccagccag ctatcaactc gcgccttgga agggattttt gaagcaactc atcagattgt 8520
 ttaacgcgct aaggatgaat ctggtccagc ataccctgac tgcctggac accgtgcgcg 8580
 tgcggagacc gcgcgagata tggccgcgcg tggagtcca ataccggaga tcatgcaagc 8640
 tgggtgctgg accaatgtaa atattgcat gaactatata cgtaacctgg atagtgaac 8700
 aggggcactg gtgcgcctgc tggaaagagg cgaataagc 8739

<210> 128
 <211> 13975
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Plasmid pZKSL-5S5A5

<400> 128
 cgaatgttaa tgcctctgta tgcctctgct gctgtgtgtt tggcgcctat tgttcgctta 60
 tgcagcgtac accacaatat tggaaactta ttagctcttc tattttttcg tttgcagagg 120
 ttaacacact tgcctgtgag agggatgggg atatggaggc cgtcggaggg agtcggagag 180
 gcgttttggg gcggtttggc ctggcgccca gctcgggaaa cgcacctagg accctttggc 240
 acgcggaat gtgcacattt tcagtctagt aacgccttac ctactctatt ccatgcgtgc 300
 atgttttgcg cttttttccc ttgcctctga tgcacacaca gtacagtcca ctgtacagt 360
 gaggttttgg ggggttctta gatggagctt aaaaagcgcc tagcggatca ctagtgggat 420
 tcatggaggt ggcattggag ctatgtggag cctgacagga cgcacgacg gctagccgt 480
 gacagacgat ggggtgctcc tgtgttccac cgcgtacaaa tgtttgggccc aaagtcttgt 540
 cagccttgcg tgcgaacta attcccaatt ttgtcacttc gcaaccccat tgatcagacc 600
 ctaacccctg cccatcagga aatccaaata agctcgcatt gtcctgcctg tttagttttg 660
 ctctcgcccg tttcggctgc caettgcaca aacacaaaca agcattatat ataaagctcg 720

tototocoto ccaacccaac tcaotTTTTt gccgtotto cettgctaac acaaaagtoa 780
 agaacacaaa caaccacccc aacccocctt cacacaaagc atatctacag caatggccat 840
 ggcctcagct cttacccacg aacagctggt agaacgcctt gatttggttg cgaatgtagg 900
 catctctacg gaecttgaag ggottgcaa agttcatoca ggaatcgatt tgattctogc 960
 ttctgggtgc totgatgoot cccctctctt ttatctaatg catccatacg tcaaacccga 1020
 gaactccaaa ttgcttcaac agtttgtccg aggaagcat gccgcacot cgaaggacat 1080
 tgtctacacg tatgatcttc ccttgcaca agacgttaag ccgcaaatgc gcaagtgat 1140
 gaagggagg aactggtagc caacccclgg ctictggctg ccacccgttg ggaatctcgc 1200
 cgtgaaggcc ttttgagagt ggcactgggg taccacgggy atgtgtcgtt ggggcctgtt 1260
 gactggattc atgcacatgc agatggctt atccatccag catgatgagt cccacggggc 1320
 catcagcaag aagccttggg tcaacgcctt cttegctac ggaattgacg taatgggac 1380
 gtccgggtgg atttggtgc agtcgcacat catggggcac cacacotaca ccaacccaca 1440
 ccgcctcagc clggtatccg agtcgcaga gccgttcttg ctgttccaca actacccgcg 1500
 ccgaacaccc gcccgaaagt gcttccacgc cttccaggct tggtaaatgt acctgtgct 1560
 gggggcatac ggggtatagc tgggtacaa cccgtctac attttccgga tgcagacaaa 1620
 tgacacatc ccagagcttg tccggccat gccgaaaaat ggccttctgc ggcgtacag 1680
 cacacttcca ttctgtatgc gacttttctt catcttcgg accgacttct tgccttgcta 1740
 cctgaatggg acctaatgc tgatcccat tctctgggt cccacccgaa ctggtgcctt 1800
 ctgaagttc ttctcattt ttgtccaca ttittgagc tccgaaggga tcccgacaaa 1860
 gaactgcaag gtttaagcat ctgagaagga cgttgagct gaccaaatg actgtatcg 1920
 ggcagagtg gagaagctt ccaatacagc tggcccatc gcaatgttt tcaatggcgg 1980
 tctcaattc cagatagagc aacacotctt tcccgagtg tctcttgga actaacctt 2040
 cgtccagcag ccgttcaggg agtgttgcga accaatgga gtgcataag ttcttacc 2100
 taccatgtc ggcacacata tctccacctt gaaatcacat cataaggttg gtctgtcca 2160
 ctgctgaag gccgcacagg attcctaagc ggcgccttg atgattggaa acacacacat 2220
 ggggtatato taggtgagag ttatgtggac agttatatat taatcaagat atgcaacgg 2280
 taacttcatt cagtccaag aggaacaggt gactgcaagt aatatagaat ttgacacat 2340
 tgcattctc ttgcactctt ttactatato tcaattattt ctatataca aatcaattct 2400
 tcttccagc atcgagctcg gaacactcat gagaataaac atcgtgato tctgaatag 2460
 agggctttt ggaactcctg ctgttggcca ccttgccctt gctgtcggc taattctgtt 2520
 tcaacgcctt ttaattaaat ctctgcagc ttctccagc agacactgct cctgtctgga 2580
 ccagagotaa ggcctctgtc gacaaacagc ttctgggaat tggcgagtaa tttaattat 2640

gcaaatagac gtgtatttga aaaggagagc atgatgtag cagtaattt cgtgtttga 2700
 atagacaaat tatatatata aaattacat gatgaatgag tgaatgaat gatgtctaa 2760
 tatccagag gagagccggc atttcaata ttattgtcc taagaagcta aatatgctg 2820
 ttccgttat agcagtgctt gtttgagat caatggcaga accggagga gtattgttag 2880
 ttgggtatcg gtttactga catcgatgag gggcgagag agaatatc cgaatccat 2940
 tgaattgtt attagagat tcaaacaca actccotata gaotcgggg agtgtgta 3000
 gaagaatgac ctggagcaag agtgttcaag ttgatcgtt caattttca agattaact 3060
 tgggtgtgt ttggaatca catgatctg accaatatg catctgagcg ctggaaagt 3120
 ctcccttcc cttcagctt atcagagtg attccataag ccgaacotc cgaattgct 3180
 tgggtgtga agccatagc aaagtgtgg ctatggaatg catgacagct ccggcccaat 3240
 tgcctcata gtgatctga ttacattca ctggcgtcg tttaacagc tctgactg 3300
 gaaaacccg ggttaacca acttaatgc cttcagac atcccttct ccgaagctg 3360
 cgtaatagc aagagcccg caccgatgc cttcccaac agttgcgcag cctgaatgg 3420
 gaatggagc gcccttagc ggcacataa gccggcggg tgtgtgtt acygcagcg 3480
 tgacgctac ccttgcagc gccatagcg ccgtcttct cgttttct cttcttct 3540
 tgcacagct cgcggctt ccccgtaag ctataatcg ggggtccct ttgggttcc 3600
 gatttagtg tttaaggac ctgacccca aaaaattga ttagggtgat gttcaagta 3660
 gtggccatc gccctgata ccggttttcc gcccttgac gttggagtc aagttcttta 3720
 atagtgaat ctgttccaa actggaacaa cactcaaccc tatctogtc tattctttt 3780
 attataagc gatttgcgc atttcgctt attgttaaa aatgagctg atttaacaa 3840
 aatttaagc gaatttaac aaattataa cgttaaat ttctgatgc ggtatttt 3900
 cottaagat ctgtgggta ttacacag cactagtg cactttcgg ggaatgtg 3960
 ggggaaccc tatttgtta ttitttcaa tcaattaaa tatgtatcg ctatgagac 4020
 aataacctg ataatgctt caataatatt gaaaaggaa gattatggt atcaacatt 4080
 tccgtgagc ccttatccc ttitttgcg catittgctt tctgtttt gctacccag 4140
 aaacgtggt gaagttaaa gctgtgag atcatttgc tcaacagtg gttacatc 4200
 naatggatc caacagcgt aagatcctg agagtttgc ccccgagaa cgttttcaa 4260
 tgatgagac ttttaagtt ctgtatgtg gccggtatt atccgtatt gaagccggg 4320
 aagagcaat ccgtgcgcg atacactatt ctcaaatga cttggttgag taatcaacg 4380
 tcaacgaaa cactctaac gctgcatga cgttaagaa attatgcat gctgccata 4440
 ccatgagta taacactgc gccacttac ttctgacac gatcgagga cgaaggagc 4500
 taacgcctt ttgcaaac atgggggac atgaactcg ccttgatctg tgggaacgg 4560
 agctgaatga agcactaca aacagagagc gtgaacacac gatgcctgta ccaatggca 4620

caacgtttgag caaaactatta actggcgagac taactactct agcttccggg caaacaattaa 4680
 tagactggat ggaagcggat aaagtctgag gccactttct gcgctcgccc ctccggctg 4740
 gctggttat tgcgtataaa tctggagccg gtgagcgtg gtctcgcggt atcatctgag 4800
 cactggggcc agatggttag cctcccgta tctagtatt ctacaagagc gggatcagg 4860
 caactatgga tgaacgaat agacagatcg ctgagatagg tgcctcaatg attaagcatt 4920
 ggtactgtc agaccaagtt tactcatata tacttttagt tgatttaaaa ctctatttt 4980
 aatttanaag gatctagggt aagatccttt ttgataatct catgaccaaa atccttaac 5040
 gtgagtttc gtccactga gcgtcagacc ccgtagaaaa gatcaagga tcttcttgag 5100
 atccttttt tctgcggtta ctctgctgct tgcacacaaa aaacccacgc ctaccagcg 5160
 tggttgttt gcggatcaa gactaccaa cctcttttc gaaggttaact gcttcagca 5220
 gagcgagat accaataact gttctctag ttagcggta gttagggcac caottcaaga 5280
 actctgtagc acgcctaca taactcgtc tgcatactt gttacacgtg gctgctgca 5340
 gtggcgataa gtcgtctct acccggttgg actcaagacg atagttaccc gataagccgc 5400
 agcgctcggt ctgacggggg gcttctgca cacagcccg ctggagagca acgactaca 5460
 cogaactgag atactcaag cgtgactat gagaagcgc acgcttccc gaaggagaa 5520
 agcgagacg gtatccggtt agggcaggg tgggaacag agagcgacg agggagctc 5580
 caggggaaa cgcctggat ctctatagtc ctgtcggtt tgcacacct tgactttagc 5640
 gtccatttt gtgctgctc tccgggggga ggaactatg gaaaaacgc agcaacggg 5700
 ccttttacg gtccctggc ttgtctgga ctttctgca catgtcttt cctcgctat 5760
 cccctgatt tgtgataac cgtattacgc cttttagtg agctgatac gctcgcgca 5820
 gccagacac agagcgacgc cagtcaatga gccaggaagc ggaagagcgc caatacga 5880
 aacgcctct ccccgcggt tggcgatct attaatgca ctggcgcgcc gaggttcca 5940
 acagataac atcgtggcag ctgcacccat gccgatatg cctgaataa gtgaattgac 6000
 catcagagt tctgattcca aatagaacta tatcatatt ccgtcagttt ttcaataat 6060
 agtccatcc ctgaactccc tatcaaat ggtcctgac cagccacatc catgatcatg 6120
 actgctgac aggaacagtc attcgtgga tgaacactt taactcagg tactgtaat 6180
 ctctgaag ggcagacaa caaccattg agtaacctg gagactgaa acgtaatg 6240
 acttcacac caaactcact tgcctcaac gtgctctca attgcacaaa atcactctg 6300
 actaatctat tgcagaggtc aggtatgaa caactagac agcgtactt atgtgtata 6360
 tagaggaata taaaaaatc taagaataat cataagtta caaatlaatt atcagttc 6420
 gagccacag taccctcaa catgtgttat tgcacacaa caactcctag ctgatgca 6480
 tttaattct ccagcacaca tctctctc caacctaat ggcacattt atttgttg 6540
 agcgtact ctactcaac tgcacatac ttaacacaa ttgtccccc gctgatac 6600
 acacacatc catgtcaag cttaacaca taacaacaa cagctacgc aactacatg 6660
 aatgaatcg atatactca aagctatga tcccgatct tgcacactgt acgataaga 6720
 gcactagcca ctgcactcaa gtgaacccgt tgcgggga cagtatgag tatgtacgt 6780
 atgtttatga ttgtactgg acagtctgt tatcgtact tctcagttgt caaacataa 6840
 tatcgttgc tatatctcg caacacacag tagctcgata tatccctgt ttgaatcat 6900
 caactctgga ttgcacattg tgcacacaga accgggcaat caattccca tccacactg 6960
 cggccgctta ggaatcctga gcgtccttga caactgaac caacacgaat ttgtcatgt 7020
 acttgaggt ggaatgatg ttgcacaaa tggtaggta gaagacgtac cgaactcgt 7080
 gtctctgca acactctcg acagctgct gccgaaggg atagtccaa gaacacatc 7140
 gaggaagag gtgatctcg atctggaagt tgaacccgc agtaagaac atggcaatg 7200
 gtccacgta ggtggaag gtctccact gactctgta ccagtctat tcatcgctt 7260
 caacgtctt cctggagtc ttgaccttgc agttcttgc ggggattcgc tccagacat 7320
 cgaagtttg agacagatg aaaaagaag tgaagaagc accgtagca gtggcacca 7380
 gaggaatgt gatgacgg ggggttcag tgaatacaa gggcaagag cgggttcga 7440
 agatgaaga agctcgata acgaatgca ggttcggta cgttcgaga aagcgttct 7500
 ctgcctgga tgtgacagac tgggaatgg tctctgtg ctgcattcga aagatgaga 7560
 gaggttcta caccagoga acccgtagg ctccagaca gagtcatg taccaaggct 7620
 ggaatcgtg aaacacattt cgaacaggt tgcacagag ctagtctgc aacacaaqa 7680
 atgtcttgc gaactcgc tccagctga gccatgct atttgttag gtgtatgt 7740
 gcatgatg agactcaga cagatcata tgaacatcc aatgaatcg atgcgtagg 7800
 caaagagag gtgacccag ggccttttgc tctggcacc atgagagga tctgtgaa 7860
 tggacagcc gatctgat tgcctgata cgtcaagag accccacga accattcgg 7920
 tagtagcaca gtgcacctcg caaagcggg tgaacgaat gatgcaagc gttcagac 7980
 acaatccag tctgacatac ccgttcgac ctctcatgac ctctcgata ttctcgttga 8040
 cgtctgtgc aagggagag tctagatgt agcaatgtc cttgaggtt cgtctgtct 8100
 tgcctgcac gaactgtga agcagcttg agttctcgg cttgaactaa ggtgcatg 8160
 ntagacag aggnagaca tccagagcc cgaagcgag gatcaatcg gatccggat 8220
 ggaacttgc aagaacttcc agatctaga gaatgcgtc gatggcaac aggtcgggt 8280
 gctcagagag ctgctgcta gtaaggaga gacacatga gagctgggt agtttgtta 8340
 gagatgtgt gtgctagc acttcggat tgtgtatta caaacaagc gctctctga 8400
 cactgatct gtctgata ctacggctc ggaatcgtc gccagatg acccggact 8460
 ttctctaa gactcagta acagcattg tctgatgt agtttagatt tcaatctgt 8520

ggggaagaa aggaaaaaa agactggca cagattggga gagcaactgt ttatatatac 8580
 cctagacaag ccccccgctt gtaagatggt gtcactgta aaccagattt aaggttgga 8640
 agtcaaggag aaqcaaggtg tgggtaccga gcaatggaaa tgtgcggaag gcaaaaaaat 8700
 gagggcaagg cctattgtcg ggggtatata cagggggcga ttgaagtaca ctaaatgac 8760
 atgtgtccac agacccotaa totggcctga tgaagcaaat ccaatcoggo ttogcaagct 8820
 ctaaaagcta taacaagta cacaacccgt ctgacotca gogccotcac tttttgtaa 8880
 gcaaaactgt acacgtctgt ccagcgtttt ctgcctgac ctggctggac atttggtgca 8940
 acctaaagt ctgcgaacct ctgttgtct cagatcagcg cagcagttcc gaggtagttt 9000
 tgaggccctt agatgacga atgtgtcag tgcctggat acagagctta atggcagiat 9060
 tctgtcttat ttgtgcact gagcccggtt atcccgat ctctaccoc ccatccate 9120
 cctttgttgg tgaacccca cccatttttt gttgggtgca gcccaacga cgtggagagc 9180
 ttggcttggc catataaaaa gggccccccc tagtggcaat ggcagaaggt cagctgtgag 9240
 ttgtgaatt tgcactcag cccgctggc cgtctctcc cggcgaattg ggggtgttt 9300
 ttgggacaca aatacggcg caacccgggt tctctgaat tctgcagatg gctgcagga 9360
 attcgtcgt cgcctgagtc gactcaact ttcaacctg agcgtaaaat gtggagaaga 9420
 aatcggcact aaaaagtcag gtgactgga aatgcgcga tgaatgaat atctctgtc 9480
 cagtatatg ccagcatoga ggggtatgt gtacccaaca ctatagtgga agctgaagcg 9540
 ctctgtattg tagtatagtc cttaattgtt gatgggaaga gttactotaa tatctcgtt 9600
 actgcacaaa caccacgcta atggccaga caccatggt gtatagcaac aagctgtgga 9660
 atgtattcg agctaaaatg ccaatgggt gtgaaggag tagttctgt cgaattcgt 9720
 cgtgcctga gtaactttt attaccagt tggccacaaa ccttgagga totgtatgt 9780
 cccctccgac atactccgg cggctgggg tactgtcat agcgtatcg gcatcgaaa 9840
 ggttggttc ctagagcgat acgcactac ctgactaca atcttcggag gttatgctt 9900
 ccacataga cgggcaaaag tgcgtatata tacaagagtg ttgcacgac acagatttcc 9960
 actccacaca ccaactotaa catacaacca ccaactcca caatggaaac cgaactaag 10020
 aagaacaaga ctgactotaa gaagattgtt ctctcggcg cgcactctg tggccccag 10080
 gtgattgccc aggcctgcaa ggtctcaag tctgttgtg aggcctccg caccagatt 10140
 gtgtttgag accgactcat tggagagct gccattgaga aggagggcga gccatcac 10200
 gacgtactc togactctg ccgaaggct gactctatta tgcctgttg tctcggagc 10260
 gctgcacaca cgttatggac cactcccgac ggcgaaccc acgtgcgac ccagcagggt 10320
 ctcccaagc tgcgaagga cctgaacctg tacgcacac tgcacccctg ccagctcgtg 10380
 tgcacaaag tgcgactct ctcccccac cgaacgttg agggcaca cttactat 10440
 gtcagagag tctcggagg tatctactt ggcagcgaa aggaagtga cggatctgg 10500
 gtgccttcg acacagagac ctactcgtt cctgaggtg agcaattgc cgaatggcc 10560
 gcttcttgy ccttcagca caaccccc ctcccggtg ggtctctga cagcgcaac 10620
 gtctggcct cctctgact ttggcaag actgtactc gactctcaa ggcgaattc 10680
 cccagctcg agctcaaca ccaactgac gactggcg ccatgactc catcaagag 10740
 cctccaaaga tgaatggt catctacac accaactgt ttggcgat catctccag 10800
 gagcctcog tcatcccg ttctctgggt ctgtgcct cgcctctct ggtctctg 10860
 ccgacacaa accagcggtt cgtctgtac gacccctg accgactct cccgactct 10920
 ggcagcaga aggtcaacc ctgcccac attctctg ccgcctgat gctcaagtc 10980
 tctcttaaca tgaagcccg cgtgacgct gttgagctg ccgtcaaga gtcctcag 11040
 gctgtatca ctaccgcga tatcgagc tctctcca cctccaggt cggagactg 11100
 ttgcacaca ggtcaagg ctgtcaaga aggaagta gtttctac agcatgat 11160
 ggaaggagca aactgacg cctgcgggtt ggtctacgg cagggctcgc tagtgtata 11220
 gactctata aaagggcct gctctgata tgaatgat atttata ttaccgtgt 11280
 gcaacttga ctgaagaag cagattgggt gtttgttag tggagcgag tggtaagtt 11340
 tggaaacagt ctcttgaaa gtgttgtg tacagtatat tcatctaa cctcaatga 11400
 aaaggtgta gtcgttat taaggaag ggttgtgga tgatgtggt agtatcttt 11460
 aagctggcga ctgacccaa cgaagtgtgt gtagctgt ttaaacagag tgtgaagac 11520
 tcaatagct cgggcttat ctgacaaat agccaaatc tggagtttt gagaqaaaa 11580
 gcaagatac gtttgtaca aaqgcacga tgtacaata ataccgagc catgtatcat 11640
 agaggttag tggctcgt atggcactg tgcctggt gacttatac ggtgactac 11700
 atattgtcc tcagacatac aattacagto aagcaattac ccttgacat ctgtaggtac 11760
 ccccgccga agacatctc agcgtgtgt atctggatt ggcgtagct cctcgtcgt 11820
 caattggct cactctact tctctgctt ggtacacac agcagctct ctatgctcg 11880
 ttctgtcc ttatctatc tccagttat accaactca aatgactga tctgattgg 11940
 tctacacttt catctagag ataccgcta gcccgctgc ataaagcc caactctat 12000
 cagctcttc ctctctgta attagtaaa aggtgattg cgaactctg aagctagtt 12060
 ggcctaaaa aaatcaaaa aagcaaaaa cgaanaaga naaacccag ttttgagac 12120
 agggaggtta cgaaggtcg tatatatata tatatatata tacaacacg gatccagga 12180
 cggccttgg attctctct acacacac attctacac cctaatcca caaccatgg 12240
 caccatccc ctgactacg agcagctct ggaacaccc gactcgttg ccatcagcg 12300
 agtctgtac gatctctcg gtcggcaca ggtccatgoc ggaagcaacc tcatcagac 12360
 tgcggtgca tccagggaa cgtctctgt ctactcatg catctggag tcaagcaga 12420

```

gaactogaag ctctgcage aatttgccc aggcagcac gaacgaagct ccaaggatcc 12480
cgtgtaoacc ttgaactctc cctttgctca ggaactgaag caagtcogttc gagaggtcat 12540
gaagggttga aactggttaag ccaactcctgg cttctggctg agascogcac tcatcatcgc 12600
ttgtactgco attggcgagt ggtactggat cacaacogga gcagtgatgt ggggtatctt 12660
tactggatac ttcaactcgc agattggctt ggcactcaaa caagatgctt ctacoggage 12720
catcagaaaa aagcctcggg tcaacgcctt ttgccttat ggcctgaagc ccattggttc 12780
ctctcgttgg atctggctgc agtcccaaat tatgcacat caacttaaa ccaacacaga 12840
tggcctogac ctggtatctg cctcggcaga gccttcate ttgttccact cctatcctgc 12900
taccacogcc tctcgaaggt ggtaccaccc atttcaggcg tggtaaatgt acatcgttct 12960
gggaatgtat ggtgtctcga tgggtacaaa tcccatgtac ctcttcacaa tgcagcacaa 13020
cgacacaaat cccgaggcca cttctctcag accaggcagc tttttcaatc ggcagcagag 13080
tttcgcogtt tcccttcgac tgccttccat ctccgaaaa gcctttcttc cctggtacat 13140
tgctgggtgco tctcctctgc tcaccattct tctggtgccc acggtcaacg gcattctcct 13200
cacctttgtg ttcgtttctg cccataactt cgaggagacg gaacggaccc cagagaagaa 13260
ctgcaaggcc aaacagacta aggaagcaca ggaggtcaga gacgtggaag aggatcaggt 13320
cgactggtac cgagcacagc ccgagaactgc tgcacacatc ggtggcagcg tgggaatgat 13380
gattacaggo ggtctcaacc tgcagatcga gcactcattg ttcccccga ttctcctttg 13440
gcactatccc ttcatccaag acacccctcg ggaagtgtgc aagcgacatg gcgtccgtta 13500
cacatactat cctacacatc tcgagaacat catgtccact cttcgatada tgcagaaggt 13560
gggtgttgcct cscacacatc aggatgccca ggagttctaa ggggcgcgat gtacatacaa 13620
gattatttat agaaatgaat cgcgactcga caaagatcac gagtgtacga gtaggggatg 13680
atgataaagc tggagaagct tccgcattct tggatttata aagctgtagg acgatacttc 13740
ctgtaaaaat gcaatgtctt taccataggt tctgctgtag atgttataa ctaccattaa 13800
catgtctact tgtacagttg cagaccagtt ggagtataga atggtacact taccaaaaag 13860
tgttgggtgt tghtaactacg atatataaaa ctgttgacgg gaccccgct gatatgccta 13920
aggaacaalc aaagaggagc atattaatlc agaattctag tatacagttt gggat 13975

```

<210> 129

<211> 13066

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pZP3-Pa777U

<400> 129

```

tctcgggtcta ttcttttgat ttataaggga ttttgcgat ttccgctat tggtaaaaa 60

```

```

atgagctgat ttaacaaaa tttaacogga attttacaa aatattacg attacaattt 120
cctgatggcg tattttctcc ttacgatct gtcggttatt tcaacacga tcaagtgcca 180
cttttcgggg aaatgtgcgc ggaaccccta ttgtttatt ttctaaata catccaata 240
tgtatccgct catgacacaa taacccgat aaatgcttca ataatttga aaaggaaga 300
gtatgagtat tcaacatttc cgtgtgcgcc ttattccctt ttgtggcca ttttgcttc 360
ctgtttttgc tcaacacaga aagctggtga agtaaaaga tgcctgaat cagttgggtg 420
cacgagtggt ttacatcaga ctggactcga acagcggtaa gatccttgag agttttcgcc 480
ccgaagaacg ttttcaatg atgagcaact ttaagttct gctatgtgc ggggtattat 540
ccctattga cgcggagcaa gacgaactcg gtcgcgcgat acactattct cagaatgact 600
tggttagta ctacacagtc acagaaaagc attttacga tggcatgaca gtaagagaat 660
tatgcagtcg tgcataacc atgagtgaat acactgcgga caacttactt ctgacaaaga 720
tcggaggacc gaaggagcta accgttttt ttcaacaaat gggggatcat gtaactcgcc 780
ttgatcgttg ggaacgggag ctgcatgaag ccaatacaca cgaacagcgt gacacacga 840
tgcctgtago aatggcaaca acgttgogca aactattac tgggaaccta ctactctag 900
cttcocggca acatttaata gactgcatg agccgataa agttgcagca ccactctgc 960
gtccgcctct tccggttgc tggttattg ctgataaata tggagccgtt gagcgtgggt 1020
ctcgcggtat catgtaacga ctggggccag atggttaagc ctcocgata gtatctatct 1080
acacgacggg gagtacgcca actatggatg aacgaatag aagatcgtct gagataggtg 1140
cctcactgat taagcattcg taactgtcag acaagttta ctcatatata ctttagattg 1200
atttaaaact tcatitttaa tttaaaaga tctaggtgaa gatccttttt gataatctca 1260
tgacaaaaat cctttaactg gegtitttgt tcaactgaga gtcagacccc gtgaaaaaga 1320
tcaaggatc ttcttgagat ccttttttlo tgcogtgaat ctgctgcttg caacaaaaa 1380
aaccaacgct accagcggtg gtttgttgc oggatacaga gatacaaat ctttttcga 1440
agttaacttg cttcagcaga ggcagatcac caaatactgt tcttctagtg tagccctagt 1500
taggcacaca cttcaagaac tctgtacac ogctacata cctcgtctcg ctaatcctgt 1560
taccagtgcc tgcctcagc ggcataaagt cgtgtcttac cgggtlgcac tcaagcagat 1620
agttacogga taaggccgag cggtcgggct gaaocggggg ttcgtgcacn cagccacgat 1680
tggagcgaac gactaaccc gaactgagat acctaacgag tgagctatga gaagcgcca 1740
cgcttcocga agggagaagc gggacaggt atccgttaag cggacgggtc ggaacaggag 1800
agccacagag ggaacttcaa ggggaaaaag cctggttatct ttatagcctc gtcgggttca 1860
gcaactctg acttgaagct cgaattttgt gatgctctgc agggggcgag aguctatgga 1920
aaaaactcag caacgggccc ttttaacgt tcttgccctt ttgctgccc ttgctccca 1980
tyttctctcc tgcgttatcc cctgatctct tggataacag tattaacccc tttagagtga 2040

```

ctgataccgc tgcgcgacgc cgaacgacgc agcgcagcga gtcagtgagc gaggaagcgg 2100
 aagagcgccc aatacgcaaa ccgcctctcc ccgcgcgttg gccgatccat taatgcagct 2160
 ggcgcgcacc caatacaaat tctgaanaag acatcttgat ctctccattg cggcgagtc 2220
 aacggtgggc ttattccccc gaatttcgcg ctcaatctcg ttcagacacg acccgacac 2280
 agtgcttaac gcggttcoga aantctacgc cagatattgt ccaacggact gggctgcata 2340
 gatgtatcc tgcgttggc gaattggata aaagccggcc azaaaaaag cggaaaaaag 2400
 cggaaaaaaa gagaaaaaaa atcgcaaaat ttgaaaaata ggggaaaaag acgaaaaaac 2460
 gcaagagggg gggagtatat gacactgata agcaagctca caacggcttc ttatttttt 2520
 ttctccatct tctgcctagg ttcccaaaat cccagatgct tctctccagt gccaaaagta 2580
 agtaccacc acgtttttcg ccgaaaaatc cactgcgcgc aacgtctgtg ggggtgttaa 2640
 aatgtgggg gggggaacca cgaacagagg ctcttggggg agcgaatga gacacaaaag 2700
 cggcggggtg tgataagggc atttttgccc atttccctt ctctctcttc tccgaagggtg 2760
 atgcctgtgt cgcctcctta ttctctttta ttctcttttg ttktatttct ctgactacgc 2820
 atttggttg atttcccaa ccccacaca ataaagctcg gccgaggaat atatataac 2880
 acggaacacg tgcgcctgtg gacaaacagt cactacctct acgatacaca ccgtactgtg 2940
 tctggaagct tctgagcgga taacatttc acacaggaaa cagctatgac catgattacg 3000
 ccaagctoga aattaaccc cactaaaggy aacaaaagct ggcgtccac ccgcgacaca 3060
 atactgtgtc aaatttcagt ttctgtacat ttaaatctct tcacttcag ttctctctc 3120
 atctgtctct gttttactt gacaggcaaa tgaagacatg tgaacattg atggaaggca 3180
 agaagcccat ttcccccoga gacacagaaq tgcctgaatt cctgcctcgc cccatigata 3240
 acatcgaaa ctacggtatt ccggaagtg tatatagaac ctttcccaag cttgtgtctg 3300
 tggatatgga tgggtgaac cctttgagt actgctctg gctctctcc gagcagatg 3360
 aggcctctca atctagcgca tttaatatct caatgtatt atatattat cttctctgc 3420
 ggcgccttag ttgcttttg tcttgacgc cttggctctc ttgagggtaa acatctgtgc 3480
 atctgtctgc accacgcagt ccttggcgtc cttaagacca attcggatga aggtgggaat 3540
 gatggagaaa gcgcacttc gccacagtc ggaagagcc tgagcgaggg cagagatggc 3600
 ctgcctgagc ttgtagttag gaatgtgg aaacagatgg tggatctgat gttacaaat 3660
 gttgtggac aggttgcga tgaaggctcc gtagcttcgg tccacagagg acaagttgcc 3720
 ctgacatag gtccactccg aatcgcgcta ccagggaggt tctctctgt tctgatggag 3780
 gaaggtagt ccaacccaga tcttgccgaa tccaaagaaa ggtgcgaagt aatacagagc 3840
 catgctcttg aggcctaga cgtaggttaa gtaggcgtac agacacagaa agccacagag 3900
 agagccgagg gaattgatga ccgcagacat tcttcgagc tagagaggtt cccagggatt 3960

 gaagtggttg acctttcgg gaggaatcc agcaacagag taggcanaac aagccgaacc 4020
 aaggagatg accatgtgtc ggaacagggy atgagagtcg gcttctcgtg gaggttagaa 4080
 gatctcatcc ttgtgatgt tgcgggtgtt cttgtgatgg tgcgatgac tgatcttcaa 4140
 cgaactgtag ggaatcagaa tgatgagtg aatgagtgct ccaacagaga agttgacag 4200
 gtgggatgc gagaaagcac catgtccaca gtctgacac atgttaagaa atcccagaa 4260
 acgataccc tggagcagaa tglagccagt gcaagagag ccactagcga gtgcnaaato 4320
 ctgcacgata gcaagggctc gagcatagta cagtccagga gcaagggaa cggcaatgcc 4380
 cagaactcgc accgtatagt agaggaacca ggaacagag gcttcnaagc agtgggcagg 4440
 caggatcgc ttgatctcg tgaagtagg gaactcgtag ggaagcgcaa cgttagagga 4500
 agccctggtt gtgaattagg gtgtgagaaa tggttggtg tagggaagaa tcaaaagccg 4560
 gtctcggaat ccgtgggtat atatataat atatataat acgatccctc gttacccccc 4620
 tgttctcaa actgtgggtt ttogtttctc gtttttgtct tttttgatt tttttaggg 4680
 caactaaagt tccagatttc gtaattccac ttgtactaa ttacagaaa ggaagaagt 4740
 gattagagtt gggcttitta tgaactgtg ctactcctta tctctgatat gaagtgtag 4800
 acccaatcac atcattgtat tttaggttgg taactcggg aggatagata aggcacgcaa 4860
 acagacata gcagacatgc tgggtgtagc caagcagagc aagtagatg ggcgcacatt 4920
 gacagcgag ggaactagc caatccgaca taagacagcg tgagatcgtc ttggccgggg 4980
 ggtactcaaa gatgtccaa ggttaagtgt tgaactgaat tgtatgtctg aggacaaata 5040
 tgtatgcag cgtataaagt catcacagcg accagtccca taatccaaac actaaactct 5100
 tatgatcat cctcccgcta ttattgtacc atcctcgtct ttgtacata cgtatcttgc 5160
 cttttctct cagaacccc agactttggc tatgttgcga gataagcccg gaacatagt 5220
 agtctttcac actctaaatt tctcccttgc tcaaatata gatgtgtct taataactat 5280
 cgtacgataa ctctgtatg caataattat acgaattat ccgctcgaag agtatctgtc 5340
 tgaactgta tgcgccttt tggagtacga ctcaactat gactgtcttt ggaactcttt 5400
 gacatatac tcttcttgg agctgttgg tctgaagct cgtttctcgg ccggttgcg 5460
 ccacacaaat atcagctga acgtcattgc tggcttcat catgatcaca tttttctcg 5520
 caaagcgac gccacagag ccattgaagt tctttctaat ttgacccagt agcctatag 5580
 tccagtctat ctataagtc aactaaactg taactattac caataaatac actccatgc 5640
 cccagataag gttccgataa aaagttctgc agactaaatt takttaagtc tctcttca 5700
 caccaaaatg cctctcaagc aagctcagac taactccac aagtcacact tgcgcctcgc 5760
 agtctcaag ctctgtggcg ccaagcaaac caactgtgt gcttctctgg atgttcaac 5820
 caccagagag ctactgagc ttgcgataa ggtccgacat tatgtgtcga tgaacaaac 5880
 ccatatgac atcattgagc acttcaacta cgcgggaact gtgtctcccc tcaagggaact 5940

tgcctttaag caoggtttct teotgtctg ggacagaag ttgcagata ttggaacac 6000
 tgcacagac cagtaacoggt gtcaacogat cgcagagtg tccgatata ccaacgccac 6060
 cgggtgaccc ggaacaggaa tcattgtctg cctgcagct ggtgcagag aaactgtctc 6120
 tgaacagaag aaggaggagc tetctgacta agagaactcc cagtaacagg agtctcagt 6180
 cccctctccc aacagagaagc tggcagaggc tetgtcatg ctggcagagc tctcttgcaa 6240
 gggctctctg gccactggcg agtaactcaa gaagcaatt gagcttgccc gatccagccc 6300
 cgaattttg ttggtttca ttgcacaga ccagactaag ggcagctctg agcactggct 6360
 tattctgacc cccgggttgg gtcttgacga caaggagac gctctggagc agcagtaacg 6420
 aactgttag gatgtcatgt ctacccgaac ggaatcata attctggccc gagctctga 6480
 cggccagac cagatcata ttgaggagcc caagcgatc cagaaggctg gctgggagcc 6540
 ttaccagaag attactgtt agaggttaga ctatggatat gtaatttaac tctgtatata 6600
 gagagctgc aagtatggag cgtttgttca gttgtatga tggcagagc acctgtctga 6660
 tcagttatgt atgatactgc acaactctg tatcccatg atctgtccaa tggccatgt 6720
 tgttgtttt ctgcatacgg agatgctggg tacagtcta atacgttga ctacttacc 6780
 ttatatgagg ctgagaanaa gctgacttgt gtatgacta ttctcaata catccacagt 6840
 cacatacca ccaactgact accactaac caaaccatg atcaaaccc cactggactt 6900
 cctgagcca gaagaacttg ttatgaaaa gctcaagaga gagatcataa ctctgtatag 6960
 catacttat acgaagttat cctgcaggtt aaggaattca tctgttcat cgtgtttaa 7020
 gctgctgtg gctgtgtgtg ttgtgtgtt ggctctcat gtgctgtat gcaagtaaca 7080
 ccacatatt ggaagcttat tagctttct attttttgt ttgcaagctt taacaacatt 7140
 gctgtgaga gggatggaga tatggagccc gctggagaga gtggagaggc cgttttgag 7200
 cggcttgccc tggcggccag ctgcgaaac gcaactaga ccccttgcca cgcgaaatg 7260
 tgcactttt cagtctagta aegccttacc taacttacc catgctgca tgtttgcgc 7320
 tttttccct tgccttgat cgcacacag tacagtgcac tgtacagtg agctttggg 7380
 gggctcttag atggcagcta aagcggcct agcgtcac tactgggatt gtatggagt 7440
 gcatggagcc tagtgagagc ctgaagagac gcaagccgg ctgacccgtg acagagatg 7500
 ggctgctct gtgttcacc cgttacaat gtttgcca aagctctgc acctgtct 7560
 ggaacctaa ttccaatit tctacttgc caccocatt gatcgagccc taacccctgc 7620
 ccatagcca atcaattaa cctgcattg tctgctgtg ttagtgtgg tctgcctct 7680
 ttgcggccc acttgacaa acacaaacaa gattatata taaggctctg ctctccctc 7740
 caacacacat cactttttg cccgtcttc ctgttaaca caaagctaa gaacacaaac 7800
 aacacccca acccccttac acacagaca tatcataga aatggccatg gcttttcca 7860

ctgttctgc gcgtacagc ttccgaagc tgacggagat caagcctgc ctgcagagc 7920
 actgcttga ggcctagtc cagtgttgc tctactaac cgtgcgcgc ctggcctag 7980
 cggctcgtc cgcctgcgc ctactacgc cgcgcgcct cgcgactgc caggagttg 8040
 cctgtgga tgcgtgctc tgacggggt acattctgt gacgggctc gtattctgg 8100
 gttcttcc catcgccat gactggccc agcgcggtt ctgcgctgc cactgttca 8160
 actcagct cggcagctc attactoga tcatctcac gcgtacagc tcatgaga 8220
 tctgcagcc caccacac aagacacgc gaaactoga caaggagagc atttttacc 8280
 cgcagcaga ggcgactgc caccactgt ccgaacact ggtgatctgc ctgcctgc 8340
 cctgttgc gtactcgtt gggccttc ctctgcga ggtgaacac ttcaacct 8400
 ggaacccgt gtactgcgc cgcgtgtct cgtctcat ctactgcgc tgcctgtg 8460
 cgttcggcg ctgtatgc tatctact agctctatg ccttaagacc atggcctgt 8520
 actacttgc cctctctt gggctgcga agatgctgt ggtactaac tttttgccc 8580
 acatgacga ggaacgcga tggtagccg actcgagtg gactagctc aaggccacc 8640
 tctgtcct ggaacgcgc taagcgcgc tcatgacaa cctgagccc aactagcca 8700
 cgcacagat caccacctg ttctgcata tccgcacta caagctaac gaagcagcg 8760
 cagccttgc gcaggcgtc ccggagctgc tgcgaagag cgcctgcgc atactccga 8820
 cgttcatcg catcggttc atgacgcga agtaagcgt cgtgacaaag gaagcaga 8880
 tgtttacgt caaggagcc aagcgcga agacaaagc caactagcg gcgcattga 8940
 tgaattgaaa cacacacat gtttatatc aggtgagagc taattgaca gtttatatt 9000
 aactagata tgcacacgt aacttctc atctaacga gaacacagc actgcaagta 9060
 atatagaat tgacacact gcaattctt tgaactctt tactatatc ctttattt 9120
 ttatatcaa atcaattct ctccagca tgaagctgc aaactcag agnatcaa 9180
 tctgtgact cgtcaatga ggcctttt gatccttgc tgttgccac ctgtcctt 9240
 ctgtttaa acgttagca gttactact agagaaact ttaattgcg ccggagaga 9300
 cgcgaagcc gctagatga caatttacc actcacagc taacttctc caattgccc 9360
 taggggggg cttttata tggcgaagc aactctcca cgtcgttgc gctgaccca 9420
 caataaat ggtagggtt ccccaacaa ggaatggat gggggtaga agtaagcg 9480
 ntaaggggc taatggcac aataaagac gaatactgc attaagact gtgactcag 9540
 gactgacac attgactat ctgaagcct caaaactacc tgggaactgc tgcctgact 9600
 tggacacac agaggttcc agcactttg gttgaccaa atgtccacc agtgagagc 9660
 agaaacgt ggaacagct gtacgttt tcttaacaa aagtgaagcc gctgagtcg 9720
 agcaggttg tgtactgt tatagcttt agactcga aagcagatg ggtttggt 9780
 catcagcca gattagggc ctgtgacac atgctatgt agtgtactt aatgcaccc 9840

```

tggatatago ccagcaaatga ggcgtggcc tcattttttt gocttcogoa catttccatt 9900
gctcggtacc ccaacttgc ttctactga cttgocaaac ttaactctgg ttactattga 9960
ccaacatctt acaagcgggg gcttctcta ggttatatat aaacagtggc tctccaatc 10020
ggttgccagt ctctttttt ctttcttcc ccacagatto gaattctaaa ctacaactca 10080
cagaattccg agcgtgagt atccacgaca agatcagtg cagagcagcg cgttttctgt 10140
aatgacacaa tccgaagtc cttagcaaca ccaactctct acacaaacta acccagctct 10200
ggtaccctgg ctctctccac tcttctgcgc ccgtacagat tcccgacgct gacggagatc 10260
aagcgctcgc tgcacgcga ctgctttgag gctcgggtcc cgtggtcgct ctactacac 10320
gtcgcgcgcc tgggcacgc ccgtcgcctc gcgtcgggac tctactaagg gcgcgcgctc 10380
gcgactctgc agagctttgc cctgctggat gcggtgctct gcacggggta cattctgctg 10440
cagggcctcg tttctgggg cttcttccac atcggsctatg actgcggcca ccgcgcgttc 10500
tcgcttctgc acctgctcaa cttcagcgtc ggcacgctca ttcactcgat cactctcag 10560
ccgtacgagt catgcaagat ctgcacgcgc caccacccca agacacgcgc caacatcgac 10620
aaggacgaga tttcttacc ccagcgcgag gcgactctgc acccactctc ccgacacatg 10680
gtgatctcgc tcggctcggc ctggtctcgc tactcgttg cgggtctccc tctctcag 10740
gtgaacctct tcaaccttg ggaaccttg tactcgccc gcattgtcgc cgtctatc 10800
tcactcgcct cgtcctggc gttcgcggc ttgtatgct atctcaacta cgtctatggc 10860
cttaagacca tggcctgta ctactcgcgc cctctcttg cgttcgcacc gatgctcgtg 10920
gtcaactcct ttttgacaca caatgacgag gaacgcact gttacgcgga ctggagctgg 10980
acgtacgtca agggcaacct ctgctcgtg gacgctcgt accgcgcgct catcgacac 11040
ctgagccaca acctgggac gcaccagatc caccactctt ttcgactat ccgcactac 11100
aagctgaacg aggcgaaggc agcgtctcgc caggcgttcc cggactcgt gcgcagagc 11160
gcgtgcgga tcaatccgac gttctccgc atcgggctca tgtacgcaca gtaacggctc 11220
gtggacnagg acgcacagat gtttaccctc anagagccca aggcgcgcaa gaccaaggcc 11280
aactaggcgc ccgcctggag cgtgtgtct gactcgctgt tttctatgga gttgtgagt 11340
ttagtacaca tgatgggttt atatatgct aatgataga tgtgatcttg attgacaga 11400
tggatctgag aactttgaa acgtacatgg gaatgatga atgtgggggt tttgtactg 11460
gataactgac ggtcagtga ccgcgttgt caaatatcca agagatcgga gaactttgg 11520
gtcaagtcaa catgtcctct ctgttcaagt aaacacatca ctatgggtag tatatttagt 11580
aaggacaaga gttgagatc ttggagtcgc tagaaactga tttcgcgtt ccaagatcaa 11640
attagttagg taatacgggc accggaatcc attcatagc tcaactctgc agctgagta 11700
attaagatga ccgaatttgc gactcgagc agaatagat ggagcgtgtg ttctgagtcg 11760

aigtcttcta tggagtgtg agtgttagta gactgatgg gtttatatat gatgaatgaa 11820
tagatgtgat tttgatttgc acgatggaat tgagaacttt tgaacgac atgggaatgt 11880
atgaatgtgg ggtttttgt actggataac tgaacgtcag tggacgcgt tgttcaata 11940
tccagagatg gcgagaaact ttgggtcaag tgaacatgic ctctctgttc aagtaaaaca 12000
tcaactatgg gtagtatatt tagtaaggac aagatgtgag attctttgga gtccagaaaa 12060
cgtattttcg cgttccaga tcaaatagat agagtatac gggcacggga atccattcat 12120
agtctcaatt ttcccatagg tgtgctacaa ggtgttgaga tgtgttacag taacaccatg 12180
attcgaggta aagagccagc aagtctatga tgaggtcaag aaataacag atctacagct 12240
caatacaatg aatatctct tcatattct tcagtgaca ccaaggtgtg ctattttccc 12300
cagaatgagc tgaaggcg cgtgtttagc gtggagatg ggttcggttg gcgtatcctt 12360
catatctcga cgaatatga tggcagagga tgacaaaaag tatctatag tagacagcgt 12420
agaatctgga tttgatttgt ataatctat ttattcgtg tctcaaaat actctcgata 12480
agttgggggt aaactggaga tggacaatg tcatatctc gacgcatgc acgtcgggcc 12540
caattcgccc tatagttagt cgtattacaa ttcactggcc gtcgtttac aactcgtga 12600
ctgggaaaaa cctggcgtta ccaacttaa tgccttgca gcacatccc ctttcgcag 12660
ctggcgtaat agcgaagag ccgcacacga tgcgccttcc caacagtgc gcagcctgaa 12720
tgccgaatgg acgcgcctg tagcggcgca ttaagcgcg cgggtgtgtt ggttacgcgc 12780
agcgtgacgc ctacacttgc ccgcgcctca gcgcgcgctc ctttcgctt cttcccttcc 12840
ttctcgcaca cgttcgcgc ctttcgcgt caagctctaa atcgggggct ccotttaggg 12900
ttccgattta gtcgtttaag gcactctgac cccaaaaaac ttgattaggg tgatggttca 12960
cgtagtgggc catgcgcctg atagacggtt tttgcgcctt tgacgttga gtccacgttc 13020
tttaatatgt gactcttgtt ccaacttga acacacctca acctca 13066

<210> 130
<211> 4313
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZKUM

<400> 130
taatcgagct tggcgtaac atggtcatag ctgttctcg tgtgaattg ttatccgctc 60
acaattccac acaacatacg agccggaagc ataaagtga aagcctgggg tgctaataga 120
gtgagctaac tcacattaat tgcgttgcgc tcaactgcgc ctttcaagt gggaaacctg 180
tcgtccagc tgacttaatg aatcgcccaa ccgcggggga gaggcggttt gcgtattggg 240
cgctcttccg cttctcgtct caactgaatg ctgcgcctcg tcgttcgct gcggcgagcg 300
gtatcagctc actcaaaagg cgttaacgg ttatccacag aatcagggga taacgcaagg 360

```

```

aagaacatgt gagcaaaaagg ccagcaaaaag gccaggaacc gtaaaaaggg ccgcttgcgt 420
gcgtttttcc ataggctccg cccccctgac gagcatcaca aaatccgacg ctcaagtcag 480
aggtaggcgaa accccacagag actataaaga taccaggcgt ttcccctcgg aaqctccctc 540
gtgcgctctc ctgttcagac cctgcgcgtt acccgatacc tgtccgocct tctccctctg 600
ggaagcgtgg ccgtttctca tagctcagc tgtaggtatc tcaagtcgtt gtatgtcgtt 660
cgctccaaag tgggcgtgtg gcacgaaccc cccgttcagc ccgacccgctg ccctttatcc 720
ggttaactatc gctctgctc caaccgcgtc agcacagcat tatccgacat gccagcagcc 780
actgttaaca ggtatagcag ccgagagtat gtacggcggt ctacagagtt cttgaagtgg 840
tggcctaact accgctacac tagaaggaca gtatttgata tctgcgctct gctgaagcca 900
gttacctctg gaaaaagagt tggtagctct tgatccggca aacaaaccac ccctggtagc 960
ggtggttttt ttgtttgcaa gcagcagatt accgcagaa aaaaaggttc tcaagaagat 1020
ccttgatctc ttctacggg gtctgacgct ccgtggacg aaaaactcag ttaagggatt 1080
ttggtcatga tatlatacaa aagcatcttc acctagatcc ttttaatta aaatgaagat 1140
tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc 1200
agtgaaggcc ctatctcagc gatctgtcta ttctgtctc ccatagttgc ctgactcccc 1260
gtcgtgtaga taactacgat accggagggc ttacatctg gccccagtcg tgcgaatgata 1320
ccgcagagcc ccgcctcacc ggtctcagat ttatcagcaa taaccacagc agccggaagg 1380
gccgcagcca gaagtgttcc tgcacttta tccgcctcca tccagttatc taattgttgc 1440
cgggaagcta gactaagtag ttgcagatt aatagtttc gcaacgttgt tgcactgtc 1500
acagcctcgc tgggttcacg ctgtctgctt ggtaggctt caltcagctc ccgttcccaa 1560
cgataaggg gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccctcgtt 1620
ctccgcagct ttgtcagaga taagtggccc gcaggttatc cactcaggtt tatggcagca 1680
ctgcataatt ctctactgtc catgcatccc gtaagatgct ttctgtgac tggtagtacc 1740
tcaaccagat cactctgaga atagtgtatg ccgcagccga gttgctcttg cccgggtcca 1800
atacgggata ataccggccc acatagcaga actttaaagc tgcctcatat tggaaaacgt 1860
tcttcggggc gaaaactctc aaggatctta ccgcctgtga gatccagttc gatgtaaccc 1920
actcgtgcac ccaactgac ttacgacatc ttactttca ccagcgttcc tgggtgagca 1980
aaaacaggaa gccaaatgc ccgaaaaag ggaataaggc ccacacggaa atgttgata 2040
ctcatctctc tctttttcca atattatga agcatttata agggttatgt tctcatgaga 2100
ggatacatat ttgaatgat ttgaaaaat aaacaaatag gggttcocgc cacatttccc 2160
cgaaaagtcg caactgacgc ccctgtgagc ggcgcattaa ccgcgggggg ttggttggtt 2220
acgcgagcgc tgaacgttac acttgccagc gccctagcgc ccgtctcttt ccgtttcttc 2280

```

```

ccttccttcc tgcgcaggtt ccgcgcttt cccgtccag ctctaantgc ggggtccct 2340
ttaggggtcc gatthagtgc ttacggcacc ctgacccca aaaaacttga ttagggtgat 2400
ggttcacgta gtgggcctac gccctgagag accgttttc gcccttgac gttgagtc 2460
acgtcttcta atagtggact ctgttcccaa actggaacaa cactcaaccc tatctcggtc 2520
tactcttttg attataagg gattttgccg atttccgctt attggttaaa aaatgagctg 2580
atttaacaaa aatttaacgc gaattttaa aaatatata ccgttataat ttcaattgca 2640
cattcaggct gcccaactgt tgggaagggc gatcgttgcg gccctcttgc ctattaagcc 2700
agctgcgcaa aggggagatg gctgaaggc gattaagttc ggttaacgca ggttttccc 2760
agtcacagcg ttgtaaaag ccggccagtg aattgtaata ccactcaata tagggcaat 2820
tgggtacccg gcccccctc gaggtcagc agtatctgtc tgactcgtca ttgcgcctt 2880
tggagtacga ctccaactat gagtctgctt ggtacacttt gacatataat tctctgtg 2940
aggctgtggg tctgacagct gcgttttggc ccggttggc ccacacaaat atcagctgca 3000
acgtcattgc tggcttcat catgatcaca tttttgtcgg caaaggcgac gccacagag 3060
ccattagcgt tcttttaaat ttggacgat agccgtatag tccagttcat ctataagttc 3120
aaactaactg taactatcac caaacatat acttcaactg ccagataag gttccgataa 3180
aaagttctgc agactaaatt tatttcagtc tctcttccc caacaaaatg cctctctacg 3240
aagctagatg gctcaagctc gtggcagcca agaaaaccaa cctgtgtgct tctctggatg 3300
ttaaccacac caaggagctc attgagcttg ccgataaggt ccgacattat gttgtcatga 3360
tcaaaaacca tatcgaatc attgaacgat tcaactaagc ccgaactgtg ctcccctcca 3420
aggaacttgc tcttaagcac ggtttcttcc tgttcagaga caaaaagttc gcaatattg 3480
gcaacactgt caagccacag taacggttgc accgaatgcg ccagtgttcc gatataacca 3540
accgccaggg tgtacccgga accggaatgc attgtggcc tgcagatggg tgcgtacag 3600
gaaactgtct ctgaacagaa gaaggaggac gtcctgact accgaactcc ccagtacag 3660
gagttcttag tcccctctcc caacgagaag ctggcagag gtctgctcat gctggccag 3720
ctgtcttcca agggctctct gccactgac gactactcca accagaccat ttgacttgc 3780
cgatccgacc ccaggtttgt ggttgccttc attgcacaga accgaactaa ggcgactct 3840
gaggaactgc ttattctgac ccccggggtg ggtcttgacg acaagggaag ccgtctcgga 3900
cagcaatacc gaactgttga ggtatgcat tctaacggaa ccgatatcat aattgtcggc 3960
ccaggtctgt accgcagaa ccgagatctt attgaggagg caaagcgata caagaaggt 4020
ggctggggag cttaccagaa gattaactgt tagaggttag actatggata tgaatttaa 4080
ctgtgtatat agagagcgtg caagtatga cccttcttc acgttgtatg atgtcagac 4140
gacctgtctg atcagatag tatgatctg caaacctgt gtatccgat gatctgtcca 4200
atggggcatg ttgttgtgt tctgctacg gagatgtgg gtacagtctc aatcgttga 4260

```

```

actactata cttatatgag gctcgaagaa agctgacttg tgtatgactt aat 4313

```

```

<210> 131
<211> 15991
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZKL2-5mB89C

<400> 131

```


gtacgttato atttgaacag tgaagggota cagtaacaga agcagttgta aacttcattc	60
cgttgattct gtaactaacgt accccaatcc gcgccttcog ctgacactgt toaacccaaa	120
aactaacatct gcgtgcgctg tgtgaaggota teatcagata catactgtag attctgtaga	180
tgggaacotg ottgtatcat atacatcccc ctcccotga octgcaacag caagcaaatgt	240
gacattgata ttgtctgcta totagtgcog aggatgtgaa agccagagact caaacatttc	300
ttttactctc ttgtctctga ccagacotgg oggagattac gccagtatga ttottgcagg	360
totgagacaa gcctgcaaca gccaaattt attttogaa gcgacaacaa tgcacaaccc	420
cgccacgttc agagatgcac atgatttgtt tticagataa cagtaacccc cccccccccc	480
ccaatgaaac cagttattact cacaccatcc tcattcaaac cggtacactg attacgcgcc	540
catcaacgac agcatgaggg gactgctgat ctgatctaact caaatgacta caaaaatgc	600
aataatgaag agcaaacgac aaaaaagaaa caggttaacc aatccogctt caatgtctca	660
ccacaatcca gcactgtttc tcattacctc ctcccctaaa tticagagtt gcatcaggg	720
cottgatggc gcgcacagctg caattaatgaa tcggcaacog cgcggggaga ggcggittgc	780
gtattggggc ctcttccogct tctctgctca ctgactogct gcgcctggtc gttcggctgc	840
ggcgagcggc atcagctcac tcaaaagcgg taatacgggt atccacagaa tcaggggata	900
aogcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc cagggaacgt aaaaaggcog	960
cgttgctggc gtttttccat aggtctccgc cccctgacga gcatcaaaa aatcgacgct	1020
caagtcaagc gtcgcgaaac ccgacaggaac tataaagata ccaggcgctt cccctggaa	1080
gctccotcgt gcgctctctt gttccgaacc tgcgcgttac cggatacctg tcgcgtcttc	1140
tcctctgggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg	1200
aggtcgttgc ctccaagctg ggcctgtgtc acgaaccccc cgttcagccc gacgcgtgcg	1260
cottatccgg taaatctcgt ottgagtcga acccggttaag acacgactta tcgccactgg	1320
cagcagccac tggtaacagg attagcagag cagagtatgt aggcggtgct acagagttct	1380
tgaagtgggt gccataactac ggcataacta gaagaacagt atttggcttc tgcgctctgc	1440
tgaagccagt taccttcgga aaaaagcttg gtagctcttg atccggcaaa caaacccacg	1500
ctgtagcggc tggttttttt gtttgaagc agcagattac gcgcagaaaa aaaggatctc	1560
aagaagatcc ttgatcttt tcacggggc ctgacgctca gtggaacgaa aactcacgtt	1620
aaggatltt ggtcatgaga ttatcaaaaa ggcattctca ctgatcttt taaattaaa	1680
aatgaagttt taatcaatc taagcttat atgaatgaa ttggtctgac agttaccat	1740
gcttaactag tgaggcacct atctcagca tctgtctatt tgcgtctccc atagttgct	1800
gactcccgct cgtgtagata actcagatac gggagggctt accatctgc ccagtgctg	1860
caatgatacc gcggaacca cgtccacccg ctccagattt atcagcaata aaccagccag	1920
ccggaagggc cgaagcgaga agtgcctcgt caactttatc cgcctccatc cagtctatta	1980
attgtgcog ggaagctaga gtaagtgtt cgcacgtaa tagtttgcgc aagttgtgt	2040
ccattgctac aggcctcgtg gtcctcogct cgtcgtttgc tatggcttca ttcagctccg	2100
gtcccacag atcaagggca gttacatgat ccccctggt gtcgaaaaa cgggttagct	2160
cottcggctc tcgcctcgtt gtcgaagta agttggcgc agtgttatca ctcatggtta	2220
tgcgcagact gcataattct ctactgtca tgcctccgt aagatgcttt tctgtgactg	2280
gtgagtactc aaccgaatca ttctgagaat agtgtatgcg ggcacagagt tgcctttgac	2340
cgcgctcaat accggataat accgcgcacc atagcagaa tttaaaagt ctcatcattg	2400
gaanaacgtc ttccggggca aaactctcaa ggcctttacc gctgttgaga tccagttcga	2460
tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg	2520
ggtgagcaaa aacagggaag caaatgcgc caaaaaaggg aataaggcg acccggaat	2580
gttgaatct catactcttc ctttttcaat attattgaag cattttcaag ggttatgtc	2640
tcattgagcg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcga	2700
catttcccg aaaaagtcca ctgctgcgg tgtgaatat cgcacagatg cgttaaggga	2760
aaataccgca tcaggaaatt gtaagctta atatttgtt aaatttcog taaattttt	2820
gttaaatcag ctcatttttt aaccaatagg ccgaatcgg caaatccct tataaatcaa	2880
aagaatgac cgaataggg ttgagtgttg ttccagtttg gaacagagct ccatattaa	2940
agaacgtgga ctccaaagtc aaaggcgcaa aaacgctata tcaggcgcat ggccacatac	3000
gtgaaccatc accctaatac agtttttttg ggtcagagtg ccgttaagca cttaaatgga	3060
accataagc gaccccccga tttagagctt gacggggaaa gccggcgaaac gtcggcagaa	3120
aggaaggaaa gaagcgaaaa ggaagcggcg ctaggcgctt ggaagtga cgggtcaagc	3180
tgcgcgttao caccacaccc gcgcgctta atgcgcogct acaggcgcg toaatcgcc	3240
attcagctg cgaacatgtt gggagggcg atcgtgcgc gcctcttcga tattacgca	3300
cttgcgcaaa ggcqatgtg ctgcaagcc attaaattgc gtaacccacg qdttttccca	3360
gtcacgcgt tgtaaaaa cggccagtg attgtatca gactcactat agggcgat	3420
ggccocgagc tcgcctgctg gtttcgattt gtttagagc aacgaataa cagtatacat	3480
agagaataaa cgtattctat ttattaaat agatagttga ggtagaagt gtaaaagtg	3540

ataatagct tagataaac agaacoccto ggtgaagaag tactgaagat ggtttocaaat 3600
 cacattgacc tgcotggagca gagtgttacc ggcagagcac tgtttattgc tctggccctg 3660
 gcaatagaca acgttggaga gaggaggggtg gatcaggggc cagtcacataa agaacocacac 3720
 agagcagtcg tggtaacocgt oocagaaggg cacttgaggg aogtatotot ctcogtgggt 3780
 gattoggttag agotttogggt ctttggacac cttggagaca togggggtct cctggccaaa 3840
 gaagagttta tgcacocaggt tagcaaaagoc agcgttacgc acactgggtt gacaaagagt 3900
 aacacagagg ggaatcgtgc cgttaacott gagggtgatt ccgaacagaa ggctgcagac 3960
 tcttcogaga gsgtgaocgg tgaagacaat ctggtagtgc ggtactgctt caatcacaga 4020
 gtgcagcttg gggccgatct gattgaggtt gttgtgttag gactggatga agccattgtg 4080
 gacaagacag tcatcacag tagcagtaga agagatgta ccagcaagat caaagttaat 4140
 taactaacct gaggattga gactatgaat ggtatocgtt gcccgtatta ctatactaat 4200
 tgcattctgg aacgcgaaaa taagtctcta ggcacccaaa gaatctcaac tctgtctctt 4260
 actaaatata ctaccatag ttgatggctt actgaacag aagagacatg ttcaattgac 4320
 ccaagatttc tgcactctct tggatatttg aacacagggc tccactgacc gtcattatc 4380
 cagtcacaaa acccccacat tcatcattc ccatgtactt ttacaaagtt ctcaattcca 4440
 togtgcacat caaaatocaa tctaticatt cctcatatat aaacccatca tgtctataa 4500
 cactcacac tccatagaaa acatgcactc agaacacag ctccatggcg ccgttagga 4560
 atcctgagcg tcttgcacac agtgaacac accgactttg tgcattgact tgaaggtgga 4620
 aatgatgtg cccacaatgg taggttagaa gacgtacga accctcgtgc gtgcgcacaa 4680
 ctctcgaca gcttgcga cgaaggata gtgcacagac gacattcgag gaagaggtg 4740
 atctogato tggaggtga caccgccagt aaagaacatg gcaatgggtc caccgtagt 4800
 ggaagaggtc tccacctgag ctctgacca gtctatcta togggtcaa cgtctctc 4860
 ggaactcttg accttgcagt tcttgcggg gattogtcc gagccatoga agttgtgaga 4920
 caagatgaaa aagaaggtga ggaagcacc ggtagcagtg ggcacacag gaatggtgat 4980
 ggcacgggag gttccagtga cataccagg cagaagggcg gttcgaaga tgaagaagc 5040
 tgcataacg aatgaaggg ttcgtacgc tgcagaaag ccgttctctc gcatggtgt 5100
 gacagactcg ggaatgggtt cgttgcgtg cattcggaag atgtagagag ggtgtcacac 5160
 cagcgaacg ccgtaggctc caagcacgag gtacatgtac caggccctga atcgtgaaa 5220
 ccaatttoga gcaatgttgg cagcagggta gtttggaaac caaaggaatg gttctgcga 5280
 ctgcgcaccc aggtgcagac catgctgatt ggtgtaggtg tcatgtcga tcatgtgaga 5340
 ctgcagcag atccatctgg acatccaat gacgtcagtg ccgtagcga aagagagctt 5400
 gaccacaggg tttttgctga tggcaccatg aagagcatcg tcatgaatgg acagcccat 5460

 ctgcattgct atgaatccag tcaagagacc cccagacacc attccggttag tagcccatg 5520
 ccactcgcaa aaggcggtag cagcaatgat gccacgggtt cgcagccaga atccaggtgt 5580
 ggcataccag ttccacatt tcatgcctc tgcatagtt cgttgcagc cctgtcaca 5640
 gggagagtcg taggttaga caatgtcctt ggaagttcgg tctgtcttc ctcgcagaa 5700
 ctgttaagc agcttcagat tctcggcctt gacgttaagg tgcattgagt agaacagag 5760
 agaagctcg gaggccacag aagcagagat caagtcgat ccgggatga ccttgcaag 5820
 acctccaga tcttagagaa tgcctcgat ggcacacagg togggtcgt cgaacagctg 5880
 ctctgtatga agggagagag ccaatggcag gactctgtt agtaattgt ccggagctca 5940
 tcaattggt cgaaggttg tgcactgta gctagctc aattggctc tgttggctg 6000
 atgacattg tcatctgtt ctgttgggtc atgttccat cactctctat ggtactaca 6060
 attcgtcga tgcgcagaa tccgttaata ccgaattga tggccatgt gatgtgtt 6120
 taattcaga atgaatatag agaaagag aagaaaaag attcaattga gcggcgatg 6180
 cagaccctta tataatgtt gcttggaca gacggagcaa gccgcaccaa acatcgttc 6240
 ggtataatg gtttaagctt ttaacacaaa ggttggctt ggggttaact gatgtgtg 6300
 aagagacgg gcgttggga gccattgcgc ggcgaatgg ggcgtgaat cgtctcaat 6360
 toagggggt gctcaattc gtgcacccgt ggtttttcc cgcgtttcc gcccgcttg 6420
 caccatgca gcctctctt tggtcggac accttgctg gagctaggt ccttgtgta 6480
 cttaaaaagt ggcctcccaa caccacatg acatgagtc gtcggccag aacgttggc 6540
 gggctcagc tgcctcaat gccccgaaaa aaacgtctc ggaatgctt ccgaacagct 6600
 ccgcgcgcg ctatgatat ccgcgaagtt ccatagccc attgcctcc ctcggctct 6660
 atccgcacac ctctaatag agcgggaata taaccaaagc ttctttttt tctttaaca 6720
 cgcacaccc caactatcat gttgtctg ctgttgaat ctactctg gagggtgt 6780
 cccacacac ccaactaca ggtggtccg gcctgtgat tggatgata gtctctatc 6840
 cggactaat ctgacaaatg ggcattgoc gacagacaa aatgcgcaa ttacgtaac 6900
 ccaacgaat gctacccct ctttgagcc cagcggccc aaatccccc aagcagccc 6960
 gttctacgg ctctctctc caagcacaag cagcccgaaa ttctctctt tgaattttc 7020
 catacaagt tctctgctt ccaagaaagc cctgggtgt tcatctggg tttgtgtg 7080
 gtggtagtgc agtgggtgta ttgtcactg ggtatgatt gagaataag catcacaaag 7140
 tcaagttct toagcctca tataagtata agtagtcaa cgtattaga ctgtaccaa 7200
 catctccta tgcgaacaa caacacatg cccatttga cagatctgc ggtacacag 7260
 gttgtcagt atatacata ctgatacga caggtctct gacacata caagctgaac 7320
 aagcctca tacttgacg ctctctatat acacagtaa attacata catagtcaa 7380
 cctcaacag ttaattctt ggttaagctc ccaagacgc ttctgtatc gcttgcctc 7440

ctcaatagga tctcgggttct ggcgtacag acctcgccg acaattatga tatcagttcc 7580
 ggtagacatg acatctccaa cagttcggta ctgcctccg agagcgtctc ccttgtctc 7580
 aagaccacc ccgggggtca gaataagcca gtctccagag tccctctag gtcgggtctg 7620
 ggcattgaag ccaaccacaa actcgggggt ggcctgggca agctcaatgg tctgtctgga 7680
 gtactcgcca gtggccagag agcctctgca agacagctcg gcacacatga gcagacctct 7740
 ggcagcttc tctgtgggag aggggaactag gaactccttg tactgggagt tctcgtagtc 7800
 agagagctcc tctctctctc gttccagagc agttctctcg gcacagctgc gcaggccagc 7860
 aatgattccg gttccgggta caccgtgggc gttgttgata tcggacacct cggcgattcg 7920
 gtgacacggg tactcgtctc tgacagtgtt gccaatctc gcgaacttcc tctctcga 7980
 caggacgaaa ccgtgcttaa ggcgaagttc cttgaggggg agcacagctgc cggcgtaggt 8040
 gaagtctgca atgattctga tatgggtttt gatcatgac acaataagtc cgaacttctc 8100
 ggcagctcca atgagctcct tgggtggtgt aacatccaga gaagcacaca ggttggtttt 8160
 ctggcctgcc acgagcttga gcactccagc ggcacagccg cacttgtgga cgttagctcg 8220
 agctctgtag gaggcctatt tgggtctgaa gaggagactg aaataaattt agtctgcaga 8280
 actttttatc ggaactctat cttgggctgt gaagtatatg ttatggtaat agttaagagt 8340
 tagttgaact tatagataga ctggactata cggctatcgg tccaaattag aaagaacgtc 8400
 aatgctcttc tgggctctgc ctttgcgcac aaatagtga tcatgatgaa agcagcaaat 8460
 gacgttgcaac ctgattattg tctcggccaa ccgcgcgcaa aacgcagctg tccagccacc 8520
 agctcccaac gaagaatga tctcaaatg gatccagca cactcagatg tggagtctga 8580
 ctccaaaggc ggcattgacg agtcagacag atactcgtcg acctttctc tgggaaccac 8640
 caccgtcagc cttctgact cactgattgt agccacgac acaggcaaca gtcctggat 8700
 agcacatatg gttctgtcgg tccatttctc acaacttta ggcgtcaagt gaattctga 8760
 gaagaagtat gtccttctat tgagacatcg tgttgtgat tccaataag tcttgatc 8820
 agttggcca gtcattgtgt cgggggtaat tggattgagt tatcgcctac agtctgaca 8880
 ggtatactcg ctgcacactt tatacttttt gattcctgtg cacttgagc aatctgttt 8940
 accaaagtg agaatgctcc acagaacaca cccacgggta tggttgagca aaaaaaac 9000
 actccgatac ggggaataga acccctgtct ccaaggttct caagaagtat tcttgatgag 9060
 agcgtataga tcaggaaga ccacacagcg ctgcttctta agtttgtgac atcagtatcc 9120
 aaggacacct tcaaggaatt caagctttt aaccctgat ttgcattcg taaagctgt 9180
 agacaggctg atcgttccc taaggctccc acctgtgca atcttctcaa gctgcctgac 9240
 tatcagacca ttgatcaact tcggaagaaa cttttgtatg caatctgac acatgctgt 9300
 ttgatttgtt cttagaggaa cgcatacaca gtaactatg agaataaaag atattcatt 9360

atnaagtag atagttgag tagaattgt aagagtgtat aaatagcgcc cgtcactga 9420
 atcttttttg ctcccttgtg ctttcggag atgtagtgtt gacgtagaa gttgaggaa 9480
 agcacagga cagtaacca gtagaagtag ttgaacac ccacacacat tctcattcca 9540
 tctgtcgtg agcagggaa gttcgtgtac ttcacagaca tgaagaagcc aagttgaac 9600
 tgaatgactr gcatagaagt aatcagggac ttggcctag ggaacttag ctgtactagt 9660
 cgggtccaat agtagccta catgatccag tgaatgaag cgttagacag caaaagato 9720
 caaacgctt cgtttcggtg gttgtgagac agcacatgt ccataggagc tccagatag 9780
 tgaagaacct gcaacaggt cagagcgttg ccatgaggg gcagatagaa ggaactaat 9840
 tactcagga acctgtctag gtacacagc tgaattgtga ttoggagac atgttgtctg 9900
 aaagccttct cgcagtgtc ggcactgaca ccaatggtgt acatggccta ggcctagag 9960
 aggaagagc ccagcagta gatggacat agcaggttgt agttgtgaa caaaccttc 10020
 attcagactr gaccttggg tccagagga ccaagggtga acttcagat gacgaagcg 10080
 atggagaggt acagcactc gcagtcgag gcatcagacc agagctgagc atagtcgac 10140
 ttgggaagaa ctctctggcc aalgagagc atttcttca cgaactccat ggttgtaat 10200
 taggttgtg apaalgggtt gttgtaggga agaatcaag gcggtctcg ggtcgttg 10260
 gtatatatat atatatatat atatacagc cttcgttacc tccctgtctc caaacctgt 10320
 gtttttgtt ttctgttttt tgcctttttt gatttttta gggcaacta agtttccga 10380
 ttctgctaat cacttttga ctaattacaa gaaggaaga agctgattag agttgggtt 10440
 ttatgcaac tctgctactc cttatctctg atatgaagt gtagaccaa tccatcactg 10500
 tcaattagag ttgttaatac tggagagata gataagcac caaacagagc catagcagac 10560
 atgctgggtg tagcaagca gaagaagta gatggagcc aattgacgag cggggagat 10620
 acgcacatcc gacataagac agctgagat cgtcttgcc ggggggtacc taagatgtc 10680
 caagggtta gttctagt taattgtatg tctgagaca aatagtatg cagcgtata 10740
 aagtcataac aggcacagt gcatcctag aacactaac tctctatgat acatgctcc 10800
 ggtattattg taacctgctt cgttttga ctaactgac ttgcttttt ctctcagaa 10860
 ctccagactt tggctattg tccagataag cccggacact agtagcttt tccactctg 10920
 ttaaacacc actaaaccc ccaaaatat atctacaga atatacagat ctactatga 10980
 ggaacattg ccccgagaa gaaggcagc cgcctagat gacaattcc acaactcacc 11040
 gctgacttcc tgcactgac actaggggg ggccttttta tatggocaa cgaactctc 11100
 cagctcgggt gggctgacc caacaataa tgggtagggt tgcacacaa aagggtggg 11160
 atgggggtga gaagataga ggaacagcg gctcaatgg ccaataaga cgaactctg 11220
 caataaagc tctgtatcca ggcactgaca cactgcatc atctagggc ctcaaacct 11280
 cctcggaact gctgcgtgta tctggacac acagagttc cagacattt agttgacac 11340

aaatgtccca ccaagtgcag gcagaanaag ctggacagag gtgtacagtt tgttttaaca 11400
 aaaaagtgaag gcgtcagagt cagcagaggt ggtgtgacct gttatagcct ttagagctgc 11460
 gaanaagcgt atggtatttg ctcacacagc cagattgaag gtctgtggac acatgtcatg 11520
 ttagtgtact tcaatcgccc cctggatata gcccgagcaa taggcctggg cctcattttt 11580
 ttgctctccg ccaatttcca ttgctcggtt cccacaactt gcttctcctg caattgcca 11640
 ccttaactat ggtttacatt gacaaacatc ttacaagcgg ggggtctgtc taggtatata 11700
 ataaacagtg gctctcccaa tgggttgcca gttctttttt tctttctttt ccccaagat 11760
 tcgaactata acatacact cacaacagc ctgtactga cgtccttaag cgaagctcg 11820
 gtgtcatcgt cggcgacagt gtccgagcgg tgaatatcca cgaacagatc agtgtcagaa 11880
 cgaacggttt tgtgtaatga cacaatccga aagtcgctag caacacacac tctctacaca 11940
 aactaaacca gctctccatg ctgaaggttt ctcgacaggg totgcacctc gtcctcagag 12000
 gaagagtgta cgaactctcc gcttgggtga acttccaccc tgggtggagct gaactcattg 12060
 aqaactacca ggaacgagat gctactgacg ccttcactgt tatgcactct caggaagcct 12120
 tcgacagctt caagcgaaatg cccaaagatca accaggtctc cgaagctgct cccagggctg 12180
 cgttcaacga agctcaggag gatttccgaa agctcagaga agagctgato gccactggca 12240
 tgtttgagc cctccacctc tggtaactgt acaagatctt gacacactgt ggttttggg 12300
 tgcctgctt cttcatgctg gtcacgtacc acctgtactt catlgtgtct cctgtctcg 12360
 gtatgcacta ccaagcaaatg ggtgtgctgt ctcactgact ctgcacacac cagaacttca 12420
 agaacggaaa ctggaataac gtctcgggtc tggcttttg caacgactc cagggtctt 12480
 ccgtgacctg gtggaagcag agacacacg cccatcttc tgcacacac gttcagggtc 12540
 aagatccaga catgtataac ctgcctctgc tgcctgtgtc cgaaggaagt gtcactcag 12600
 cttctccact ccccgaaag ctcactcagt tccacaagta ctatttctgt gtcactgta 12660
 ttctctcag attcatctgg tgtttccagt ctgtgtgac cgttcagatc ctcaggaacc 12720
 gagaacacca gttctacaga tctcagtaaa aqaagagag catlgtgact gctctgact 12780
 ggaacttcaa gacctgttc caactctct ttatgcctc catctcagac tcatgtctgg 12840
 tgtctttgt ttcagagctc gtctgtggtt tcggaattgc catcgtggtc ttcactgaac 12900
 actacctct ggaagagatc ggtgattcgg tctggagcgg acatggtctc tctgtgggtc 12960
 agatccatga gacctgaac attcagcag gcatcattac tgaactgttc ttgggagctc 13020
 tgaactacca gatcagagac catctctgga ccaactgccc tcgacacaaa ctcactcag 13080
 ttcttacca ggtggaacag ctgtgcagaa agcacacact cccctacaga aacctctgc 13140
 cccatgaag gtctcctacc ctgtccagat acctgtccaa gttcgtcaga atgcacagaa 13200
 agcagcccg tgcacagct cagtaagcgg ccgactgaga agataaatat ataaatacat 13260
 tgagatatta aatgcgtag attagagagc ctcactatgc tcggaagcaa gccaaagca 13320
 gtaactcaag ggaattacac catcctatcc caacagacaa agctggggaa aggtcttata 13380
 tcaactttcc ggaatacgtt agtttcagat gttatcaatg gggcgaccaa ggtttcag 13440
 caactcgtgt tctcgggtgt aatggcgtt cttggctccc atcaagctgt accatgctt 13500
 ctttgcgtg tcaagtaaa acgaagcag atgaagaagt aactgaagt gaaagattt 13560
 aaatagttg agcaagggag aaatgtagag tgtgaagaa tcaatctgt cgggttat 13620
 ctgcaccaat agccaagtc tggagtttt gagaagaaaa ggaagataac gtatgtaaa 13680
 aagcagcaa tggtaacata atacagagag catgtatcat agagaattag tgttcagatg 13740
 atgcactgg tgcctgtat gactttatac ggtgactaac atattgtcc tcagacatac 13800
 aattcactgc aagcacttac ccttgacat ctgtagttac ccccgagcca agacatctc 13860
 agcgtgtgt atgtcgtatt ggcgtagctc cctcgtctgt caattgctc cactctact 13920
 tctctgttt ggtcacccc agactgtct ctatgtctg ttttgtccc ttatctatcc 13980
 tccagttat accaactata aatgacatga tgtgttggg tctacattt catatcagag 14040
 ataaagatga gcaagttgc ataaagccc caactctaat cagctcttct ctttcttgta 14100
 atagtacaa aggtgattag cgaactctgg agcttattt ggcctaaaa aactcaaaa 14160
 aagcaaaaa cgaanaacga aaaccccag ttttgagaa agggaggtaa cgaagagatc 14220
 tatatatata tatatatata tatacccag gatccagaga cgggctttg attcttcct 14280
 acaacacac attctacaa ccttaattca aaacatggg cgtattcatt aaacaggaga 14340
 agcttccggc tctcaagaag tacaagtact cgcgcagaga tcaactgttc atctcaaca 14400
 acattctgag cccctctgag cgaagtttg tcaaatctt cccctgtgag atgcccacca 14460
 aactgtgac totgtggg tcttctttg tctttgtgaa cttctcacc atgtcattg 14520
 ttgatccaa ccaagacgc ggcctccaa gatgggtata cctcaactac gctcgggt 14580
 tgttcttta ccaagacttt gatgtctg aggatccaa tgcgcagca actggcaga 14640
 gtgaacctt tggagagctg ttgacact gtgtcagac catgataacc tctctgatt 14700
 tcaagtggt gttgtccacc accatctag gatatacat gaagctactg atgtcaga 14760
 ttgcgcctct cgaacactc taactgtcga cctggagac ctaacatacc ggaactctg 14820
 accttctg ctctctgtgt cctgtgaag gtatctgt totgtgtgt ctttctgccc 14880
 tcaactctt cactgtccc aaggtgtac ctctgacgt ctaagaggt cttccaggt 14940
 ccatcactc gctgtgct gccagcttcc tgaagctaac catcaccag atctacattg 15000
 gattcggagt gctggcag gtgtcaaca tcaagggc cgcggaac gtgatcaagt 15060
 actacaaca caaggcag agcgtctccc ccgcattct cgaactccc cctttggca 15120
 tctctactg cgggtcttt gctgggccc atgtgtccc tctgtctc tcaagtagc 15180
 cactgtata tctgttccc atggggctg cctttgcac gaaagtcga cagatgatt 15240
 ttgcacatct cgtcttgc cctttccc actggaact gctgtcttc ttcctcttg 15300
 tgggaactgg agtgacatac attgcacccg tgtttggctg ggaagcag atcgtgtcg 15360
 ttaaacactct ctcaactgt tttggcgcca cctctcat ttacgcttc tttgtgttg 15420
 agatcatcga cgaactcacc aactacctcg atctctgtg tctgcagac aagtaacctc 15480
 aggaagaaa gacgaataa cgcgcgcagt ggaagctgtg tctgagctg atgtttcta 15540
 tggagttgt agtgttagta gacatgaggt gtttatatat gatgaatga tagatgtgt 15600
 ttgtattgc acgatggaat tgagaactt gtaaacgtac atgggaatgt atgaatgtg 15660
 ggtttttgt actggataac tgaagctcag tggagcctgt tgttcaata tccaagagt 15720
 gcaagaaat tgggtcaag tgaactgtc ctctctgttc aagtaaaaa tcaactatg 15780
 gtatgtatt tagtaaggac aaggttgtg attctttgga gtcctagaa cgtattttg 15840
 cgttcaaga tcaaatagt agagtaatac ggcacggga atccattcat agtctaatc 15900
 ctgcaagtg gtttaatat agacttggc gtaactatg tcatagctgt tctctgtg 15960
 aattgttat ccgtcaaca ttcacaca c 15991

<210> 132
 <211> 14554

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pZKL-1-2SR9G85

<400> 132

```

cgatagtgg agcaaggag aatgtagag tgtgaagac tcaatatggt cogggttat      60
ctcgacaat agcaaaagtc tgggtttct gagagaaaaa ggcaagatac gtatgtaaca      120
aagcgaagca tggtaacata ataccggagg catgtatcat agagagttag tggttogatg      180
atggcaactgg tgcctgggtat gactttatac ggctgaactac atattttgtcc tcagacatac      240
aattacaagtc aagcaacttac ccttggaacat ctgtaggtac ccccgagaca agacgatctc      300
agcgtgtcgt atgtcggatt ggcgtagctc cctcctcgt caattggctc ccactactct      360
tctttgtcgt ggctacaccc agcatgtctg ctatggctcg ttttctgccc ttatctatcc      420
tcccagtatt accaactcta aatgacatga tgtgattggg tctacacttt catatcagag      480
ataaggagta gcacagttgc ataaaaagcc caactataat cagcttcttc cttctctgta      540
attagtacaa aggtgattag cgaatctggy aagcttagtt ggcctaaaaa aatcaaaaa      600
aagcaaaaaa cgaaaaaaca aaaaaccacg ttttgaagac agggagggtaa cgaaggatcg      660
tatatatata tatatatata tataccacag gatcccgaga cgggcctttg attcttcctc      720
acacccaacc attctcaaca cactaatcca caaccatggc tgcctctatc gaggtggcca      780
aagagtctgt cgtctatcact gccgaagccc tcccaaggt ggactatcag cgaactctgc      840

gagacatcta ctctcgagag ctctgtact tctccattg tttgtctatc ctcaagtta      900
cccttggaga gctctggat tctggcaaaa agattctgag agtctgttcc aagtggtaca      960
acctctctat gtcctctttt tgcctgggtg cctctctctg tatgggttac gcatctaca      1020
ccgttgacct gtaaccaac gaatcgaca gaetcttga caacagcttg ttcctgttg      1080
ccaccaaggt ctctactat tccagtttc tggagtaaat cgaetcttcc taactctccc      1140
tcattggcaa gccctctgcc tttctcagc tcttctacca ctggggagct cctatggaca      1200
tgtgctctct cgtcgagcac tctggcgaat ccatttggat ctttgtgttc ctgaacggat      1260
tcattcactt tgcctagttac ggcatactatt ggaacaggct gatgaagttc aacttctcca      1320
tgcccaagca gctcattacc gcaatgcaqa tcaccaagtt caacgttgcg ttctactcgt      1380
tgtgttggtc caaggacatt cccgttacc gaagatccc catgcaaatg ctggcttga      1440
tctcaacata ctggtacgtc ggtacogtto tctgtctct cctcaacttc ttgtcaagt      1500
cctacoggtt tcccaagcct aagactgong acaaaaaggt ccaggcgccc ggtcccgctc      1560
gaactgcagg acttctccc gctactact acgaactctc ggcctctatg ggtacogtga      1620
aggtctctcg acaggctctg cccctctgta tcgaaggaaa ggtgtacgac gtcctcogtt      1680
gggtgaactt ccaacctggt ggaegtgaaa tcattgagaa ctaccaggga cgaatgcta      1740
ctgaacogct catggttatg cactctcagc aagccttoga caagctcaag cgaatgccc      1800
agataacca ggtctcagc ctgctccccc agctgcogt caacgaagct caggagatt      1860
tcgaagagct cagaagaag ctgctcgaca ctggcatgtt tgaogctctc cccctctggt      1920
actgtacaa gatcttgacc acctgggtc ttggctgtct tgcctctctc atgtgggtcc      1980
agtaacacct gtaactcatt ggtgctctcg tgcctgttat gcactaccag caaatgggat      2040
ggctgtctca tgacatctgc caccacacga ccttcaagaa cgaaaactgg aataacgtcc      2100
tggtctggt ctttggcaac ggaactcagg gcttctcgt gaactggttg aaggacagac      2160
acaaagccca tcttctgct accaacttgc aggttcaaga tccgacatt gataactgc      2220
ctctgtctgc ctggtcagc gacgatgta ctgagcttc tccatctccc cgaagatcca      2280
ttcagttcca acagtaatat ttctgtgta tctgtattct cctcgatttc atctggtgtt      2340
tccagtctgt gctgaacgtt cgaatccctca aggaacgaga caacagttc taacgatctc      2400
agtacaagaa agaggccatt ggaactgctc tgaactggac tctcaagacc ctgttccccc      2460
tcttctttat gccctccatc ctgaactoga tgcgtgtgtt ctttgtttcc gagctctcgt      2520
gtgctctcgg aattgcacac gtggtcttca tgaacacata cccctcgag aagatcgttg      2580
attcgtctg ggaacgacat ggcctctctg tgggtcaaat caatgaagac atgaacatto      2640
gacggagcat cactactgac tgyttctttg gaggcctgaa ctaccagatc gagacacatc      2700
tctggccccc cctgctcoga caacaacctca ctgcogtttc ctaccaggtg gaacagctgt      2760
gccagaagca caactccccc taacgaacac ctctgcccac tgaaggctct gtaactctga      2820

```

tccgatacct gtcccagttc gctcgatagg ccgagaagca gcccggtgco aaggctcagt 2880
 aagcggccgc atgagaagat aaatatata atacattgag atattaaatg cgtatagatta 2940
 gagagcccca tactgctcgg agagaagcca agacgagtac tcaaaagggg ttacacacac 3000
 catatcccca gacacagct ggggaaggt tctatataca ctttcoggaa taccgtagtt 3060
 tccgatgtta tcaatggggg cagccaggat ttccagcaat tccgtgtctc ggggtgaat 3120
 ggcgttcttg gccctccatca agtcgtacaa tgtcttcatt tgcctgtcaa agtaaaacag 3180
 aagcagatga agaatgaact tgaagtgaag gaatttaaat gtaacgaac tgaatttga 3240
 ccagatattg tgtcccggtt gtagctccag ctttctgtcc ctttagtgag ggttaatttc 3300
 gagcttggcg taatcaggt catagctgtt tctgtgtga aattgttato cgtccacaag 3360
 cttccacaca acgtacgata gttagttagc acaatcaga acatctccct ccttatataa 3420
 tccacacagg ccagaacggc taaactaaag cgttttggac actatgttac attggcattg 3480
 attgaactga aacacacgtc tccctgcctt gaacagagca atggatgttg tccgaagtca 3540
 acttccatag aagacgggtt ctatgctctg tcaagatcat atcataaact cactctgtat 3600
 taccocctct atagaacact tcttatgaat ggcgggaac attccgctat atgcaccttt 3660
 ccacacaaat gcaagatgt gactcttcaa cgggtagtaa gactggtcc gacttccgt 3720
 gcatggagag caatgaacct gataatgaga acatccccc catatacaat cttaacacag 3780
 ccaataaat ctgtgctatt actaaatatt taagctctat cactgtcttg atgaagcgc 3840
 gcaacggatg tatcatotag ttccggccaa tccagaaccc aactgttg gcaagtgtgt 3900
 agccatggc acacagacca aagatgaaaa taccagacac ggcggctoga gccgtgtgc 3960
 ctcgagcaac acccttgtaa tgcacaaagc gaggttaat gtacaccca ggcacacatg 4020
 caaacgaccc ggtgagagcg aagaacccat ctagatctgc ggcacccccc catgaacaa 4080
 aggcgggtgac aaacacacag aagaacggg aaatgtctt ctgcacattg atgttagagt 4140
 tgaacttgcc tgatcgggtg aagagccact tctcgatgat tggatggcg cgcagctgc 4200
 attaatgaat cggccacacg cggggagagc gcggtttgcg tattggcgcc tcttccgctt 4260
 cctcgtccac tgaactcgtg cgtcggctcg ttccgctgcg ccgagcggtta tcaactcaat 4320
 caaagcggtt aatacgttta tcccaagat caggggataa cgcaggaag aacatgtgag 4380
 caaaaggcca gcaaaagccc aggaacgcta aaaaaggccc gttgctggcg ttttccata 4440
 ggtccggccc ccttcacggc catccacaaa atccagcttc aagtccagag tgcgcgaacc 4500
 cgcacggagt ataaagatcc caggcgttct cccctggaag ctccctcgtg cgtctctctg 4560
 ttccgaacct gccgcttacc cgtacactgt cgccttttct ccttcggga agcgtggcgc 4620
 tttctcatag ctcaacgtgt agtatctcca gttcgtgtga gttcgttgc tccaaagtgg 4680
 gctgtgtcca cgaacccccc gttcagcccg accgttggcg ctatccggtt aactatctc 4740

 ttgagtccaa cccgttaaga cagacttat cgcacatggc agcagccact ggttaacaga 4800
 ttacagagcg gaggtatgta ggcgtgtcta caggttctt gaagtgttgg cctaactag 4860
 gctacactag aagaacagta ttgttatct cgtctctgtc gaagccggtt accctcgaa 4920
 aaagagtttg tagctcttga tccggcaaac aaacacccgc tggtagcggt ggtttttttg 4980
 ttgcagaaga ccagattacg ccggaaaaa aagatctcca agaatctct ttgatctttt 5040
 ctacgggttc tgagctcag tgaacgaaa actcaogtta aggtatttg gtcatgagat 5100
 tatcaaaaag gatcttccc tagctctttt taaattaaaa atgaagtttt aaatcaatct 5160
 aaagtatata tgagttaact tgggtcagca gttacaaatg cttaacagct gaggcaccta 5220
 tctacagat ctgtctattt cgttccatcca tagtgcctg actcccgctg gtttagataa 5280
 ctacgatagc ggcgggctta ccatctggcc ccagtgtgc aatgatagcg cagacccccc 5340
 gtcacccgc tccagattta tcaagataaa accagccagc cggaaaggcc gagcgagaaa 5400
 ggggtcctgc aactttatcc gccctcatcc agtctattaa ttgttcggcg gaagctagag 5460
 taagtgttc gccagttaat agtttgccca acgttgttgc cattgtacaa ggcctcgttg 5520
 tgcacgctc gtcgttgggt atggtctcat tcaatccggg ttcccaacga tcaagcgag 5580
 ttacatgato cccatgttg tgcacaaaag oggttagctc ctccgtctcc cagatcgttg 5640
 tcaagaagta gttggccgca gtttatccc tcatgttat ggcagccctg cataatctc 5700
 ttacgtcat gccatccgta agatgctttt ctgtgaatgc tgagtactca accaagtcat 5760
 tctgagataa gtttatcgcg ccagccagtt gctcttgcgc ggcgtcaata cgggataata 5820
 ccgcgccaca tagcagaact ttaaaagtgc tcatattgg aaaaacttct tccggcgcaa 5880
 aactcctaaq gatcttaccg ctgttgagat ccaatctgat gtaacccact cgtcaccoca 5940
 actgatcttc agcatcttt acttcccca ggtttcttg gtagonaaaa accggaaggg 6000
 aaaaatgcgc aaaaagggg ataaaggcga ccgggaatg ttgaatact atactctcc 6060
 ttttcaata ttattgaagc atttccgg gttattgtct catgagcgga taactatttg 6120
 natgatatta gaaaataaaa caataagggg tcccgccgac atttcccca aagtgccac 6180
 ctgatccgt gtaaaatacc gccagatgc gtaagagaaa aataccgcat caggaaattg 6240
 taagcgttaa tatttgttta aaattccgtt taaattttt taaatccgc tcaattttta 6300
 nccaatagc cgaatccgc aaatccctt ataataaaa agaatagacc gagataggt 6360
 tgagtgtgt tccagtttg aacaaagctc ccatattaaa gaactggac tccaaagctc 6420
 aaggcgaaa aacgtctctat caggcgatg gcccaactag tgaacactca ccatatcaa 6480
 gttttttgg gtcaggtgc cgtaaaagac taaatcgaaa cctaaaggg agccccgat 6540
 tttagacttg accgggaaa cggcggaac tggcagaaa ggaagggaag aagcgaaaag 6600
 gagcgggcg tagggcgtg gcaagtgtg cggtaacgt ggcgttaacc accacacccg 6660
 ccgcgttaa tgcccgcta caggcgagt ccatcgcca tcaaggtgc gcaactgttg 6720

ggaaggggga tgggtggggg cctcttgggt attacggcag ctggcgaaag ggggatgtgc 6780
 tgcaggcgga ttaagtgggg taacggccagg gtttcccg tcacgaagtt gtaaaagac 6840
 ggcagtgaa ttgtaatacg actcactata ggcggaattg ggcgcgaagt cgcagtgtta 6900
 gaagtggaga ttacaagaag cctctgggta tcaatgatga aogtaotoag oggtgggtca 6960
 agcatttoga cgtcgatcgc gaagaggtgt tcacotttga caagcgaggg ttcccaattg 7020
 atcactgtct ogagtgttc aaattttctc tcaaatcttc tctgatgaa ctatctctgt 7080
 tgacgaagct gtcaaccact gttcctcgaa cgcctttctc ogagtttggt ctgaacatct 7140
 tggatctcaa actgaacccc gcaagtatca atagtgccat gcaactggcg atgoggtgog 7200
 aacatccctg gagggaattct oggagctcta cacaatgcag attctgttgt ogagtactct 7260
 ctacactgct cgaatgactt atttgctcac tactgcactc atgcttgat catgtgccc 7320
 actgcacccc aaatttggtg atctgattga gacagagtao cctcttcago tgaatagaa 7380
 gatcatcago aacatgaatg atgtggttga ccaggcgagg tgttgtagtc actgcttca 7440
 ctccaagtic attcttcaat tctttctgtt ttactttgac agccaatga agacatgcta 7500
 cgaattgatg gaggccaaga accgcatttc acccgagac accgaagtgc ctgaactcct 7560
 ggtgcocccc attgataaca tgggaacta cgttattoog gaagtgtat atagaacct 7620
 tccccagctt gttgtgtgtg atatggatgg tgaatcccc ttaattaaot cactgcagg 7680
 attgagacta tgaatgactt cctgtgcgcg tattactcta cttaattgat ctgggaacgc 7740
 gaaatacgt ttctaggact ccaagaatc tcactctgtt tcttactaa atatactacc 7800
 catagtgat ggtttacttg aacagagagg acatgttca ttgacccaaa gtttctgca 7860
 tctcttggt atttgaacaa cgcgtccac tgccgtcag ttatcagtc acaaaacccc 7920
 cacattcata cttcccatg taagtttaca aagtcttcaa ttccatgtg caaatcaaa 7980
 tcacatctat tcaattcata tatataaacc catcattgtc actaacaact acaactcat 8040
 agaaaacatc gactcagaac acacgtctca tgcggcgctg taggcaagcg gcttgatgc 8100
 agcgggaagga gtgccacat tgtttcagtt tcgaagaac aggaacccct tgcagctcc 8160
 ctccggacca cggcagggtt caacccactg gcaactcgt gcaatcgtt acatggtcg 8220
 aatgaatcct caggaacgt cctggacatc agctcgatag tgcctgcoca tgaatagttt 8280
 gatgcctcg gtatcttgt cgcatttga gaagggaatg gaagagagct agtgatcag 8340
 gacgtgagtc tggataatgc cgtgagcag atgaacttca atgaagccca tctctggtc 8400
 gatgttgca cgggcacatc gcaaaaagtt ccaactcgtg ttggttagt ggggaagat 8460
 aggatctgtg tgcctcagaa agttaatggo gacgagccag tggtaaccc acaagtaggg 8520
 aacgaagtao cagatgcaca tgtttagaaa tccgaacttc tgaacagaaa agtaacagac 8580
 gttgcacata agacaatgc caatgcgga gacacagatg agottgggt cgtgttctc 8640

gtacagagga gatcgggat cgaatgtgtt aactccacg ccaagacgct tgtcttcc 8700
 ctgtcctga cctctcgtg cgcgtcagtg gtatgttgtt cagtaaacgt tggtaatgag 8760
 atagtggg caaccgacca gttgtgag ccaagcagtg agcaggtga aagcagagt 8820
 ttctcgtgta agatgggga gttcgtgggt catcttcgag agtcagtag ctgctgtctc 8880
 tgggttoga ggaacgaaga ccaatctctg ctccatgttt ccaatggcct tgtgatgctt 8940
 cgggtggag atttgcacg tgaataggg aacaaagagg gaagatgaa gcaaccaagc 9000
 agtaatctg ttgatgattc gggatcoga gaaacacca tgcacacat optggccat 9060
 gacccacagt ccaatcaga agatccctg aagaacgttg tacaagccc accgaacggg 9120
 tcaagcaga gtcaggga tgaactcagg tgcacaaag ttgtacaga tctgaagt 9180
 ggtagttag aggaacatg ctogaagaat gtagccgat ccttgagag cagatcgctt 9240
 gaagcagtc tgggaatag cgttgtatg gtcttgatg gtgaagtcg gaacttoga 9300
 ctgggtcgg taggtatca gactgaacac gtaactcagg ttgggttgg caatgtcac 9360
 ctggacatg gaagcagcg atgtagagga ggcagagtg ctgggagaat cggaggga 9420
 gacggcaga gactccaggt cgttcacagt ggtggaagt accgttctgc ggaaggcag 9480
 gtctgtctg ggcagagcg agtggaagc cactggcact gctgtagata tgtcttgt 9540
 gtaaggggtt tgggtgtgtt gtttgttgc ttgaatttt tgttagcaag ggaagcggg 9600
 caaaaagtg agtgtgtgtt ggaaggagag acgaactta tatataatg ttgttgtgt 9660
 ttgtgcaat ggaacogaa accggcagga gccaaactaa acaaggcaga caatgcaga 9720
 ttaattgcat tgcctgatg gcaaggttta ggcctcagat aatgggggtg cgaatgaca 9780
 aaattggaaa taagttcgc aaqcaagct gacaaactt tggcccaaac atttgtacg 9840
 ggtggaacac aggaacacac catgtctgt caggggtag accgttctgc gtcctgtag 9900
 gctcacata ggtccatgc cactccatac atcccaata gtgtaacgt aggcgcttt 9960
 tagctccat ctgaacccc cccaaact ccaatgtaca gtgaactga ctgttgag 10020
 atcaaggga agggaaaaa ggcgaacaa tgcacgcatg gaatgacta ggttaaggct 10080
 tactagctg aaaaatggca ctttccgag tgcacaaagg tcaatgtgc gtttccag 10140
 ctggcgcca gccaagcgc ctccaaacg cctctccagc tccctccagc ggcctccata 10200
 tcccatccc tctccacag aatgttgtta agccttgaac acaaaaaat agaaagcta 10260
 ataatctcc aatattgttg tgaactgag ataaogaaac atagagcgc aacaaacac 10320
 caccacagc accacagag attaacacag atgaacagca tgaattctct ctcttgag 10380
 ttccataac aagtcttctt gctccagga agtcctagg tggttgata atggttttg 10440
 tgaatgta gtgaatgtt ggtatttga ctgggatgt agttagaat aagtcata 10500
 caagcagct ttcttcagc ctcatatag tataatagt tcaactatt agcactgac 10560
 cagcatctc cgtatcgaga aacacacaa catgcccac ttgacagatc atgcygata 10620

```

acagggttggt cagtatcata catactegat cagacaggto gttgacocat catacaagct 10680
gaacaagagc tccatacttg caagctctct atatacacag ttaaatbaca tatccatagt 10740
ctaaectcta acagtttaac ttctgttaag cctccacagc agctctcttg tatcgcttgg 10800
cctctcatat aggatctagg ttctggcctg acagacctcg gcgcacaatt atgatatcgg 10860
ttccggtaga catgaacatcc tcaacagtto ggtactgtcg tccgagagcg tctccttgt 10920
cgtcaagacc caaccocggg gtoagaataa gccagtcttc agagtcgcoc ttaggctcgt 10980
ttcgggcact gaagccaacc caaaactcgg ggtcgatcg ggcacgtca atgctctgct 11040
tggagctact gccagtggcc agagagccct tgcagacag ctccgcocag atgagcaagc 11100
ctctggccag ctctctcttg gtagagggga ctaggaaact ctgttacttg gagttctcgt 11160
agtcaagagc gtctctcttc ttctgtcag agacagtttc ctccgcacca gctccagagc 11220
cagcaatgat tccgcttcgg cgtacacccg ggcgttgggt gatctggac cactcgccga 11280
ttcgggtgaca ccggtactgg tgcttgacag tgttgccaat atctcgcaac ttctgtctct 11340
cgaacagcaa gaacccgtgc ttaagaacaa gttccttga cgggagcaca gtcocggcgt 11400
aggtgaagtc gtcaatgatg tccatagggg ttttgatcat gcacacataa ggtccgacct 11460
tatccgcaag ctcaatgagc tcttgggtgg tggtaacatc cagagaagca cacaggttgg 11520
ttttcttggc tgcacagagc ttgagcactc gagcggcaaa ggcggacttg tggagttag 11580
ctcagacttc gttagggggc attttggtgg tgaagagggg actgaataaa atttagctcg 11640
cagaactttt tatcggaaac ttatctgggg cagtgaagta tatgttatgg taatagttao 11700
gagttagtig aaactataga tagactggac tatacggcta tccgtccaaa ttgaaaaaa 11760
cgtcaatggc tctctggggc tcccttgc cgacaaaaat gtagctatga tgaagccag 11820
caatgaagtt gaagctgata ttgttgtcgg ccaacccggc cgaataagca gctctagac 11880
ccacagcttc caacgaagaa tctatctgca aagctgacca agcaactcca tagttggagt 11940
cgtactccaa aggcggcaat gacagtcac agagatactc gtccacattt aattctcag 12000
tgacacagat tattaacgtc tcttaccac caccagttac gaacctctcg cagtccagct 12060
tcaatagggt ttgggttgcg tccgttgaga gccgtttggt ttaaaccttc tccatgtct 12120
cactcaggtt ttgggttcag atcaaatcaa ggcgtgaacc actttgtttg aggaataatg 12180
tgacacaacc aaccagtgc aggggcaagt ccgtgacaaa ggggaagata caatgcaatt 12240
actgaagtt acagactgcc tccatgcctt aaccttgccc caaataaaga caactgtct 12300
cgtttaagcg caacctattt cagctcacc tccataatag gtttgatag cactagtcta 12360
tgaggagcgt ttatgttgc cgtgagggcg attgtgtctc atatgggttc aattgagtg 12420
gcggaacgag cttagcttc aattgaggtg cgaagcagac aattgggtgt caactgctct 12480
aattgaactc gggctgtgga gtcccaagtt atacagaaca caagaggtgc atgggttaga 12540

```

```

gacgcacca gacaataggg ttttttttg actggagagc gttgggcaaa agcgtccaac 12600
ggcgtgttgg gggagctgtg ggggagaatt tggcatatt tgtgaggtta accgctccga 12660
tttgcgtgtt ttgtctctcc tgcctctccc catcacata tcttccctcc cccctcttt 12720
ccacgataat tttaaggatc agcaataagg tctctctccc tagtttccac gtccatatat 12780
atctatgtcg cgtctctctt ttctgacat caccaaaaaa catcaacaaa tggctctctc 12840
ccttaactac gagcagctgc tggagcagac agacctggtt gcaatcgagc gaattctcta 12900
cgatctggaa ggtcttgcca aggtccatcc cggatccgac ttgatctctg cttctggtgc 12960
ctccagatgt tctctctctg tctactccat gcaacttacc gtcaagcccg agaacctgaa 13020
gctcttcaa cagttctgc gaggaagca cgaacgaac tccaaagaca ttgtctaac 13080
ctacgactct cctttgcac aggaactcaa gcaactatg cgaaggttca tgaaggtcg 13140
gaactgttat gccacactg gatctggtt ggaacccgtt ggcactattg ctgtcaocgc 13200
ctttgcag tggacctgg ctactccgg aatggctgtg tgggtctctt tgaactgatt 13260
catgcacatg cagatggcc ttctcattca gcaagatgoc tctctgttgg ccatcagcaa 13320
aaagcccttg gtcaacgttc tctttgcta agcatcgac gtcatlggat cgtccagatg 13380
gatctggtg cagttctcaa tcatcgaca tcaacttcc acaactcagc atggtctoga 13440
cctggatgcc gactccgag aacattctct ttgttccac aactaacctg ctgcacacac 13500
tgtccagaa ggtttccac gattccagc ctgtacatg taactgtgct ttggagccta 13560
cggcgtttcg ctggtgtaca accctctcta ctcttccga atgcagcaca aagcaacat 13620
tcccgatct gtccagacca tgcagagaaa cggctttctg cgaaggtaac gaacacttgc 13680
attcgttatg cagactttct tcatctttcg aaccccttc ttgcctgtgt atctcaactg 13740
aacctccctg ctcatcaaca tctctgtgt gccactgtt aacgtgtgct tctcaactt 13800
cttttctac ttgtctcaaa acttcgatgg ctggagaga atcccgaca agactgcaa 13860
ggtcaagagc tccgagaagg acgttgagc cgtatagata gactgttaca gactccagt 13920
ggagactctt tccactacg gtggaccatc tgcactgttc tttactggcg gttctcaact 13980
ccagatcgag catcactct tctctcaat gtcgtcttgc caatatctct tctgcagca 14040
agctgtcga gactgttgc aacgacagc agttcggtac gttcttacc ctaccattgt 14100
gggcacatc atttccacc tcaagtaact gccacagtc ggtgtggttc actgtgtcaa 14160
ggacgtctag gattcctaag cggccgcatg tccatcacag attatttata gaattgact 14220
gcgtccgac aaagaactag agtgtacagc taggggatga tgataaaagt ggaagaagtt 14280
ccgatcttt ggatttaca acgttagga cgtacttcc tgaataaatg caatgtcttt 14340
acctaggtt ctgtgtaga tgttattaac taccattaac atgtctactt gtacagltgc 14400
agaccagltg gacttagaa ttgtacactt accaaaaagt gttgatgtt gtactcaga 14460
tatataaac tgttagcggg atcccctgtg atatgctaa ggaacatca aagaggga 14520

```

```

tattaattca gaatgctagt atacagttag ggat 14554

```

```

<210> 133
<211> 15119
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZSCP-Ma83

<400> 133

```


gtacgacccc tctcaggcca agcagaaggc tgagtcacac aagaaggcca acgctatcat 60
 tgtcttoaac ctcaagaaca aggtcggona gacogagtct tggtaocttg aotcaagaa 120
 cgacgggtgac gtocgcaagc gcaacaagtc ccccaagggt gatgctgaca tccagtcac 180
 tctctctgac gacocactcc agocagctgt tgagggttaag gctaacgccc agcagactctt 240
 catgacoggc aagctcaagc ttaagggcaa cgtcatgaag gctgcogoca ttgagggtat 300
 cctcaagaac gctcagaaca acctctaagc gcactattta ttgattaatt gatgatttac 360
 tataattgatt togoaactgt agtgtgattg tatgtgatct ggtcgttagg ctccagttaa 420
 tactagaagc gtatcctaag tagttgtatc atacatcgag cctgtgggta cttgtacaat 480
 aattcgtaat gttagagata cctctgatcc attgccgtgt tctaacatac aatgactcc 540
 acgcataat cccactcttg actaaaagtt gctactcttg caggtttacc toggcatagt 600
 cagcctctct tgtctctgtc tcgaacgcac aaagtoaat gacaacgcca ctcaactgag 660
 tgtgccccaa cagggcacaa tatcgactaa tttaggcca actagggtga ttttggatgg 720
 aatttgatcg gaaaaaatag ctgcagaat tcttgagag aaaaattgac cgcatacaca 780
 tggtttgacc aaaaaatcgt cctcatctct gtgctaacct ctctgaaga gatatgcgcg 840
 cgcaccccca catgatgta ttgatctcaa caaactcac ccagacccct atcttcccg 900
 gaaacttact gtataagtg tggtgcgaac agaaagtggt cgcactttag gtgtctagat 960
 ccgattgttc togtctgat aatgagccag ccccgcgagg caatgtttt tacaattgaa 1020
 aactctgta accactcaaa ttaccgtttt tgcccctat ttaccctctg gtacactccc 1080
 tottgatcac acaacactg caagtcaaat gcactcogtt agacogcttg tgattgttc 1140
 agggcacagc tttgttggtt taaggcgcaa ctacatcaat atgaaaacag gagacgtga 1200
 aaagggttaa tatcgactg ctgctatggt gtatgtactg catgacgaat tgggttatt 1260
 caagacogtg gcaacggttg ctgcggtacg agacgtgta gcttctcaa accgcatgct 1320
 taggtggcgc gccagctgca ttaattgaat ggcacacgcg cggggagagg cggtttgct 1380
 attggcgct ctcccgcttc ctgcctcact gactcgtgc gctcgtgctg tggctgcgg 1440
 cgagcggtat cagctcactc aaaggcgcta atacggttat ccacagaatc aggggataac 1500
 cgaggaaaga acatgtgagc aaaggcgacg caaaaggcca ggaacgtaa aaaggccgcg 1560

ttgctgggtt tttctcatag gctccgccc cctgaagaga atcaaaaaa toagcctca 1620
 agtcagaggt ggcaaaacc gacaggaata taagataacc agcggttccc cctggaaag 1680
 tccctcgtgc gctcctcgtg tccagacctg ccgcttaccg gatacctgct cgcctttctc 1740
 ccttcggaaa cgttgcgct tttctatag tcacgctgta ggtatctcag ttgggttag 1800
 gtgcctgcct ccaagctggg ctgtgtgca gaacccccc ttacgocga cgcctgcgc 1860
 ttatccgcta actatcgtct tgatgcacac ccggttaagc aagacttacc gccactgcca 1920
 gacccactg gtaacaggat tagcagagcg aggtatgta gogtgctac agagttcttg 1980
 aagtgtggc ctacatcag ctacactaga agaacgtat ttggtatctg cgcctgctg 2040
 aagccagtta ccttcggaaa aagcgttggt agctcttgat ccggcaaaaa aaccacogct 2100
 gtagcggctg gttttttgt ttgcaagcg cagattaccg gcagaaaaaa aggtctcaa 2160
 gaagatcctt tgatctttc taagggtct gcgctcagt ggaacgaaaa ctacgttaa 2220
 gggattttg tcatgagatt atcaaaaaag atctcact agatccttt aaattaaaa 2280
 tgaagttta aatcaacta agtatatat gattaaactt ggtctgacag ttaccaatgc 2340
 ttaactcgtg aggcactat ctacgcgac tgtctatct gttcatccat agttgctga 2400
 ctcccgctc tttagataac tacgatacgg gagggcttac catctggccc cagtctgca 2460
 atgatacgc gagacccag ctacccgct ccagatttat cagcaataaa ccgccaagcc 2520
 ggaagggcgc agcgcaagc tggctctgca actttatccg cctccatcca gtctattaat 2580
 tgttcgggg aagctagagt aagtagtctg ccagttata gtttgocaa cgttgtgca 2640
 attgctacg cactcgtggt gtacgcctc tcttttgta tggctcatt cagctccggt 2700
 tcccaacgat caaggcgagt tacatgacc cccatgttgt gcaaaaaagc ggttagctcc 2760
 ttggtctct cgtatgttgt cagaagtaag ttggccgcag tgttatcact catggttatg 2820
 gcgcactgc ataattctct tactgtcatg ccatccgtaa gatcttttc tgtactggt 2880
 ggtactcaa ccaagtcttt ctggaatat tgtatgggc gaccgagttg ctcttgccc 2940
 ggtcaatac gggataaac ccgcacacat agcagaactt taaggtgct catcattgga 3000
 aaactctct cggggcgaaa actctcaagc atcttaccgc tgttgagatc cagttcgatg 3060
 taacccactc gtgcacccaa ctgactctta gcactttta ctctcaccag cgtttctggg 3120
 tgagcaaaa cagggaagca aaagccgca aaaaaggga taaggcgac ccggaatgt 3180
 tgaactcaa tactctctct tttcaatat tattgaagca tttaaggg ttattgtctc 3240
 atgagcgat acatattga atgtatttg aaaaaaac aaatagggt tccgcgaca 3300
 ttcccccga aagtccccc tttatccgtg taaataccg caccagtcg taaggaqaaa 3360
 atcccgcatc aggaattgt aagcgttaat attttgtta aattcggtt aaattttgt 3420
 taatcagct catttttta ccaatagcc gaatcgga aaatccctta taatcaaaa 3480
 gaatagacg agatagggtt gagtgttgt ccagtttga acaagagtc acatataag 3540

```

aacgtggact ccaacgtcaa agggcgaaaa acgcctctatc agggcgatgg cccactacgt 3600
gaaccatcac cctaactcaag ttttttgggg tcgaggtgcc gtaaaccaact aaatcggaac 3660
cctaagaagg gcecccagatt tagagcttga cggggaagc cggcgaaagt ggcgagaaag 3720
gaagggaaga aagcgaaaag agcgggcgct agggcgctgg caagtgtagc ggtcaocgtg 3780
cggttaacca ccaacccggc cggccttaac gcgcgcctac agggcgcgctc cattcgccat 3840
tcaaggtcgc caactgttgg gaaggcgcat cggtcgggc ctcttcgcta ttacgcagc 3900
tggcgaaaag gggatgtgct ccaagcgcat taagttgggt aaocgcaggg ttttccaggt 3960
cacgcgcttg taaaacgacg gccatgtaat tgaatacga ctacatag ggcgaaatgg 4020
gccgcagctc gcactgcgtca ctaactcaag atactacaa tgcactatg atgtttgcag 4080
gaggtgtacc tcggcagctca tcaaaaaatg gaactactgg cttagatctc tgttgtatgg 4140
catcgcgctt aaaaaagaaa cccctctcca gcagctact acaagtagtt gtagttcgg 4200
ggtttggata cggaaaagtc caagcacatg tcgaagctct catctgaac accgacagtc 4260
gtctgcaccc cgcaggtctc gtttgtacc agcaaccaat ttacgcagaa ctatacaca 4320
gagggcggac gatcacttcg cgttaggca actgaagctt attttgggt gttacttag 4380
gggaacaggg aaacgcgaat gattagtaa tcggataata ggctgttag tttaacgaaa 4440
tggtggggga ggggttcogt ggtatcttg aagttatgga ggtgatogt tatlttgtgg 4500
gatgatatc atgttatgga catactgag ctacgtata aacaaoygat cttaacactg 4560
ctcttgtat gccacttgct tgaatctca tctgttact gtacatatac aatagatata 4620
gggaagaaaa gcgggaagta gagaacatag tcggcgagaa gtaacgctt cggctcgaga 4680
gaactataac aaagtcaaac cggcggtctt agaatagccc caaggatcac acagttccgc 4740
aatccagttt cacatgttcc gttgcaatga ctittgcatg totactgttg ctacgatcc 4800
cccattgcaa cccagtttg ggttaacccc gcattatatt agcatgata cgaagagat 4860
aagtataca tgggaacatgt gaaggtagt atgcaggtcc ggcggagaaa gagaatgacg 4920
ttttcattaa gcgattogct tggcgcttg tgggggatgt gaacatactt accgtaaaga 4980
ccctgtgcga gagctgttac tgcctgtta ctctcgtgat ctgttgggac gtcactgaa 5040
tctgtggaaa ctgcattct tottaactgt gtcatacaca gaacactaat gaaaataca 5100
agctacgaaa atcattttac tctactgac cggtaacgta ctgcacaa tagtgaact 5160
tcgaaaata gccagctcca tgcactatcg cttaacccct totgttgacc taaaaagcat 5220
tccaaoygta aaaaattata aocgcgcaaa ctggatggtt gtaacgoggt tgaacacaa 5280
tgctggggg ctggcgtag gaccgagctt attcgtccca atagctctt tggattgat 5340
tctttgggt gtgtgtaaa attcacatgg gaaagaaac ggtggcagtt tgaagcagag 5400
gccacgctg tagttccag cgcataata taccgaactc atggcgaga attgactga 5460

atgcgcasaa agctacagga tcaacgcgt tagaatguc gcaaatguc actaatccc 5520
cgactgttc caaatgattc tgtgggata aatcnaaac tgggttagc tttgtcaat 5580
ttcttttgt cgtctcgtt cgtccgggc aatgtgcaca cgttgcgtc totccatac 5640
ctcggtaaa aactatcaca tgcctgccct ctgcagcag cattaaatgc atatatcaa 5700
tctaacgaca tatatatag taggtgcat cctccggttt agctcccag aatatctctt 5760
atcattaca caaaaaaac aatgtctct aagtgagag gottacttc ttaattagt 5820
tgcacacat gtttgtatg tatctgcat tctctctct gagctttcc ataacagatt 5880
ctctcctcc caggaagtc atgggtggt tgaataggt tttgtgtag tgytagtca 5940
gtgcgttat tgtactcgg gattagttg acaataagc atacacagc cagcttctt 6000
cgagctctc ataatataa gtactcaac gattagcac tgaacccag atctcctat 6060
cgaagaaac aacaacatgc cccattggac agatcagcg gatacacagg ttgtcagta 6120
tcatacaca tgcatacag aggtcgtctg accatacac aagctgaac agcctccat 6180
actgcacgc tctctatata ccaagttaa ttacatccc atagtctaac ctataacgt 6240
taactctcty gtaagctccc cagucagctt tctgtatcg ctggcctcc tcaatagat 6300
ctcgtttcty gcgtacaga cctgggcaga caattatgat atcgttcog gtacacatga 6360
catctcaac agttcggta tgcgtcaga gacgtctcc ctgttgtca agscccaacc 6420
cgggggtcag aatagcagc tctccagat cgccttagg tgggtctgg gcaatgaag 6480
caacacaaa ctcggggctg gatcggcna gctcaatggt ctgcttagg tactcgccag 6540
tgcacagaga gcccttgcaa gacagctcgg ccaagctag cagacctctg gccagctct 6600
cgttgagaga gggcctagc aactcctgt actcggagt ctctagtaa gacagctct 6660
ccttctcty ttaagagaa gttctcogy caacagctog caggcagca atgattccg 6720
ttccgggtac accgtggcg ttgtgatat cgaacacac ggcattcgy tgaacoggt 6780
actggtgtt gacagtgtt caatatctg cgaactttct gtcctcgaa aggaagaaac 6840
cgtcttaag agcaagttcc ttgaggggga gcaagtgcc ggcgtaggtg aagtctcaa 6900
tcatctcat atgggttttg atctgcaca cataagctcc gacattatcg ccaagctcaa 6960
tgaactctt gttgttgta acatccagag aagcacacag gttgttttc ttgctacca 7020
cgaacttag cactcgagc gcaagggcg acttgtagc gttagctcga gcttctagg 7080
agggcatttt gttgtggaag aggaactga aataattta gctcgaagaa ctttttatcg 7140
gaacottatc tgggcaagt aagtatatgt tatgtaata gttacagagt agttgaact 7200
atagatagac tggactatac ggtatcgggt ccaattaga aagaacgta atgctctct 7260
ggcgctgcc ttggcgaca aaatgtgat catgatgaa gccagcaatg acgttgacg 7320
tgatattgt tgcggcaac cgcgccgaaa accgagctgt cagacccaca gccctcaacg 7380
aagaatgat cgtcaagtg atcaagcac actcatagtt ggaactgtac tcaaaagcg 7440

```

gcaatgaaga gtaagaagaga taactgttga cettttcett gggaaocaaa accgttaagcc 7580
 cttctgactc acgtatttga gccacagaca caggcaacag tccgtggata gcagaatatg 7580
 tcttgcgggt ccaattttca ccaactttag gcgtcaagtg aatgttgcaag aagaagtatg 7620
 tgccttcatt gagaatcgggt gttgcgatt tcaataaagt cttgagatca gtttggccag 7680
 tcatgttgg gggggttaatt ggtatgatt atgcctaca gtctgcaag gtatactcgc 7740
 tgccaccttt atactttttg attcgcgtgc acitgaagca atgtcgttta ccaaaagtga 7800
 gaattgtcca cagaacacac cccagggtat ggttgagca aaaaataaca ctccgatagc 7860
 gggaaatgaa ccccggtctc cccggttctc aagaagtatt ctgatgaga gcgtatcat 7920
 ggaagccggg agaacccggg tcttctgctc tggagatgga agccggtaga accgggctgc 7980
 ttgggggggt ttggggccgc tgggctccaa agagggttag gcatttcgtt ggggttacgt 8040
 aattgggcca ttgggtcct gcgcgcatgt cccattggto agaattagto ccgattaggag 8100
 acttatcagc caatcacagc gccggatcca cctgtaggtt ggggtgggtg ggaacacccc 8160
 tccacagagt agatccaac agcagcagca acatgatagt tgggggtgtg cctgttaaaag 8220
 gaaaaaaaag aagcttgggt tatattccgc ctctatttag aggttgcggg atagaccccg 8280
 acggagggca atggcgctat ggaacttgc ggtatccat accgcggggg ggaatgcgtc 8340
 cgaaccagct ccagcagcgt tttttccggg ccattgagcc gactgcagcc ccgcaacgt 8400
 gctctggccc acccaactcat gtcattgttg tttgggagc ccacttttta agtagacaaa 8460
 ggcacactag tccagcagag gttccagacc caaagagcgc gctgcagtg tgcacacggg 8520
 cggaacacgg cgggaaaaag ccccgggggc acgaattgag ccacgccttc gaatttga 8580
 cagtcacgg cccacttcgc ccgcgcaatg gctgcacac ccccggtttt ttgcacaca 8640
 tcaggttacc ccaagcamaa ccttctgttt aaaaagttta acatattata ccgaacgtag 8700
 gtttgggggg gctgtctcgc tctgcacag gcaacattta tataagggto tgcctgcgcg 8760
 gtcacattga atctttttt ttcttctct ctctattto attcttgat taaacacaca 8820
 tcaacacttg tcaagcagcc cgtctgcct ctacacgttg accgtgcac ctacgacgtt 8880
 tctgctggc tcaacacaca tcccgaggt gcgcacatta tccgaacta ccgaggtcgg 8940
 gatgtacgc acgtcttcatt gtttatgac tccagaaagc ccgtgcacaa actcagacga 9000
 atgcacatca tggaaacctc ctctccctg actccacac ctcccaagcc aaactccgac 9060
 gaacctcagc aggtattccg aaagctgcga gaccagctca tigtgcagc catgttcgat 9120
 gctctccaca tctgttaagc ttacaagacc ctctgcacgc tggagctggg tctcctgccc 9180
 gtctgttga tgaccacagtg ccaactggtac ctgggttgat ctatcgtcct ccgcatcacc 9240
 tttaacacga tgggttggt ctgcacagac atttgcacac accagctgtt caaggacga 9300
 tccatcaaca atgcacatgg cctgtctctc ggaacgtgc ttaaggctt ttctgtcact 9360
 tgggtgaagc accgacaaa cgtctcacc tccgcacaa accgtcaggg tcaagctcc 9420
 gacatgaca accgtcctct cctggcgtgg tcaagaggag accgtcagag agctggccc 9480
 tttctcga ggtatgaca gtaacacag tatctctct tttctctg tgcctctg 9540
 cgtactcat ggtgttcca gtcactcat actgcacagc gtcacagga tccagcaat 9600
 cgtactatc gaagacgta ccgaagagg tccgtcgttc tggcactcca ctgggtctc 9660
 aagccttgt tctactatt ctacatgccc tcttctcca ccgaactcat agtgttctt 9720
 gctccagagc tcttgggtg cttcggaatt gccatcgtt tcttatgaa ccaactccct 9780
 ctggagaaga tccagactc cgttggggt ggtcatggt tctgtcgtg accagttcac 9840
 gacacatga accgttcagc aggcctcgtc acgaactgt ttttggtg cctcaactac 9900
 cagatcgac atcactctg gctactctt cccagacaca accctacccc tgcctccatc 9960
 aaagtgaagc agctgtgcaa gaagcacaac ctgcctacc gatctccctc catgtcgaa 10020
 gggtcggca tttctatct ctaactggc accctgcgc gaattggtc caaggagac 10080
 aagcctaa cggccgctt gatgattga aacacacaca tgggttatat ctaggtaga 10140
 gttagtga cagttatata ttaactcagc tatgcacagc gtaactcat tcatgtcac 10200
 gaggacacag tgaatgaa taatatgaa ttgaacac ttgcactct ctgcactcc 10260
 ttactatat ctacttat tttatatca aaacactc tcttccag catcgactc 10320
 ggaacactca tgagcaataa cactcgtgat ctgtacata gagggttt ttgactcct 10380
 gctgttgcc accctgtct tctgtttta accagctgtg aaagactcac tatgtccg 10440
 gttatctcg accaatagcc aaagtctgga gttctgaga gaaaaagca agatacgtat 10500
 gtaacaaagc gacgatgtt acataatc ccgaagcatg tatcatagc agttatgct 10560
 tcatgatgg cactgtgccc tggatgact ttatacgtt gactaatat ttgtctcag 10620
 acataaat accgtcagc actacactt ggcactctgt aggtaccccc ccgcaagac 10680
 gatctcagc tgtctatgt cggattggc tagctccctc gctgtcaat tggctccat 10740
 ctacttctt ctgttgggt acaacacga tctctgctat ggtcgtttt cgtgccttat 10800
 ctactctcc agtatacca actctaatg acctgatgt attgggtcta cautttaata 10860
 tccagagata ggaatgac agttcactaa aaagccacac tcaatcagc ttctctctt 10920
 ctgttaata gtacaaagt gtttagcga atctggagc ttatgtggcc ctaaaaaat 10980
 caaaaaagc aaaaaagaa aaacgaana cccagtttt gagaacggg aggtacaga 11040
 ggtatgata tatatatata tatatatata cccacggatc ccagacacgg cctttgatto 11100
 ttccctaca ccaacactt tcaacacct aattcaaac catgtctccc aacacactgt 11160
 tgcacgcat gcagctgcc gctcccgag acccaactt catcgactc gacacgctc 11220
 ggaactgac ttacgatgc gccattgctc tttccgctg aatcgtgga gctatggagc 11280
 cactcggcat tccacccgga gacagcttg ccttgcaggt ccgaagctc gccaggggt 11340

```

tgatttctta cctggcctgt cttogaacgg gagctgtcta cctgcctctc aacactgcct 11400
acacccctggc cgaactcgac taattctatcg gcatgcccga accgcgtctg gtgctcgttg 11460
ctcccgacgc tctaggttggc gtcggagacaa ttgccaaagg acacggctgct atcgtcgaaa 11520
ccctcgacgc cgaagggaaga ggtcacttgc ttgacacttgc tagagatgag cctgcgcact 11580
ttgtcgtatg ttgcgcatct gcgcgcgato ttgctgctat tctctaacct tccggtacaa 11640
ccggacgctc gaagggtgcc atgcttactc atggcaatct gctctccaac gctctaacct 11700
tgacagacta ttggagagtt accgcagacg atcgactcat ccatgccttg ccaatcttct 11760
acactcattg tctgttctgt gctacgaacg tccactgct tgcaggagcc tcatgttct 11820
tgctctccaa gttcctatgc gacgaggtcg ttctctcat gccacaggcc accatgttta 11880
tgggcgtgcc cacttcttac gttcgtatgc tgcagatct tgcactcgag aagggtgctg 11940
tgccagcat cagaactgtc attctggat cagctcctct gcttgcgaa acccaacgg 12000
agttctatgc tctactggt cagccattc tgcagcgata cggcatgag gaaaccaaca 12060
tgaatacttc caaccctac gaggccaagg ctattgcggg aacgctggt ttctctctgc 12120
ccgacgtcac tctgcgagtc accgatcccg ccacggctct cgttcttcca cctgaagaga 12180
ctggcatgat cgaatcaag ggaaccaacg tctcaagggt ctattggaga atgcgcgaaa 12240
agacgcctgc cgaatttacc gcgaacgggt tctttatctc tggagatctc ggcgaagatg 12300
accgaagagg ttactctcac atttgaggac gaggcaagga cctggctatt tccggtgctg 12360
acaactctc tcccaagag gtcgaaggcg agatcgacca gatcaggggt gtcgtcagat 12420
ctcgtctcat tggcttctct cactccgatt tgggaagag tgcacccgt gttctgtctg 12480
gcaaacctgg tgccttctc gacgaaaaga ccatcgtct tgcctcgag gacgctctg 12540
ccgatacaa gcaaccaag cggattatct ttgcgacga tctgcctoga aacactatgg 12600
gaaggttca gaagaacct cttgcacagc aatcgccga tctctacac agacgataag 12660
cgccgcgatg agagataaaa tatataata cattgagata ttaaatggcg tagattagag 12720
agcctctac tctcgtgaga gaagccagac cgaactctc aaggcgatta cactctcat 12780
atccacgac accagctggg caaaggttct atatacctt tccggaatac cgtagtctc 12840
gatgtatca atggggcgag ccagatttct agcactctcg gtgtctcggg gtgaataggc 12900
gttcttgccc tccatcaagt cgtacatgct ctctatttc cgtctcaagt aaaaacgaag 12960
cagatgaag atgaacttga actgaaggaa tttaaatgat gtcgacgag taggatgtcc 13020
tgcaagggtc ttttttggg gtttgagaga aggggtgctt gtagatggaa gccgtagaa 13080
ccggcgtctg tctgttga gatgaagcc ggtagaacgg ggtcgttgg ggggtattgg 13140
ggccgctggg ctccaaagag cggtaggcct ttctgtggg ttacttaatt gcgcatttg 13200
ggtctgccc gactgtccca ttgtcagaa ttactcggg taggagactt atcagccaat 13260

```

```

cacagccggc gatccactg taggttgggt tgggtgggag caccctcca cagagttagag 13320
tcaaacagca gcgcaacct gatgttggg ggtgtcgtg ttaaaagaaa aaaaagaagc 13380
ttgggtkata tcccgctct atttagagt tgcgggtag acgcgcagcg agggcaatgg 13440
cgctatggaa ccttgcggat atccatcgc cgcggcgagc tgcgtccgaa ccagctccag 13500
cagcgttttt tccgggccat tgagccgact gcgacccgc caactgtct tggccacgc 13560
actcattgta tgttgtgtt gggagggcac tttttaagta gcaaaagca cctagctcgc 13620
agcaaggtgt ccgaacaaa gaagcggctg cagtgggtca aacggggcgg aacggcggg 13680
aaaaagccac gggggcagca attgaggcac gccctcgat ttgagacgag tcaaggcccc 13740
atccgcgcg caatggctc gccacggccc ggtcttttc accactcag gtiaccocaa 13800
gccaaacctt tgtgttaaaa agcttaacct attatacga cgttagttt gggcgggctt 13860
gctcgtctg tccaaaggca catttatata aggtctgca tgcggcctc aattgaatct 13920
ttttctctt tctctctct atattcttc ttgaattaaa cactctcaa ccatggagtc 13980
tggaaccaa cctgtggca ttccctccc tgagtaatat gactcttlla tggactggaa 14040
gactccctg gccatcctg ccactacac tctgcctc gctctcttca accccaagt 14100
tgccaaggtc tcccgagtgg ttgccaaagc ggttaacgca aagcctgcgg agcgaacca 14160
gtccgagat gccatgact ccttctctt tctgcacaa ctcattctgt gttctactc 14220
tggtacacac ttctactaca tgttctctg tatgtcagg aacttccgaa cccacacact 14280
gcacgaagcc tactgcgaca cggatcagc cctctggaa acgcacattg gctactggg 14340
ttactcttct tactgtcca agttctacga ggtcattgac accatctca tcatctgaa 14400
gggacagcg tctcgtcgc ttacagccca ccacatgct ggaacatga ttacatgtg 14460
gtctgcatc aactaccaag ccactcccat ttgatcttt gttgtcttca actcctcat 14520
tcacaccctc atgtactgtt actatgctt cactctctc ggttccctc ctctggcaa 14580
aaagtactg acttcatg agattaacta gttctggtc ggtacacca ttgcctgtc 14640
ctactcttct gttctggct gcattccgaa acccgtgct cagatggctg tctggatcaa 14700
cgtcgctac ctgttctct tgacctatct gttctggac ttgcgaagc gaactactc 14760
caagcgtat gccattgcgc ctcaaaaaa ggtcagtaa cgggcgcgaa ggtgtgagtg 14820
ggaagttagt gcccggttct gttgcacaa ttgcaatcc aagatggatg gattcaaac 14880
agggatatag cgaactacgt ggtgtgcga gatatagca accgatatt atgttgaca 14940
cttggaatg tactatacca gcactgcca agtacaatc taaacactc gtactactc 15000
atactctac ccgggcacg gtttacttg agtcagtggt ctagtctct tactctaca 15060
gttgcgaata ctgcgtatca tagcttttga tctatatct atcattcat gttagtgc 15119

```

<210> 134

<211> 14623

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pZKL4-398F2

<220>

<221> misc_feature

<222> (1048)..(1048)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (1053)..(1056)

<223> n is a, c, g, or t

<400> 134

taaaagggct tgaacacagaa tgagccagac agcaaggaca aggtggccaa cagcaaggag	60
tccaaaaagc cctctattga cgaatccac gatgtattg ctcatgaggt ttccagagctc	120
gatgtggga agaagaagtg atttgtatat aagaataaa tgagatatag taaagagtg	180
caagagaatg gcaaggttgt caaatcttat attacttga gtccactggt cctcgttgc	240
atgaatgaag ttaccgttgg catagctgat ttaatatata actgtccaac taactctcac	300
ctagatataa cccatgtgtg tgtttccaat catcaatgag gcgcgcctta cttaagcaac	360
gggtttgata acagcggggg gggtgcaccg gttgttggg ttgcggaaga acagaacacc	420
cttaccagca ccttcggcac cagcgtggg ctcaaccac tggcacatag gcgcactgag	480
gtacatggcg cggatgaagc caagaggaac atcctggaca tcagcccggt agtgcctgac	540
catgatggg ttaatggcct oggtggcctc gtccggttg tagaaggga tgctgtgac	600
gtagtgtgt aggacatgag tctcgtgat gccgtggaga aggtggcggc cgaatgaagc	660
catctcaagg tcaatgtag cagcggcacc acggaacga ttccactgt cgttgggtga	720
gtggggaag gtatgggtcgt tgtgtggag gaaggtgatg gcaacgagcc agtggttaac	780
ccagaggtag ggaacaaagt accgatggc catgtgttag aaacgaact tctgaacgag	840
gaagtacaga cagtgggcca tcagacgat accaatatg ctgaggacga tgagcttagc	900
gtcacgttc tegtacagag ggctcgggg atcgaagtg ttaacaccac cgcgagggc	960
gttatgcttg ccttgcgcg gaccctcacg ctggcgctcg tggtagtgt ggcggtaac	1020
attggtgat aggtagtgg gccagccnac gannnctca tgaagtagg cagctcgtg	1080
ggtcatctt ccagacgag tagctgtgt ctgcggggt cggggaacga agaccatgtc	1140
acgtccatg ttgcagtggt ccttgggtg otttcggtg gagatttgc agotgaagta	1200
ggggcaaggt agggaaaggt gaagaaccca gccagtatg tctgtgatga tgcgagatc	1260
ggagaagca cagtacccg actcatggg aataacccag agaccatgc cgaagaagcc	1320
ctgaagaacg gtgtacacg ccacacagac agcgcgggg gggttgagag ggaatatctc	1380
gggggtcaca aagttgtacc agatgctgaa agtggtagtc aggaggacaa tgtcggggag	1440
gatataacc tatccottga gagcggagcg ctggaagcg tgettaggga tggcattgta	1500
gatgtccttg atgttaagt cgggaacctc gaactggttg ccgtaggtgt cgaagatgac	1560
accatactcg gacttgggtc tggcgatata aacctcgac atggaacaga gcgatggga	1620
agagccagag tggcggggag agtctgaag agagaagcg gcagactcag aatcgtcac	1680
agtatttgag tgaaggtgc gtctaacgc aggtttctgc ttgggcagag ccgaagtga	1740
cgcatgggc aggaactgtg ttatgacatt gtccgggagt catcaatttg ttgaaaggt	1800
tgtcactgt tagtatgag tcaattggc tctgtgggt cgaatgacat tctcatctgt	1860
ttctgttgg tcatgttcc atcaactct atggtactca caattcgtcc gatcgcag	1920
aactcgttaa taacgacttt gatggcatg ttgaatgtg ttaactcaa gaatgaatat	1980
agagaagaga agagaanaaa agattcaatt gacccggcga tgcagacact tatataaatg	2040
ttgcottgga cagaaggagc aagcccgccc aaactaagt tgggtataat atgttaagt	2100
ttttaacaca aagttttggc ttggggtaac ctgatgtgt gcaaaagacc gggcgttggc	2160
gagccattgc gcggcgcaat gcggccgtga ctgcctcaa attcagggc gtgcctcaat	2220
tcgtaccccc gtggtttttt ccggccgttt ccgcccgtt tgcacacatg cagccgttc	2280
tttgggtcgg caaacttgtc gcgagctagg tgccttgtgc taactaaaaa gtggcctccc	2340
aacaccacaa tgacatgagt gcgtgggcca agacacgttg gcggggtcgc agtcggtca	2400
atggcccgga aaaaacgtg ctggagctga ttggacgca gtccgcgcgc gcgtatgat	2460
atccgcaag ttccatagcg ccattgcctt ccgtcgggt ctatcccgca accctaaat	2520
agagcgggaa taaaacccaa gcttcttttt ttctcttaa cagcgcaccc cccaactatc	2580
atgtgtctgc tgtgtttga ctctactctg tggaggggtg ctccacccca acccaacta	2640
cagttggtc cggcgtgtg attgctgat aagtctcta tccggactaa ttctgaccaa	2700
tgggcatgc gcgcaggacc caatgcgcg aattacgtaa ccccaacgaa atgcctaccc	2760
ctctttggag cccagcggcc ccaaatcccc ccaagacgac oggtttaccc ggcttcacat	2820
tccaagcaca agcagcccggt aattctctct cttagcttt tccataacaa gttctctgc	2880
ctccaggaag tccatgggtg gtttgatcat ggttttgggt tagtggtagt gcaatgggtg	2940
tattgtgact ggggatgtag ttgagaataa gtccatacaca agtcagcttt ccttagacat	3000
catataagta taagtatttc aactatttag cactgtaccc agcatctccg tatcgagaaa	3060
cacaaacaca tgcgccattg gacagatcat gcggatacac aggttgtgca gtatcataca	3120
tactogatca gcaaggtcgt ctgacatcca tacaagctga acagagctgc catcaattga	3180
cgtctctat atacacagtt aaattacata tccatagctt aaactctaac agttaacttt	3240
ctggtaaagc tccagccag ccttctggta tgccttgccc tccatcatag gatctgggt	3300
ctggcgtgac agacctcggc cgaacattat gatctcgtt ccggtagaca tgcactctc	3360

aacagtctgg tactgtgtgc cagagagctgc tcccttgctg tcaagaccca ccccgggggt 3420
 cagaataagc cagtccctcag agtcgccttt aggtcggttc tgggcaatga agccaaccca 3480
 aaactcgggg tccgtatcggg caagctcaat ggtctgttg gactactcgc cagtgcacag 3540
 agagcccttg caagacagct cggccacgat gacagacact ctggccagct tctcgttggg 3600
 agagggagct aggaactcct tctactggga gttctcgtag tcaagagcgt cctcctctct 3660
 ctgttcagag acagtttctt cggccacagc tgcacggcca gcaatgatto cggttcgggg 3720
 taccocgtgg cggttgggca tatcggaaca ctccgggact cggtagacac ggtactgggt 3780
 cttagacagt tgcaccaat ctgcgaactt tctgtccctg aacaggaaga aaactgtctt 3840
 aaagcacaat tctttagggg ggaacacagt cccggcgtag gtcaactcgt caatgatgc 3900
 gatctgggtt ttgatctgc acacataagg tccgaactta tgggcaagct caatgagctc 3960
 ctgtgtgttg gtaacatcca gagaagcaca caggttggtt ttcttggttg ccccgagctt 4020
 gagccactga cggccaagg cggactgttg gactttagct cgaactctgt aggagggcat 4080
 ttgtgtgttg aagagagac tgaataaat ttagtctgca gaacttitta tgggaactt 4140
 atctggggca gtgaagtata tgttatggta atagttaaga gttagttaga cttatagata 4200
 gactggaata taccgtctgc ggtccaaatt agaaagaagc tcaatggctc tctggggctc 4260
 gctcttgccg acaaaatgt gatcatgat aaagccagca atgaacttgc agtcatatt 4320
 gtgtcgccgc aacgcgcgcg aaacgcgagc tctcagaccc aagctctcca acgaagatg 4380
 tatcgtcaaa gtgatccaaag caactctata gttggagtcg tactccaaag cggccaatga 4440
 cgaactcaga agatactcgt cgaacttttc cttgggaacc accacagcca gacctcttga 4500
 ctcaactatt gtagccacgc acacagccaa cagtccttgc atagcagaat atgtcttgc 4560
 ggtcacttgc tcccaactt taggctgcaa gtagctgttg cagaagaagt atgtcactc 4620
 attagaactc ggtgtgtgtg atttcaata agtcttgaga tontttgga cagtcatgtt 4680
 gtggggggtt attgattga gttatcgctt acagtctgta caggtatact cgtgcacca 4740
 tttatcctt ttgatctcgc tgactctgaa gcaatgtcgt ttacccaag tgaagaatgt 4800
 cccagacaca caaccacagg tatgtttgag caaaaataaa caactccgat ccgggaactc 4860
 gaaccccggt ctccacggtt ctccagaagt attcttgatg agagcgtatc gatcagagga 4920
 gaggacagc ggtcgtctct taagtcttg acatcagat ccaaggccac cttcgaagga 4980
 tccaaggctt tgaacccgtc atttgcactt cgttaagctg gtagacaggt tgcctggctc 5040
 cctacggctt ccaactgtgt caactcttcc aagctcctgc actatcagga cattgatcaa 5100
 ctccggaaga aacttttgta tgcacttoga tcaactgctg gtttcatgtt gcttagaggg 5160
 aacgtatata cagtaactat agagaataaa cgtatctcat ttattaaagt agatagtga 5220
 ggtagaagt gtanaagtg atanaatagc cgcgcttagc cttgtctgc cttggcaacc 5280
 attcagcga aggtgcacag gtagggagata agaatgcga caactcagag catgggagga 5340
 gatcggtagg caggttgttg cttcttgcaa agctgtctca ctttgatgga ggcagcgtg 5400
 aggttgtctc tgggaagagt aggcacagag tcatgtctga tctgtcagtt gaggcacgc 5460
 aaaaacagct ctgtcagcag cctcctgctga acgtctcaggt tctcgtcaat ctgtccagca 5520
 cagaagccat gaacatccca cagggaatcc tgaactctct ccaagaggtt gatgtctatg 5580
 aagaacaga tggcaattcc gaagccacca agcagctcgg agccaagaaga caaatgagt 5640
 ccggttagaa accagggcat gtagaataag tagaacaagg ccttagagac ccagtggagt 5700
 gccagacaga cggactcctt ctgtactctg cttcagatgt actgatgtct tcatccttg 5760
 agacccgttg ccgtatgaat ggaactgaag caccagatga atcgcagagag ggcacagatg 5820
 aaaaagaagt aatactgttg gtaacttgatc atccgtcag aaacccggcc agctcgtcgc 5880
 acgtcctcct tggacacagc caggagaggg aggtctgctg tctcgggctc gtaacctcgc 5940
 acgttgtctg cggagtcatg agcgttgtgt cggctccttc accaagtgac agaaaagccc 6000
 tgaagcagct ttccgaagag caggccaagt gcatgttgta tggatcggct cttgaacagc 6060
 tggtagtcgc aaatgtcgtg cgaagacat ccaactcgtt gaaagtgaat gccagggacg 6120
 atagacacca ccaggtacca gtcacactgg gtcatacaca gccagggcaag gacacccagt 6180
 ccgagagtcg acaggtcttt gtaagcgtac caactgggag aggaatcgaa catgcctgca 6240
 gcaatgagct cgtctcgcag ctttgcgaaa tctcctgag gttctcggga gtttggttg 6300
 ggaggtgttg ggtcagggg agaggaaggt tcaatgatgg gcaatcgtct ggttttgac 6360
 accggttctt cggagtgaat aacatgaag acgtcggtag catcccgacc tctgtatctc 6420
 tcatatgtt cggcaactcc cggatgggtg ttgagccagg cagaaaagtc gtaagtgaca 6480
 ccgtccacgg tgaagggcag agcgggtcgc ttgaacctgg atcccatgac gccagagagag 6540
 tcttagtagg tagcgggagg aagtcggcca ggtcagaggg gacccggccc ctccttttg 6600
 cctcgttctt ttccaggtct cgaatgtatg gttcggagct agaggttgag gaacagaagc 6660
 agcacccgtc cgaacataca gtagttgaag atccaggaaa acattcgaat accgtcctgc 6720
 cggtagcagc gcaactttcg gtaactccag acgatgtaga atccgaagtt gaactggata 6780
 atttgcagtg aggtaatgag gttctgggcg atagggaagt tcaagttgat ggtcgcgtc 6840
 caatagtaac cgtacatgat ccagtgaatg aagcgttga gccagaacaa gatccagagc 6900
 ccttcgttcc gatacttga gaacagccac atgtcagatg gactccagag atgttgaaag 6960
 aactgcagaa accacaaggg ctgtccatg aggggaaggt agagggagtc gatgtactgc 7020
 accaacttgc tgaagtagaa cagctgagtg gtagttogga acccgtcaatt gttgaaggca 7080
 gttcccaaat cgcgggaag aataccagtg acagagagag cgtaggccat cgcagggag 7140
 gagccggaag agtagatgga catgagaag ttgtatgagg tgaacagctt tttaggggtg 7200
 gtttgcgtc tgaagtcag aggtogaag atgaacttga tggccaagaa agcagaccag 7260

aggtacagaa cctcgcaaga ggaaggcgtcc tgcacagatc gggcgtatgc caacttggga 7320
 agttcctcct ggaacgatga gaaacagctcc ttggcagcct caatggtaac agagctgggt 7380
 tagtttctgt agagagtggt tgttctcaga gactttcgga ttgttcaatt acacaaacag 7440
 cgtcctctcg acactgatct tgtcgtggat actcaaggct cggaaattctg tcatgtgtag 7500
 tttagatttc gaactcgtgg gaaagaaga gaaaaaagag actggcaacc gattgggaga 7560
 gcaactgttt atatatacc tagcaagcc ccccgcttgt aagatgttgg tcaatgtaaa 7620
 ccagrattaa gtttggcaag tgcaggagaa caaagggtgt ggtatcgagc aatggaaatg 7680
 tgcggaaggc aaaaaaatga ggcacagccg tatgtcgggc gatataatca gggggcagtt 7740
 gaagtacact aacatgacat gtgtccacag acctcaact tggcctgatg agccaaatcc 7800
 atacgcgctt tgcagcctct aaaggctata caaagtcaac caacctgtct cgaacctcagc 7860
 gccctcaact ttgttaaga caaactgtac acgtgttcc agcgtttctt gcutgcaact 7920
 ggtgggacat ttggigcaac ctaaagtctt cggacactct gtgtgttcca gatcagcgca 7980
 gcagttcga ggtagtttt aggcacttag atgatgaat ggtgtcagtc gctgatacac 8040
 gagttctaat ggcagtattt gttttattt gtgcattga gcccgittat cctgttatct 8100
 tctacccccc atcccatccc ttgttgggt caacctcaac catltattgt tgggtgcagc 8160
 caaacgagcg tggagagctt ggttggcca tataaagag ccccccccta gtggcaatgg 8220
 cagaagatca gctgtgatt gttgaattg tcatctagga ggcctggcgg tcttctcagg 8280
 ggcatttaa atgttctct atagtagatc tgcataact gtttaaacag agtgtgaag 8340
 actcaatag gtccgggctt atctcgaaca atagcaagc totgagttt ctgagagaaa 8400
 aaggcaagat acgtatgtaa caaagcagcg atgtgtacaa taataccgga ggcattatc 8460
 atagagagt agtgttcca tgaatgaact ggtgcctgt atgactttat aagcctgact 8520
 acatatttgt cctcagacat acaatcaag tcaagcaact aacctggac atgtaggt 8580
 acccccccga caagacgac tcaagctgtc gtatgtcga ttggcgtagc tccctcgctc 8640
 gtcactggc tcccatctac ttctttctgc ttgctacac caagcagtc tcatgtgct 8700
 cgtttctgt ccttatctat cctcccgcta taccacact taatgacat gatgtgattg 8760
 ggttaccact ttcatctag agataaggc tagcacagtt gcatcaaac cccaactcta 8820
 atcagctct tccctctgt taattagac aaagtgtatt agcaaatct ggaagcttag 8880
 ttggccctaa aaaaatcaaa aaagcaaaa acgcaaaaac gaaaaccac agttttgaga 8940
 acagggaggt aacgaaggt cgtatatata tatatatata tatataacca cggatccaga 9000
 gaacggctt tgattctcc ctacacccaa ccatctcacc caacctaat caaacactg 9060
 gagtctgac ccatgcctgc tggatccccc tccctgagc actatgact ctttatgac 9120
 tgaagactc ccttggcact cgtctcacc tcaactgct cgttggtct ctcaacccc 9180
 aagtttgca aggtctcccg agtgttgcg aatcggtca cgcgaagcg tgcagagca 9240

accagtcag gagctgaat gaactcctc gtctttgtgc acacactcat tctgtgtgc 9300
 tactctggca tcaactctca ctactgttt cctgtatgg tcaagaactt ccgaacccac 9360
 acactgcag aagcctactg cgaacaggtc cagtcctct cgaacacgc acttggctac 9420
 tggggttacc tcttcaact gtccaagtc taccaggtca tgcacacat catctatc 9480
 ctgaaggagc gacgtcttc gctgtctag acctacacc atgtgtgag catgatacc 9540
 atgtgtctg gatacaata caagccact ccaatttga tctttgtgt cttaactcc 9600
 tcatctacca ccatctgta ctgttaact gcttcaact ctatcgagtt ccatctct 9660
 ggcaaaagt acctgaact gatcagatt actcatttc tggctgtat caacttgc 9720
 gtgtctacc tcttcttccc tggctgcat cgaacacccc gtgtcagat gctgtctgg 9780
 atcaacgtcg gctactgtt tccctgacc tctcgttct tggactttgc caagcaacc 9840
 tactcaagc gatctgcaat tgcctctcag aaaaaggctc agtaagcggc ccatgagaa 9900
 gataaatata taatacact gagatattaa atgcgctaga tttagagagcc tcatactgt 9960
 cggagagaag ccaagacag tactcaagg cgaattacac atccatctcc acagacaaa 10020
 gctggggaaa gttctatat acactttccg gaataccgta gtttcagatg ttatcaatgg 10080
 gggcagccag gattcaagg acttggtgt ctcgggtga aatggcgttc ttgctcca 10140
 tcaagtctga ccatgttctc atttgcctgt caaagtataa cagaagcaga tgaagaatga 10200
 acttgaagt aggaattta aatgtaacga aactgaatt tgaacagata ttgttctgc 10260
 ggtgagctc cagcttttgt tccctttagt gaggttaat tttagccttg gctaatcat 10320
 ggtcatagct gttctctgt tgaatttgt atccgtcac aagcttccc acaagctacg 10380
 taggatacag gtcttagga agctcagcca accacgaga ctgttgaac tggatgtcgg 10440
 taacagcacc tggatgctg aatgtctct gaataaana atatttctcc ttgttaggt 10500
 gatcataagc tatgtatcc gtgattgag tgaatagaa gtctctccg aagactgag 10560
 ccaacgtcat gttcgggaaa taacgacac tctctccaca tgaataatca gttcgtagag 10620
 gagtactg cgcattgaca cagttagcga tgtttgcaat ccgagaaaac ttgctcgtaa 10680
 agttgtacag ctccgtggag gcttgaact gacttttga aagtgcctg ggtgctcgc 10740
 cgaagagga ggcataagg taacgaacca cttgcocgag cgtgaagtc atgatgcaa 10800
 tagtgaatg catctatcac cgtactgc agtatatta tagggctcat cgttcatgt 10860
 atagatctgt ccaattatga caaccccatg tctattat gtgtaaggt gcaagcgggt 10920
 gggatcacg tcaagagttg gaggaatca ggtatgtgg gtttaagact gctcagatc 10980
 caaatttcaa ctctccatg tcaagacac cttggttca gattattac agatctagc 11040
 ctgttccag gtgaggggac ctatctgga tgggcaagac gatctcacc ttacagagga 11100
 cgtctgtcag agggaaaggt catgtggcgc gccagctgca taatgaatc ggcacacgc 11160

```

cgggagagag cggtttggt atggggggt ctccggctt ctcgtcaat gactgggtg 11220
gtcgggtgt tgggtgggg cagcgggtat cagctcaatc aaaggcggt atacgggtat 11280
ccacgaatc aggggataac gacggaaaga acatgtgagc aaagggccag caaaagacca 11340
ggaacggtaa aaaggcggcg ttgtggcgt ttctcatag gtcggccccc cctgacgagc 11400
atcacaaaa tgcacgtcca agtcagaggt gggaacccc gacagacta taagatacc 11460
aggcgttcc ccttgaagc tccctgtgc gctctcgtg tccgacccg cggttacg 11520
gatacctgt cgccttttc cctcgggaa gctggcggt ttctacagc tccgctgtg 11580
ggtatctag ttgggtgag gtggttggc ccaagtgagg ctgtgtgca gaacccccc 11640
ttcagcccg cgcctggcc ttatccgta actatcgtt tgaqtccaac cggtaagac 11700
acgacttat gccactggca gacggcact gtaacaggat tagcagagcg aggtatgag 11760
gcgggtcac aggtttctg aagtggtag ctaactacg ctacactaga agaacagtat 11820
ttggtatcg cgtctgtg agcgggtha ccttcggaaa aagagttgt agtcttgat 11880
cggcacaaca aaccacggt ggtagcggg gtittttgt ttgaagcag cagattacg 11940
gcagaaaaa aggtatcaca gaagatcct ttatctttc taagggtgt gacgtcagt 12000
ggaacgaaa ctacgttaa ggttttgg tctgagatt atcaaaaa atcttcact 12060
agatccttt aattaaaa tgaagttta atcaatata agtatatat gattaaact 12120
ggtctgacg ttacaaatg ttaactcgt aggcacat ctacagatc tgtctattc 12180
gttccactc agttgctga ctcccctgg tgtagatac taagatacgg gagggttac 12240
ctctcggcc cagtgtgca atgatacgg gaaacacac ctacacggt ccaagtttat 12300
cagcaataa cagccagcc ggaagggccc agccagagag tggctcga actttatcc 12360
ctccatcaa gtctataat ttgtcgggg aagctagat aagtgttgc cagttata 12420
gtttcgcaa cgttgttgc attgtacgg gctcgttgt gtcacgtcg tggttgtg 12480
tggctcatt cagtcgggt tcccaagat aaaggcggt taactgtcc ccaatgtgt 12540
gcaaaaaac ggttagctc ttggctccc cgtcgttgt cagaagtag ttggccgag 12600
tgttatcct catgttatg cgcgcctgc ataatctct tactgtcgt ccatcgtaa 12660
gatctttc ttgtactgt gactatcaa ccaagtcatt ctgagatag tgtatcgcc 12720
gacggagtg ctcttgcgc gctcaatac ggaataaac cgcgcacat agcagaact 12780
taaaagtgt catcattga aaactgtct cgggggaaa actctcaag ctctacccc 12840
tgttagatc cagttgatg taaccaact gtcaaccaa ctgattcca gacttttta 12900
cttcacacg cgtttctgg tgaacaaaa caggaagga aatgcgcga aaaaaggaa 12960
taaggcgac cgggaatg tgaatactc tactcttct ttctaatat tattgaaga 13020
ttatcaggg ttattgtct atgacgggt acatattga atgtattag aaaaataac 13080
aaatagggt tccggcaca ttcccggaa aagtgcacc tgaatgggt tgaataacg 13140

cacagatgg taaggagaaa ataccgcat aggaatatg aagcgttaat attttgtaa 13200
aattcgggt aatttttgt taatcagct cttttttta ccaataggcc gaactcgga 13260
aaatccctta taatcaaaa gaataagcc agatagggt gagtgttgt ccaatttga 13320
acaagatcc actattaaag aactggact ccaactcaa agggcgaaa accgtctac 13380
agggcgatg cccactacgt gaaccatcc cctaatcaag ttitttggg tcaaggtgc 13440
gtaaacact aaatcggaac cctaaaggga gcccccgat tagagcttga cgggaaagc 13500
cggcgaact ggcgaagag gaagggaaga aagcgaagg agcggcgct agggcgctg 13560
caagtgtag cgtccgctg cgcgtaacca ccacacccc cgcgttaat gcgcgcctc 13620
agggcgctc ctttcgccat tcaaggtcgc caactgtgg gaaggcgat cgtgcgggc 13680
ctcttcgct ttacccgac tggcgaaagg gggatgtgt gcaaggcgt taagctgggt 13740
aacgcaggg ttctccagt caccagctg taacacgag gccagtgat tgtaatcga 13800
ctcactatg ggcgaattg gcccgacgt gcctgccaa gcttatatta gttgtggga 13860
aacagttag gacgggtcc tgtataaca gttcagtg cactctgta ggagtgttc 13920
cgtgtgatt tgggcagtc taagaactt cgcgaacca tgttaagtc acaatagcc 13980
atgtagtag tgtaagccag atgttttgg agcagtcga ttcgagaca cagattgaa 14040
gtgcctgat ggcagcctc gtttttct cttcacaga cagacacttt gtcagtgca 14100
gggttagacg tttctgttc aatgtacat tctccatgg cgtgtgtgta attgaccaa 14160
tctggcagtc caacgaagat atcgttccg tgggtgagc tgaagttcg tgcgtgtcg 14220
tttaactcaa ggcctgtgc attaaacac agctgttga tgaagttgc aaactctgag 14280
ttctactc tgggtggcc aaagttagc atgttaggt catgtccaa taatttgaag 14340
tgagttgca aaagaagag ttgacagcg cccagcaggt gaccagtaac atacatttt 14400
tagtcagtc ggttgtgag gaactttta aactgaggag cagcattgac catagtttc 14460
tgaaggcct tggcgaaccc atcagtagc atcagcctt tgcactcact agtttgatg 14520
ggaataagac ggggtcttc aaccaacga gcccgctct tgaagtgct aagaccttg 14580
ttctcactt ccaagctcg tggactgac catctcgtt aat 14623

<210> 135
<211> 14619
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pZF2-85m68F

<400> 135
cgatggaagc cgttagaac cggctgcttg tctttgaga tggagccgg tagaacggg 60
ctgcttggg ggatttggg ccgttggct ccaaaaggg taggcattt cgttggggt 120

```


aogtaattgc ggcatttggg tccgcgcgcg atgtcccaatt ggtccgaatt agtccggata 180
 ggagacttat cagccaatca cagcgcgcga tccactgtga ggttgggttg ggtgggagca 240
 cccctccaca gagttagtgc aaacgcgcgc agcaaatga tagttggggg tgcgtgtt
 aaaggaazaa aaagaagctt ggttatatt cccgtcttat ttagaggttg cgggtagac 360
 gcgcgcggag gcaatggcg ctatggaacc ttgcgatat ccatagccgc cggcggaactg 420
 ogtccgaacc agtccacga ggtttttt cgggcattg agcgcactgc gaaccgcaca 480
 aogtgccttg gccacgcac tcatgcatg ttggtgttg gaggccact tttaaatga 540
 aaagggcacc tagctgcag caaggtgtcc gaacaaaga agcggtgcga gtggtgcaaa 600
 cggcgcgaaa ccgcgcgaaa aaagccacgc cgcgcgaat tgaagccgc ctcgaattt 660
 gagaagctc acggcccat tgcgcgcgc aatgctgc caacgcgcg tctttgcac 720
 cactcaagt taccacagc caacacttg tttaaaaa cttacatat tatccgac 780
 gtaggtttg ggggttgc tccgtctgc caaggaaca tttataga ggtctgact 840
 gcggctcaa ttgaatttt ttctctctc tctctctat attctctt gasttaaca 900
 cactcaacc atgttcagc gaacgcctc gctctcacc gtggaagtg tccctaca 960
 cgtttgcg ttgctcaac accctccgc aggtgcgcg attatagaa actaccgag 1020
 tgggtagct accgagctc tcatgttat gactccgag aacgcgtgt ccaactcag 1080
 aogaatgcc atcatggaac ctctctctc cctgactca aactccca agcaaatc 1140
 cgaacacct caggagatt tccgaagct cgcagcagc ctcatgttg caggcatgt 1200
 cgtgctctc ccatgtgt acgtttaaa gacctctgc actctggac ttggtctct 1260
 tgcgtgttg ttgatgacc agtgcactg gtactgtt ggtctatag tccctgcat 1320
 tccatttcaa cagatggat ggtctgcga cgaacttgc cctccagca tgttcagga 1380
 ccatctatc aaacatgcca ttggtctgt cttcggaac gtgctcag gctttctgt 1440
 cacttggtg aaagacgac acaacgctc tccctccgc acaactgc aggttcacga 1500
 tcccgacatc gacaaactgc ctctcttgc gtgctcag gaggatgag agcgagctg 1560
 cccgtttct cgaagatga tcaactcga acgtattac ttcttttca tctgtctct 1620
 tctgccttc atctgtgtc ttctctcat tcatctgac accgtctca aggtctcag 1680
 caatcagta tctgcagac agtgcagaa ggaactgct ggtctggaac tccctgggg 1740
 tctcaaggcc ttgtctact attctcaat cccctgttt cctccagga tcatgtgtt 1800
 cttgtctcc gactgcttg gtggtctgg aatgcatc gttgtctca tgaacacta 1860
 cctctggag aagattcag actcgtgtg ggaatgcat ggtctctg ctggacgat 1920
 tccagcacc atgaacttc agcagagct cgtccagac tggttttg gtggtctaa 1980
 ctaccagtc gaacatcac tgtgctcac tctccaga caaacctca cgtctctc 2040
 ctcaaatgt gacagctgt gaaagagca caactccgc taccgctct ctcctgct 2100
 ogaaaggtgc ggcattotta tctctactc ggcacactc gctcgaatg ttgcaagc 2160
 agacaagcgc taagggcgc cattgatgt tgaacaca cactgggtt atactaggt 2220
 gagatagt tggacagta tatataaat cagctatgc aacggaact tcatctagt 2280
 caacgggaa cagtgactg caagtatat gaatttgac caacttgca tctcttga 2340
 ctcttact atactact tatttttat atacaatac ctctcttca ccagactga 2400
 gctcgaaac ctcatgaga ataatctgt ggtctctgc aatagaggg ttttggact 2460
 ccttctgtt gccacactg tcttctgtt ttaacatag tggtaatgc tctgtgtg 2520
 tgtgtgtg tgtgtttg cgtcatgtt tgccttgc agctacacc acaatttg 2580
 aagttatta gctttctat tttttgtt gaaagotta acacacttg tctggaggg 2640
 gatgggata tggaggggc tggaggggt cggagaggg ttttggagc gctggcgt 2700
 ggcgcagct cgcgaagc acatagacc ctttggcag cgaatgtg caactttca 2760
 gctagttaac gcttacta cgtctctca tctgtcatg tttgcgctt tttcccttg 2820
 ccttgcag ccacacagta cagtcaact taccgtgag gttttgggg gctcttagt 2880
 gggagctaaa agcgcctag cgtacacta gtcggattg atgagtgcc atgagccta 2940
 ggtgagcct gacagagc agacagct agccgtgac agagatggg tggctcgt 3000
 tgcacacgc gtaacatgt ttggccaaa gtttgtcag cctgtgttc gaactaat 3060
 ccaattttg tcaacttga ccccaattg tgaagccta accctgcgc atcaggaat 3120
 caattaaag tgcatttgc tgcctgttt agtttggct ctcgcgttt cgggtccac 3180
 ttgcacaaac acaaaagc attatata aggtctgtt ctcctccca accacacta 3240
 ctttttgc cgtcttct tctaacaca aaagtcaag acacaaaca ccacccaac 3300
 cccctaac acacacata tctacagaa tggcatgga tctctcctt actacagag 3360
 agctgctga ggaacgac ctggttcca tgcagcgt tctctacat ctggaagtc 3420
 ttgcaaggt ccatccgga tccacttga tctcgttc tgggtgttc gatgtctc 3480
 ctctgttcta ctcatgac cttactgca agccagaga ctggaagct cttcaagct 3540
 tctgtgagg caagacgac cgaactcca agacattgt ctacactac gatctcct 3600
 ttgcaagga cgtcaagga actatgagc aggtcatga aggtcgaac tggatgca 3660
 cactggatt ctggtcga accgttgca tcatgtgt caccgcttt tgcagtgga 3720
 actgggtac taccgaatg gtgtgtgg gttcttgac tggattcat cactgcga 3780
 tgggtgtc cttcagac gatgtctc atgggtcat cagcaaaaag cctgggtca 3840
 agctctct tgcacagc atgacgca ttggatgta cagatggat tgggtcagt 3900
 ctacatcat ggaactaac acatcaaca atcagatgg tctcagctg gatgcagct 3960
 cgcagacac attcttgt tcccaact accctgtgc caaacctgt cgaagtggt 4020

ttacacgatt ccaggccgtg tacatgtacc tegtgttgg agcctacgga gtttcgtg 4080
 tgtacaaccc tctctacatc ttccgaatgc agcaacaaga caccattccc gagtctgtca 4140
 ccgcctatgc agagaaacgg ttctgtcgac ggtacogaac ccttgctctc gttatcgag 4200
 cttttctcat ctttcgaacc gctctcttgc cctgggtatct cactggaacc tccctgtcca 4260
 tcaccattcc tctgggtccc actgtacacg gtgccttctc cactctcttt tctatcttgt 4320
 ctacacattt cgtgggtcgc gagggaatcc ccgaacaaga ctgcaaggtc aagagctcgc 4380
 agaagagcgt tgaagccgat cagctcgact ggtacagaga tcaggtggag aactcttcca 4440
 ctaacgttgg acccattgoc atgtctctta ctgaggtctc caacttccag ctacagactc 4500
 aactcttccc tcaatgtctc tcttggaact atccttctgc ccaagcaagt ctccgaagct 4560
 gttgcgaacg acacggaggt cgttactgtc tctacccctac cacttggggc aacatcattt 4620
 ccccccctca gtacatgcac aagctcgttg tggctcactg tgtcaaggac gctcaggt 4680
 cctaagccgc ccctagagaa gataatata taatatactt gatatata atgcgtctga 4740
 tttagagacc tcatctctgc ccgagagag ccacagacag tactcaagg ggttaccacc 4800
 atccatctcc acagacacaa gctggggaaa ggtctctatc aacttctcgc gaatacgtc 4860
 gtttcgtatg ttatcactgg gggccacccg gatctcagga acttcgttgt ctgggggtga 4920
 aatggcgttc ttggcttcca tcaatctgta ccaatgtctc atttgcttgt caagtaaaa 4980
 cagaacagaa tgaagatga acttggaagt aaggaatta aatgtaacga aatgaattt 5040
 tgacagata ttgtgtccgc ggtggaagtc aacttttgtc tccctttagt gaggttact 5100
 ttcaagcttg cgttaactat gctcactgct gttctctgtg tgaattgttt atccgttccc 5160
 aagcttccc acacgttacc ggcctccttg cttgttgtat ttctgagac cactcccttt 5220
 ggtatataag tctactctgg ggttaaggtt gccctgttag tctaggttat agtttctctg 5280
 tgaataacgc agagccaggg ggaataaacc gggggtattt ggaactgttt ttctcggga 5340
 aaagcgttga atcaacctg ccggccttgc aactgtccc cgaactgttt ctgcgccaa 5400
 ttcccccctt gcaactcaaa attagcctcc cactagacac cctccataac atgtactgtc 5460
 ggggaagagt taaggggaaa ccttgcaacc atagacagc tgaagacgc gagggaacca 5520
 ctggagccca aagaactggg gtacacacgc tccagagacc agccagagag ccagagagag 5580
 ggcgcccaaa agatccgaaa accaactatg tccaattggg gactgacgc aaacgaacgc 5640
 gcaacttcag taactatgac ccacactctc aaagccaggg ttattccgtt atccacttgg 5700
 ccgttaacct ccgttgttac ctgtatattgt acacgttacc atcaatata ctctcagcta 5760
 caataagaga ggtctgttgt ccggcacttg tgcctgtct atgggtgtct gtcagagggc 5820
 gaaattgctt acagcttaa ctctggcgtc tgcctagtat gaatagcaaa gtccagacca 5880
 ggggtgcat gattacacag gaggtaacag acttcagac tcaagcttta ctggacttg 5940
 ttgcgtatga acagacgggc gatcgtctt cccctgattt gccggcgcgc cagtgcatt 6000

aatgaatcgg ccaagccgcg gggagaggcg gtttgctat tgggcgtctc tccgcttccc 6060
 cgtcactga ctgcctgcgc tegtgtcttc ggtgcggcgc agcgtctaca gctcactcaa 6120
 aggcggtaat acggttacc accaactcag ggaataacgc aggaagaacc atgtgacaa 6180
 aaggccagca aaagccacag aacgttaaaa agcccgctgt gctggcgttt tccatagcc 6240
 tccgcocccc tgaagcaact caaaaaaat gaactcaag tcaaggttgc cgaacccga 6300
 caggactata aagataacag cgtttccccc ctggaagtc cctcgtgcgc tctcgtctc 6360
 cgaacctgcc gttacccgga taactgtccg cctttctccc ttccgggaag gtggcgttt 6420
 ctcatagctc acgtctaggg tatctcagtt ccgttgtagt cgttcgtccc aagctgggt 6480
 gtgtgcagca acccccctgt ccgcocagcc gctgcgcttt atccgttaac tatcgtctg 6540
 agtccacccc ggttaagcac gacttatcgc cactggcagc agccactggt acacagatta 6600
 gcagagcagc gtatgtagcc cgtctacacg agttcttga gtggtgctt aactacgct 6660
 acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc ttccgaaaaa 6720
 gagtgtgtag ctcttgatcc ccgaacaaa ccccgctgc tagcgtggt tttttgttt 6780
 gcaagcagca gattaacgc agaaaaaag gactctcaga agatcctttg atctttcta 6840
 cgggtgttga cgtcactgg aaaaaaact caacttaag gattttgttc atgagattat 6900
 caaaaaagat ctccactag atccttttaa attaaaaag aagttttaaa tcaatctaaa 6960
 gtatatatga gtaaacctgg tctgacagtt accaactgct aatcagtgag gcaactatct 7020
 cagcatctgc tctatttctt tcatctatag ttgcctgact cccctcgtg tagataacta 7080
 cgatacggga gggcttacc tctggcccca gtctgcaat gataccgga gccccagct 7140
 caccgctcc agatttaca gcaataaacc agccagccg aaggcccgag ccgaagaagt 7200
 gtccctgcaac ttatatcgcc tcatccaggt ctattaatg ttgcgggaa gctagagtaa 7260
 gtgctgcgc agttaaatgt ttgcgcacg ttgttgcct tgcacagcg atcgtgtgt 7320
 caoctcgtc gtttggtatg gcttcatca gctccggttc ccaagatca agcagagtta 7380
 catgacccc catgttgtgc aaaaaagcg ttactcctt ccgtcctccg atcgttgtca 7440
 gaagttaagt ggcgcagtg ttataactca tggctatgga agcactgact aattctctta 7500
 ctgctatgca atccgttaaga tgccttctg tgaactgtga gtaactaacc aagtactct 7560
 gagaatagtg talyccgcga ccagtgctc ctggccgcgc gtcaabacgg gataatccg 7620
 ccacacatag cagaacttta aagtgtcca tcatggaaa acgttcttcg ggcgaaaaac 7680
 tctcaaggtat cttaacgctg ttgagatcca gttcagatga acccactcgt gccacaaact 7740
 gatctcagc atcttttact ttaacagcg ttctgggtg agcaaaaaa ggaagggcaaa 7800
 atgcgcgaaa aaagggata agggcgacac ggaatgttg aatactcata ctctctctt 7860
 ttcaatatta ttgaagcatt tacaaggtt attgtctat gaggcgatac atattgaa 7920

gtatttagaa aaataaataa atagggttgc ogogocatt tcccgaataa gtgcacatg 7980
atgoggtgtg aaataacgca cagatgogta aggaanaat accgatcag gaattgtta 8040
gcgttaata: ttgttaaaa ttogcgttaa atttttgtta aatcagtcac ttitttaaco 8100
aataggccga aatgggcaaa atccctataa aatcaaaaga atagacogag atagggttga 8160
gtgtgtgtcc agtttggaa aagagtcac tatataaaga cgtggactcc aagtcaaa 8220
ggcgaataac cgttatcag ggcgatggcc aactacgtga aactatccac taatcaagtt 8280
ttttgggttc gaggtgcgt aaagactaa atcggaaccc taaggggag cccgatatta 8340
gagcttgacg gggaaagcgc ggaacgtgg ggaanaaga agggaaagaa ggaagagag 8400
cggcgcctag ggcgcgtgca agtctagccg tccgcgtgc cgtacacac accccgcgc 8460
cgttaatgc gccgcacag ggcgcgtcca ttgcacatc aggtgcgca actgttggga 8520
agggcgatcg gtgcgggctc ctgcctatt accgcagtc ggaagggg gatgtgtgc 8580
aaggcgatta agttgggtac ogccagggtt ttccacgtca cgaactgtta aaacgcgcgc 8640
cagtgatlg taatcagact cactatagg gaaattgggc cagacgtgc atgogctgat 8700
gacactttg ttggaagag atgcatttly aatcccaac ttgcagtgc cagtgacat 8760
acatccgcg gtttggaaa atgttcagaa acagttgatt gtgttgaat ggggaatgg 8820
gaatggaaa atgactcaag tatcaattcc aaaaacttct ctggcgtgca gtactactg 8880
tcatactac tgcattttct ccagtcaggg cactctatac tgcagacac agtagtaaa 8940
cccagataat ttgcacataa accgaanaa agaccataa atattatata atagtcagcc 9000
gtttgtccag ttcagactgt aatagccgaa aaaaatcca aagtttctat tctaggaata 9060
tatattcaa tatittaat tottaactct attatatta ttctagcga atactttca 9120
gacttgtag acatgtgata ccccaaaatc ggtatggag tgggtgttc agagagcat 9180
atggcatcg tgcctgctg ttacgtatt ctccctgttc cactcttgg cagacatac 9240
cacaanaag ggggttttt ttatattca atgttactt accgaanaa tgcgataag 9300
cagacacgt attcactaat caccagaaga gtcttgggc gtcttaata actcactgc 9360
aggttgaga ctatgaatg attccctgc ccgtatctat ctactaatt gatcttgaa 9420
cgcgaataa cgttttagg actccaaaga atctcaact ttgtcttca taatattct 9480
accatagtt gatgtttac ttgaacagag agaatgttt cactgaccc aagtttctc 9540
gcactcttg gatatttga ccaaggcgtc cactgcctg cagttatcca gtcccaaac 9600
cccacattc atacattccc atgtacgttt acaagttct caattacatc gtgcaatac 9660
aatcacatc tattactca tcatataaa accactatg tcatataaa ctccaaactc 9720
catagaatac atgactcag aacacacgt cactgcgggc gcttactgag ccttgccc 9780
ggcgtgttc tgggcattc ggcgaactg gacagatg cgaagcaga tgcagacc 9840
ttcatgggc agagggttc ggtaggggag gttgtgttc tggcacagt gttccactg 9900

gtaggnaac gcatgaggt tgttcogag cagggtgggc cagagatggt gctogctg 9960
gtagtccag cctccaaaga accagcagc aatgatcct cgtcgaatgt tcatgtctc 10020
atgactctga cccacagaga agccatgtcc gtccagagc gaataccga tcttctccag 10080
aggttagtgc ttcatgaaga cccagatgga aatccgaag cccacgaaga gctgggaac 10140
aagaaaccc agcatcaggt tccagtgga gggcataag aagagtgga accaggtctt 10200
gagatccag tgcagagaga gtccaatgga cttcttctg tactagatc ggtagaactg 10260
gtgtctcgg tcttgaggg atcgaaaggt cagcacagac tggaaacac agatgaatc 10320
caggaataa cagtgacaa ggaatagta ctgttggaac tgatgagct ttccggagat 10380
gggaagact cagtgacat cgtctcggga caggcgagc agggcgaggt tatcaatgtc 10440
gggatctga cctgaacgt tggtagaga atgaggggc ttgtgtctg ccttcacaca 10500
ggtcagggag aagccctgga gtccgttgc aagacacga cccaggaagt tattcagtt 10560
tgggtcttg aaggtctgt ggtggcagat gtcatgagac agccatccc ttgtctgta 10620
gtcataccg agcagagag caccatgaa gtacaggtg tactgacaa gcatgaagaa 10680
ggcagacac ccaagaccca ggtgttcaa gactctgac gactacaga gggagaggc 10740
gtcaaatag ccagtgga taagctcttc tcggagctt cggaaatct cctgagctc 10800
gttgacgga cctcggggag ccagctcga agcctgttg atcttggca ttogcttag 10860
cttgtagag gcttcttag agtgataac catgaaggg tcaatgacat ctctccctg 10920
gtagtctca atgatttca ctcacacag gtggaagtc acccaaggg agacgtctg 10980
cactttccg tcatgacga gggcagagc ctgtcagaa gcttccagc atccatgac 11040
ggcagagag tctagtagg tagcgggag aagtcggca gctcagcgg gacggcgcc 11100
ctgaatctt ttgctccct tctgtcttc gacgatgag gtctgaact agaagttgag 11160
gaacagacac aggaactac caacgtaga gtatgtgaa aaccagcaa acattctcat 11220
tcatcttgt cgttagaag gaatgttcg gtacttcag acgatgaga agcaacgtt 11280
gaactgaatg atctgcata agtaactag ggcctgggc atagggaact tgaactgat 11340
cagtcgggtc caatagtag ctaaatgat ccagtgaat aagccgttg gacgaacaa 11400
gatcaaacg gcttcttgc ggtagtgtg gaacagcaa atgtccatag gactccag 11460
atggtgaag aactgaacc agtcaagag cttgccatg agggcgagat agaagagtc 11520
aatgactcg aggaacttc tgagtagaa cagctgagtg gtatttoga agaatgtt 11580
gtcgaagcc ttctcagat tctcgacat gacacatg gtgtaatg cgttagccat 11640
agagaggaag ggcgccgc agtagatga catgagcag ttgtagtgg tgaacaaaa 11700
cttcaatga gactgaact tgggtccag aggaacag gtgaactta ggttagcga 11760
ggcatggag agtcaagca cctgcagtg cgaagcaca gacagagat gacuatgtc 11820

```

gacottggga agaacctcct ggcacatgga gcagatttgc ttacagaccc coattgttgt 11880
gaattagggt ggtggaatg gttggttga ggaagaatc aaaggccgtg ctoggatcc 11940
gtgggtatat atatatat atatataac gatccttgcg taactccctg ttutcaaac 12000
tgrggttttt cgtttttgcg tttttgcttt ttttgatttt tttaggggca actaagcttc 12060
cagatttgcg taactacctt tctactaatt acaagaaggg aagaagctga tttaggttgg 12120
gottttttat caactgtgcg actccttalc totgataga aagttagaac coactccat 12180
cctgcatttc agagtggta atactggggg gatagataag gcacgaaac gagccataga 12240
agacatgctg ggtgttagca agcagaaaga agtagatggg agccaattga cagcagaggg 12300
agctacgcaa atccgaacta cgaacacgctg agatcgtctt ggcggggggg taactacga 12360
tgtccagggt taagtcttg actgtaattg tatgtctgag gacaaatag tagtcagccg 12420
tataaagctc taccaggcac cagtgccttc atcgacccac taactctcta tgatacctg 12480
ctccggattt atgtacact ggtcgtcttt gtacatacag tatcttgctt tttctctca 12540
gaactccag aattctctct cttagctttt tccatacaaa gtctctctgc ctccagggaag 12600
tccatgggtg gtttagcat ggttttgggt tagtggtagt gcaagtgttg tatttgtact 12660
ggggtgttag ttggaataa gtcatacaca agtcagcttt cttagagcct catataagta 12720
taagtgttc aacgtattag caactgaccc agcatctccg tatcgagaaa caacacacaa 12780
tgcccaattg gacagatcat ggggataccc aggttgtgca gtatctatca taactgata 12840
gacagtggtc ctgacactca taacagctga acaagcgctc catacttgca cgtctctat 12900
atacagctt aaattacata tccatagctc aactctaac agttaactct ctggttaagcc 12960
tccacgacag cttcttgcta tccgttggcc tctcaatag gatctcgtt ctggccctac 13020
agacctcggc agacattat gatctcgtt ccgttagaca tgacactcct aacgtctgg 13080
tactgtgtc cagagcgtc tccctgtgc tcaagaccca ccccggggtg cagaataag 13140
cagtctcag agtcgcctt agtcggttc tggcaatga agcacaacac aaactcgggg 13200
tcggatcggg caagctcaat ggtctgctg gatctatgc cagtgcacag agagcccttg 13260
caagacagct cggccagcat gacagaccc ctggccagct tctcttggg agagggaact 13320
aggaactcct tctactggga gttctcctg tcaagacgct cctctctct ctgttcagag 13380
acagttctct cggcacacag tcgcagacaa gcaatgctc cgttccggg taacccgtg 13440
gcgttggtga tatcggaaca ctggcgactt cgtgacccc ggtactggtg cttgacagt 13500
ttgcaaatc ctgcgaactt tctgtcctg aacaggaaga aacgtgctt aagagaagt 13560
tcttgagggt gtagcacagt gccgcctag gtagagctt caatgatgc gataggggt 13620
ttgatctgc acacataag tcgcactta tcggcaagct caatgagctc cttgttgtg 13680
gtacatcca gagaagcaca caggttgtt ttctgtgtg cccagactt gacactcga 13740
gcggcaagg cggactgtg gactgtagt cagactgtt aggagggcat ttgtgtgtg 13800

aagaggagac tgaataaat ttagtctgca gaactttta tcggaacctt atctggggca 13860
gtgaagtata tgttatggtc atagttaaga gttagtgaac cttatagata gactggaata 13920
taccgctatc ggtcccaatt agaaaagacg tcaatgctc tctggcgctc gctttgccc 13980
acaaaatgt gatcatgag aaagcgaaga atgacgttgc agctgataat gttgcggcc 14040
aacgcgcgcg aaacgcgcgc tgcagacccc acagctccca acgaagatg tatcgtcaa 14100
gtgatccaaq cacactcata gttggagctg tactccaaag gcggcaatga agagtacag 14160
agatactgt agacotttct cttgggaacc aacacgctca gccctctga ctacagctt 14220
gtagccaccc acacaggcaa cagtccgtgg atagcgaat atgtcttctc ggtccatttc 14280
taacacactt taggcgtcaa gtagaagctg cagaagaagt atgtgccttc attgagaac 14340
ggtgtgtctg attcaataa agctctgaga tcaagttggc cagtcatgtt gggggggta 14400
attgattga gttatgcct acagtctga caggtatact cgtgcccc ttataactt 14460
ttgattccgc tgcacttgaa gcaatgtctt ttacccaaaag tgagaatgtc ccacagacaa 14520
caccocaggg tatgtgtgag caaaaataa acactccgat accgggaata gaaccccggt 14580
ctccacgggt ctcaagagat attcttgatg agagcgcat 14619

```

<210> 136

<211> 15262

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pZK16-ML8N

<400> 136

```

gtacgcttc ttgacttgcg cccacgttcg cctttttggt cacaactcag cccgcactct 60
ctccccaaag aggagatgga cttacaatag atacagacaa acagatattt tcggccacct 120
ggtgcattgc gttctgttta ttaccoccaa taatatccac agcgtccgaa tcaactcagc 180
aaactcttct tcaactact gatcttacc cataccctcc agacagcggg gccatattca 240
taccgcactt aactgagaca gtaaacctcg ggaaggagc ggtcttggtc catcgcgacg 300
ctgtaagctt tgcctttgog ttttgtgac caacacacgg cagtcaaggt ctcggtgaa 360
agatgggtct ggaatgtgg atttgtgaat cattgacata taataacag tgatcctgcc 420
atttcccaat ctttagaaat catcttttct tcaacatac aaaaaacaca catcaaatg 480
cagttctcgc ccgcgcgcgt ccttgccttc gcgccttcgt tctctgcaca ggcgcgcctc 540
gctgctctg aagctgggtg tgcgcttgc tccggtgctt ccggtgcgcg tgcgcgcctc 600
tattctgctg taattggctc agagtctcag gcgccttctt ctgctgcctc ttctgctgct 660
gagtcgtgct ctccgcctgc ctctgagtc gccctccgc ctccgcctc agagtcacag 720
gctctgagct ccgcttccaa gtcgcctcc aaggtctctg agtctgcttc caaggccgag 780

```

tottctgtgt ccaaggttc etcogctgt ggttctgttt ctccgcgcgc tgcctgtgca 840
 acaaggtctg gtgatgcttc cgggtcttcc aagctcttcg gctctgcttc ctctacoggt 960
 gacaagtcog gtgtgtgtat tgttgcctcc gctgttgtgt ccatgttgtt cgtgtcttt 960
 gctcagctca tctaagcggc ctttgtacaa ctaccacact aatatttata ttatgatata 1020
 gactgcataa ctattgcatg tgaatgctt cgtttgaagc togtagtccg gccattcatg 1080
 gcaagtgtat gcaatltgc aacaaatgg cttaataaca tatacaaatg ggaatttgt 1140
 tgttgaac acacacacac ccaactattt tgacattgtt acaagtcttc gaattgtctt 1200
 toggcgcgac gacatgcatt gtagacaga caattgcaga agcttgccca atggttaggt 1260
 tqccctttaa ccagtaaac caataatagc ttgcacata acaaacacac cagccacaga 1320
 ctgggtgaaa gggcttggtt ttgtgcaat caactcatg actcaactcc ccaatgctc 1380
 catgaaccca tgaatccata gtacatccat ccaacttcag agaggttcag aacacgcga 1440
 ttggcacaag taacacaagg tctgcacac ttaagggtgt cagltgtgca aagttaaat 1500
 taactgtgt gatgtgttc cactcaact gtggttgtat cagacgttgt caataacac 1560
 agatcaata tgcctacaa gtacgtgtcc atgagtgta taatacaca catgatcaga 1620
 gtaaggctt ggaacaaa gtctgcgcgc tatgtaacac tgcgtgaatc agtgtgtgca 1680
 gtltccaggt actggaatga atgactcaca ccgttgtgta gctaaactcc caactatgt 1740
 actcaactt actgtagcta ctctaatgac gataatgaca taatgactca caactttat 1800
 ctcccccac gaagtgaag gatgaagag aggttaacag cttttgtaac aaatccaca 1860
 aatcaacta ccgtattaca aaagtacag tatcaatgac cagcgccgac agtgcatta 1920
 atgaatgccc caacgcgcgc ggaagacgcg ttgcgtatt gggcgctctt ccgttcttc 1980
 gctactgac togtgtgct cgttgtgtg gtgcgcgga ggggtatcag ctactcaca 2040
 ggcgttaata cgtttatca cgaatcagc ggaatacga ggaagaaca tgtgacaaa 2100
 agccacgaca aagccacaga accgtaaaaa ggcgcgtgtg ctggcgtttt tcaataggt 2160
 ccgcaccttc gaagagctac acaaaaatg acgtcaagc cagagtgccg gaacccgac 2220
 agactataa agatccagc cgtttccccc tgaagctcc ctgtgcctt ctctgttcc 2280
 gacctgtcc cttaacggat accgtccgc ctcttccct toggagacgc tggccttcc 2340
 tcaatgctca cgtctaggt atctcagtc ggttaggttc gttcgtccca agtgggtgt 2400
 tgtgacgaaa ccccccgttc agccgcagc gtgcgcctta tccgttaact atgtcttga 2460
 gtccacccgc gtaagacacg acttatccg actggaaca gacatgcta aacgattag 2520
 cagagcgagg tatgtaggc gtgtacaga gttcttgaa tggtagctta actacgcta 2580
 cactagaaga ccgtatttg tlatctgca tctgtgaag ccagttactc toggaaaaa 2640
 agttgtagc tctgtatcc gcaacacac ccccgctgt agcggtgtt ttttgttgt 2700
 caagcagcag attacgcga gaacaaaag atccaaaga gatcctttga tctttctac 2760

ggggtctgac gctcagtgga acgaanaact acgttaaggg attttggtca tgagattatc 2820
 aaaaaggttc ttcactaga tctttttaaa ttaaaaatga agttttaaat caactaaag 2880
 tatatatgag taaacttgtt ctgacagtta ccaatgctta atcagcgagc cactatctc 2940
 agcagctctg ctatttgtgt cctcactagt tgcctgaact cccgttgtgt agataactac 3000
 gatacgggag ggttaacct ctggccccgc tgcgtgaatg ataccgcgag acaaacgctc 3060
 accgcctcca gatttatcag caataacaca gccagccgga agggccgagc gcagaagtg 3120
 tctgcaact ttatccgctt ccatccagtc tattaattgt tgcggggaag ctagagtaag 3180
 tagttcgcca gttaatggt tgcgaacgt ttgtgcatt gctacagcca togttgttc 3240
 accgtctgac ttgtgtatgg ctctatccag ctccggttcc caacagatca ggcagttac 3300
 atgactcccc atgttgtgca aaaaacggc tagctccttc ggtcctccga tctgtgtcag 3360
 aagttaagtg gcgcagtgct tatcactcat ggttatgcca gactgcata attctctac 3420
 tgcctgcca tccgttaagt gctttctgtg gactgggtgag tactcaacca agtctctgt 3480
 agaatagtgt atgcgcgac cgaattgttc ttgcgcgcgc tcaatccgag ataatacgc 3540
 gccacatagc agaactttaa aagtgtcat ctttgaaaa cgttctctgg gcgaaaaact 3600
 ctcaaggatc ttaacgctgt tgagatccag ttctgttaa ccaactcgtg caacaaactg 3660
 atcttcagca tcttttactt tccacagcgt ttctgggtga gcaaaaaacg gaaggcaaaa 3720
 tgcgcacaaa aagggtaaaa ggcgcacgc gaagtgtga atactcactc tcttctttt 3780
 tcaatattat tgaagcattt atcagggtta ttgtctcatg agcgatata tatttgatg 3840
 tattagaaaa aataacaaaa taggggttcc gcgcacattt ccccgaaaaa tgcacactga 3900
 tgcgtgtgta aataccgcac agatgcgtaa ggaacaaaaa ccgcctcagg aatttgttaag 3960
 cgttaatttt ttgttaaaat togcgttaaa tttttgttaa atcagctcat ttttaacca 4020
 ataggccgaa atcgcacaaa tcccttataa atcaaaaaga tagacgcaga taggggtgag 4080
 tgttgttcca gtttgaaca agagtccact attaaagaac gtggaactcca acgtcaaaag 4140
 gcgaacaccc gtctatcagg gcgatggccc actaogtga ccatcaccct aatcaagttt 4200
 ttgggggtcg aggtgcogta aagactaaa tcggaaacct aaaggagagcc cccgatttag 4260
 agcttgacgg ggaagccogg cgaactggc ggaaggaaga ggaaggaag cgaagggagc 4320
 gggcgctagg gcctgcgcaa gtgtagcgtt cagctgcgc gtaacacca caccgcgcgc 4380
 gtttaatgac ccgtacagc gcgcgtccat tgcacattca ggtgcgcgaa ctgttgggaa 4440
 gggcgatcgg tgcgggcctc ttctgattta cgcacagctg cgaagggggg atgtgtgca 4500
 aggcattaaa gttgggtaac gccaggggtt tccagtcac gactgtgtaa aacgaagcc 4560
 agtgaattgt aatacgaact actatagggc gaattgggac cgaactgcga tctatgttac 4620
 actgtgtgag tgtagagggc aatgtgtgtt acaactgagc taactaacac acaagcttg 4680

tactgtata agttgtgagc tacagtaaca gtaatactac aaatttcagt atgtacaagt 4740
 atgcoggtac ggcagtagca tgtatcaaa agagacagaa agaaagctgg aaagcgaag 4800
 ccagctgagc gtgggaaccc cagtaactgg asgtgocaaa cctgcgcagc aatgacatgy 4860
 tcttgacatt tttaacogta ccggtacagc acgtacacgg taagcacttg ctacacatga 4920
 tcatagacca ccagaaatac cctgcattag cgcacacttg ccaagccctc aactttcttt 4980
 gaagctaac ctttttaacc ttttaggctg ccaggaagag caaaatgaa ttgtactlgy 5040
 ggcacagaa tttaggtga agtgttoctg ggaagaatg tgcttgctg tatagtggg 5100
 tgcctctcct tcaatgtaga tgtgagctta cagcacttg agattgtcta ccggtctctg 5160
 ttaagtggt accagtatta tgggttgcg tgaagaagc cgaagagcca cgggtaccc 5220
 aagagtgtgg gcgcacactc acacatgaag caagaccttg ttctcatit ctatttgitt 5280
 acaaggccg acataatata cttaggcagc cctgggcoga ccggcagga taactctcaa 5340
 acacactgt atgtggcgt cccctctcga tccagtagt tgcatagtt agcagtctga 5400
 gagtttgct gtgacgcaag accagagag agacagagag ttgtccagc ttgactact 5460
 gtacgtact ccagagtaaa tcaaatoga cgtctcact tctgacttt gtgaagcgg 5520
 agatcaac atttgtctc agcccttctc cgttgaagc gacttgacca cgaatgttg 5580
 gaataaagct gtgttaata ccagacactac agttagagtg tggatctgca caactgaca 5640
 ggttagatga aagttaacta acacacagag agcggacacc aactgcgaa aatgtaagga 5700
 tgaacccatt gctgtccgc atagcaatct gttttgocaa gaagtcttg tctgttgga 5760
 ttctgacct tgaactatt tgaactttgt caattaacac tggactcttg ctacacgatt 5820
 agcagacct cctgtgctc actgagcttg tctgttcaag cgtggaatg atgacaaag 5880
 cgtactgaa ggggaggtg aaagcggta tgaactaact gctgtctat agtgagggg 5940
 aggatatga ggaacgccc ggttcagtg gggatttgg gagctactca gtaactgta 6000
 ctctgtcgg aaagccctg tgaaggcaaa tatgcaaaa ccgaacatg tggagcaacc 6060
 ttctcatttc gagagatac acaggggttg cactggggcc tgggcatccc gtgcaataac 6120
 ccagacatg tgaaggtgca cgtcgcacac ttatccagcc agtcaactg ggtgtgcac 6180
 gctaaccaac aatgggattc ccgaactctc gtgggggacc gtgcacggg gaactgaac 6240
 aaacagggtc aggggtctct tgaataatga gagaatata acaaaaatc caaaaacgc 6300
 accaaaata tgcacaaact gtaacacgtg gaactcagc gccctggct cgaactccca 6360
 ctgacatacc cccctcacc caataccta tccacactag caactcgtt aattaaagt 6420
 ccacacatg cttgataga tctgaactc tctctctgca gctttccact acaaatctc 6480
 tctgtccca ggaagtccat ggttggttg atcatggtt tggtagtag gtgtagcag 6540
 ggtgtatg tgactggga tgaagttag aataagctat acacagctca gctttctg 6600
 agctcatat aagtataagt agtcaactg attagactg taccagcat ctccgtatcg 6660

agaaacaaa caacatgccc cactggacag atcagcggg taacacaggtt gtgaagtac 6720
 atacatactc gatcagacag gtgtctgac cctcataca gctgaacaa cgtccatcac 6780
 tgcacgctc tctatataca cagttaaatt acatatccat agtcaactc ctaacagtta 6840
 atcttctggt aagcctccca gccagcctc tggtagctc tggcctcctc aataggatct 6900
 cgttctggc cgtacagacc tgggcgcaca attatgatat ccgttcoggt agaatgaca 6960
 tctcaaacg ttggtactg cgtgcagaga gctctcctc tgtgtcgaag acccaaccg 7020
 ggggtcagaa taagccagtc ctcagagtcg ccttaggtc ggttctggg aatgaagcca 7080
 acccaactc cgggttcgga tgggcgaag tcaatgctc ccttgagta ctcgcagtc 7140
 gccagagag ccttgcaaga cagctcggcc agcatgagca gactctcgg cagttctcgg 7200
 ttggagagag ggaactaggaa cctcctgta tgggagttc cgtagtcaga gactcctcc 7260
 ttcttctgt cagagacagc tctcctggga ccagctgcca ggcacgaat gattccggt 7320
 ccgggtcac cgtggcggt ggtgatctg gccactcgg cgttcoggt acacoggtac 7380
 tgggtctga cagcttgc ccaatctgag aacttctgt cctcgaacg gaagaaccc 7440
 tgcataagag caagtctct gagggggagc acagtcccg cgtaggtgaa gtcctcaatg 7500
 atgtgatag ggttttgtc atgcaacaca taaggctcga ccttatcgg aagctaatg 7560
 agctccttg tgggtglaac atccagagaa gccacaggt tggtttctt gctgcaag 7620
 agcttgagca ctogagcggc aaaggcggac ttgtggactg tagctoyagc tctgtaggag 7680
 ggcattttg tgggtgaagag gagaactgaa taatttagt ctgcagaact ttatatcgya 7740
 acctatctg gggcagrgaa gtatagtta tggtaagt taccagtag ttgaacttat 7800
 agatagactg gactatcgg ctatcgtcc aaattagaaa gaacgtcaat gctctctgg 7860
 gctgacctt tgcgcacaaa aatgtgaca tgaagaaga cagcaatgac gttcagctg 7920
 atattgtgt cggcacaacc cgcgcaaac gcagctgcca gccccacag ctcacagaa 7980
 gaatgtatg tcaaatgtat ccaagcaca tcatagttg agtctactc caaaggcggc 8040
 aatgacagc cagacagata ctgtcagcc ttittcttg gaacccacc cgtcagccct 8100
 tctgactcc gtaattgagc caacgacaca ggcacacgtc cgtggataga agaatatgtc 8160
 ttgtcgccc atttccacc aacttaggc gtaagtga tgttgcaaa gaagtatgt 8220
 ccttcattga gaatcgtgt tctgatttc aataaagctc tgaagcagc ttggcagtc 8280
 atgtgtggg gggtaattg attgattat cgcctacagc ctgtacagct atactcgtc 8340
 ccaacttat acttttgtat tccgtgcac ttgaagcaat gtcgttaac aaaaatgaga 8400
 atgtccaca gaacacacc cagggtatg ttgacaaaa aataaacact ccgatacgg 8460
 gaatgmacc ccgtctcca ccgtctcca gaagtattct tgaagagag gtatogtag 8520
 ttgagcaag ggaqaatgt agtgtgtaa agactcacta tggtcoggg ttatctgac 8580

caatagccaa agtctggagt ttctgagaga aaagggcaag atactgtatgt aacaaagoga 8640
 cgcattgtac aataatccg gaggcatgta tcatagagag ttatgggttc gatgatggca 8700
 ctggcgctg gtaatgaatt atacggctga ctacatattt gtctccagac atacaattac 8760
 agtccagcac ttacccttgg acatctgtag gtaccocccg gccaaagaga tctcagcgtg 8820
 tcttatgtcg gattggcgtg gtccctcgcg tctcgaattg gtcccatctc actttctctc 8880
 gcttggttac acccagcatg tctgtatgg ctgttttgcg tgccttatct atctccocag 8940
 tattaccacg tctaatgac atgtgtgat tgggtctaca ctttccatcc agagataagg 9000
 agtagccag ttgcataaaa agcccacatc taatcagctt ctctctttct tctaattagt 9060
 acaaggtga ttacggaatt ctggaagctt agttggcctt aaaaaatca aaaaaagcaa 9120
 aaaaagaaa acgaaaaacc acagtittga gaacggggag gtaacgaagg atcgtatata 9180
 tatatatata tatataacc cagcgctccc gageccggcc ttgtattctt cctacacacc 9240
 aacctttctc accacccctc ttccacccc tgtctattgg ttcttccacc cccgtctctc 9300
 tggctgcgat tccctctgtc tccctgtttg tctccocagc agtctggctt tctctcctc 9360
 agaaggtcac gtctctggcc aaatgtattg tggctctgat tgcacggctt atcatgtccg 9420
 ttgcaggctg ctctcctcgc atcgtgtgag ctctcttga caagagatcc gtctccatt 9480
 acgttgttgc cagattgttc tctctcttgc cctctcagcc gtgtgtgtgt acatataga 9540
 ttgttgttga ggaacccctc gataagtacc ctgtatctgt ggtctgtaac catcaatctc 9600
 ctatggatat gatgttttg ggaagatttt ttccaaagca ctgctgtctc atggcagaag 9660
 aggaactctc gtaacttccc tttttggaaa tgtttatgaa actgagcaac gctatctcca 9720
 tgcacgggaa gaaacacaaq aaagccctgc agtctccacc ccaagccctg cgcgacatga 9780
 agaagccaaa ctctgaatc tggattttcc cagaaggcac cccgtctaga ctggacaagg 9840
 cagactctgt gcccttaag aaagtgctt tctatcttgc aattcaggcc cagctccta 9900
 ttctcccatc tatctgcag ggcattccc atctataga ctcttgaag ccgtacttcc 9960
 ccggtggaga gctcagatc agagtctcgg agcccttccc tacaactggc ctactactgt 10020
 atgatgtgaa cgcactgatg gccagccacc gaacacttat gctcagcac ttgaaggaga 10080
 tggattcccc gtattctgc agccctcgtg aaatggatc cccgcacatc gacgcagata 10140
 ttgcacatgc tacaagccac agcattggca acctgtgaga cgcataacca aaagctgcta 10200
 cccctaagga ataaagccgc gccattgtga ttggaacacc aacactgggt tatatctagg 10260
 tgaagattag ttggaagatt atatatata tcaagctatgc caacgttaac tctatctagt 10320
 tcaacagaga accagtgact gcaagtata tagaatttga ccaactggcc atctctctgc 10380
 actctttacc tatatctcat ttatttctta tatcaaatc actttctctc cccagcctgc 10440
 agtctgaaa cctcatgagc ataacatgc ttgatctgt caatagcggg cttttgga 10500
 tcttctgtgt tggcaactt gctctgtgt tttaacctc catcaagggt ctcacaaact 10560

acctcgaac tgcctgcgtg atctggacac caccagaggt ccaggaactt taggttgca 10620
 caaatgtccc accagtgcca ggcagaacac gctggaacag cgtctcagtt ttgtctaac 10680
 aaaaagttag ggcctcagag tccagcaggg tggctgact tgtatagcc tttagagctg 10740
 cgaagccgcg tatggaattg gctcactcgg ccagatttag ggtctgtgga cactgtcat 10800
 gttagtgcac ttaaatgcc cccgtgatat agcccagaca ataggccgtg gccatattt 10860
 ttgtctccc gcaatttccc attgtcgtt acccaacact tgcctctctc gcaattgcca 10920
 acctaatc tggtttacc ttgcaaacat cttacaagcg ggggcttctg ctagggtata 10980
 tataaacagt ggccttccca ctggtgtccc agtctctttt tctcttctt tcccaacaga 11040
 ttgcacatct aaactacaca tcaacccatg gtctccaccc accgtctaga cgcactgga 11100
 gctccctctc cccgagacgc accttctact cgaactgaca accctcggac ctggaattac 11160
 gatgaagcca ttgctcttcc cggctgaact cctggagata tggacgaact ccgaattcca 11220
 cccgagaca gaggctccgt gccagtcgag aagctcgcg aggcgtgtat tctctacctg 11280
 gccctcttcc gaacccgagc ttgtctacct ccttcaaca ctgctcaac cctggccag 11340
 ctgactact tcatcggcga tcccgacccg cgtctggttg tcttctctcc cgcagctcca 11400
 ggtgctgttg agacattgc caagcagacac ggtgtatctg tcaaaacact cgaagccgat 11460
 ggaagaggtc ccttgcttga ccttgtcaga gatgacgtg ccgactttgt ccatgtctg 11520
 ccatctgcgc acgatctggc tctattctc tacaattccg gtacaacccg acgatcgaag 11580
 ggtgcactgc ttactcctgg caactctgct tcaaacgtc tcaacttgcg agcatattg 11640
 agagttacgc cagaagatgc actcactcat gcccttgcac tctttccacc tcatgtctg 11700
 ttcttgtcta cgaacgtcac actgcttcca ggaagctcga tgtttctgt ctcacagttc 11760
 gatgcagcac aggtcttctc tctcattgca caggcaacaa tcttattggg cgtccacaca 11820
 ttctacgttc gattgtcga gattctcga ctgcagaagg gttgtgtggc cagcatcaga 11880
 ctgttaattt ctggtatcagc tccctgttct gccaaaaacc acccagagtt tcatgtctgt 11940
 actgtctccg ccatctcga cgcataccgc atgaaggaac ccaacatgaa tacttccac 12000
 cactaagagg gcaagcttat tgcaggacac gttgttttct ctctgcocga cgtcaactgt 12060
 cagatcaccg atcccgcacc cgtctcgtt cttccactg asgagactgg catgatcag 12120
 atcaagggac ccaacgtctt caaggcttat tgcgaatgc cgaaaaagac cctgtccag 12180
 tttaacgcag ccgttttctt tatctctgga gattctgga agatcagacc agaaggttac 12240
 gttcaacttg tgggaagagg caaggacctg gtcatttccg gttgctacaa catctatccc 12300
 aaagaggtgc aaggtcagat cgaacagatc gagggtgttg tcaagctcgc tctaatgtt 12360
 gttctctacc ccgatttggg agaaggtgct accgtgttg tctgtgcaaa accgtgtgac 12420
 gttctgacgc aaaaagcaact cgtctctgtc ctgcaagcac gttctgtccc atacaagcaa 12480

```

ccccagcgga ttatctttgc cgaagatctg cctcgaaaca ctatgggaaa ggttcagaag 12540
aacattcttc gacagcaata cgcagatctc tacacacagc gataagcggc cgcctagaga 12600
gataaatata taataacatt gagatattaa atgcgctaga tttagagccc tcatactgct 12660
cgagagagag ccaagacgag tactcaaggy ggattacacc atccatatcc acagacacaa 12720
gtgggggaaa ggtctatata acattttccg gaataacgta gtttccgatg ttatcaatgg 12780
ggcgagccag gatttaaggc atttcgggtg ctcggggtga aatggcggtt ttggcctaca 12840
tcaagtctga ccatgtcttc atttgcctgt caaagttaaa cagaagcaga tgaagaatga 12900
acttgaagtg aaggaattta aatagttgga gcaagggaga aatgtagagc gtgaaagact 12960
cactatgctc cgggttcttc tcgacccaata gccaaagtct ggaatttctg agagaaaaag 13020
gcaagatcgc tatgtaacaa agcagacgcct ggtaacataa taccggagcg atgtatcata 13080
gagagttagt ggttcgatga tggcactggt gctcggtatg actttatagc gctgactaca 13140
tatttgtctc cagacataca attacagtca agcaacttacc ctgggacatc tgtaggtacc 13200
ccccggcaaa gacgatctca gcgtgtcgta tgcctgattg gcgtagctcc ctgcctcgtc 13260
aatgtctccc catctacttt ctctctgctg gctaacacca gcatgtctgc tatggctcgt 13320
tttcgtgctt tatctatctt cccgctatta ccaactctaa atgacatgat gtgattgggt 13380
ctacacttcc atatacagga taaggcgtag ccagtttga taaaaagccc aactctaatc 13440
agctctctcc ttitttgtaa ttagtacaaa ggtgattaga gaactctgga aguttagtgt 13500
gacctaaaaa aatcaaaaaa agcaaaaaaa gaaaaacgaa aaacacagct tttagaagaa 13560
ggaggttaac gaaggtatgt atatatatat atatatatat ataccacggc atcccgagac 13620
cgccctttga ttcttcccta caacaaccca ttctcacacc cctaattcac aacatagtg 13680
aaggtcttc gacagctctt gccctcgttc atgcacgaaa aggtgtccga cgtctccgct 13740
tgggtgaact tcacacccgg tggagctgaa atcatttaga aacacacggg acagatgct 13800
actgaagcct tcattggtat gccctctcag gaagccttgc acaagctcaa gcgaatgcc 13860
aagatcaacc aggtctccga gctgcctccc caggtcggcg tcaaacgaag tcaggagat 13920
ttccgaagac tccgagaaga gctgactccc actggcattg ttgacgcttc tcccctctgg 13980
tactctgcca agatcttgac caccctgggt ctggcgtgc ttgcctcttt catgctggtc 14040
cagtacccc tgtaactcat tggctctctc gtctcgtgta tgcactccca gcaaatggga 14100
tggctgtctc atgaactctg ccacacccag accttcaaga aacgaactcg gaataacgtc 14160
ctgggtctcg tctttggcaa cggactccag ggtctctcag tgacctggtg gaaggacaga 14220
caacaagccc atcattctgc taccacagtt cagggtcagc atccagacat tgataacctg 14280
cctctgctcg cctgttccga ggaagctgtc actcagactt ctccactctc ccgaagctc 14340
attcagttcc aacagtacta ttctcgttc atctgtatc tctgcatc catctggtg 14400
ttccagctg tctgacccgt tcgtccctcc aaggacggag acaacagtt ctacagatct 14460

```

```

cagtacaaga aagagggccat tggactcgtc ctgcactgga ctctcaagac cctgttccac 14520
ctctctctta tgcctccact cctgaactcg atgctggtgt tctttgttcc cgaagctcgc 14580
ggtgcctcgc gaattgccat cgtgtctctc atgaacact accctctgga gaagatcggt 14640
gattcctctc gggacggaca tggcttctct gttggctcaga tccatgagac catgaacatt 14700
cgaacaggca tcattactga ctggttcttt ggaagcctga actaccagat cgaacacct 14760
ctctggccca cctgcctcgc acacaacctc actgcctgtt ctaccaggt ggaacagctg 14820
tgccagaaga acaacctccc ctaccgaac cctctgcctc atgaaggtct cgtcactcgt 14880
ctccgatacc tgtccaggt cgtctgaatg gccgagaagc agcccggtgc caaggtccag 14940
taagcgccgc caagtgtgga tggggcaagt agtcgccggt tctgtgtgca caattggcaa 15000
tccaaatgag atgattccaa cacaaggata tagcagacta cgtgtgtgtg cagagatata 15060
gcaacggata ttatctgttg acacttgaga atgtacgata caagcactgt ccaagtacaa 15120
tactaaacat actgtacata ctcatactcg taccggggca acggtttcac ttgagtcag 15180
tggtagtctc tctactcgt acagtgtgca atactcgtga tcatactctt tgatgtatat 15240
cgtattcatt catgttagt gc 15262

```

<210> 137
 <211> 13075
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Plasmid pZK16-ML

```

<400> 137
gtacgcatc ttgacttctt ccccaattgc ccttttttgt cataactcag ccccgactct 60
ctccccaaag aggagatgga cttacaatag atacgagaca acagatattt tcggccacct 120
ggtgcgaattg gtctgtttta ttaccoccaa taatactcac agcgtccgaa tcagctcgac 180
aagctctttc tcagcatact gatcttacc cataccctcc agacagcggt gccatattca 240
taccgcactt acctgagaca gtaaacctcg ggcaaggagc ggtcttggtc catcgcgacg 300
ctgtaagctt tgcctttgog tttttgata caacacacgg cagtcaaggt ctcggtgaa 360
agatgggtct ggcaatgtgg atttgtgaat cattgacata taataacgag tgatcctgcc 420
atttcccaat ctttagaaat catcttttct tcacacatac aaaaaacaca catcacaatg 480
cagttctcgc ccgcgcgcgt cttgcctcgc gcgccttctg tctctgccca ggccgcctcc 540
gctgctctg aagctgggtg tgcgcttgc tcgggtgcat ccggtgcgcg tgcgcgcctc 600
tctctctgct tcattggctc cagctctcag gcgccttctt ctgctgcctc ttctgctgct 660
gagtcgtctt ctccgcctgc ctctgagctc gctcctcgc ctccgcctc caggtccaaag 720
gctctgagt ccgcttccaa gtcgcctccc aaggtctctg agtctgcttc caaggccgag 780

```


tottctgtgt ccaaggttc ctcogctgt ggttctgtt cttccgcgc tgcctgtgc 840
 acaaggtgt gtgatgttc cgggtcttc aagcttcgc gctcgttc cttacoggt 960
 gacaagtgc gtgtgtgt tgtgtccgc gctgtgtgt ccatgtgtt cgtgtctt 960
 gctcagctca tctaagcgtt cttgtacaa ctaccactt aatattata ttatgatata 1020
 gactgcata ctattgcat tgaatgctt cgttgaag togtatgcg gccattcat 1080
 gacgtgatg gacgtatgc aacacttg cttataaca tataaaggt ggaattgt 1140
 tgttgaac atcacacac caccatatt tgacattgt acaagtctt gactgtct 1200
 tggcgcgac gacatgatt gtacagaca caattgcaga agttggcca atgttaggt 1260
 tgaccttaa ccagtaaac caataatag ttgcacata acaaacaca cagccacaga 1320
 ctgggtaaa gggctgggt ttgtgcaat caactcatt actcactcc ccaatgct 1380
 catgaaccca tgaatcata gtacatcat ccaattcag agaggtcag aacacgcaa 1440
 ttggcacaag taacacaag ttctgacac ttaagggtg cagtggtgc aagttaaat 1500
 taactgttg gatgtgtt cactcaact gtggtggtt cagacgttg cactaaaca 1560
 agatcataa tgcctacaa gtacgtgct atgagtgta taatacaca catgatcaga 1620
 gtaaggctt ggaacaaa tgcgcgcgc tatgaacac tgcgtatca agtggtgc 1680
 gtttccagc actggaatg atgaactcaa ccgttgtta gataactcc ccaactatg 1740
 actcacttt actgtatga cttcaagta gataatgaa taatgataa ccaacttat 1800
 cttccacac gaagtgaag gatgaagag aggttaacg ctttgttaa aaatcacaa 1860
 aatcataa cgtattaca aagtacag tatcagtaac cagcgccgc agtgcatta 1920
 atgaatgcc caacgcgcg ggaagcgcg ttgcgtatt gggcgtctt cgtctctc 1980
 gctactgac togtgtgtt cgtgtgttg gctgcgcga ggtgtatcg ctactcaaa 2040
 ggcgtataa cgttatcaa caaatcagc ggaacgca ggaagaaca tgtgacaaa 2100
 agccaccaa aagccacga accgtaaaa ggcgcgttg ctgggtttt taataggt 2160
 cgcgcacct gacagcact acaaaaatg acgtcaagt cagagtgcc gaacccgac 2220
 agactataa agatccagc cgttccccc tgaagctcc ctgctgcct ctctgttc 2280
 gacctgcg cttacggat accgtccgc cttctcctt tgggagcgc tggccttc 2340
 taatagctc cgtctaggt atctcagtc ggttaggtc gtcgtccca agtgggtg 2400
 tgtgacaaa ccccccgtc agccgcagc ctgcgccta taaggtaac atgtcttg 2460
 gtccacccc gtaagacac acttatccc actggaaca gactagcta aaggttag 2520
 cagagcgag tatgtagcg gtgtacaga gttctgaag tggtagcta actacgcta 2580
 cactagaag acgtatttg gtatctgcg tctgtgaag caagtacat tggaaaaa 2640
 agttgtagc tctgtatcg gaaacaaac caccgtgtt agcgtgtt ttttgttg 2700
 caagcagag attacgcga gaacaaaag atccagaa gatccttga tctttctac 2760

ggggtctgac gctcagtgga acaaaaact acgttaaggg attttgtca tgagattac 2820
 aaaaaggtc ttcactaga tctttttaa ttaaaaatg agtttaaat caactaaag 2880
 tatatatag taaacttgt ctgacagta ccaatgcta atcagcagc cactatctc 2940
 agcgtctgt ctatttgtt cttcactgt tgcctgact cccgtgtgt agataactac 3000
 gatacggag ggttaacct ctggccccc tgcgtgaat atccgcgag acaaacgct 3060
 accgcctca gattatcag caataacaa gccagcggg agggcgcag cagaagtg 3120
 tctgcaact ttatccgct ccatccagtc tattaattg tgcgggaag ctacagtaag 3180
 tagtgcgca gtaaatgtt tgcgaacgt ttgtgcatt gctacagga cgtgtgtc 3240
 accgtgcgc ttgtgtatg cttcattcag ctcgcgttc caacgataa ggcgagta 3300
 atgactccc atgttgtca aaaaagcgt tagctcctc ggtcctcga tctgtctag 3360
 aagtgaagt gcgcaggtt tatcactcat ggttatgca gactgata attctctac 3420
 tgcctgcca tccgtgaat gctttctgt gactggtgag tactcaaca agtctctg 3480
 agaatagtt atgcgcgac cagttgtc ttgcgcgcg tcaatccgc ataatacgc 3540
 gccacatag agactttaa agtgtctat ctttgaaaa cgttctcgc ggcgaact 3600
 ctaaggatc ttaacgctg tgagatcag ttctgttaa ccaactcgt caacaaactg 3660
 atctcagca tcttttact tccacagct ttctgggtg caaaaaacg gaaggcaaaa 3720
 tgcgcacaaa aggggaataa ggcgcacgc gaagtgtg atactcact tcttcttt 3780
 tcaatattt tgaagcttt atcaggtta ttgtctcat agcgtatca tattgaatg 3840
 tattagaaa aataacaaa taggggttc gcgcacatt ccccgaaaa tgccactga 3900
 tgcgtgtga aatacgcac agatgctaa ggaacaaa cgcctcagc aattgtga 3960
 cgttaattt ttgttaaat togtttaa ttttgttaa atcagctat ttttaaca 4020
 atggcgcga atcgcacaaa tcccttata atcaaaaag tagacgaga taggtgtg 4080
 tgtgttcaa gtttgaaca agagtccat attaaagac gtggaatca acgtcaaa 4140
 gcaaaaaac gttctcagc gcatggcc actaagtga ccatcactt aatcaagtt 4200
 ttgggtgtg aggtgcgta agcactaaa tgggaacct aaaggagac cccgattag 4260
 agctgaagc ggaagcgcg cgaagtgcc ggaaggaag ggaaggaag cgaaggaag 4320
 ggcgcctagc gctgtgcga gtgtagcgt cagctgcgc gtaacacca caccgcgc 4380
 gttatgtgc cgcctcagc ggcgtccat tgcctatca gctgcgcaa ctgttgga 4440
 ggcgcctgc tgcgcgcct ttcgtatta cgcagctgc cgaaggggg atgtgtga 4500
 aggcattaa gttgtgaac ccaaggttt tccagtaac gactgtga aacgaagcc 4560
 agtgaattt aatacactc actataggg gaattgggc cagctgcga tctatgtac 4620
 actgtgtag tgaagaggg aatgtgttg acaactgac taatataac acaagcttg 4680

tactggtata agttgtgagc tacagtaaca gtaatactac aaatttcagt atgtacaagt 4740
 atgcoggtac ggcagtagca tgtatcaaa agagacagaa agaaagctgg aaagcgaag 4800
 ccagctgagc gtgggaaccc cagtaactgg asgtgocaaa cctgcgcagc aatgacatgy 4860
 tcttgcaactt tttaacogta ccggtacagc acgtacacgg taagcaactg ctacacatga 4920
 tcatagacca ccagaaatac cctgcattag cgcacagttg ccaagccctc aactttcttt 4980
 gaagctaac ctttttaacc ttlttaggcl ccaaggaagc aataaatgaa ttgtactlgy 5040
 ggcacagaa tttaggtga agtgttoctg ggaagaatg tgcttgctg tatagtggg 5100
 tgcctctcct tcaatgtaga tgtgagctba cagcaactgt agattgtcta cggctctctg 5160
 ttaagtggt accagtatta tgggttgcg tgaagaagc cgaagagcga cgggtacccg 5220
 aagagtgtgg gcgcacactc acacatgaag caagacctgt ttctcatit ctatttgitt 5280
 acaaggscgc acataatata cttaggcagc catgggcoga ccggcagga taactctcaa 5340
 acacactcgt atgtggcgt cccctctcga tccagtagt tgcatagtit agcagtctga 5400
 gagtttgct gtgcagcaag accagagagc agacagagag ttgtccagat ttgagtaact 5460
 gtacgtact ccagagtaaa tcgaatatga cgtctcact tctgcactit gtgaagcaga 5520
 agcatcaac atttgttcc agcccttllc cgtltgaagc gacttgacca cgaatgttg 5580
 gaataaagct gtgttaata cagagacctac agttagagtg tggatctgca caactgaaca 5640
 ggttagatga aagttaacta acacacagag agcggacacc aactgcgaa aatgtacgga 5700
 tgaacccatt gctctcgcgc atagcaatct gttttgocaa gaagtcttg tctgttgga 5760
 ttgctgcact tgtactattt gtaactttgt caattaacac tggatcttg ctacacgatt 5820
 agcagcactc ctgtgcctc actggccttg tctgttcaag cgtggaatgc atgacaaag 5880
 cgtactgtaa ggggaggttg aaagcagta tgaataact gctgtctat agtgagggg 5940
 aggatatga ggaacaccc ggttcagtg ggaatttgg gagctactca gtacactgta 6000
 ctctgtcgg aaagccctgc tgaaggcaaa tatgaacaaa cgcacacatg tggagcaacc 6060
 ttctcatttc gagagataat acaggggttg cactggggcc tgggcatccc gtgcataaac 6120
 cgaagcactg tgaaggtgca cgtcgcacac ttatccagcc agtcatcgt ggtgtgcaca 6180
 gctaaagca aatgggattc acgatctat gtgggggacc gtgcacaggg gaatatgaac 6240
 aaacagggtc aggggtcttt tgaatatga gagaatata acaaaaata caaaaaagc 6300
 accaaaata tgcacaaact gtacacagtg gaatatcaga gccctggct cgaactccca 6360
 ctgacatacc cccctcaca caatatcata tccacactag caaatcgttt aattaaagtg 6420
 cgaacatgt cttgataga tctgtaactc tctctcttga gctttccact aacaaagttc 6480
 tctgtcccca ggaagtcact ggttggttg atcatggtt tggtagtagt gtagtcagc 6540
 ggtgtatig tgaatggga tgaatggga aataagtcac acacagtcac gctttctcg 6600
 agctcatat aagtataagt agtcaacgt attagactg taccagcat ctccgtatcg 6660

agaaacacaa caacatgcgc cactggacag atcagcggga taacacaggtt gtgaagtac 6720
 atacatactc gatcagacag gtgtctgac cctcatacaa gctgaacaa cgtccatcac 6780
 tgcacgctc tctatataca cagttaaatt acatatccat agtctaacct ctacagtta 6840
 atcttctggt aagcctcaca gccagcctt tggtagctgt tggctcctc aataggatct 6900
 cgttctggc cgtacagacc tgggcgcaca attatgatat ccgttcoggt agaatagaca 6960
 tctcaacag ttggttaact gtgtcagaga cgtctcctc tgtgtcgaag acccaaccgc 7020
 ggggtcagaa taagccagtc ctacagagtc ccttaggtc ggttctgggc aatgaagaca 7080
 acccaacat cgggttcaga tgggcagag tcaatgtct ccttgagta ctgcacagtc 7140
 gccagagagc ccttgcaaga cagctcggcc agcatgagca gacctctggc cagttctcgc 7200
 ttggagagag ggaactaggaa ctctctgta tgggagttct cgtagtcaga gacgtctccc 7260
 ttcttctgtt cagagacagt ttctcggga ccagctgca ggcacgaat gattccggtt 7320
 ccgggtacac cgtggcggtt cgtgatctgc gaccactcgc cgttcoggt acacoggtac 7380
 tgggtctga cagtttgc acatatctgc aacttctgt cctcgaacag gaagaacccg 7440
 tgcataagag caagtctct gagggggagc acagtccgc cgtaggtgaa gtcctcaatg 7500
 atgtogatat ggggtttgat catgaacaca taaggctcga ccttatcgga aagctcaatg 7560
 agctccttg tgggtglaac atccagagaa gccacacaggt tggtttctt gctgcacag 7620
 agcttgagca ctogagcgcc aaaggcgac ttgtggagct tagctogagc ttctagtag 7680
 ggcattttg tggtagaag gagactgaaa taatttagt ctgcagaact ttltatcgya 7740
 acctatctg gggcagrgaa gtatatgta tggtaatagt taccagtag ttgaacttat 7800
 agatagactg gactatcgg ctatcgtcc aaattagaaa gaacgtcaat gctctctgg 7860
 gctgcctct tgcacacaaa aatgtgacca tgaagaaagc cagcaatgac gttcagctg 7920
 atattgtgt cggcacaacc cgcgaacaa cagctgtca gccccacagc ctccacagaa 7980
 gaatgtatcg tcaaatgtat ccaagcacac tcatagttg agtctactc caaaggcgcc 8040
 aatgacagat cagacagata ctgtcagacc ttttccttg gaacccacac cgtcagccct 8100
 tctgactcac gtattgtagc caacgacaca ggcacacgct cgtggataga agaatatgtc 8160
 ttgtcgcccc atttccacc aaacttaggc gtcaagtga tgttcagaa gaagtatgt 8220
 ccttcattga gaatcgtgt tctgatttc aataaagctc tgaagacagt ttggcagtc 8280
 atgtgtggg gggtaattg attagttat cgcctacagt ctgtacaggt atactcgtc 8340
 ccaactttat actttttgat tccgtgcac ttgaagcaat gtogtttaac aaaaagtaga 8400
 atgtccaca gaacacacc cagggtatg ttgacaaaa aataaacact ccgatacgg 8460
 gaatgmacc ccgtctcaca ccgtctcaca gaagtattct tgaagagagc gtatogtag 8520
 ttgagcaag ggaagaatgt agatgtgaa agactcata tggtcogggc ttatctcagc 8580

caatagccaa agtcgtgggt ttctgagaga aaagggcaag atactgtatgt aacaaagoga 8640
 cgcattgtac aataatccg gaggcatgta tcatagagag ttatgggttc gatgatggca 8700
 ctggcgctg gttatgaatt atacggctga ctacatattt gtctccagac atacaattac 8760
 agtcaagcac ttacccttgg acatctgtag gtaccccccg gccaaagaga tctcagcgtg 8820
 tctgtatgcy gattggcgtg gtccctccgc tctcgaattg gtcccatctc actttctctc 8880
 gcttggttac aaccagcatg tctgtatgg ctgttttctg tgccttatct atcctccag 8940
 tattaccac ttcaaatgac atgtgtgtat tgggtctaca ctttccatcc agagataagg 9000
 agtagccag ttgcataaaa agcccacac ttatcagctt ctctctcttc tctaatagat 9060
 acaaggtga ttacgcaaat ctggagcctt agtggccct aaaaaatca aaaaaagcaa 9120
 aaaaagaaa acgaaaaacc acagtittga gaacggggag gtaacgaagg atcgtatata 9180
 tatatatata tatataacc cagcgctccc gagccggccc ttgtctctt cctacaccc 9240
 aacctcttc accaccccaa ttccaccca tgtctatttg ttgtccacac cccgtctct 9300
 tggctgcgt tccctctgc tccctgttg tctcccaag agtccggct tctctccct 9360
 agaaggtcaa gtctctggcc aaatgtatg tggctgtat tgcacggctt atcatgtccg 9420
 ttgcaggctg ctctctctg ctctgtctg ctctcttga caagagatcc gtctccatt 9480
 acgttgttc gcaatgttc tctctctg ccgtccgccc gtgtgtgtg acatataaga 9540
 ttgttgtga ggaacccctc gatagtacc ctgtatctgt ggtctgtaa cctcaatct 9600
 ctatggatat gatgttttg ggaagattt ttccaaaga ctgctgttc atggcaga 9660
 aggaactct gtaattccc ttcttgga ttcttatgaa actgagcaac gctatctca 9720
 tgcacggaa gaaacacaa aaagccctg agtctccac ccaagccctg ccgacatga 9780
 agaagccaa ctctgaatc tggatttcc cagagggccc cccgtctga ctggaacgg 9840
 cagactctt gcccctcaag aaagtgctt tctatcttga aattcggcc cagctccta 9900
 ttctcccat tatctccag gctattccc atctcaaga ctctcgaag cgtacttcc 9960
 ccggtggaga gctcagatc agagtctgg agccctacc tacaatggc ctactactg 10020
 atgtgtga cagctgtg gacagccac gaaccttat gctcagcac ttgaaggga 10080
 tggattccc gtattctg agcctctgt aaatggatc cccgacatc gacccgata 10140
 ttgcacatc tacaacccc agcattgga acatctgga cgaattaca aaagctgta 10200
 cccctaaaga atagccgccc gcaattgga ttggaacac aacatgggt tatatctgg 10260
 tgaagttag ttggaagtt atatatata tcaagatgc caacgtaac tctatctat 10320
 tcaacgaga accagctact gcaatata tgaattga ccaactgac attctctgc 10380
 actcttacc tatctctat ttattctta tacaacata acttctctt cccagctgc 10440
 agtccgaaa cctcatgagc ataacatgc tgaatctgt caatagcgg ctcttgga 10500
 tctctgtgt tggcaactt gctctgtg tttaacct catcaaggg cctcaaac 10560

acctggaa tctctgtgt atctggaac cagagaggt ccagcaact taggtgca 10620
 caaatctcc accagtgca ggcagaacac gctggaacag cgtctcagt ttgtctaac 10680
 aaaaagtga ggcctgag ttgagcagg tggctgact tctatagcc tttagagctg 10740
 cgaagccgc tatgatttg gctatccag ccagattgag ggtctgtga cactgctat 10800
 gttatgtac ttaaatgccc cccgtgatc agcccgaca atagccggtg gctatattt 10860
 ttgtctcc gcaatttcc atgtctgtt accacaccc tctctctt gcaatgca 10920
 acctaatc tggttacat tgaacacat cttacaagc ggggcttgt ctaggata 10980
 tataacagt ggtctccca ctggttgc agtctctt ttctctct tcccaacga 11040
 ttgcaatct aaactacaa tcaacccatg gtctccacc accgtctga cgcctcga 11100
 gctccctcc ccgagagcc accttccat cgaatgaca accctggac ctgacttcc 11160
 gatcagcca ttgtcttcc cgtctgact gctggagata tggacgact ccgacttga 11220
 ccggagaca gactgctgt caggtctgag aagtctgag aggcctgtat tctctactg 11280
 gctctcttc gaacggagc ttctctact cctccacaa ctgctcac cctggccag 11340
 ctgactact tcatggcga tcccgaccc cgtctgttg tctgtctcc cgcagctga 11400
 ggtgctgag agcaatgg caagagacac ggtgtatct tgaacaccc cgaagcgt 11460
 ggaagaggt cctgttga cctgtgaga gatgagctg ccgacttgt cgtgtctg 11520
 cgtctgccc agatctggc tctattctc tacaatccg gtacacccg acgatcgag 11580
 ggtgcatgc ttactctgg caatctgct tcaacgctc tcaactgag agcatattg 11640
 agagtacc cagagatgc actatccat gcttgcaca tcttccac tcatgtctg 11700
 ttctgtga cgaagctac actgttga ggaagctga tcttctgt cctcaagtc 11760
 gatcagcac aggtcttcc tctatgca caggaacaa tcttctgg cgtccaca 11820
 tctcagctc gattgtgca gattctgca ctgaggaag gttgtgtgc cagatcga 11880
 ctgttatt ctggtatg tctctgctt gcaaaaacc accgaggt tcatgtctg 11940
 actgtctcc cactctcga cgtacagcc atgaggaac ccaacatgaa tcttccac 12000
 cctcagagg gcaagctat tgcaggaac gttgttttc ctctccga cgtcaatgt 12060
 cagatccgc atccgccc cgtctctgt cttccactg agagactgg catgatcag 12120
 atcaaggac ccaagctct caaggtctat tgcgaatgc cgaagaag cctgtccag 12180
 tttaacgag ccgtttctt tatctctga gatctgga agatcgac agaaagttac 12240
 gttcaatgt tgggaagag caagacctg gtaatttcc gttgctaca catctatccc 12300
 aaagagctc aaggtgagat cgaacagat gagggtgtg tcaagctgc tctaatgt 12360
 gttctctcc cagatttgg aagaaggtc accgtgttg tctgtgcaa accgtgtgc 12420
 gttctgagc aaaaacact cgtctctgt ctgaggaac gttctgccc atcaagcaa 12480

cccaagoga ttatcttgc cagagatgc cctcgaaca ctatgggaaa ggttcaag 12540
 aacttcttc gacagcaata cgcagatct taccacagc gataagccc cgtatgaa 12600
 gataatata taaatcatt gatatatta atgogctga tttagagcc tcatactgt 12660
 cggagagag ccaagacag tactcaagg gattacac atccatctc acagacaca 12720
 gctgggaaa gttctatat aacttccc gaatacga gttctgat ttatcaatg 12780
 ggcagagcc gatttgaag acttctgtgt ctcgggtga aatgggttc ttgctcca 12840
 tcaagtga ccatgtct atttgctgt caaagtaaa caagaagca tgaagaatga 12900
 acttgaagtc aaggaattta aatgaacga aactgaatt tgaccagata ttgtctcc 12960
 ggtggagctc cagctttgt tcccttagt gagggttat ttccagctg gctaatcat 13020
 ggtcatagct gttctctg tgaattgtt atccctcac aagttccac acac 13075

<210> 138

<211> 15799

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pZKL2-5n89C

<400> 138

```

gtacgttata atttgaacag tgaaggcta cagtaacaga agcagttgta aacttcattc      60
cgttgattct gtactacagt accccactac gcgccttcog ctgacactgt tcaacccaaa      120
aactacatct gcgtgcgctg tgaaggcta tcactagata catactgtag attctgtaga      180
tgccaacctg cttgtatcat atactacccc ctcccctga cctgcacaag caagcaatgt      240
gacattgata ttgttgctta tctagtcccg aggatgtgaa agccgagact caaacatttc      300
ttttactctc ttgttctcga ccagacctgg cggagattac gccagtatga ttcttgcaag      360
tctgagacaa gccctggaac gccaacatbt atttttcgaa gogagaacaa tgcacacccc      420
cgccacgttc agagatgcac atgatttgtt tticgagtaa cagtaacccc cccccccccc      480
ccaatgaac cagtattact caccacctcc tcattcaag cgttaacctg attacgcgac      540
catcaacgac agcatgaggg gactgcctgt ctgactcaat caaatgacta caaaaatcgc      600
aataatgaag agcaaacgac aaaaaagaaa caggttaacc aatcccgctt caatgtctca      660
cccaaatcca gcaactgttc tcattacctc ctcccctcaa ttccagagtt gcatcagggc      720
ccttgatggc gcgcagcctg cattaatgaa tggcccaag cgcggggaga ggcggtttgc      780
gtattggggc ctcttcctgt tctcctctca ctgactcgtt gcgtcgtctc gttcggctgc      840
ggcgagcggc atcagctcac tcaaaaggcg taatacgggt atccacagaa tcaggggata      900
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaagc caggcaacct aaaaaagccg      960
cgttgctggc gtttttccat aggtctccgc cccctcagca gcatcacaaa aatcgacgct      1020
caagtcaagc gtggcgaaac ccgacaggac tataaaagata ccaggcgctt ccccttgaaa      1080

gctccctcgt gcgtctcctt gttcgaacc tcccgcttac cggataacct tcccgcttc      1140
tcccttcggg aagcgtggcg ctttccata gctacgctg taggtatctc agttcgggt      1200
aggtcgttcg ctcaagctg gcctgtgtgc acgaaccccc cgttcagccc gaccgctgcg      1260
ccttatccgg taactatcgt cttagctcca acccgtaag acagantta tgcacactgg      1320
cagcagccac tggtaacagg attagcagag caggttatgt agcggtgtct acagattct      1380
tgaagtggtg gcttaactac gcttaacta gaagacacgt atttgctate tgcctctgc      1440
tgaagccagt taccttggga aaagagttg gtactcttg atccggcaca caaacaccg      1500
ctggtagcgg tggttttttt gtttgaagc agcagattac gogcagaaaa aaaggtctc      1560
aagaagatcc ttgtatcttt tctacggggt ctgacgtca gtggaacgaa aactcacgtt      1620
aagggatttt ggtcatgaga ttatcaaaaa ggtattccac ctgactcctt ttaattaaa      1680
aatgaagttt taaatccatc taaagtatat atgagtaac ttggtctgac agttaaccaat      1740
gottaatcag tgaggcaact atctcagcga tcgtctatt tcttccatcc atagtctct      1800
gactcccgct cgtgtagata actaagatac gggagggcct accatctggc cccagtctg      1860
caatgatacc gogagaccca cgtccacgg ctccagattt atcagcaata aacagccag      1920
ccggaagggc cgaagccaga agtggctctg caactttatc cgcctccatc cagtctatta      1980
attgttgccg ggaagctaga gtaagtagtt cgcagtttaa tagtttgccg aactgtgttg      2040
ccattgctac aggcctcgtg gtgtcacgct cgtcgttttg tatggcttca ttcagctccg      2100
gttcccaacg atcaagcaga gttacatgat cccccatgt gtgcaaaaaa gcgttagct      2160
ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcaatggtta      2220
tgccagcact gataattct ctactctca tgcctccgt aagatgctt tctgtgactg      2280
gtgagtactc aaccaagtca ttctgagaat agtgtatcg gogacogagt tctcttgc      2340
cggcgtcaat acgggaiaat acgcgcacc atagcagaac tttaaaagtg ctcatcaltg      2400
gaaaacgttc ttccggcgca aaactctcaa gcatcttacc cctgttgaga tccagttcga      2460
tgtaacccac tctgtccacc aactgatctt cagcatcttt taacttccac agcgtttctg      2520
ggtgagcaca aacagggaag caaaatgcgc caaaaaagg aataagggcg acacggaaat      2580
gttgaatact catactcttc ctittcaat attatigaag catttaoag ggttatgtc      2640
tcattagcgg atactatbt gaatgtatbt agaaaaataa acaaataggg gttccgcgca      2700
catttccccg aaaaagtcca cctgatcgcg ttgaaatac cgcacagatg cgttaaggaga      2760
aaataccgca tcaggaattt gtaagcgtta atattttgtt aaatttcgcy ttaaattttt      2820
gttaaatcag ctcatttttt aaccaatagg ccgaatccg caaaatccct tataaatcaa      2880
aagaatagac cagatagggg ttgagtgttg ttccagtttg gaacaagagt ccaatttaa      2940
agaaactgga ctccaaogtc aaaggcgcaa aaacgctcta tcagggcgat ggcacactac      3000

```

gtgaaccatc accctaatac agtittittgg ggtcgaggtg ccgtaaagca ctaaatcgga 3060
 accctaanaag gagccccga ttttagagctt gacggggaaa gccggcgaaac gtggcgagaa 3120
 aggaaggga gaaagcgaaa ggaagcgggc ctaggggctt ggcagtgta gcggtcaagc 3180
 tgcgcgtaac caccacaccc gccgcgctta atgcgcgcgt acagggcgcg tocatctgcc 3240
 attcaggtcg cgcacactgt ggaagggcg atcgtgcgg gcctcttcgc tattaagcca 3300
 gctggcgaaa gggggatgt ctgcaagcg attaagttg gtaacggcag ggttttccca 3360
 gtccagactc tgtaaacga cggccagtg attgtatac gactactat agggcgactt 3420
 gggcccgacg togcctgctg gtttcgattt gttctagagg aacgcataa cagtaactat 3480
 aqagaataa cpatattcat ttattaaqt agatagttg qgtagaqgt qtaaaagtg 3540
 ataatagct tagataccac agacacccct ggtgacgaag tactgcagat ggtttccaat 3600
 cacttgccc tgcctgagca gegtgttacc ggcagagcac tgtttattgc totggccctg 3660
 gccctgaca acgttgaga gaggagggg gtccggggc ccgtcaataa agccctccac 3720
 agacagtgcc tggtaacgt cccgaaggg cccttgaggg acgatctct cctgggggt 3780
 gattcgtag agctttcgtt ctlttgacac ctggagaca togggtttc cctggccaaa 3840
 gaagagltta tgcaccaggt tagcaagccc agcgttacc acatgggtt gaccaagcgt 3900
 acaacagag ggcctgtgg ccgttaacct gaggtgatt ccgaacgaa ggcctcggc 3960
 toctccgaga gaagtgcagg tgaacgaat ctggtagtc ggcatactc caataccga 4020
 gtccagcttg gggccgactt gatttaggtt gttgtgtag gactgatat agccattgtg 4080
 gacaagacag tcatccaaag tagcaqtga agagatgta gacgaagat caaagtaat 4140
 taactacat gccggattga gactatgat ggttcocgt gccctatata ctctactaat 4200
 ttgactttg aacgcgaaa taogtttca ggcctcaaa gactccaac tottgtctt 4260
 actaatata ctacccatg ttgattggtt acttgaaag agggcctg ttccattgac 4320
 ccaagtttc togcactctt tggctattg acaacgggg tocatgacc gtcgattat 4380
 cagtcacaaa acccccacat tcatcaatc ccatgactt ttacaagtt ctcaattcca 4440
 tctgcaaat caaatccca tctattcatt cactatata aaacctca tctctactaa 4500
 cactcacaac tcatagaaa acatcgactc agaacacac ctccatgcg ccgttagga 4560
 atcctgagcg tccctgacac agtgaacac accgaattg tgcagtcat tgagggtgga 4620
 atgactgtg cccacacatg taggttagaa gcgtgcaga atccgtctc gttcgacaa 4680
 ctctcgaca gcttgctga cgaaggata gtccaaagc gacattcag gaaagagtg 4740
 atcctgact tggagttga gcccgccagt aagaacatg gcaatggct caccgtagt 4800
 ggaagagtc toccactgag ctctgacca gtcatctga toggttcaa cgtctctc 4860
 ggaactctg accctgcagt tctgtcggg gattcgtcc gacccatga agttgtga 4920
 caagatgaa aagaagtgga ggaaggccc ggtagcagtg ggcacagag gaatggtat 4980

gagcagggag gttccagtga gataccaggg caagaaggcg gttccgaaga tgaagaagc 5040
 tgcataaac aatgcaggg ttcgtaccg tgcagaaag ccgtctctc gcatggtgt 5100
 gacagactcg ggaatggtt cgttgtctg ccttcggaag atgagagag gctgtacac 5160
 cagcgaacg ccgtaggctc caagcagag gtacatgta caggccctga atcgtgaa 5220
 caacttga gactgttg cagcagggtt gttgtggaac acagggaat gttcgtgga 5280
 ctgcgcacc aggtcgagac catgtgatt ggttaggtt tgalgcga tcatgtgaga 5340
 ctgcagcag atccatctg acgatccat gactcgatg ccgtaggcaa agagagcgt 5400
 gacccagggc tttttgtga tggccacct agaggcctg tgcagatg acaggccat 5460
 ctgcctgac atgaatccag tcaagagacc ccacagcacc attccgtag tagccagt 5520
 ccactcgaa aaggcggtga cagcaatgat gccacaggt ccagcagaca atccaggtt 5580
 ggcataccag ttccgacctt taatgacctc tgcatagtt cgttgactt cctgtcaaa 5640
 gggagctcg taggttaga caatgtcctt ggaagttcg tctgcttcg ctgcacgaa 5700
 ctgttgaag accttcagt tctcggtctt gacgtaaag tcatgagat acaacagag 5760
 agaagctcg gagccaccg aagcgaggat caagtggat ccgggatgga ccttggaag 5820
 accctcaga togtagagaa tgcctcgat ggcacacag togggtcgt ccagagcgt 5880
 ctggtagta agggagagag ccatgttgtt gaattagggt ggtgagatg gttgttgt 5940
 gggagaatc aaaggccgtt ctgggatcc gttgggtatat atatatatat atatatatac 6000
 gatcttgt taactcctt ttctcaaaa tgtgttttt cgtttttgt ttttgttt 6060
 ttttgttt tttagggcca actaagcttc cagatttgc taatccctt tctactaat 6120
 acaagaagg aagaagtgga tttagttgg gctttttat caactgtct actccttcc 6180
 tctgatata aagttagac caatccatc catgtcattt agagttgta atactggag 6240
 gatagataa gcaacaaaac gagccatag agcactgtg ggtgtagcca agcagaaga 6300
 agtagatgg agccaatga ccagcgaggg agtcaagcca atccgacata ccacacgtg 6360
 agatgtctt gccgggggg taactacaga tgtccaaagg taagtgttg actgtaatg 6420
 tatgtctag gcaaatatg tagtcagcc tataaagta taacaggac cagtgcctc 6480
 atcgaccaa taactctta tgatacctg ctccgtattt attgaacat gctgtcttt 6540
 gttacatac tabcttgctt ttttctcca gaaactcag aattctctt cttagcttt 6600
 tccatacaa gttctctgc ctccaggaag tccatgggtt gttgtcat ggttttgtg 6660
 tagtgtagt gactgtgtg tattgtact gggagtag ttggaataa gtcatacaa 6720
 agtcagctt ctccagctt cataaagta taagtgtt aacgtattg cactgtacc 6780
 agcatctcg tatcgagaaa cacaacacaa tgcacattg gacagctat gggatcac 6840
 aggttgtca gtaataca taactgata gacagtgct ctgacacaa taacagtg 6900

acaagcgctc catacttgca cgcctctctat atacacagtt aaattacata tccatagctc 6960
 aacctctaac agttaacttt ctggttaagcc tccacagcag ccttctggta tgccttggcc 7020
 tctccaatag gatctcgggt ctggccgtac agacctcgcc cgacaattat gatctcggtt 7080
 ccggttagaca tgcactcttc aacagttcgg taactgctgtc cgaagcgctc tcccttgtcg 7140
 tcaagaccga ccccgggggt cagaataagc cagtcctcag agtcgcctct aggtcgggtc 7200
 tggccaatga agccaacac aaactcgggg tggatcggg caagctcaat ggtctggttg 7260
 gagtactcgc cagtggccag agagcccttg caagacagct cggccagact gagcagaact 7320
 ctggccagct tctcgttggg agaggggaat aggaactcct tgcactggga gtctcgttag 7380
 tcaagacgtc cctccttctt ctgttcagag acagttctct cggcaacagc tggcagacca 7440
 gcaatgatic ccgttccggg taccocgtgg cgttctgtga tatcgacca ctggcgatt 7500
 ccgttcaccc ggtactgggt ctggacagtg ttgccaatat ctgcgacctt tctgtcctcg 7560
 aacagggaag aacctgtctt aagcgcaagt tcttgaggg ggaagcagct gccggcgtag 7620
 gtgaagtctt caatgatgc gatctgggtt ttgatctgc acaataaag tccgacctta 7680
 tggccaagct caatgagctc cttgtgggtg gtaacatcca gagaagccaa caggttggtt 7740
 ttcttggctg ccccgagctt gagcactcga gcggcaaggg cggacttggt gacgttagct 7800
 cgaagcttct agaggggcat ttgtgtgggt aagaggagac tgaataaat ttactctgca 7860
 gaacttttta tcgggaactt atctggggga gtgaagtata tgttatggta atagttaaga 7920
 gttagttaga cttatagata gaactgaata taccgatata ggtccaaatt agaaagaagc 7980
 tcaatggctc tctggcgctc gcctttgccc acaaaatgt gatactgat aaagccagca 8040
 atgacgttgc agctgatatt gttctcggcc aaccggccgc aaacgcagc tctcagacca 8100
 aacgcctcca acgaagaatg tatgtccaaa gtgatccaa ccaactcata gtggagtcg 8160
 tactccaaag gcgggaatga cgaagtcagc agtaactgt cgaacttttc ctgggaaccc 8220
 acaacgctca gcccttctga ctacgtatt gtgcacccg acaagggcaa cagtcctggtg 8280
 atagcagaat atgtctgttc ggtccatttc tcaacaactt taggcgtcaa gtgaatgttg 8340
 cagaagaagt atgtgccttc atgagaatc ggtgttctgt atttcaata agtcttgcga 8400
 tcagtttgc cagtcctgtt gtgggggcta atggatga gttatcgctt ccagctctga 8460
 caggtatcat cgtgcgccac ttataacttt ttgattccgc tgcacttga gaactgtctg 8520
 ttaccaaaag tgaagtgtgt cccagagaca ccccccaggg tatggttagc caaaaaataa 8580
 acaactccat accgggaatc gaaccccggt ctcaaggtt ctcaagaagt attcttgatg 8640
 agagcgtatc gatcgaggaa gaggacaagc ggcctgttct taagtctgtt acatcagtat 8700
 ccaaggcacc attgcagga tccaagctt tgaacccgtc atttgcatt cgttaacgtg 8760
 gtgaacaggt tgatcgttcc cctacggcct ccaactgtgt caactcttc aagtcgctg 8820
 actatcagga cattgatcaa cttcggaaga aacttttcta tgcacttcga tccatgtgtg 8880

 gtttcgattt gtcttagagg aacgcatata cagtaactat agagaataaa cgtatattcat 8940
 ttattaaagt agatagttga ggtagaagtt gtaagagtg ataatagcg gccgtctact 9000
 gaactctttt ggtccctctg tcttttcgga cgaatagagt ctgcacgtag aagttgagga 9060
 acagaccacg gacagtacca aagtagnaag agttgaaaaa cccgccaacc attctaatc 9120
 catcttctgc gtacagggga atgtcccggt actccagac gatgtagaag ccaagctga 9180
 actgaatgat ctgcatagaa gtaactcaggg acttgggcat agggaaacttg agcttgaca 9240
 gtccgggtcca atagtagccg taactgatcc agtgaatgaa gccgttgagc agcaaaaaga 9300
 tccaaacggc ttgctttagg tagtttaga acagccacat gtccatagga gtcctagagt 9360
 ggtgaagaa ctgcaacccg gtacagggct tgcccattgag gggcagatag aaggagcaa 9420
 tgaactcgag gaacttctgt aggtagaaca gctgagtggt gattcggaaag acattgttgt 9480
 cgaagccttt ctgcagttg tggacatga caccaatggt gtacatggcg taggcacatg 9540
 agaggaaaga gccacggag tagatggaca tgaagcaggt gtagtctgtg aacacaaact 9600
 tcattcgaga ctgacccctg gttccgagag gaccaagggt gaacttcagq atgacgaag 9660
 cgaatggag gtacagcacc tgcagtgccg aggcactcga ccaagactga gcatagtcca 9720
 ccttgggaag aaactccgtg ccaatggaga cgaattctgt caagacccct atggttgtga 9780
 attagggctg tgaagaaggt ttgtttagg gaagaatcaa aggcctgtct ccggtccgt 9840
 gggatatata atatatata atatataga tctctgttta cctccctgtt ctcaaaaactg 9900
 tggttttctg tttttgttt ttgtttttt ttgattttt tagggccaac taagcttcca 9960
 gattctgata atcaactttg tactaattac aagaaggaa gaagctgatt agagttgggc 10020
 tttttatga actgtgtac tcttatctc tgatagaaa gttagaccc aatcaatca 10080
 tgcactttag agttgtaaat actgggagga tagataagga accaaaaaga gccatagcag 10140
 acatgctggg ttagaccaag cagaagaaag tagatgggag ccaattgagc agcaggggag 10200
 ctacgccaat ccgaacatac acacgtcgag atcgtcttg ccggggggta cctacagatg 10260
 tccaagggtt agtcttgac tgaattgta tgtctgagga caaatatgta gtcagccgta 10320
 taaagtcata ccaggccaac gtgcacatct cgaacacata actctctatg atacatgct 10380
 ccggtattat tgaacatgc gtgcgtttgt tacatacgtt tcttgccttt ttctctcaga 10440
 aacccagac ttgtgtattt ggtcgagata agcccgagcc atagttagtc tttaacatc 10500
 tgtttaaaca ccaataaac ccccaaaat atacttacc gaataaacg atctaatata 10560
 gaggaacat tgcocggag aagaaggcca ggcgcctag atgcaaaatt caaaoactca 10620
 cagctgaact tctgcacatg ccaactaggg gggcctttt tatatggcca agcgaagctc 10680
 tccacgtcgg ttgggtcgca cccacaata antgggtag gttagccaa caaagggtatg 10740
 ggaatggggg tagaagatca caggataacg ggcctcaatg gcaaaaaaa gaacgaatac 10800

tgcattaaag actcgtgac cagcgactga cccatttga tcatctaaag gctcaaaaac 10860
 tacctcgaaa ctgctgogct gatctggaca ccacagaggt tocgagcaat ttaggctgca 10920
 ccaaatgtcc caccaggtgc agcgagaaaa cgtctggaaca gogtgcacag ttigtcttaa 10980
 caaaaagtga gggcgctgag gtccagcaggt gctgtgtgac ttgttatagc ctttagagct 11040
 gcgaagagcc gtatggattt gctcatcag gccagattga gggctctggg acacatgtca 11100
 ttttagtgta cttaaalgc cccctggata tgcctccgac aataggccgt ggctcaatt 11160
 ttttgcttc cgcacttc catctctggg tacccacacc ttgctctcc tgcacttgcc 11220
 aacctaata ctgggttaca ttgaccaaaa tcttanaagc ggggggtctg tctagggtat 11280
 atataaacag tggctctccc aatgggtgc cagctctctt ttctctttct ttcccacag 11340
 attcgaaatc taactacac atccacacat gctgttact gactcttia agcgaaagtc 11400
 cgggtgtcat gtccggcagc atgtccagag cgtcagtcac cagcaacaga tccgtctga 11460
 gacgacggt ttgtgtaat gcacacatcc gaagtcgct agcaacacac actctctaca 11520
 caaactaac cagctctcca tggtaaggo ttctcgacag gctctgccc tctctatga 11580
 cggaaaggtg taccagctct ccgttgggt gaacttccc cctgggtgag ctgaatcat 11640
 tgagaactac caggacagag atgtactga cgtctcatg gttatgcat ctccggaagc 11700
 ctccagaaay ctcaagoga tgcacagat caacaggt tocgagctgc ctcccaggg 11760
 tgcctcaac gaagctcagg aggtttccg aaagctcaga gaagactga tgcacatgg 11820
 catgtttgac gctctcccc ttgttactc gtacaagtc ttgacacccc tgggtcttgg 11880
 cgtgcttgc ttctctatgc tggctcagta ccaactgtac ttcatgtgt ctctctgt 11940
 cgttatgac taccagocaa tggatggct gctctatgac atctgacac accagacatt 12000
 caagacaga aactgaata acgtctggg ttgtgtttt ggcacaggac tccgggctt 12060
 ctccgtgac ttgtggaag acagacacaa cgcctcatat ttgtacaa acgttcagg 12120
 tccagatccc gacattgata acctgctct gctcgtctg tocgaggaay atgtcatcg 12180
 agctctccc atctccagaa agctcatcca gtccacacag tactattccc tggctatctg 12240
 tattctctc cgtactcatct ggtgttcca gctgtgctg accgttcat cctcaagga 12300
 ccagacacac cagttctcc gctctcagta caagaaagc gccattgac tgcctctga 12360
 ctggaactcc aagacccgt tccactctt atttatgccc tcatcctga cctcgatgt 12420
 ggtgtcttt gttccagag tctcgtgtg attcggatc gccatgtgg ttctcatga 12480
 ccactaccc ctggaagaag tctgtatga cgtctggac ggcactgct ttctctggg 12540
 tccagatccat gacacatga acattcagc agcatcatt actgactgt tctttgagg 12600
 ctgaactac cagctcagc accctctctg gccacccct cctgcacaca acutcatgc 12660
 cgttctac cagtggaac agctgtgca gaagacac cctccctacc gaacccctct 12720
 gcccatgaa ggtctgtca tctgtctcg atactgtcc cagttctct gaatggcga 12780
 gaagcagccc ggtgcaagg ctcaatgagc ggcagatga gaagataat atataatac 12840
 attgagatat taactgctct agattagaga gctcactact gctcggagag aagccaaag 12900
 gactactcaa aggggattac accatccata tccacagaca caagctgggg aaaggttcta 12960
 tataccattt ccggaatacc gtatttccg atgttatcaa tggggcagc caggatttca 13020
 ggcactcggg tctctcgggg tgaatggcg ttcttggct ccatcaagtc gtaacctgt 13080
 ttcatltgc tgtcaaatga aaacagaagc agatgaagaa tgaactgaa gtgaaggat 13140
 ttaaatgatc tgaagcagc aggtgtctct gcacgggtct ttttggggg tctgggaaa 13200
 ggggtgttg gatcagaga agcgtgaga accgggtgc ttgtgcttg agatgaagc 13260
 cgtgaacac gggctgttg cggggtttg gggcgtgtg gctccaaga gggtaggca 13320
 ttctgtggg gttactaat tgggcattt ggtctctgc cgcattgccc attgtcaga 13380
 attagtccg ataggagact tatcagccaa tccacagccc ggtacacact gtaggttgg 13440
 ttgggtggga gcccccctcc acagagtaga gtccacacag agcagcaaca tcatagtgg 13500
 ggtgtgtgt gttaaaagaa aaaaaaagc cttgggttat attccogctc tatttagagg 13560
 ttgaggata gacgcagacg gaggcactg gctctatgga accttgogga tatccatag 13620
 ccgggggga ctgctcaga aonagctcca gcaggtttt tccgggcca ttgagcgac 13680
 tgcaccccc ccaagtgctc ttggccacag cactcatgto aigtgtgtg tgggagcca 13740
 ctttttaagt agcacaagc acctagctcg cagcaaggtg tccgaaccaa agaagcgct 13800
 gcagtggtgc aaacggggcg gaacggggcg gaaaagcca cgggggcagc aattgagca 13860
 cgcctcgaa tttagaaga gtacagccc cattcgccc cgaatggct cgcacagcc 13920
 cgtcttttg caccacatca gttacccc agccaaact ttgtgttaa aagttacaa 13980
 tattatacc aagtaggtt tgggggggt tctcctgtc gtccaaagca acattatat 14040
 aaggtctgc atcgcggtc caattgaata tttttctc ttctctctc tatattcat 14100
 ctgaattaa acacacatca accatgggc tattcattaa acaggaagag ctccggctc 14160
 tcaagaagta caagtactcc gccagagtc actcgtctat cccacacac attctgcgc 14220
 cttctggcg acagtgttc aaatcttcc ctctgtggt ggcacacaa atggtgactc 14280
 tctgtgggt ctctttgtc attgtgaat tcatcaaat gctatgtt gatccaccc 14340
 acgacggga gctccacaga tgggtctacc tcaactagc tctgggtctg ttctttacc 14400
 agacattga tgcctgtgac cgtaccatg cccagagaa tggccagag gcaacccctg 14460
 gacagctgt tgaacctgt gtccagcca tgaatactc totgattct acgtgtgtg 14520
 tgcacacac ccataggga tataacatga agctactga tgtcagatt gccgtctcg 14580
 gaacttcta cctgtgacc tggggaact accatacgg aactctgac cttctgtgt 14640
 tctctgtcc tgttgaaggt atctgatto tgtgtgtct ttctgctcc acatttcta 14700

```

ctggtcccaa cgtgtacgct ctgacogtct acgaggctct tcccagatcc atcacttgcg 14760
tgetgectgc cagcttctcg gaactcaaca tcaccacagat ctacatgga ttccgagtcg 14820
tgggctgtgt gttcaacatc tacggcgect gggaaaacgt gatcaagtac tacaacaaca 14880
agggcaagag cgtctctccc gccatctctg gaatgcoccc ctitggcatic ttctacgtcg 14940
gogtctitgc ctgggcccac gtigtctctc tgcctctctc caagtacguc atogtctatc 15000
tgtttgcacat tggggctgoc ttgtccatgc aagtcggcaa gatgatcttt gccaatctcg 15060
tgettgtgoc otttccocac tggaaogtgc tgetctcttt cccctttgtg ggaatggcag 15120
tgcactacat tgcacccctg tttgtgtggg acgcogakat cgtgtcgggt aacactctct 15180
tccactgttt tggcgccaac ctctcaattt acgctctctt tgtgcttgag atcatcgagc 15240
agatccacca ctactcgat atctgtgtgc tgcgaatcaa gtaccctcag gagaagaaga 15300
cogaataagc ggcogaatgc aqctgtgttt ctgactcgat gtittctatg gatttgtgag 15360
tgtagtaga cgtgatgggt ttatatatga tgaatgaata gatgtgattt tgatttgca 15420
gatggaattg agaactitgt aaacgtacat gggaaatgat gaatggggg gtitttgac 15480
tggataactg acgtcgatg gcgcgcgttg ttcaaatc caagagatgc gagaactttt 15540
gggtcaagtg aacatgtctt ctctgttcaa gtaaacatc aactatgggt agtatattt 15600
gtaagagcaa gagttgagat tctttgagat cctagaacag tattttcgcg ttccaaagtc 15660
aaattagtag agtaatacgg gcacgggaat ccaattcatg tctcaatctt gcaggtgagt 15720
taattaatcg agcttggcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc 15780
gctaccaatt ccacacaaac 15799

```

<210> 139

<211> 9641

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pY201

```

<400> 139
gtacgataac ttcgtagac atacattata cgaagtatc gcgtgaaga gtatctgtct 60
gaactgcat tgcgccttt ggagtaagac tccaaatag agtgtcttg gatcaacttg 120
acgataact ctctgttga ggtctgtggc ctgacagctg cgttttcggc goggttggcc 180
gacaaacata tcagctgcaa cgtcattgct ggttttcate atgatacaat ttttgcggc 240
aaaggcgagc ccagagagag cattgcagtt ctctcaatt tggacgata gcgctatag 300
cagctctata tataagttca actaaactgt aactattac ataacatata ctcaactgac 360
ccagataaag ttccgataaa aagtctctga gactaaatt atttcagct ctctctcaac 420
acaaaaatgc cctctcaaga agctcgagct aactccaca agtccgcctt tgcgcctoga 480
gtgctcaagc tctgtgcagc caagaaaaac aaactgtgtg ctctcttga tgttaccac 540

```

```

accaaaggag tcattgagct tgcgataag gtccgaact atgttgtcat gatcaaaac 600
catatcgaca tcattgaaga cttaacctac gcgcgcctg tgcctccct caaggaaact 660
gtcttcaagc acggtttctt cctgttcgag gacagaaagt tgcagatat tgcgaacact 720
gtcaagaccc agtaacgggt tcacgaato gcgagtggt ccgataaac caagcccaac 780
ggtgtaccgg gaacgggaat cattgtggc ctgcagctg gtgcagagga aactgtctt 840
gaacagaaga agagagagct ctctgactac gagaactccc agtacaaga gtctctatgc 900
ccctctccca acgagaagct ggcacaggtt ctgctcatgc tggccagact gcttgcagc 960
ggctctctgg ccactgggga gtactccag cagacactg agcttgcocg atccgacccc 1020
gagtttgtcg ttgcttcat tgcacgaac cgcactaag ccactctga ggaactgctt 1080
attctgaccc cgggggtggg tcttgagac aaaggagag ctctcggaca gcagtacga 1140
actgttgagg atgtcatgtc taacggagc gatataaa ttgtcggcg aggtctgtac 1200
ggccagaacc gagatcatat tgaggaggcc aagcgatac agaaggtcg ctgggaggt 1260
taacagaaga ttaactgtta gaggttagac tatggatat taatttaact gtgtatatg 1320
agagcgtgca agtatggagc gcttgtcag ctgttatgat ggtcagaaga cctgtctgat 1380
cgagtatgta tgatactgca caacctgtgt atcccgatga tctgtccaat gggcctggt 1440
gttgtgttcc tcgatacga gatctgggt acagtgctaa taegttagac taactatact 1500
tatatgagc tcgaagaagc ctgacttggt tatgacttat tctcaactac atccccagtc 1560
acaataccac cactgcata ccactacac aaaaactga tcaaacacac catggacttc 1620
ctggaggcag aagaacttgt tatggaagc ctcaagagag agatcataac ttcttatagc 1680
atcattata cgaagttatc ctgcaggtaa aggaattcag gaggagacgg gttgcgcgc 1740
tatttgtgtc ccaaaaaaca gcccaaatg cccaattga ccccaattg acccaatgac 1800
gggccccacc cggcgagag ccccttccac cccacatata aaactccccc cgtttccac 1860
acttgccttt acggtcgtag gtaactgcag tctgaactct acgttgttc agacttgtta 1920
ctagtttctt tgtctggpca tccgggtaac ccatgcggga cgcataatag actactgaaa 1980
attttttgc ttgtgtgtg ggaactttag caagggtata aaagacacac gtcccgaat 2040
taacttctct ctcttttct ctctctctt gtcaactcac aacggaaato gtttaagcatt 2100
tctcttgag tataagaatc attaccatct gacttctggt aggcagaaga actgtttatg 2160
gaaaagctca agagagagaa gccaaatcac tatcaagaca tgtgtcgcga cttaatttag 2220
atgacagcat ttgcagctg gacaggaat agatggagcg tgtgtctga gtcatgttt 2280
tctatggagt tgtgagtggt agtagacatg atgggtttat atatgatgaa tgaatagatg 2340
tgattttgat ttgcagatg gaattgagaa ctttgttaac gtacatggga atgtatgaat 2400
gtgggggttt tbtgaactga taactgagcg tcagtggagc cagtgttcaa aatatcaag 2460

```


agatgogaga aactttgggt caagtgaaca tgtctctct gticaagtaa aacataaact 2520
 atgggttagta tattttagtaa ggaacaagagt tgaattcttt tggagtctta gaacagtatt 2580
 ttccgtttcc aagatcaaat tagtagagta atccggccac gggatccat tcatagtctc 2640
 aattttccca taggtgtgct acaaggtgtt gagatgttgt acagtaccac catgattoga 2700
 ggtaaagagc ccagaagtca ttgtagggt caagaatac acagatctac agtcaatac 2760
 aatgaatato ttcttccata ttcttcaggt gacacaagc gtgtctatit tccccagaaa 2820
 tgcgcgaaaa ggcgcgtgtg tagcgtggag tatgggttgg gttgggttat cttcatata 2880
 tcagcgaat agtagggcaa gagtagacaa aaagtatcta tatgtagaca gogtaganta 2940
 tqcatttgat tggataaat tcaatttttg cgtgtctcac aaatactctc gataagttcg 3000
 ggttaactg gagatggaac aatgtcgata tctcgacgca tgcgaagtgc ggcacatto 3060
 gccctatagt gagtcttatt acaattcaat ggcgtcgtt ttacaagtc gtgactggga 3120
 aaacccctggc gttacccaac ttaactgact tgcagcaaat ccccttttg ccagctggcg 3180
 taatagcga gaggcocgca ccgtctgcac ttcccaacag ttgcgcgcac tgaatggga 3240
 atggacgcgc cctgtagcgg ccgattgaag cgcgcgggtg tgggtttac ggcagcgtg 3300
 accgtacac ttgcacgcgc cctagccccc gctcttttg ctttttccc ttctttctc 3360
 gccagtttg ccggtttcc ccgtcaagct ctaactggg ggtctcttt aggttccga 3420
 tttagtgtt taaggcaact ccagcccaaa aaacttgat aggtgtatgg ttoactagt 3480
 gggcactogo cctgtagaac ggttttggc cctttgaagt tggagtccac gttctttaat 3540
 agtgaactct tgttccaaac tggacaacaa ctaaacctta totcgtctta ttctttgat 3600
 ttataagga tttgcogcat ttgcctcat tggttaaaa atgaactgat ttaacaaaa 3660
 tttaacgca attttacaa aatttaacg cttaacattt cctgacggt tattttccc 3720
 ttacgcctc gtgcgttatt tcaacgcga tcaagtgga cttttggg aaatgtgga 3780
 ggaaccccta tttgttatt ttctaaata cttcaaaaa tgtatccgt catgagaca 3840
 taacactgat aaatgcttca ataatatga aaaggaaga gtatgagat tcaacattc 3900
 cgtcgcgc ttattccctt tttagcgcga tttagcttc cgtttttgc tccccagaa 3960
 ccgtgtgtg aagtaaaaga tgcgaagct ccgttggtg ccgagtcgg ttacacgaa 4020
 ctgagctca acagcggtaa gatccttgag agttttgcgc ccagaagagc ttttcaatg 4080
 atgacactt ttaaggtct gctatgtgga ccgttatat cccgtatga ccgcgggaa 4140
 gagaactgc gtgcgcgat acaattatt cagaatgact tggtagata ctaacagtc 4200
 acagaagag atcttaagga tggcatgaca gtaagagaat tatgcagtc tgcataacc 4260
 atgagtata acactgcgc caacttact ctgacacga tggaggaac gaagagcta 4320
 accgtttt tgcacaaat ggggctcat gtaactgoc ttgatcgtg ggaacggag 4380
 ctgaatgaag ccaataccaa ccagcagcgt gacacacga tgcctgtaga aatggcaaa 4440

acgttgccca aactattaac tggcgaacta cttactctag cttcccgcca acaataata 4500
 gactggtgg aggcgataaa agttgcagga caactctgc ctcgcgcct tccgctgga 4560
 tggtttatt ctgataaatc tggagccggt gacgtgggt ctcgcgctat catgcaaga 4620
 ctggggccac atgtgaagcc ctcocgtatc gtatgtatct acacagcggg gatcagga 4680
 actatggat acagaaatag acagatcgt gagataggt cctcactgat taagactgg 4740
 taactgcaag acaagttta ctaataata ctttagattg atttaaaact tcaatttaa 4800
 tttaaaaga ctatgtgaa gatccttttt gataactca tgacaaaaat ccttaactg 4860
 gagtittgt tcaactgac gtaacacccc atgaaaaga tcaagagatc ttcttgagat 4920
 ccttttttc tgcgcgtaat ctgcgtctg caaacaaaa aacacacgct accagcgtg 4980
 gttgtttgc ccgataaaga gtaacaaact cttttccga aggttaactg cttcagcaga 5040
 ggcagatac caataactgt tctctagtg tagcgtagt taggcaccca cttcaagaa 5100
 tctgtagac ccctacata cctcgtctg ctaactcgt taccagtgg tctgcccagt 5160
 ggcataagt cgtgtttac ccggttgac tcaagacat agttacccga taagcgcag 5220
 cgtcgggtt gaacggggg tctgtcaca ccgcacagt tggagcgaac gactacacc 5280
 gaactgagat acataacgc tgagctatga gaagcgcca ccgttccga agggagaaag 5340
 ggcgcaggt atccggaag ccgcagggtc ggaacaggag agcgcacgag gtagcttca 5400
 gggggaacgc cctgtatct ttatagctct gtcgggttc gccacactg acttgagcgt 5460
 cgtttttgt gatctctgc aggggggcg agcctatga aaacgcacg caacgggccc 5520
 tttaacgtc tctggcctt ttgtggcct ttgtcaca tgttttcc tgcgttacc 5580
 cctgattct tggataacgc tattaacgc tttagtgag ctgataccgc tgcgcagc 5640
 cgaacgaac agcgcagcga ctaagtgaag gaggaaagg aagagccccc aataacaaa 5700
 ccgcctccc ccgcgcgtt gccgattcat taatgcagct ggcgcgcac caatacaat 5760
 tctgaaaagc acatctgat cctcctatg ccgggagtc aaagctggtc ttattccccc 5820
 gaatttccc ctaactctc ttccagacgc acccgacac agtgcctaac ccgttccga 5880
 aactctacc cagatagct ccaacggact ggcctgata gatgtatcc tgggttga 5940
 gaattgata aaagccggcc aaaaaaaag cggaaaaaa cggaaaaaa gaaaaaaa 6000
 atcgcaaat ttgaaaaata ggggaaaag ccgcaaaac gcaagaggg gggagtatat 6060
 gacactgata agcaagctca caacggttcc tctatttt ttctcactc tctgcctagg 6120
 ttcccaaat ccaagatgct tctctccagt gccaaaaagta agtacccccc aggttttcg 6180
 ccgaaattc caagtgaagc aagctcgtg ggggtgtta aatgtgggg ggggaaaca 6240
 ggacaagag cttctgtgg agcgaatga gacacaaaag ccggcgggtg tgaataaggc 6300
 atttttccc attttccct cctcgtctc tccagcgtg atggcgtgt gctctctca 6360

```

ttttttttt ttttttttt ttttttttt ctgaactacg atttggtttg attttctcaa 6420
ccccacaaa ataatgtcgg gongaggaa atatatatac aaggacacag tggcctgtg 6480
gacaaacagt cactacotct acgtacaca ccgtacgata gttagttag acaatogat 6540
agttggagca agggagaat gtatgtgtg aaagactcac tatggtccgg gttatctcg 6600
accaatagcc aaagtctgga gttctcgaga gaaaaaggca agatacgtat gtaacaaagc 6660
gacgcatggt acataatcac cggagggaag tatcatagag agttagtgt togatgagg 6720
cactggtgac tggtagact ttatacgggt gactacatat ttgtctccag acatacaatt 6780
anagccaagc aactaacctt ggaactctgt aggtaccccc cggcaagac gatctcagcg 6840
tgtctatgt cggattggcg tagotccctc gctcgtcaat tggctcccat ctactttctt 6900
ctgcttggt acacccagca tgtctgtat ggtcgtttt cgtgccttat ctatctccc 6960
agtattacca actctaatg acatgtgtg attgggtota cacttcata tccagagata 7020
ggagttagcc agttgataa aaagcccaac tctatcaga ttcttcttt ctgtatata 7080
gtcaaaaggt gattagocaa atctgggaag ttagtggcc ctasaaaaa caaaaaaagc 7140
aaaaacgaa aaacgaaaaa ccaactgttt ggaacaggg aggtacgaa ggtatctata 7200
tatatatata tatatatata cccccggalc ccgagacggy cctltgatic ttcccacaa 7260
ccaacattc tcccccacct aatccacaa ccgtacaaac cccgtggagc cagtgttag 7320
taagattatt acaaactacg gaattgctc tttaacctg cgtatgcaa ttgtctgtt 7380
gggactttt cctctaacg ctattctgaa cgggattcct gaaaagcga tggcctgaa 7440
gtgtgtttt atcatttota tctccatgtt ttatctctc cggctctga atctctgag 7500
cggatttga accctctca ttcccaaat gttcaatac ctatctctc gttctacg 7560
atcaagtt atgcacac tcaactcct gttctcagc ggcacattg ctatcaaca 7620
cattcatgt cagtcttga acgacacaa tcaacagac gtcatatta cactctoga 7680
gatgtctg gtatgaagc tgacaagctt tgcctggtt tactatgac gttctgtac 7740
gagcgatgc gacttcaag acctacaga acacagaag tcccgagcg tccgaggca 7800
tctccctt ctgaatttt tggcttaag cttttctac tctaccttc tccaggtcc 7860
ctccttgat taactgatt tgcctcttg gctgaactgc gaattgtcc gggacttcc 7920
cggctcaag aaacccatg gaagacata tctgtgtgag cggctcga ttcccaaga 7980
cggcaagtc gctctgga aggtgtgaca ggcctcgcgc tggatgattc tgagcaagtt 8040
gggtatgaag cacttcccg tgaagtaagt cctggacag gaggatttc ctacccgttc 8100
ctttatttc cgtattcatt atctgtttt gctggattc atccacgac ttaagtatta 8160
cgtcgttg agattagc aggttctg ctctctgt ggtctgtt ataatgata 8220
cgtatcaag acccagaaga tccgtggga tccagtcgc aatatgata ttggacgt 8280
ggagactgca caaaacccc gagatgct ggaagcgtg acatgata ctacaaatg 8340

gctgaagat agcgtgtac ttatagtgac taagaagggt aagaagcag gtttctgac 8400
taactgttt acctctctga cctccgctt ttggcaaggt accgtctctg gatactact 8460
taactctga actggtgccc tgtaccaa ccgtggaaag atctatagac gaaacttctg 8520
tccatcttt ctgagagaag atggcgtgac acctctcccg tccaaaga tttacgaact 8580
ggtcggcatt taactatata agctgctt tggttacat gttcaacct taattact 8640
tgactgaag cctctcttta tggtttggg atccgtgtat ttctaactgc atattattg 8700
ggcctctct tctttctgt tccaggaac atacgttaag caggttaact aattttcaa 8760
aagcaagcaa ccgaaggaga tcttctccg aaagcagaag aagttgaaa aagacatctc 8820
tgctcttcc cccaactcg gaggtattt taaggcaaaa atcgacatg agaagggaaa 8880
gagcgacag gaggaagaga tgaacttggg cactccacc atcgaaactg agaagtgga 8940
caacgcaag gaggaactgg aggtattctg caaggaactac aaggatgccc ggaacaagaa 9000
cggactggaa attgaagagg gaactctgc caaggcctc gagcgattta agcaggaatt 9060
ttccaaact cgtctgggt ctggtgaac ggttcgaaa atgtcttct cgggatattc 9120
tctaaaacc atctcaga aagaagaata ggcggcogca tgagaagata aatatataa 9180
tacattgaga tattaatgc gctagattag agagctcat actgtcga gagaagocaa 9240
gacgaactat caaaggggat taacccatcc atatccacag acacaagctg gggaaaggt 9300
ctatatcac ttccggaa acctgattt ccatgttat caatggggg agcagagatt 9360
taaggcaact cgtgtctcg ggtgaaatg gctgtcttg cctccatcaa gtcgaacct 9420
gtttcattt gctgtcaaa gtaaaacaga agcagatgaa gaatgaact gaagtgaag 9480
aatttaaat taacgaact gaattttgac cagatattgt gtcgpggtg gagtccagc 9540
tttgttccc ttatgtgag gtttaatttc agcttggtt aatcatgtc atagctgtt 9600
ctgtgtgaa attgttatcc gctcaacagc ttccacacaa c 9641

```

<210> 140

<211> 34

<212> DNA

<213> Escherichia coli

<400> 140

ataactctgt ataatgtat ctatcgaag ttat 34

<210> 141

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer 798

<400> 141

ctaattcaca accatggcct ttcatgggc agataaglgg 40

<210> 142

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer 799

<400> 142

ctatcgggc cgttacttg gcttgaagg tgcct 36

<210> 143

```

<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 800

<400> 143
cgatagttag tagacaacaa tcgatgttg gagcaaggga      40

<210> 144
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 801

<400> 144
gaaaggccat ggttglgaat tagggtgglg agaattg      36

<210> 145
<211> 9320
<212> DNA
<213> Artificial Sequence

<220>
<223> Plasmid pY168

<400> 145
gtacgataac ttcgataga atacattata cgaagtatc gctgtaaga gtatctgtct      60
gactcgtcat tgcgccttt ggaagtcagc tccaactatg agtgcgttg gatcactttg      120
acgatacatt cttcgttggg ggtctgtggg ctgacagctg cgttttcggc gctgttgacc      180
gacacaataa tcagctgcaa cgtcatgtct ggttttcata atgatacat ttttgcggc      240
aaagcgagac cccagagagc cattgacgtt ctttctaatt tggacogata gccgtatagt      300
ccagttatct tataagttca actaacctgt aactattacc ataacatata ctccactgcc      360
ccagataaag ttccgataaa aagtctctga gactaaattt atttcagttt cctcttcacc      420
acaaaatgca cctctacaga agctcgagct aactgcacaa agtcgcctt tgcgcctoga      480
gtgctcaagc tcgtggcagc caagaaaacc aactctgtgt cttctctgga tgttaccacc      540

acaaaggagc tcattgagct tgcgataag gtcggacott atgtgtcat gatcaaaacc      600
catatcgaca tcattgacga cttcacctac gccggcactg tgcctccctc caaggaaact      660
gtctttaagc acggtttctt cctgttcgag gacagaaagt tcgagatat tggcaacact      720
gtcaagaccc agtaccggtg tcaacgaatc gccagtggtt ccgatatacc caacgcacac      780
ggtgtacccc gaaccggaat cattgtctgg ctcgcagctg gtgcogagga aactgtctct      840
gaacagaaga agggggaagt ctctgaactc gagaactccc agtaacagga gtctcagtc      900
ccctctccca acgagaagct gccacgaagt ctgctcatga tggcogagct gtcttgcaag      960
ggctctcttg ccaatggcga gtactcaag caagccattg agcttgccc atccgaaccc      1020
gaatttatcg ttgccttcat tgcacagaac cgaactaagg cgaactctga cgaactgctt      1080
attctgaccc cgggggtggg tcttgacgac aaggagagcg ctctgggaca gcagtacoga      1140
actgtgaggt atgtctatgc taacggagag gatatacaa ttgtgcggcg aggtctgtac      1200
ggccagaacc gagatootat tgaggaggcc aagcgatacc agnaggctgg ctgggaggt      1260
taccagaaga taaactgtta gaggtagac tatggatatg taatttaact gtgtatatag      1320
agagcgtgca agtatggagc gctgttcag cttgtatgat ggtcaagaga cctgtctgat      1380
cgaatatgta tgatactga caactctgt atcccatga tctgtccaat gggcatgtt      1440
gttgtgttct tcgatacaga gatgtgggt acagtgttaa taacttgaac taattatact      1500
tatatgaggt tcgaagaaag ctgacttgtg tatgaattat totcaactac atcccaagtc      1560
acaataccac caactgaact caactaaccc aanaacttga tcaaacccac catggacttc      1620
ctggaggcag aagaacttgt tatgaaaaag ctcaagagag agatcataac ttcttataga      1680
atacattata cgaagtctac ctgaagttaa aggaattcag caagaccccg gttgcgcggc      1740
tatttgttc ccaaaaaaca gcccaattg cccaattga cccaatttg acccagtcg      1800
gggcccaacc cggcgagag ccccttcac ccaacatata aaactccccc cgttcccaac      1860
acttgcctgt aaggcgtag ggtactgag ctggaattct aogctgttc agactttgta      1920
ctaagttctt tgtctggcca tccgggtaac ccatgcccga cgcataatag actactgaaa      1980
atctttttgc tttgtggtg ggaactttag caagggtata aaagaccacc gtcccgaat      2040
tacctttcct ctcttttct ctctctcct gtcaactcac acccgaatc gttaaagcatt      2100
tctctctgag tataagaatc attaccatg gacttctcgg agggagaaga aactgttatg      2160
gaaaagctca agagagagaa gccaaatcac tacaagaca tgtgtcgcaa ctaataaag      2220
atgacgacat ttgcagctg gaacaggaat agatggagcg tgtgtcttga ctgatgttt      2280
tctatggagt tgtgaggttt agtagacatg atgggtttat atatgatgaa tgaatagatg      2340
tgattttgat ttgcacgatg gaattgagaa ctttgttaac gtacatggga atgtatgaat      2400
gtgggggttt tgtactgga taactgacgg tcaatggagc cgttgttcca aatataccag      2460
agatcgagga aactttgggt caagtgaaca tgtctctct gtccaagtaa accataact      2520

```

atgggttagta tatttagtaa ggaagaaggt tgagattott tggagtoota gaaaagtatt 2580
 ttocggttcc asgatcaaat tagtagagta atacgggcac gggatccat tcatagtctc 2640
 aattttccca taqgtgtgct acaagtggtt gagatgtggt acagtaccac catgattcga 2700
 ggtaaagagc ccagaagtcn ttgatgaggt caagaataac acagatctac agtcaaatc 2760
 aatgatatcc ttctttcata ttcttcaggt gacaccaaag gtgtctattt tcccaagaaa 2820
 tgcgtgaaaa ggcgcgtgtg tagcgtggag taigggcttg gttggcgtat ccttcataa 2880
 tcgacgaaat agtagggcaa gagatgacaa aaagtatcta tatgtagaca gctagaata 2940
 tggatttgat tggataaaat tcaattatg cgtgtctcac aaatactctc gataagtgg 3000
 ggttaactg gagatggaac aatgctgata tctcgagcga tgcgaactg ggcacaatc 3060
 gccctatagt gagtgtatt acaattcact ggcgtctgt ttacaacgtc gtgactggga 3120
 aaacccctgg gttacccaac ttaactgcct tgcagacact cccctcttgc ccagctggcg 3180
 taatagcga gaggcccgca ccgatgcgcc ttcccaacag ttgcgaagcc tgaatggcga 3240
 atggacccgc cctgtagccg ccgattaaag ccggcgggtg tggtagttac gcgcagcgtg 3300
 accgtcacac ttgcagcgc cctagcgccc gctcctttcg atttcttccc ttctttctc 3360
 gccaggttcg ccggctttcc ccgtcaagct ctcaactggg ggtctccttt aggttctga 3420
 tttagtgctt taaggcaact cgaacccaaa aaacttgatt aggttgatg tcaagtatg 3480
 gggccatcgc cctgatagac ggtttttcgc ccttgacgt tggagccac gttctttaat 3540
 agtggactct tgttccaaac tggacaacca ctcaaccta tctcgttata ttctttgat 3600
 ttataagga ttttgccgat ttggcctat tggtaaaaa atgacgtgat ttaacaaaa 3660
 tttaacgcga attttacaaa aatattaacg cttaaatlt cctgatggg tattttctc 3720
 ttacgcatct gtgcgtatt tcaacccgca tcaagtgcca ctttggggg aaatgtgcgc 3780
 ggaaccccta ttgtttatt ttcttaata cttcaataa tgtatccgt catgagacaa 3840
 taacccgat aatgttcca ataatttga aaaaggaaga gtatgagat tcaacattc 3900
 cgtctgcgcc ttattccctt ttgtgcgca ttttgcttc cgtttttgc tcaacagaa 3960
 accgtgtcga aagtaaaaga tctgaagat cagtgggtg caagagtggg ttacatcga 4020
 ctgatctca acagcgttaa gatccttgg agttcttgc ccgaagaag ttttccaatg 4080
 atgagcact ttaaaattct gctatgtgc gcgtattat cccgtattg cgcgggcaa 4140
 gagcaactgc gtcccgcat acactattct cagaatgact tggtagta ctcaacgtc 4200
 acagaaagc atcttaagga tggcatgaca gtaagagaat tatcagctgc tgcataacc 4260
 atgagtata caactgcggc caacttactt ctgacaacga tggagagac gaaggagata 4320
 accgttttt tgaacaaact cgggcatcat gtaactgcg ttgatgttg ggaacggag 4380
 ctgaatgaag ccataccaaa ccagcagcgt gacacaacga tgcctttag aatggcaaca 4440

acgttgcgca aactattaac tggcgaacta ctactctag ctcccgcca acaattata 4500
 gactggatgg agcgagataa agtgcagga ccaattctgc gctcggcctc tccgctggc 4560
 tggtttatg ctgataaact tggagccgt gagctgggt ctccggtat cattgcaga 4620
 ctggggccag atggttaagc ctcccgatc gtgttatct acacgcggg gactagcca 4680
 actatggat aacgaatag acagatcgct gagatagtg cctcactgat taagcatttg 4740
 taacgtcag acaagttta ctcatatata cttagatg atttaaatc tcatlttaa 4800
 tttaaaagga tctaggtgaa gatccttttt gataatctca tgaacaaact ccttaactg 4860
 gagttttgt tcaactgagc gtcacacccc gtgaaaaaga tcaaaagata ttcttagat 4920
 cttttttt tgcggtaat ctgtgctg caaacaaaa aacacacgt acagcgggtg 4980
 gtttgttgc cggatcaaga gtaaccaact cttttcga aggttaactg cttaacgaga 5040
 gcgcagatc caaatactgt tctctagt tagcgtagt taggcacca cttaagaac 5100
 tctgtagcac cgcctacata cctcgtctg ctactcgt tacaagtgc tctgtccgt 5160
 ggcataagt cgtgtctac cgggttgga ctaagacgat agttacgga taaggcgag 5220
 cgttcgggt gaacggggg ttctgtgaca cagccagct tggagcgaac gactacacc 5280
 gaactgagat acctacagc tgaacttga gaagcgcca cgtttccga agggagaaag 5340
 gggacaggt atccggtaa cggcaggtt ggaacaggag agcgacagag ggccttcca 5400
 gggggaacg cctgttatct ttatgtct gtccggttc gcaacactg acttgagct 5460
 cgtattttgt gatgctgc agggggcggt agctatgga aaaaacgag caacgcggcc 5520
 ttttaaggt tctggtctt ttgtgagct ttgtctaca tgttcttcc tgcgttacc 5580
 cctgattctg tggataacc tattacccc tttaagtga ctgataccc tgcgcgaag 5640
 cgaacgacg agcgacgga gtcgtgagc gaggagcgg aagagcgccc aatacgaaa 5700
 ccgctctcc ccggcggtg gcgattcat taatgaagt ggcggccac caatcaaat 5760
 tctgaagaag acatctgat ctctccttg cggggagtc aacgtggtc ttattccccc 5820
 gaatttccc ctcaactcgc ttacagacg acccgacac agtcttaac gcggttcca 5880
 aactctacc cagatatct ccacgcact ggcctgata gatgtatcc tgggttga 5940
 gaactgata aaagccgccc aaaaacaaag ccgaacaaag ccgaacaaag gagaacaaa 6000
 ttcgcaaat ttgaanaa gggggaag agcgaacac gcaaggggg ggggtatat 6060
 gacactgata agcaagctca caacggttc tttattttt tctctatct tctgctagg 6120
 ttcccaaat ccagatgct tctctccagt gccaaaagta agtaacccac aggttttcg 6180
 ccgaanaac cactgcagc aactcgtgt ggggtgttaa aatgtgggg gggggaaca 6240
 gacaaagag ctcttgagg agcccaatga gacacaaa cggcggggtg tctaaaggc 6300
 attttgcc atttccct ctctgtctc tccagcgtg atggcgtgt gcgtctcta 6360
 ttcttttta ttctttttg ttattttct ctgactaacg atttgggttg atttctca 6420

```

ccccacacaa ataatgctgg gcggaggcat atatatatac acggacaacg tgcgcctgtg 6480
gacacacagt cactacctct acgatacaca ccgtacgata gttagtagac aacatcgat 6540
agttggagca agggagaaat gtatagtggtg aaagactcac tatggtccgg gcttatctcg 6600
accatagccc aaagtctgga gtttctgaga gaaaaaggca agtaactgat gtaacaaagc 6660
gcgcctaggt acataatac cggaggcctg tatctatagc agttagtggt tgcgtatggg 6720
cacgtgtgcc tggtagtact ttatacggct gactacatat tigtcttcag acatacaatt 6780
acagtcaagc acttaccctt cgaactctgt aggtaccccc cggcaagac gatctcagcg 6840
tgtctgtagt cggattggcg tagtcccttc gctctcaat tggctcccat ctactttctt 6900
ctgcttggtc acacccagca tgtctgctat ggtctgtttt cgtgccttat ctatctccc 6960
agtattacca actctaaatg acatgatgtg attgggtcta cactttcata tcagagataa 7020
ggatagcacc agttgcataa aaagcccacg tctaatsago ttcttctttt cttgtaatta 7080
gtacaaaggt gattagcgaa atctggaagc ttatgtggcc ctaaaaaat cnaaaaaagc 7140
aaaaaacgaa aaacgaaaaa ccacagtttt cgaacacggg agttaacgaa gcatcgtata 7200
tatatatata tatatatata cccacgggac ccgagacggg cctttgattc ttccctcaa 7260
ccacacatto tcacacacct aattcacaa cctggccttt cctggggcag ataagtggcg 7320
agccgatgcg tctgcctata cagggctgcc tcgggacctc ctcaagattg catioactct 7380
ggtcatgtct tatccgctga gttctctcat gaacggctg ccagatgacg ccaaaaacct 7440
caagatcttc tatatctct cctgtctcat ctctacatg gtgggtgtct tctccctcta 7500
tggcggagct gcaactctgc tcttctcttc aatgggtacc ttcttctata ccaaaagtaa 7560
gagcccttcc atgcccgggg tcaattttgg tttgtctatg accatctct tcttcaatca 7620
cctgcttctc cagtttttcc ccgaacata cgaacccaat gtcattgaca tcacggagc 7680
acagatgggt ctgtgtatga agtatctgtc ttttggatgg acgtctacg atggtatgca 7740
gattgagaag ggtgagcagc tcagcgagtt ccagactaaa agggctgttc tcaagacccc 7800
cagtcctatg gactctctag cttttgtgtt ctactctctc tccatctga cagtccttc 7860
ttaacactat atggagttcc ataactggct cgtatctcagc ctgttcaagg agctggagaa 7920
agataaggac ccaagcgagc ctgctcagcg aaagcgacac aagatccccc gatctggaat 7980
cgctcttccc aagaaactcg ccgtcggtat ctctcgggac gttctgtgga cccaggtgga 8040
ctctcgatcc tcacacgctc acgttacttc agacgccttc accaaggagc acacatctt 8100
tggagaattt gtgtacctct acatgctcgg ttctatgtac cgtactaagt actaaggagc 8160
ctggtcattt tccgagggag cctgcacttt gtctggcttc ggttccatg gcgtggaacc 8220
caaaactggc aagtaacaag cggacccgtg ccagaaactg gaacccgtgg gattcgaaac 8280
tggtcaaaac acaaggcttc tcttggaagg ctggaacagc aacactaaca agtggctacg 8340

```

```

aaactatgtg taactccgag tgggtcccaa aggcocaaaag cctggattcc gageccactat 8400
cttcacattt gtggtttccg ccttctggca tggaaactcga cctggctact atctacactt 8460
tgtcacogct gccatttacc agtctgttgg taagtcttc ccagcatacc tgcgaacctt 8520
cttcctggag tctgatggaa agactcggcg tccctataag atctactacg acattgtgtg 8580
ttggatggtt gtccaaaccc catttgata cgtacccagc tcttttatga ttctagactt 8640
ctggctgtcg ctcaagtgtt ggaagaactc ctggttctcg taaccattg ctctgggcgc 8700
catctttgca atttctagcc cctacaagcg atggcgattt cccaagatca agaaaaagca 8760
ggctggagcc gtcactgaca agaaggacgc caaggaggag gtgaagaagg acaccatcaa 8820
gaccaagtaa gcggccgcat gagaagataa atatataaat acattgagat attaaatgcg 8880
ctagattaga gagcttcta ctgctcggag agaagccagc acagtlactc aaaggggatt 8940
acaccatcca tatccacaga cacaaactgg ggaaggttc tatatacact ttccggaata 9000
ccgtagtctc cgtacttacc aatgggggca gccagattt caggcacttc ggtgtctcgg 9060
ggtgaaatgg cgttcttggc ctccatcaag tctaacatg tcttctattg cctgtcaaa 9120
taaacagaa gcagatgaag atgaacttgg aagtgaagga atttaaatgt aacgaactg 9180
aaatttgacc agatatgtg tcgcgggtgg agtccagct tttgttctt ttagtgaagg 9240
ttaatttcca gcttggcgta atcatggcca tagctgttct ctgtgtgaaa ttgttatccg 9300
ctcacagctt tcacacacac 9320

```

<210> 146

<211> 8726

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pY208

<400> 146

```

gtacgataac ttogtataga atacattata cgaagtatac gogtcaaga gtaactgtct 60
gactcgtcat tgcgcctttt ggaatacagc tccaaactatg agtgtctgtg gatcaatttg 120
acgatacatt cttcgttgga ggtctggggt ctgacagctg cgttttcggc goggttggcc 180
gacaaacata tcagctgcaa cgtcatgtct ggttttctat atgataccat ttttgcggc 240
aaaggcgagc cccagagagc cattgaactt ctttctaatt tggacogata gccgtatagt 300
ccagtctata tataagttca actaactcgt aactataacc ataacatata cttcaactgc 360
ccagataaag ttccgataaa aagtcttgca gactaaattt atttcaactt cctcttcaac 420
accaaaatgc cctctacaga agctcagact aacgtccaca agtccgcctt tgcgcctcga 480
gtgtcacaag tcttggaagc caagaaaaac aacctgtgtg ctctcttgga tgttacaacc 540
accaaggagc tcaattgagt tgcgataaag gtcggacctt atgtgtgcat gatcaaaaac 600
catatcgaca tcaattgaca cttcaactac gcgggaactg tgtccctctt caaggaaactt 660

```

gotcttaago seggtttott cctgtctgag gacagaaagt tocgagatat tggaaacact 720
 gtcaagcacc agtacccgtg toacogaato gccagtggt ccgatatcac caaagccacc 780
 ggtgtacccg gaaccggaat cattgtctgc ctgcagatg gtgcagagga aaactgtctct 840
 gaacagaaga aggaggaagt ctctgaacta gagaactccc agtacagaaga gttcctagtc 900
 cccctcccca acogaagaat ggcagagagt ctgctcatgc tggccagagt gtcttgcaag 960
 ggcctctctg ccactgpgga gtactccaag ccagacattg agcttgcocg atccgacccc 1020
 gagtittgtg ttggtctcat tgcocagaac ccagctaaag ccgactctga ggaetggctt 1080
 attctgacc ccgggttggg tcttgacgac aaggagagcg ctctcgaca gcagtaccga 1140
 actgttgagg atgtcatgtc taccggaacg gatatacaa ttgtcgcccg aggtctgtac 1200
 ggcagaaccc gagatccat tgaggagggc aagcgatacc agaaggtcg ctggaggct 1260
 taccagaaga ttaactgtta gaggttagac tatggatatg taatttaact gtgtatatag 1320
 agagctgtca agtatggagc gcttctcag ctgtgatgat ggtcagacga cctgtctgat 1380
 cgaatatgta tgatactgca caacctgtgt atcccgatga tctgtccaat gggcatgct 1440
 gttgtgttcc tccatccgga gatgctgggt acagtgctaa tacgttgaa cacttatact 1500
 tatatgaggc togaagaag ctgacttgtg tatgattat tctcaactac atcccaagtc 1560
 acaataccc cactgacata cccctacac aaaaacatga toaaacccc catggacttc 1620
 ctggagcgag aagaacttgt tatggaanaa ctcaagagag agatcataac tctgtataga 1680
 atacattata cgaacttact ctgcaggtac aggaattcag gagagacogg gttggcgcg 1740
 tatttctgtc ccaaaaaaca gcccaattg ccccaatga ccccaattg acccaatgac 1800
 gggcccaacc ccggcgagag ccccttccac ccccatatc aaactcccc cggttccacc 1860
 acttgccgtt aaggcgtag ggtactgca gcttgaaact accgttgttc agactttgta 1920
 ctgcttctct tctctggcca tccgggtaac ccatgcgcga cgcnaaatag actactgaaa 1980
 atttttttgc ttgtgtgtg ggaacttagc caagggtata aaagacccc gtccccgaat 2040
 tacccttctc cttctttctt cttctctctt gtcactcac accgaaatc gtttagcatt 2100
 tccctctgag tataagaact cttaaccatg gacttctgg aggcagaaga acttgttatg 2160
 gaaaagctca agagagagaa gccaaagata tatcaagaca tgtgtcgaca ctttaataag 2220
 atgagacat ttgcgagctg gacgaggaat agatggagcg tgtgtctga gtcgatgct 2280
 tctatggagt tctgagtgct actagacatg atgggttat atatgatgaa tgaatagatg 2340
 tgattttgat ttgcacgatg gaattgagaa ctttgttaac gtacatggga atgtatgaat 2400
 gtgggggttt tctgactgga taactgacgg tcaatggagc ccgttgttca aatatccaa 2460
 agatcgagga acctttgggt caagtgaaca tctctctctt gttcaagtaa acaataaact 2520
 atgggtagta tatttagtaa cgaagaagt tgagattct tggagtccta gaaagctatt 2580

ttccgcttcc aagatcaaat tagtaggta atacgggccc ggaatccat tcatagtctc 2640
 aattttccca taggtgtgct acaggtgtt gagatgtgt acagtaccac catgattoga 2700
 ggtaaagagc ccgaaatgca ttgtgaggt caagaatac acagatctac agtcaatac 2760
 aatgaatctc ttctttcata ttcttcaggt gacaccaagc gtgtctattt tccacagaaa 2820
 tgcgtgaaaa ggcgcgtgtg tagcgtggag tatgggttgc gttggtatct ccttcata 2880
 toacagaaat agtagggcaa gagtagcaaa aaagtatota tatgtagaca ggttagata 2940
 tggatttgat tggataaat taatttatgt cgtgtctcac aaatactctc gataggttg 3000
 ggttaactg gagatgaac aatgtcata tctgcagcca tgcagctog ggcacatto 3060
 gccctatgt gactgtatt caactcaat ggcctgctt ttaaacctc gtgactgga 3120
 aaacctgpc gttacccaac ttaactgcct tgcagcaact ccccttttgc ccagctggcg 3180
 taatagcga gaggcgcga ccgctgcacc tccccacag ttgcgcagcc tgaatgoga 3240
 atggcgogc cctgtagcgg cgaattaaag gggcggggtg tgggtgttac gogagcgtg 3300
 accgtacac ttgcagcgc ctagcgccc gctcctttgc ctttttccc ttctttctc 3360
 gccagcttgc ccggtttccc ccgtcaagct cttaactggg ggtctcttt aggtttcoga 3420
 tttagtgtt tacggcaact cgaacccaaa aaacttgatt aggttgatgg ttacgttgt 3480
 gggccatogc cctgatagac ggtttttgc cttttgaagt tggagtcac gttctttaa 3540
 agtggactc tgttcaaac tggacaacaa ctcaacotta totcggtota ttutttgat 3600
 ttataaggga tttgcgat ttgcgctat tggttaaaa atgagctgat ttaacaaaa 3660
 tttaacoga attttaacaa aatattaaag attacaattt cctgatcgcg tattttctc 3720
 ttacgcactc gtgcgctatt toacacgca tcaagtgcga cttttgggg aaatgtgoc 3780
 ggaacccota ttgtttatt ttotaaata ctttaataa tglatccgt catgagaaa 3840
 taacctgat aaatgotta atantatga aaaggaaga gtaggtat toacatttc 3900
 cgtgcgccc ttattccct tttgcggca tttgcttc cgtttttgc toacccaga 3960
 accgtgtga aagtaaaaga tctgaagat cagtgggtg cagcagtggt ttacatgaa 4020
 ctgatctca ccagcggtaa gctccttga agttttgc ccagaacag ttttcaatg 4080
 atgaactct taaagtctc gctatgtgc ccgtattat ccgtattga cgcggcgca 4140
 gagaactog gtgcgcgat acactattc cgaatgact tggttgata ctocaaagt 4200
 ncagaanaa atcttaogga tggcatgaca gtaaggaat tatgcagtc tgcataaac 4260
 atgagtata acactgogg caacttaact ctgacaaaga tggagggaac gaaggagcta 4320
 accgttttt tgcacaaat gggggtcat gtaactgccc ttgatcttg ggaacggag 4380
 ctgaatgaag cctacccaaa cgcagagct gacacacaga tgcctgtgc aatgcaaca 4440
 acgttgcga aactataac tgggaacta cttactatg cttccggca acaattaata 4500
 gactggatgg aggcgataa agttgacga cactctgc gctgggccc tccggtggc 4560

tggtttattg ctgataaate tggagocggt gagcgtgggt ctgcggtat cattgoagca 4620
 ctggggccag atggttaagc ctccgtatc gtatgtatct acacgacggg gactcaggca 4680
 actatggatg aacgaataag acagatcgct gagataggtg cctcactgat taagcattgg 4740
 taactgtcag acaagttta ctoatatata ctttagattg atttaaaat boatttttaa 4800
 tttaaaagga totaggtgaa gatccttttt gataatctca tgcacaaat cctttaacgt 4860
 gagttttcgt tcaactgagc gtoagccccc gtgaaaaga tcaagagato ttcttgagat 4920
 ccttttttct tgcggttaat ctgctgcttg caaacaaaaaacccacggc accagcggtg 4980
 gtttttttgc cggataaaga gctaccaact ctttttcoga aggttaactgg cttoagcaga 5040
 ggcagataac caaatcctgt tctctagtg tagcgttagt taggcacaca cttoagaac 5100
 tctgtagcac cgcctacata cctcgtcttg cttaactcgt taaccagggc tgcctgcagt 5160
 ggcgataagt cgtctctac cgggttgga tcaagacgat agttacogga taaggcgag 5220
 cgttcgggct gaacgggggg ttctgtcaca cagccacgt tggagcgac gactacacc 5280
 gaactgagat acctacagcg tgaactatga caaacgcga cgtctccga agcgagaag 5340
 ggcacaggt atccgttaag cggcagggtc ggaacagag agcgcacag ggaactcca 5400
 gggggaacg cctggtatct ttatgtctc gtcgggttc gcaactcgt acttgagct 5460
 cgaattttgt gatctcgtc agggggggcg agcctatgga aaacgcgac caacggcgcc 5520
 tttttacgt tctcgtcctt ttctcgtcct ttgtctcaca tgtctcttcc tgcgttatcc 5580
 cctgattctg tggataacgc tattaccgac tttagtgag ctgatacogc tgcgcgagc 5640
 cgaacgcgc agcgcagoga gtcagtgaag gaggagcgg aagagcgccc aatacgaaa 5700
 cgcctctcc cgcgcgttg gcgatctat taatgcagct ggcgcgcac caatcacat 5760
 tctgaagac acatctgat ctctcattg cggggagtc aaaggttggtc ttattccccc 5820
 gaatttccgc ctcaactcgt ttccagacgc acccgacac agtgcctaac gccttccga 5880
 aactctacgc cagatatgct ccaacggact ggcctgata gatgtatcc tgcgttga 5940
 gaattggata aaagcggccc aaaaaaaag cgaaaaaag cgaaaaaaa gagaaaaaa 6000
 atcgcaaat ttgaaaaaa cggggaaaag acgcaaaaa ccaagagggg gcgatgat 6060
 gacactgata agcaagctca caacggttcc tctattttt ttctcactc ttgcgctagg 6120
 ttcccaaat cccagatgct tctctcagt gcaaaaagta agtccccac aggttttcg 6180
 cgaaaattc cactgcagc aaactcgtgt ggggtgttaa aatgtgggg ggcggaacca 6240
 ggacaagag ctcttgggg agcgaatga gacacaaag cggcggggtg tgataaggc 6300
 atttttgccc atttccctt ctctcgtctc tccagcgtg atggcgttg gctctccta 6360
 ttctttkta ttctttttg ttktattkt ctgactacg atttggttg attctccta 6420
 cccacacaa atagctcgg gcgaggaat atatatata accgacacg tgcgctgtg 6480
 gacacacgt cactcctct acgatacaca ccgtacgata gttagtacg acaactgat 6540
 agttggagca agggagaat gttagtgtg aaagactcac tatggtcccg gcttatctg 6600
 accaatgccc aaagtctga gttctgaga gaaaaggca agtatcgtat gtaacaaag 6660
 gacgtatgt acataatcc cggagcctg tatcatagag agttagtgtc tcatgatgg 6720
 cactgtgccc tggtagact ttatacgtt gactacatat ttgtctcag acatacaatt 6780
 aaagcaagc aactacact ggaactcgt agtaacccc aggaacagc gatctcagc 6840
 tgcgtatgt cggattggcg tagctcctc gctcctaat tggctccat ctactttct 6900
 ctgcttggt accacacga tgtctctat ggtcgtttt cgtccttat ctactctcc 6960
 agtatcca actctaatg acatgctg atgggtota cacttcta taagagata 7020
 ggcgtgacg agttgcataa aaagccacac tctaactcgc ttctctctt ctgttaata 7080
 gtacaaagt gattagcga atcgggaagc ttagtggcc ctaaaaaat caaaaaagc 7140
 aaaaaagaa aaagcaaaa ccaactgtt gagaacagg aggtacgaa ggcctata 7200
 tatatatata tatatatata cccacggatc ccgagacgg ccttgatc ttccctaca 7260
 ccaacattc tccacacct aatccacac catgtctatt gtttgtcca acccgtgct 7320
 ctggctgag atctccttc tctactgtt tgcctccca cagctcctg cttctgcc 7380
 tcaagaggt cagttcctg ccaactgtat tgtgtctgt atgcccgc ttatcatgt 7440
 cgttcaggg tctctactc agatcgtgt cgtctctgt gaaagagat agtctaca 7500
 ttaactgtg tccgattgt tctctctct tgcctcaga cgtgtgtgt tgaactata 7560
 gatgttgt gagaacacac tcatataag cctctctatc gttgtctga aactcaata 7620
 ctctatgat atgatgttt tggacagat ttctcaaaag cactcgttg tcatgcgaa 7680
 gaagaaact ctgactttc ccttttggg aatgttatg aaactgaca acgctatct 7740
 catgacgg aagaacacac agaaagcct aggtctaac aacaaagcg tggcggaat 7800
 gaagaaacac aactctgaa tctggtttt cccagaggcg aacggctca gactgaca 7860
 ggcagacgt ctgccttca agaaaggtgc cttctactt gcaattcag cccagctcc 7920
 tattctccc attactcgc aggtctatc cctatctac gactctcga agcgtactt 7980
 ccccgctga gactcaga tcaagctct ggaacccat cctacaactg gctcactac 8040
 tcatgtgtg aacgaactga tggacaagc acgaacact atgctcaga cctgaagga 8100
 gatgattcc cagtattgt cgaacactg tgaatatga tcaacgaca tgcagcaga 8160
 tattgccag tctacagca cagcattgg caactcagc gacgaatta caaaactgc 8220
 taccctaa gataagcgg ccgtatgaga agataaat ataatcat tgaatatta 8280
 aatgcctag attagagc ctactctgc tggagagaa gcaagagca gtaactaaa 8340
 gggattac catctatc cagacacac agctggggaa aggttctata tacccttcc 8400
 ggaatcgt agtttccgt gttactaat gggcagaca ggaattcag cacttcgtg 8460
 tctcgggtg aaatggcct ctgcctccc atcaactgt accatgtct catttgcctg 8520
 tcaagtaaa acagaagcag atgaagaat aactgaagt gaaggaaatt aatgtaacg 8580
 aaactgaat ttgacagat attgtctc cgttgagct ccagctttg ttcccttag 8640
 tgagggttaa tttagagctt ggcgtaatca tggctatag tgttctgt gtgaattgt 8700
 tatcgtctca caacttcca caaac 8726

<210> 147

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer 856

<400> 147

tcaaacaca tglcgttgc atcaagctc g 31

<210> 148

<211> 28

<212> DNA
 <213> Artificial Sequence

<220>
 <223> Primer 857

<400> 148
 tttagcgccc gctactgag tctclgg 28

<210> 149
 <211> 8630
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Plasmid pY207

<400> 149
 gtacgataac ttctgtatag atacattata cgaagtatac gcgtcgacga gtatctgtct 60
 gaactgtcat tgcgcgcttt ggagtaacga tccaaactatg agtgtgcttg gatcaotttg 120
 aagatacatt ctctgttggg ggcgtgtggg ctgacagctg cgttttcggc gcggttggcc 180
 gacaaacaata tcaactgcaa cgtcattgct ggccttccatc atgatacaat tttgtcggc 240
 aaaggcgacg cccagagaga cattgaagtt ctcttaatt tggacogata gccgtatagt 300
 coagttctat tataagttaa actaactcgt aactattac ataacatata ottcaotgcc 360
 ccagataagg ttccgataaa aagtcttgca gactaaattt atttcagctc cctcttcacc 420
 accaaaatgc cctcctacga agctcgagct aacgtccaca agtcgcgctt tgcgcgtoga 480
 gtgtcacaag tctgtgcaga caagaaaaac aacctgtgtg ctctcttgga tgttaccacc 540
 accaaggaga tcattgagct tgcgataag gtccgacott atgtgtgcat gatcaaaaac 600

catatcgaca tcaatgacga ctacacctac gcgggcactg tgcctccctc caaggacatt 660
 gctcttaagc accgtttctt cctgttcgag gacagaaagt tocgagatat tggcaacact 720
 gtcaagcaac agtacccgtg tcaaccgaac gccagtggtt cagatataac caacgccac 780
 ggtgtaccog gaaccggaaat catgctcggc ctgcgagctg gtgcggcgga aactgtctct 840
 gaacagaaga agggagagct ctctgctac ggaactccc agtacacgga gtctctagtc 900
 cctctccca acggaagct gccagaggt ctgctcatgc tggccgagct gtcttgcaag 960
 ggcctctctg ccaatggcga gtactccaa cgaacattg agcttgcccg atccgaaccc 1020
 gagtttgttg ttggttcat tgcacagaa cgaactaagg gogactctga ggaatggctt 1080
 attctgacc ccgggtggg tctgaagac aaggagagag ctctcggaac gaagtaacga 1140
 actgttgagg atgtctatgc taaccgaac gatatacaa ttgtcgcccg aggtctgtac 1200
 ggcagaaac gagatccat tgaggagggc aagcgataac agaagctgg ctgggaggt 1260
 taccagaaga ttaactgtta gaggttagac tatggatatg taatttaact gtgtatatag 1320
 agagcgtgca agtatggagc gcttgttcag ctgttatgat ggtcagacga cctgtctgat 1380
 cgaatgtta tgatantgca caactctgtt atcccgatga tctgtccaat gggcatggtt 1440
 gttgtgttc tggatacggg gatgctgggt acagtgtaa taacttgaa taattatact 1500
 tatatgaggg tcgaagaag ctgacttgtg tatgacttat tctcaactac atccccagtc 1560
 ncaataccac cactgcacta ccaactacac aaacacatga tcaaacccac catggaactc 1620
 ctggagggag aagaacttgt tatggaaaag ctcaagagag agatcaaac ttgttatago 1680
 atcaattata cgaagtatac ctgcaagtaa aggaattcag gagaacccgg gttagggcg 1740
 tatttgtgc ccaaaaaaca gcccaattg cccaattga cccaattg acccagtagc 1800
 gggcccaac ccggcgagag cccctctcac cccacatata aaactcccc cgttccac 1860
 acttgcgctt aaggcgtag ggtactgcag tctggaatct acgttgttc agactttgta 1920
 ctagtctt ttgtctggca tccgggtaac ccatgcgga cgcacaatag actactgaaa 1980
 attttttg ttgtgtgttg ggaactttag caagggtata aaagacacac gtccccgaat 2040
 tacccttct ctcttttt ctctctcctt gtcaactcac acccgaaatc gttaaagatt 2100
 tctttctgag tataagaatc attcaccatg gacttctgg aggcagaaga acttgttatg 2160
 gaaaagctca agagagagaa gcccaagtao tatcaagaaa tgtgtcgaaa attaattaag 2220
 atgacgacat ttgcgagctg gcagaggaat agatggagcg tgtgtctgga ctgatgttt 2280
 totatggagt tgtgagtggt agtgaacatg atgggtttat atatgatgaa tgaatagatg 2340
 tgaatttgt ttgcacgagb gaattgagaa atttgttaac gtaactggga atgtatgaat 2400
 gtgggggttt tgtgactgga taactgaag tcaatggagc cgttgttcaa atataccaaag 2460
 agatgcgaga aactttgggt caagtgaaca tgcctctct gttaagtaa accatcaact 2520
 atgggtagta tatttgtaa ggaacaagat tgagattctt tggagctcaa gaacgctatt 2580

ttogcgttcc aagatcaaat tagtagagta ataccggcac gggaaatcaat tcatagtctc 2640
 aattttccca taggtgtgct acaagtggtt gagatgtggt acagtaaccac catgattcga 2700
 ggtaaagagc ccagaagtca ttgatgaggt caagaaatca acagatctac agtcaaatca 2760
 aatgaatata ttctttcata ttcttcaggt gacacaaagg ggtgtctatt tcccaagaaa 2820
 tgcgtgaaaa ggcgcgtgtg tagcgtggag ttgggttctg gttggcgtat ccttcataata 2880
 tcgaagaaat agtagggcaa gagatgacaa aaagtatcta tatgtagaca gcgtagaata 2940
 tggatttgat tggataaaat tcatattatg cgtgtctcac aaatactctc gataagttgg 3000
 ggttaactg gagatgaaac aatgtcgata tctgcagcca tgcagcgtcg ggcacaattc 3060
 gccctatagt gagtctgatt acaattcact ggcgcgtggt ttacaacgtc gtagctggga 3120
 aaacccctggc gttacccaac ttaattcgct tgcagacat cccctcttcg ccagctggcg 3180
 taatagcgaa gagcccgca ccgatcgccc ttcccaacag ttgcgcagcc tgaatggcga 3240
 atggacgcgc cctgtagcgg ccgattaaag gcggcgggtg tggtagttac gcgcagcgtg 3300
 accgtacac ttgcacggc cctagcgccc cctcctttcg ctttcttccc tctcttctc 3360
 gccacgttcg ccggctttcc ccgtcaagct ctaaatggg ggtcccttt aggttccga 3420
 tttagtctt taaggcaact cgaacccaaa aaacttgatt aggtgatgg ttacagtatg 3480
 gggccatcgc cctgatagac ggttttctg cctttgacgt tggagtcac gtttttaat 3540
 agtggactct tgttccaaac tggacaaca ctaaccccta tctcgtctc tctctttgat 3600
 ttataagggg ttttgcgat ttgcgcctat tggtaaaaa atgagctgat ttaacaaaa 3660
 tttaacgoga atttaacaaa aatattaacg cttaacatt cctgagggg tatcttctcc 3720
 ttacgctct gtcgcgtatt tcaacgcga tcaaggcca ctttcgggg aatgtgcgc 3780
 ggaaccccta ttgtttatt ttttaaaaa catcaaaaa tgtatcggct catgagacaa 3840
 taacccctat aatgcttca ataatttga aaaggaaga gtatgagtat tcaacattc 3900
 cgtgtgcgcc ttattccctt ttttggcca ttttgcttc cgttttttc tcaacgaag 3960
 acgctggtga aagtaaaaaa tgcataagat cagtgggtg cagcagcggg ttacatcga 4020
 ctgctctcca acagcgttaa gatcttgg agttttgcgc ccgaagaacg ttttcaatg 4080
 atgagcactt ttaaggtct gctatgtgcg gcgtattat ccgatattga ccgcggcaa 4140
 gagcaactcg gtcgcgcat acactattct cagaatgact tggttgata ctcaacagtc 4200
 acagaaaaag atcttcyga tggcatgaca gtaagagaat tatgcagtc tgcataacc 4260
 atgagtata acactgggc caacttact ctgacaacga tcggagacac gaagagacta 4320
 accgcttttt tgaacaacat cgggcatcat gtaactgcgc ttgatcgtg gcaacccggg 4380
 ctgaatgaag ccatcaaaaa ccagcagcgt gacacacga tgcctgagc aatggcaaca 4440
 acgttgccga aactattaac tggcgaacta ctactctag ctcccgcca acaattata 4500
 gactggttg aggcggataa agttgcagga ccacttctg cctcggcctc tccgctggc 4560
 tgtttattg ctgataaat tggagcgggt gagctgggt ctcgcgggat catgacaga 4620
 ctgggcccag atgttaagcc ctcccgctac gtagttatct acagcaggg gactcagca 4680
 actatggatg aacgaatatg acagatcgtc gagtaggtg cctcactgat taagcattg 4740
 taactgtcag accaagtta ctcatatata ctttagattg atttaaaact tcattttaa 4800
 tttaaaagga tctaggtgaa gatcttttt gataatcca tgaacaaat ccttaagct 4860
 gactttctg tcaactgagc gtcagacccc gtgaaaaa tcaaggtatc ttcttgatg 4920
 ccttttttc tgcgcgtaat ctgctgctg caaacaacaa aacacacgtc accagcgtg 4980
 gttgttttc cgtatcaga gctacacact ctttttcca agtlaactg cttacgaga 5040
 gcgcagatcc caaatcctg tctctagtg tagcctgagt taggcaccca cttcaagaa 5100
 tctgagcac cgcctacata cctcgtctg ctactcctg taccagtcgc tgcctcagt 5160
 ggcataagt cgtcttacc cgggttgga tcaagacgat agttaocga taaggccag 5220
 cgttcggct gaacggggg ttctgcaca cagccacgtc tggagogaac gaactaccc 5280
 gaactgagat acctacagcg tgagctatga gaaagcccca cgttcccca agggagaag 5340
 gcgcacaggt atccggtag ccgcagggtc ggaacagag agccacagag ggccttcca 5400
 gggggaacag cctggtatct ttatagctct gtcgggttc gcaacctctg actttagct 5460
 cgtttttg gatgctctc agggggcgcg agctatgga aaaaagccag caacgggccc 5520
 tttttacggt tctggcctt ttgctggcct ttgtccaca tgttcttcc tgccttacc 5580
 cctgattctg tggataaac tattacccc tttagtgag ctgataccgc tgcgcagga 5640
 cgaacacag agccagcga gtcactgag cagcaagcgc aagagcgcac aatcacaaa 5700
 ccgcctccc ccgcgcgtg gcgatctat taatgaagt ggcgcgcac caatcaaat 5760
 totgaanag acatctgat ctctcattg cggggagtc aaaggtggt ttattccccc 5820
 gaattcccg ctcaatctcg ttccagacg acccgacac agtgctaac gcggttcca 5880
 aactcacag cagatctgt ccaacggact ggcctgata gatgctatcc tggcttggg 5940
 gaactggata aaagccggcc aaaaaaaag ccgaacaaag ccgaacaaa gagaacaaa 6000
 atcccaaat ttgaacata ggggcacaa cgcacaaac ccaagcaggg agagatatat 6060
 gacactgata agcaagctca caacggttcc tottattttt tctctactct tctgcctag 6120
 ttcccaaat cccagatgct tctctcagt gcaaaaagta agtaacccac aggttttcg 6180
 ccgaaaatto cactgcagc aactcgtgt ggggtgttaa natgtgggg gggggaacc 6240
 ggaacagag cctctgtgg agccagatga gacacaaag ccggcgggtg tgnaaaggg 6300
 attttgcc atttccct cctctctc tccagcgtg atggcgttg cgtctctc 6360
 ttcttttta ttcttttg ttattttt ctgactacg atttggttg atttctcaa 6420
 cccacacaa ataagctcgg gcagaggaat atatatac aaggaacag tgcctctg 6480

```

gacaaacagt cactaacctc acgataacac ccgtacgata gttagttagc acaaatcgat 6540
agttggagca agggagaaat gttagtggtg aaagactcac tatgtccgg gcttatctcg 6600
accataagcc aaagtctgga gtttctggga gaaaagggca agatacgtat gtaacaaagc 6660
gacgcattgt acataaatc ccgaggcattg tatcatagag agttagtgtg togatgatgg 6720
cactgggtgc tggtagact ttatacgggt gactacatat ttgtccctag acatacaatt 6780
acagtcaagc acttaacott ggaactctgt aggtaccccc ccgccaagac gatctcagcg 6840
tgtcgtatgt ccgattggcg tagctccctc gctcgtcaat tggctcccat ctactttctt 6900
ctgcttggtc acacccagca tglctgclat ggctcgtttt cgtgccttat ctatcctccc 6960
agtattacca actctaaatg acatgatgtg attgggctca cactttcata tcagagataa 7020
ggagttagac agttgcatba aaagcccaac tctaatcagc ttcttctctt ctgttaatta 7080
gtacaaaggt gattagcgaa ctctggaagc ttagtggcc ctaaaaaat caaaaaagc 7140
aaaaaacgaa aaacgaaaaa ccacagtttt gagaacaggy aggtaacgaa ggtatgtata 7200
tatatatata tatatatata cccacggatc ccgagaccgg cctttgattc ttcctacaaa 7260
ccaaccttc tcaccacctt aattcacac ccgtcccgct gcctccaaag cgtctctcta 7320
cgtcccgccg gccatcgccg tggtaactct tgcgcctgtg gccacctaag ggtgctggc 7380
gtccacatt ctacccgaca tcggcaagca ggcctggcc caatggacgg ttgcagagc 7440
ctctactac tcggtcgaca tcttctggg tatcagcac aagctcgcta gccggcaggt 7500
gaccggaaac ccgctctgg atgcctccaa gatccagtc gccacaccaa ccaagccat 7560
tgacgacac accaaacacc tgcctcgacc atgcattctg atttccaaac accagaacga 7620
aatggacatt ctggtctgct gtgcactctt cccccagtac tgcctcgta ccgcaaaaaa 7680
ggcctcaag tggtaacctc tctggggcca gtctatggcg ctgtccggca ccatctctc 7740
ggacgaaag gcccgaaaca agtccgtgca gccctcggc ggcgcgtca agacatcca 7800
gagcggcaac ggaaggcaag gccagagcgt ctctatgtc ccgagggaa ccgatctca 7860
ctccaaagac gtccgcatca tgcctccaa gaaaggctgt ttcacactgg ccgtccagtc 7920
ggcgctcccc attgtcccg tgggtgtcca gaacactccc cgaatgtttt ctttggccg 7980
aggaagctg gacgcggag agatcctgt cgaactcctg agcccattg agaaaaggg 8040
ctggaagcc agcaacgctg acgctctcat ggcacccat tataaggcca tgtcagagac 8100
tgccgacag attgctacg ctggccagaa gactcagtag ccggcccgat gagaagataa 8160
atatataaat acattgagat attaaatgc ctgattaga gagcctcata ctgctcggag 8220
agaagccaa acgactact aaaggggatt acaccatcca tatccacaga caaagctcg 8280
gaaaggctc tatataact ttcggaata ccgtagttc cgaatttata atcggggcca 8340
gccaggatt ccggaactc cgtctctcg ggtgaattg cgttctggc ctccatcaag 8400

```

```

tactacatg tcttcatttg cctgtcaag taaacagaa gcagatgaag aatgaactg 8460
aagtgaagga atttaaatgt aaagaaactg aaattgacc agatattgtg tccgggtgg 8520
agctcaagct ttgttccct ttagtgaagg ttaatttoga gcttggcgta atcatggtca 8580
tagctgttcc ctgtgtgaaa ttgttatccg ctcaaaagct tccacacac 8630

```

<210> 150

<211> 8630

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pY175

<400> 150

```

gtacgataac ttcgtatagc atacattata cgaagttatc gcgtcgacga gtatctgtct 60
gaactgcatc tgcgccttt ggagtacgac tccaaactatg agtgtcttg gatcaacttg 120
acgatacaatt ctgcttgga ggcctgtgggt ctgacagctg cgtttctggc gcggttgccc 180
gacaaacata tcagctgcaa cgtcattgct ggccttcaat atgatacaat tttgtcggc 240
aaaggcgacg ccacagagagc caattgaactt ctctcaatt tggacogata gcggtatagt 300
ccagttctatc tataagttca actaacctgt aactattacc ataacatata ctcaactgcc 360
ccagataaag ttccgataaa aagttctgca gactaaattt atttcaatct cctcttcaac 420
accaaaatgc cctcctacga agctcgagct aacgtccaca agtccgcctt tgcgcctga 480
gtgtcdaagc tcgtggcagc caagaaaaac aacctgtgtg ctctcttgga tgttaccacc 540
accaaggagc tcattgagct tgccgataag gtccgacctt atgtgtgcat gatcaaaacc 600
catatcgaca tcattgacga ctctcaactac gccgcactg tgcctccctt caaggaaactt 660
gctcttaagc accgtttctt cctgttcag gacagaaagt tcgcagatat tggcaacact 720
gtcaagccac agtaccggcg tcaccgaatc gccagtggt cagatatcac caacgccac 780
ggtgtacccc gaaccggaaat cactgtcggc ctgcgagctg gtcccgagga aactgtctct 840
gaacagagga agggagagct ctctgactac gagaactccc agtacaagga gttctagtc 900
ccctctccca acgagaagct ggccagaggt ctgctcatgc tggccagct gcttgcag 960
ggctctctgg ccactgggga gtactccag cagaccattg agcttgccc atccgacccc 1020
gagtttgtgg ttgcttcat tgcccagaac cgaactaagg gogactctga ggaactgctt 1080
attctgaccc ccgggtggg tcttgacgac aaggggagag ctctcggaca gccgtaccga 1140
actgttgagg atgtcatgtc taaccgaaag gatataataa ttgtcggcc aggtctgtac 1200
ggccagaacc gagatcctat tgaggaggcc aagcgatacc agaaggtcg ctgggaggct 1260
taccagaaga ttaactgtta gaggttagac tatgatatg taattttaact gtgtatatag 1320
agagcgtgca agtatggagc gcttctcag ctgttatgat ggtcagaaga cctgtctgat 1380
cgagtatgta tgatactgca caacctgtgt atccgatga tctgtccaat ggggcattgt 1440

```

gttgtgttcc tggataaggc gatgctgggt acagtgctaa taegttagac tactataact 1500
 tatatgagcc tggagaagag ctgactctgt tatgacttat tctcaactac atccccagtc 1560
 acatcacacc cactgcacta cccactaccc aaaaactgta tcaaacaccac catggacttc 1620
 ctggaggcag aagaacttgt tatggaagag ctcaagagag agatctaac ttoytatagc 1680
 atacattata cgaagtcttc ctgacggtta agaatctcag gaggagaagc gttggaggcg 1740
 tattttgttc ccaaaaaaca gcccaaatg ccccaattg ccccaaatg acccaatgac 1800
 gggcccaacc cggcgagag cccctctcac ccccatatc aaacctcccc cgttccacac 1860
 acttgcctt aagggcgtag gtaactgcag tctggaatct acgttgttgc agactttgta 1920
 ctagtctctc tctctggcca tccgggtaac ccatgcggga cgcacaaatg actactgaaa 1980
 attttttgc tttgtgttgc ggaattttag caagggtata aaagaccacc gtccccgaat 2040
 taccttctcc ctctttttct ctctctctct gtaactcac accgaatac gtttaagcatt 2100
 tccctctgag tataagaatc attcaccatg gactcctcgg aggcagaaga acttgctatg 2160
 gaaaaactca agagagagaa gccaaagatc tatcaagaca tgtgtcgcga cttaattaa 2220
 atgacagcat ttgcagctgc gacgaggaat agatggagcg tgtgtctcga gtccatgttt 2280
 tctatggagt tgtgagtgtt agtagacatg atgggtttat atatgagaa tgaatagatg 2340
 tgattttgat ttgcacgatg gaattgagaa cttttgaac gtacatggga atgtatgaat 2400
 gtgggggttt tgtgactgga taactgacgg tcaatggacg cgttgtctca aatatccaa 2460
 agatgcgaga aactttgggt caagtgaaca tctctctctc gttaaatgaa accatcaact 2520
 atgggtagta tatttagtaa ggcagaagct tgaattcttc tggagcctta gaaactgatt 2580
 ttgcgttccc aagatcaaat tagtagtgc atccgggcac gggatccat ctatagtctc 2640
 aatttccca taggtgtctc acaaggtgtt gagaatgtgt acagtacac catgattcga 2700
 ggtaaagagc ccgaagatca ttgatggagt caagaataac acagatctac agctaatcac 2760
 aatgaatata ttctttcata ttcttcaggt gacacaaagg gttctctatt tccccagaaa 2820
 tgcctgaaaa ggcgcgtgtg tagcgtggag tatgggtctg gttgcgtat cctctatata 2880
 tgcacgaatc agtagggcaa gatatgaca aaagtatact tatgtagaca gctagaata 2940
 tggattgatc tggataaatc taattattg cgtgtctcac aaatactctc gataatgttg 3000
 ggttaactgc gagaatgaa atgtctgata tctgcagca tgcgagctgc ggcacattc 3060
 gccctatagt gactgtattc acaattcact ggcgtgtgtc ttacaagtc gtgactggga 3120
 aaacctggc gttacccaac ttaactgcct tgcagacat cccctcttgc ccagctggcg 3180
 taatagoga gaggcccgca ccgatgcgcc ttccccacag ttgcgcagcc tgaattggca 3240
 atgagcgcgc cctgtagcgg ccattaaagc ggcgggggtg tgggtgttac ggcagcgtg 3300
 accgtacac ttgcacgcgc cctagcgccc gctccttgc atttctccc ttctttctc 3360
 gccactgtgc ccggttccc ccgtcaagct ctcaatggg ggtcccttt aggttccga 3420
 ttaagtcttc taccgacat ccacccaaa aaactgatt aggtgatgc ttacatagt 3480
 ggcacatgc cctgatagac ggtttttgc ccttgagct tggagtcac gttctttat 3540
 agtggactct tgttccaaac tggacacaca ctcaacctc tctcgtcta ttctttgat 3600
 ttataagga ttttcgatc ttgcctcatc tggtaaaaa atgagctat ttacaaaaa 3660
 tttaacgga attttaaca aattttaag attcaattc cctgatgcg tattttccc 3720
 ttaacatct gtcgggtatt tcaacgcga tcaatgggca cttttgggg aaatgtgcgc 3780
 ggaacctata ttgtttatt ttctaaata ctttaaaata tctatccgt catgagaca 3840
 taacctgatc aaatgctca ataatctga aaagaagga gtagagtat tcaacttc 3900
 cgtgcgcgc ttattctct tttgcggca tttgccttc ctgttttgc tcaaccagaa 3960
 accgtgtgta aagtaaaaga tgcctgaagc cagtgggtg caagatggg ttacatgaa 4020
 ctggatctca ccagcgtta gctccttgc agttttgcgc ccagaagc ttctcaatg 4080
 atgagcact tttaaattct gctatgtgc gggatattc cccgtattga cgcggggca 4140
 gacacactgc gtgcgcgat caactttct cagaatgact tggttgata ctcaacatc 4200
 ccgaagaagc ctcttaacga tggcatgaca gtaaggaat tatgcagtc tgcataacc 4260
 atgagtata acactgcgc caacttact ctgacacga tggaggaac gaaggagta 4320
 accgttttt tgcacacat gggggtcat gtaactgcgc ttgatgttg ggaacggag 4380
 ctgaatgaag cctacacaaa ccagcagcgt gacacacga tgcctgtgc aatggcaac 4440
 accgtgcga aactttaac tggcgaact attactctag cttcccgga acaattata 4500
 gactgagat aggcgataaa agtgcagga cactcttgc gctgcgcct tccgtgcgc 4560
 tggttattg ctgataatc tggagcgggt gacgtgggt ctgcgggtat catgacga 4620
 ctgggcgcag atggtaagc cttccgtatc gtagttatc accagcggg gactcagga 4680
 actatggatg aacgaatag acagatgct gatatagtg cctcaatgat taacattg 4740
 taactgtcag accaagtta ctcatatata ctttagatg atttaaaact tcaattttta 4800
 tttaaaagga tctaggtgaa gctccttttt gataatctca tgaacaaat cctttaact 4860
 gactttctg tccactgagc gtccacccc gtgaacaga tcaagagatc ttcttgat 4920
 ccttttttc tgcgcgtaat ctgctgctg caaaaaaaa accacccgtc cccagcgtg 4980
 gttgtttgc ccgatacaga gataccact ctttttccga aggttaactg cttcaagga 5040
 ggcagatcac caataactg tctctagtgc tagcgtagt taggcacac cttcaagaa 5100
 tctgagcac ccctacata cctcctctg ctaactcctg taacagtcgc tctgcacgt 5160
 ggcataagt cgtctctac ccggttgac tcaagacgat agttaacga taaggcgag 5220
 ccgtcgggtc gaacggggg ttctgcaca cagccagct tggagcagac gaactacac 5280
 gaactgagat acctacagc tgaatctga gaaagccca cgtttccga agggagaaag 5340

gggacaggt atccggttaag cggcagggtc ggaacaggag agcgcacgag ggagcttcaa 5400
 gggggaacgc cctggtatct ttatagtcct gtccgggttc gccacccctg acttgagcgt 5460
 cgatttttgt gatgctcgtc agggggcggc agccatagga aaaaaccccg caacggcgcc 5520
 tttttacgtt tccgtgcoct ttgtggcoct ttgtctcaaa tgtttcttcc tgcgttatcc 5580
 cctgattctg tggataaocg tattaccgco tttagtgtag ctgataocgo tccgcgcagc 5640
 cgaacgaocg agcgcagoga gtoagtgcgc gaggaaocgg aagagcgccc aatacgcaaa 5700
 ccgcctctcc ccgcgcgttg ccgcattcat taatgcagct ggcgcgcacc caatacaaat 5760
 tctgaaaagc acatcttgat ctccctcatt cggggagctc aacgttggtc ttattccccc 5820
 gaatttcccg ctcaatctcg ttccagacgc acccggaacc agtgcttaac gccgttccga 5880
 aactctaccg cagatctgct ccaacggact ggctcgata gatgtgaccc tccgcttggc 5940
 gaaatggata aaagccggcc aaaaaaaag cggaaaaag cggaaaaaa gaaaaaaa 6000
 atcgcaaat ttgaaaaaa gggggaanag ccgcaaaaa gccaggaggg gggagtatat 6060
 gacactgata aqcaagctca caacggtccc tcttatttt ttctctatct tctgcctagg 6120
 ttcccaaaat cccagatgct tctctccagt gccaaaaag agtacccccc aggttttcgg 6180
 ccgaanatto cactgtcagc aactcgtgt gggtgttaa atgtggggg gggggaacca 6240
 ggacaagag ctcttgigg agcgaatga gacacaaag cggcggggtg tgaaagggc 6300
 atttttgccc atttccctt ctccgtctc tccgacgttg atggcggtgt gcgtctcta 6360
 ttcttttta ttctttttt tttaatttt ctgactacgc atttggttg atttctcaa 6420
 cccacacaaa ataaagtcgg gccaggaat atatatata ccggaacag tccgcctgtg 6480
 gacacacgt cactacctc acgatccacc ccgtacgata gttatagac acaatcgat 6540
 agttggagca agggagaat ctgagtggtg aaagactaac tatgtccgg gcttatctcg 6600
 accaatagcc aaagtctgga gttctcgaga gaaaaggca agatacgtat gtaacaaagc 6660
 ggcgcaggtt acaataaac ccgagcgatg tatcatagag agttaaggt togtatgag 6720
 cactgggtcc tggatgact ttatacggct gactacatat ttgtccctag acatacaatt 6780
 acagtcaagc acttaacct ggaactctgt agttaccccc cggccaagac gatctcagcg 6840
 tgcgtatgt cggatggcg tagctccctc gctcgtcaat tggtcccaat ctaattctt 6900
 ctgctgggtt acaccacga tctctgctat ggctggttt cgtgccttat ctactctcc 6960
 agtattacca actctaaatg acatgatgtg attgggtcta cacttccata tcaagataa 7020
 ggagtgcac agttgcataa aaagcccaac tctatcagc ttctctctt ctgttaatta 7080
 gtacaaaggt gattagcgaa atctggaagc ttatggccc ctasaaaaa caaaaaagc 7140
 aaaaaacgaa aaacgaanaa cccaggtttt ggaacaggg aggtaaagaa ggcctgata 7200
 tatatatata tatatatata cccacggatc ccgagacggg ctttgatct ttccatcaa 7260

ccaacatttc tcaacacctt aattcacaac catggagaac ttctgttcaa tctgttgtt 7320
 cttttctgct tccattctgt tcatctctta caacatttgg acagtctgac acttactaat 7380
 ggcgaatccc ttctaactat ttacacacct gcttccagcg atggaggtgt gogttaacct 7440
 gattccctct tggctcaacg gccagggtgc cgaactacgt ttctactcgt tcttctactg 7500
 gtgcagtgag actggaagtc acaccactgt gtatgctac gagaagacc agtccgaagc 7560
 tccgtccgtg gtcacttgca accatcagtc ctgcctcgac attctgtcta tggcttccat 7620
 ctggcccaag aactgtgtg tcatgatgaa cgggattctt gcctacgttc ccttctcaa 7680
 cctgggagcc tacttttcaa acaccacttt catcgacga tacaaccgag agcgagctat 7740
 ggcttctgct gactactgt cctccagatg gaagaaccca aacctgaagc tctgggtgtt 7800
 tcccgaagcc actcggaatc gagggggtgg attcattccc ttcaagaaag gtgccttcaa 7860
 catcgtggtt cgaacccaga ttccatcat tctgtctggt ttctctgact atcgagctt 7920
 ctactccaa cctggccgat acttcaagaa cgaatggagag gtcgtgaccc gagtctctga 7980
 tgccattccc accaagggtc tgaccctcga tgacctctct gagctttcgg acatgtctcg 8040
 agacgtcatg ctggctgctt acaaggaagt taacctcgag gctcagcaac gaaacgcacc 8100
 tgaagagga gagaaccaag ccggcaagaa atccaggtaa cgggcgcgat gagaagataa 8160
 atatatatat acattgagat attaaatcgg ctatagtaga gagctctata ctgctcgagg 8220
 agaagccaa cagagtactc aaaggggatt acaccatcaa tatccacaga caaaagctgg 8280
 ggaagagttc tatatacact ttccgganta ccgtagtttc cgaatgtatc aatggggcca 8340
 gccaggattt caggcacttc ggtgtctcgg ggtgaaatgg cgttcttggc ctccatcaag 8400
 tctaacctat tcttcatttg cctgtcaag taaaacagaa gcagatgaag aatgaacttg 8460
 aagtgaaaga atttaaatgt aacgaactg aaatttgacc agatatttg tccgcggttg 8520
 agtccagctt ttgttccct ttatgtaggg ttaatttcga gcttggcgta atcatggtca 8580
 tagctgttcc ctgtgtgaaa ttgttatcgg ctcaaaagct tccacacacc 8630

<210> 151
 <211> 926
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<220>
 <221> CDS
 <222> (4)..(915)

<400> 151
 aac atg tct gtt att gga cga ttt ctt tac tac ctg aga tgg gtg ctc 48
 Met Ser Val Ile Gly Arg Phe Leu Tyr Tyr Leu Arg Ser Val Leu
 1 5 10 15
 gtc gtt ttg gcc ctc gct gga tgt gcc ttc tat ggc gtg att gcc tct 96
 Val Val Leu Ala Leu Ala Gly Cys Gly Phe Tyr Gly Val Ile Ala Ser
 20 25 30

```

atc ctg tgc act ctc atc ggc aag cag cat ctc gcg cna tgg att acc      144
ile leu cys thr leu ile gly lys gln his leu ala gln trp ile thr
35          40          45

goc cga tgc ttt tac cac gtg atg aaa ctg atg ctg gga ttg gac gtc      192
ala arg cys phe tyr his val met lys leu met leu gly leu asp val
50          55          60

aaa gtc gtc ggt gaa gag aac ctg gct aag aag ccc tat atc atg atc      240
lys val val gly glu glu asn leu ala lys lys pro tyr ile met ile
65          70          75

gct aac cac cag tcc acc ctc gat atc ttt atg ctc ggc aga atc ttc      288
ala asn his gln ser thr leu asp ile phe met leu gly arg ile phe
80          85          90          95

ccg ccc gga tgc acc gtt acc gca aag aag agc ctt aag tac gtc ccc      336
pro pro gly cys thr val thr ala lys lys ser leu lys tyr val pro
100         105         110

ttc ctg ggc tgg ttt atg gcg ctt tcc ggt aca tac ttc ctc gac cgt      384
phe leu gly trp phe met ala leu ser gly thr tyr phe leu asp arg
115         120         125

tcc aag cgg cna gag gct att gat acc ctg aac aaa gga ctg gag aac      432
ser lys arg gln glu ala ile asp thr leu asn lys gly leu glu asn
130         135         140

gtc aaa aag aat aag cga gcc ctc tgg gtt ttt ccc gaa ggt acc cgg      480
val lys lys asn lys arg ala leu trp val phe pro glu gly thr arg
145         150         155

tct tac acg tcg gag ctt acc atg ctg cgg ttc aag aag gcc gcc ttt      528
ser tyr thr ser thr glu leu thr met leu pro phe lys lys gly ala phe
160         165         170         175

cat ttg gcc cag cag gga aag atc cct atc gtc cca ctg gtt gtc tct      576
his leu ala gln gln gly lys ile pro ile val pro val val val ser
180         185         190

aac act agc acg ctc gtc agc cct aag tac ggt gtc ttc aac cga gcc      624
asn thr ser thr leu val ser pro lys tyr gly val phe asn arg gly
195         200         205

tgt atg att gtc cga att ctg aag ccc atc tog acc gag aac ctg aca      672
cys met ile val arg ile leu lys pro ile ser thr glu asn leu thr
210         215         220

aaa gat aag att ggc gag ttc gct gag aaa gtg cga gat cag atg gtt      720
lys asp lys ile gly glu phe ala glu lys val arg asp gln met val
225         230         235

gac ace ctg aag gaa atc ggt tat tcc ccc gca atc aac gac acc acc      768
asp thr leu lys glu ile gly tyr ser pro ala ile asn asp thr thr
240         245         250         255

ctt cct ccg cag gca att gag tac gcc gca ttg cag cat gac aag aag      816
leu pro pro gln ala ile glu tyr ala ala leu gln his asp lys lys
260         265         270

gtg aac aaa aag att aag aac gaa cct gtt cca tcg gtc tcc att tct      864
val asn lys lys ile lys asn glu pro val pro ser val ser ile ser
275         280         285

aat gat gtc aat act cac aac gaa gga tcg tct gtc aag aag atg cac      912
asn asp val asn thr his asn glu gly ser ser val lys lys met his
290         295         300

tag gcggccgcat g      926

<210> 152
<211> 303
<212> PRT
<213> Saccharomyces cerevisiae

<400> 152
Met Ser Val Ile Gly Arg Phe Leu Tyr Tyr Leu Arg Ser Val Leu Val
1          5          10          15

Val Leu Ala Leu Ala Gly Cys Gly Phe Tyr Gly Val Ile Ala Ser Ile
20         25         30

Leu Cys Thr Leu Ile Gly Lys Gln His Leu Ala Gln Trp Ile Thr Ala
35         40         45

Arg Cys Phe Tyr His Val Met Lys Leu Met Leu Gly Leu Asp Val Lys
50         55         60

Val Val Gly Gly Glu Asn Leu Ala Lys Lys Pro Tyr Ile Met Ile Ala
65         70         75         80

Asn His Gln Ser Thr Leu Asp Ile Phe Met Leu Gly Arg Ile Phe Pro
85         90         95

Pro Gly Cys Thr Val Thr Ala Lys Lys Ser Leu Lys Tyr Val Pro Phe
100        105        110

Leu Gly Trp Phe Met Ala Leu Ser Gly Thr Tyr Phe Leu Asp Arg Ser
115        120        125

Lys Arg Gln Glu Ala Ile Asp Thr Leu Asn Lys Gly Leu Glu Asn Val
130        135        140

Lys Lys Asn Lys Arg Ala Leu Trp Val Phe Pro Glu Gly Thr Arg Ser
145        150        155        160

Tyr Thr Ser Glu Leu Thr Met Leu Pro Phe Lys Lys Gly Ala Phe His
165        170        175

Leu Ala Gln Glu Gly Lys Ile Pro Ile Val Pro Val Val Val Ser Asn
180        185        190

Thr Ser Thr Leu Val Ser Pro Lys Tyr Gly Val Phe Asn Arg Gly Cys
195        200        205

Met Ile Val Arg Ile Leu Lys Pro Ile Ser Thr Glu Asn Leu Thr Lys
210        215        220

Asp Lys Ile Gly Glu Phe Ala Glu Lys Val Arg Asp Gln Met Val Asp
225        230        235        240

Thr Leu Lys Glu Ile Gly Tyr Ser Pro Ala Ile Asn Asp Thr Thr Leu
245        250        255

Pro Pro Gln Ala Ile Glu Tyr Ala Ala Leu Gln His Asp Lys Lys Val
260        265        270

Asn Lys Lys Ile Lys Asn Glu Pro Val Pro Ser Val Ser Ile Ser Asn
275        280        285

Asp Val Asn Thr His Asn Glu Gly Ser Ser Val Lys Lys Met His
290        295        300

<210> 153

```

<211> 7891

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pY222

<400> 153

```

tgcacgctac gatagtagt agacaacaat cgaatgttg agcaaggag aatgtlaag      60
tgtgaagac tcaataggt ccgggcttat ctgcaccaat agccaagtc tggagtctct      120
gagagaaaaa gccaagctac gtatgtacaa aagcgacgca tggtaacata ataccggagg      180
catgtatcat agagagttag tgggttcgat atggcactgg tgccctggat gactttatc      240
ggctgaatca atattgtcc tcagacatac aattacagtc aagcaatcac ccttggacat      300
ctgtaggtac cccccggcca agacgatctc agcgtgtcgt atgtcggatt ggctagctc      360
cctcgctcgt caattggctc ccatctactt tctctcgtt ggctacaccc agcatgtctg      420
ctatgctcgt ttttcgtgcc ttatctatcc tcccaattt accaaatcta aatgacatga      480
tgtgattggg tctacacatt catatcaagc ataaggagta gcacagttgc ataaaaagcc      540
caactataat cagctctctc cttttctgta attagtaca aaagtattag cgaatatctg      600
aagcttagtt ggccctaaaa aatcaaaaa aagcaaaaa cgaaaaacga aaacccacag      660
ttttgagaac agggaggtaa cgaaggatcg tatatatata tatatatata tataccacag      720
gatcccgaga ccggcctttg attcttccct acaaccaacc attctacaa cctaattca      780
caaccatgct tgttaattga cgaattctctt actacctgag atcgtgtctc gtcgttttgg      840
cctcgctcgt atgtgcttc tatggcgtga ttgcctatat cctgtgtact ctcatcgcca      900
agcagcatct ccgcacatgy attaccgccc gatgctttta ccacgtgatg aaactgatgc      960

tgggattgga cgtcaaatgc gtgggtgagc agaacctgga taagaagccc tatatatga      1020
tcgtatacca ccagtccacc ctgcatactc ttatgtctgg cagaatcttc cctccggat      1080
gcacggctac cgaagaagag agccttaagt acgtcccttc cctggctgg tttatggcgc      1140
tttcgggtac atacttcttc gacgttcca agcggcaaga ggtatgtgat accctgaaca      1200
aaggactgga gaactgaaa aagaataagc ggcctctg ggtttttccc gaagtaccc      1260
ggtctaacac gtccgagctt accatgctgc cgttaagaa gggcgctttt catitggccc      1320
agcagggaaa gatccctacc gtcacagtgg ttgtgtata cactagaacg ctgcctagcc      1380
ctaatgacgg tgtgtcaac cgaagctgta tgaattcccg aattctgaag ccatctoga      1440
ctgaacact gacaaaagat aagattggcg agttcgtga gaaagtcgga gatcagatgg      1500
tgtacacact gaaggaaatc ggttatctcc ccgaatcaca cgaacacacc ctctctccgc      1560
agggcaattga gtaacgcgca ttgcagcatc acaagaaggt gaacaaaaag attagaacg      1620
aacctgttcc atcgtgtccc atttctaagt atgtgatac tcacaacgaa ggtatcgtcg      1680
tcaagaagat gcactaggcg gcgcgatgag aagataata tataaatata ttgagatatt      1740
aaatcgctca gattagagag cctcactact ctccggagga agccaagacg agtaactaaa      1800
ggggattaca ccatccatat ccacagacac aagctgggga aaggttttat ataaccttc      1860
cggaataccg tagtttccga tgttatcaat gggggcagcc aggatctcag gaacttcggt      1920
gtctcggggt gaaatggcgt tcttggctcc catcaagtcg taacctgtctt tcaattgcct      1980
gtcaagtaa acaagaagca gatgaagat gaacttgaag tgaaggactt taagttaac      2040
gaaactgaaa tttagaaga tatitgtccc gcggcgagc tccagctttt gttcccttta      2100
gtgaggggta attcagact tggcgaatc atggtcatag ctgtttccctg tgtgaattg      2160
ttatccgctc acaagcttcc acacaacgta cgaacgggaa gcataaagtg taagcctgg      2220
ggcgcccaat gactgagcta actcacatta attgcgttgc gctcaactgoc cgttttccag      2280
tcgggaaccc tgtcgtgcca cctgcattaa tgaatcgccc aacgcgcggg gagaagcggt      2340
ttcgtatttg ggcgctcttc cgttccctcg ctcaactgact cgtcgcgtc ggtcgttcgg      2400
ctcggcgagc cgttatcagc tcaatcaag gcggtaatac ggttatccac aagaataggg      2460
gataacgcag gaagaacat gtgagcaaaa ggcacgcaaa aggcacggaa ccgtaaaaag      2520
gcgcgcttgc tggcgttttt ccataggctc cgcctccctg acgagcatca caaaaatoga      2580
cgtcaagtc agaggtggcg aaacccgaca ggaatataaa gatccaggo gtttccctc      2640
ggaagctccc tegtgcgtc tctgttccg accctgcgc ttacccgata cctgtccgoc      2700
ttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg      2760
gtgtaggctg ttgcctcaaa cctgggctgt gtgcaagaaa ccccgcttca gcccgaccgc      2820
tgcgcttat ccgttaacta tegtcttgcg tccaacggg taagaacaga cttatcgcca      2880

```

ctggcagcag coactggtaa caggatttag agagcaggt atgtaggcgg tgotacagag 2940
 ttcttgaagt ggtggcctaa ctacggctac actagaaga cagtatttgg tatctgocct 3000
 ctgctgaagc cagttacott oggaaaaaga gttgtagct ctgtatccgg caaacaaccc 3060
 acccctggta ggggtggttt ttttgtttg aagcagcaga ttacggcgacg aaaaaagga 3120
 tctcaagaag atcctttgat ctttctacg ggtctgacg ctacgtggaa cgaanaactca 3180
 ogttaaggga ttttggctat gaggattatc aaaaagatct tccctgatct ctttttaatt 3240
 taaaaatgaa gttttaactc aatctaaagt atatatgagt aaacttggtc tgcagttac 3300
 caatgcttaa tcaatgagcg accatctcca gcatctgtc tatttgcttc atccatagtt 3360
 gcttcactcc cgtctgtgta gataactacg ataccggagc gcttaacctc tggcccccgt 3420
 gctcaatga taccggcaga cccacgctca cggctccag atttatcaga aataaacccg 3480
 ccagcgggaa gggccgagcg cagaagtgtt cctgcaacct tatccgctcc catccagttc 3540
 attaatgtt gccgggaagc tagctgaagt agttgcacg ttaattgttt gogcaacglt 3600
 gttgcattg ctacaggaat cgttggctga cgtctgtgt ttggtatggt ttcattcagc 3660
 tccgtttccc aacgatccag cggagttaaa tgaatcccca tgttgtgcaa aaaaagcgtt 3720
 agtctcttgc gtctccgat cgttgtcaga agtaagltgg ccgcagtttt atccctcatg 3780
 gttatggcag cactgataa ttctcttaet gtcatgcaat ccgtaatgtt ctittctgtg 3840
 atgtgtagt actcaacca gtctctotga gaatagtta tggggcgacc gagtgtctc 3900
 tgcacggcgt caatacggga taataccggc ccaatagca gaactttaaa agtgcctact 3960
 attgaaaaa gttcttgagg ggaaaaaact tcaagatct taccgtgtt gaggatccgt 4020
 tcaatgtaac cactctgtc acccaactga tcttcagcat ctittacttt caccagcgtt 4080
 tctgggtgag caaaaacagc aagcaaaact ggcacaaaa agggataaag ggcacacagg 4140
 aatgtttgaa tactctact ctctcttttt caatattatt gaagatttta tccgggttat 4200
 tgtctcatga gggatatac attgaatgt attagaana ataacaactc aggggttccg 4260
 cgcacatttc cccgaaaagt gccacctgac ggcctctga gggcgcatc aagcggcgcg 4320
 ggtc-ggtcg ttacggcagc cgtgacgcct acacttgcca ggcgcctagc gccgcctcct 4380
 ttctcttct tccctctct tctgcacagc ttgcggcgt tcccccgtcc agtctcaaat 4440
 cgggggtccc ctttaggggt ccgatttagt gctttacggc aactgacccc caaaaaactt 4500
 gattaggggt atggttccag tagtgggcca tgcgcctgat agcgggttt ttgcctcttg 4560
 acgttgaggt ccaactctct taatatgga ctctgttccc aaactgggaa aacactcaac 4620
 cctactcggc tctattcttt tgaattataa ggaattttgc cgaattcggc ctattggtta 4680
 aaaaatgagc tgatttaaca aaatttaac ggcatttta acaaatattt aacgtttaca 4740
 atttccatc gccattcagg ctgcacact gttgggaagc gcatcgctg cgggcctctt 4800
 cgtattacg ccagctggcg aaaaaggggt gtctgcacg gcatataagt tgggtaacgc 4860
 cagggttttc cagtcacaga cgttgaana cgaagcgcag tgaattgtaa taagatccac 4920
 tataggcgga attgggtacc gggcccccgc tcaaggtcga tgggtctgat aagcttgata 4980
 tcaattctat gtcacacaaa ccgattcttg cctcaaggaa acctaatctt acctccgaga 5040
 gactgcagag atccagctca caatgattaa ttttggggcc aataatttaa aaaaatcgtg 5100
 ttatataata ttatatgat tatatatata cactatgat taactgaacg tcaatgcoca 5160
 ttgataaata gcaagaatcc atctgcgcgc tcaactgat gttctcaata tttaaggggt 5220
 cactctgat tgtttaataa taacagactc cactctacgg cctccaaatg atgtttcaa 5280
 aatatatgt atgaacttat tttaattact tagattatt agcaacttta cttgttttat 5340
 gaataacact tctatttag gaacacattt ataagggcag ttctgtcttt taacaattta 5400
 ttagaataaa atgttataaa tgcgtatggg aaactctaaa tatggatagc ataatgata 5460
 tctgacttgc ctaattogaa atcaacagca acgaaaaaaa tcccttgtae aacataaata 5520
 gtcatcgaga aatatcaact atcaaaagac agctattcac acgttactat tgaattatt 5580
 attgacagc aatcacacac tcaactgtct tctctcttc tagaataaca ggtacaagta 5640
 tgaactatcc tcaattgcca taactctagt cacttctacc caactattcc ttggattttc 5700
 ctccatgaa tgaacttota tcttgcaaat tcaacaatta taataagata taocaangta 5760
 ggggtatagt ggaactaaaa aagcttctct ggtgtgttct tctattttat tttaattcta 5820
 atgactcatt aaaggtatat atttatttct tgttatataa tcccttctgt tattacatgg 5880
 gctgataca taaaggtatt ttgatttaatt ttttgotta aattcaatcc cccctgttct 5940
 agtgcacact gtaatggtag gaatttaaca taactttgaa gaagcaaaaa aatgaaga 6000
 aaaaaaaat cgtattttcca gtttagacgt tccgcagact ctagaatgog gtagcggtta 6060
 cactgttctt cgaaggtaaa agttgcgcct cctgagatat tgaactttt tgcatttaca 6120
 agtacaagta catctacaaa ctatgtacta ctgttgatgc atccacaaca gttgttttg 6180
 ttttttttg ttttttttt ttctaatgat tcaatccgca tatgtatacc taattgtaet 6240
 tgaatgaagc cgggttatgt cgttcaatt aatcatagac ttatgaatct gcaaggtgtg 6300
 cgtcgagagt taacttttag ttatgcatgc taactgggtg taatatgggg atctgtcgg 6360
 aaatcaagc atgctcaatc gatttgcaca gtaattatb agtcatana caagtcaagt 6420
 ttcttcagc ctcatataag tataagtagt tcaagatb agcacgtgac ccagcatctc 6480
 cgtatcgaga aacacaacaa catgcccact tgcacagato atgoggatcc acaggttgtg 6540
 cagttatcta catactogat cagaaggtc gtctgacact catcaagat gaacaagcgc 6600
 tccatcttg caagctctct atatacacg ttaattaca tatccatagt ctacactcta 6660
 acagttatc ttctggttag cctccacgac agctctctgg tatcgttgg cctctcaat 6720
 aggatctgg ttctggcgtt caagacctcg gccacaaat atgatatccg ttcgggtaga 6780

```

catgacatcc tcaacagttc ggtactgctg tccgagagcg tctcccttgt cgtcaagacc 6840
caccocgggg gtcagaataa gccagtcctc agagtgcgcc ttaggtcggt tctgggcaat 6900
gaagccaacc acaaaactcg ggtcggatcg ggcaagctca atggtctgct tggagtactc 6960
gccagtgccc agagagccct tgcaagacag ctgggcagc atgagcagac ctctggccag 7020
ctctctgttg ggaaggggga ctaggaaact ctgtactcg gagttctcgt agtcagagac 7080
gtctctcttc ttctgttcag agacagtttc ctgggcacaa gctcgaagg cagcaatgat 7140
tcgggttcgg ggtacacccg ggcggttggt gatctggac caactcgoga ttcggtgaca 7200
ccggtactgg tgcttgacag tgttgccaat atctgcgaac ttctctgctt cgaacaggaa 7260
gaacccgtgc ttaagagcaa gttccttgag ggggagcaaa gtcccgcgct aggtgaaatc 7320
gtcaatgatg tcgatatggg ttttgatcat gcacacataa ggtccgacct tatcgccaag 7380
ctcaatgagc tcttggtagg tggtaacatc caagagaaga cacaggttgg ttctcttgcc 7440
tgccaagagc ttgagacact gagcgccaaa ggccgacttg tggacgttag ctcgagcttc 7500
gtaggagggc attttgtgag tgaaggagag actgaaataa atttagctcg caaaactttt 7560
tatcggaacc ttatctgggg cagtgaaata tatgttatgg taatagtac gagttagttg 7620
aacttataga tagactggac tatacggcta tcggtccaaa ttgaaagaa cgtcaatggc 7680
tctctggggc tcgcttttgc cgacaaaaat gtgacatga tgaagccag caatgacgtt 7740
gcagctgata ttgttgcgg ccaaccggcg cgaaaaagca gctgcagac ccacagcttc 7800
caacgaagaa tgtatgtca aagtgatcca agcacactca tagttggagt cgtactccaa 7860
aggcgcaat gacgagtcag acagatactc g 7891

```

<210> 154

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer 869

<400> 154

acttggcgcg ccactggcga gatccagctc ac 32

<210> 155

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer 870

<400> 155

tcagcctagg agcatccgtt gatttccg 28

<210> 156

<211> 9598

<212> DNA

<213> Artificial Sequence

<220>

<223> Plasmid pY177

<400> 156

```

gtacgataac ttcgatagc atacattata cgaagctatc gcgtcgagca gtatctgtct 60
gactgcctat tgcgccttt ggaagtcgac tccaaactatg agtgccttg gataccttg 120
acgatacatt ctctgttggg ggcgttgggt ctgacagctg cgttttcggc gcggttggcc 180
gacaaacata tcagctgcaa cgtcattgct ggccttcact atgatacaat ttttgcggc 240
aaaggcgagc ccagagagc cactgaagtt ctttctaatt tggacgata gccgtatagt 300
ccagctctat tataagtcca actaactcgt aactattacc ataacatata ctccactgcc 360
ccagataagg ttccgataaa aagttctgca gactaaattt atttcagctc cctcttcacc 420
accaaaatgc cctctacgca agctcgagct aacgtccaca agtcgcctt tgcgcgtcga 480
gtgctcaagg tcgtggcagc caagaaaacc aaactgtgtg ctctctcgga tgttaaccac 540
accaaaggagc tcattgagct tgcgataag gtccgacctt atgtgtgcat gatcaaaaac 600
catatcgaca tcattgacga cttcacctac gccggccactg tgcctccctt caaggaaact 660
gctcttaagg acggtttcct cctgttcgag gacagaaagt tcgcagatat tggcaacact 720
gtcaagcacc agtaccggta tcaccgaatc gccagtggt ccgatalcac caacgccacc 780
ggtgtaccgg gaaccggaat cattgtcggc ctgcgagctg gtgccagga aactgtctct 840
gaacagaaga aggaggaagt ctctgactac gagaactccc agtcaagga gtctctagtc 900
ccctctccca acgagaagct ggccagaggt ctgctcatgc tggccagct gtcttgcaag 960
ggctctctgg ccactgggga gtactccaa cagaccattg agcttgcgg atccgacccc 1020
gagtttgttg ttgcttcat tgcccagaac agacctaaag gcgactctga ggactggctt 1080
attctgaacc ccgggttggg tcttgacgac aaggggagag ctctcggaac gcagtaacga 1140
actgttgagg atgtcatgtc taccggaacg gatatacaa ttgtcgccg agtcttgtac 1200
ggccagaacc gagatcctat tgaggaggcc aagcgatacc agaaggtcgg ctgggaggct 1260
taccagaaga ttaactgtta gaggttagac tatggatatg taatttaact gcttatatag 1320
agagcgtgca agtatggagc gcttgttcag ctgttatgat ggtcagaaga cctgtctgat 1380
cgagtatgta tgatactgca caacctgtgt atccgatga tctgtccaat gggcgatggt 1440
gttgtgttto tcgatacggg gatgtgggt acagtgtcaa taogttgaac taactatact 1500
tatatgagcg tcgaagaaag ctgacttgtg tatgaactat tctcaactac atcccagtc 1560
acaataccac cactgcacta ccctcacccc aaaccctga tcaaacccac catggacttc 1620
ctggagggcag aagaacttgt tatgaaaag ctcaagagag agatcataac ttcgatagc 1680
atacattata cgaagttact ctgcaggtaa aggaattcag gagagccgg gttggcgcg 1740

```


tatttgtgtc ccaaaaaaca gcccaattg ccccaattga ccccaatttg acccaatgag 1800
 ggccccaacc ccggcgagag cccacttca cccacatata aaacatccac cgggtcccaac 1860
 acttgcggtt aaggcgctag ggtactgag totggaatct aogtctgttc agactttgta 1920
 ctagtcttct tgtctggcca tccgggtaac ccatgcccga cgcasaatag actactgaaa 1980
 attttttttc ttgtgtgttg gcacttttag caagggtata aagagccacc gtcccgaat 2040
 taccttctct ctcttttct ctctctcctt gtaactaca acccgaaato gttaaagatt 2100
 tctctctgag tataagaatc attcaccatg gacttctgtt aggcagaaga acttgttatg 2160
 gaaaagctca agagagagaa gccagatata tatcaagaca tctgtccgaa cttaatttaag 2220
 atgacgacat ttgcagctg qacqagcaat agatqagag tctgttctga ctgatgttt 2280
 totatggagt tgtgagtggt agtagacatg atgggtttat atatgtaga tgaatagatg 2340
 tcatittgat ttgcacgatg gaattgagaa ctttgttaac gtacatggga atgtatgaat 2400
 gtgggggttt tgtactgga taactgcgg tcaatggacg ccgttgttca aatatccag 2460
 agatgcgaga aacttgggt caagtgaaca tctctctct gttaagtaa accataact 2520
 atgggtatga tatttagtaa ggaacaagat tgaattctt tggagtctca gaacgtlatt 2580
 ttccgtctcc aagatcaaat tagtagagta atccgggccc gggactcact tcatagctc 2640
 aatttccca taggtgtct acaggtgttt gagatgtgtt aagtagccac catgattgga 2700
 ggtaaagag ccagaatcca ttgttgaggt caagaatac aagatctac agtccatac 2760
 aatgaatata tctctcata tctctcaggt gacacaaagg gtgtctattt tcccaagaa 2820
 tgcgtgaaaa ggcgctgtg tagctgtgag tatgggttgc gttgggtat cctctata 2880
 tgcacgaat agtaggcga gactgacaa aagatctata tatgtagaca gctagaata 2940
 tggatttgat tggataaat tcaattttg cgtgtctcac aatactctc gataagttg 3000
 ggtaaatg gagatggac atgtctgata tctgagcca tgcgaatgag ggcacatto 3060
 gccctatagt gactgtatt acattcact ggcgtctgtt ttaacagtc gtactggga 3120
 aaacccctgc gttacccaac ttaactgct tgcagacact cccctcttg ccagctggcg 3180
 taatagcgaa gagcccgcca ccgctcgcct tcccaacag ttgcgcagcc tgaatggcga 3240
 atgagcgcgc cctgtacgg ccctttaagc ccgcccggat tgggtttac ccgcacgctg 3300
 accgtacac ttgcacgccc cctagcgcct gctccttgc ctttcttccc ttcctttctc 3360
 gccacgttc ccggtttccc ccgtcaagct ctaaatcgga ggcctccttt aggttccga 3420
 tttagtctt taacggcact ccagcccaaa aaacttgatt aggtgtgag ttaactagt 3480
 ggccactgc cctgtagac ggtttttgc cctttgact tggagtccac gttcttaat 3540
 agtgactct tgttccaaa tggacaacca ctcaacctca tctggtcta ttctttgat 3600
 ttataaggga ttttgcgat ttgcctat tggtaaaa atgagctat taacaaaaa 3660
 tttaacgga attttaaaaa aatattaacg cttaaatatt cctgatcggt tatttctcc 3720

ttaacatct gtgcggtatt tcaacccga tcaaggcca cttttgggg aaatgtgcgc 3780
 ggaaccccta ttgtttatt ttctaaata ctttaataa tctatccgt catgagaaa 3840
 taacccctgat aatgcttca ataatttga aaaaagaa gaatagatct tcaacttcc 3900
 cgtgtgcgcc ttattccctt ttgtggcca ttttgcttc ctgtttttg tcaacaaaga 3960
 acgtgtgtga aagtaaaaga tctgaagat caatgggtg caagatggg ttaactgaa 4020
 ctgacttca acagcgttaa gatcttga agttttgc ccgaagaa gtttcaaatg 4080
 atgagcactt ttaagttct cctatgtgga ccgttattat cccgtattga cgcgggcaa 4140
 gacgaactgc gtccgcgat acactatct cagaatgact tggttagta ctcaacagtc 4200
 acagaagac alottacgga tggctgaca gtaagagaat tatgcatgic tgcataacc 4260
 atgactgata acactgggc caacttactt ctgacaaga tggaggaac gaaggagcta 4320
 accgttttt tgcacaaact cggggtcat gtaactgccc ttgatctgt ggaacccgag 4380
 ctgaatgag ccataccaaa ccagcagcgt gacacacga tgcctgtag aatggcaaca 4440
 acgttgcga aactattaac tggcgaacta cttaacttag ctcccgcca acaattaata 4500
 gactggttg aggcgataa agttgcagga ccaactctgc cctccgccc tccggtgac 4560
 tggttattg ctgataaat tggagccggt gagctgggt ctccggtat ctttgagca 4620
 ctggggccag atgttaagcc ctcccgctc gtatttatct acagcaggg gactagga 4680
 actatggatg aacgaatag acagatcgt gagataggt cctcactgat taagcattg 4740
 taactgtcag accaagtta ctcatata cttagattg atttaaaact tcaatttaa 4800
 tttaaaaga tctagtgaa gatcttttt gatacttca tgaacaaa ccttaactg 4860
 gagtttctg tcaactgag ctacagccc gtgaaaaa tcaagagatc ttcttgagat 4920
 ccttttttc tgcggtaat ctgtcgttg caaacaaaaa aacacacgt accagcgtg 4980
 gttgtttgc cggatcaga gtaacaaact cttttccga aggttaactg ctccagcga 5040
 gcgcagatac caataactgt tctctagtg tagcgtagt taggcacaca cttaagaac 5100
 tctgtagac cgcctacata cctgcctctg ctaactctgt taacagtgcc tctgcccagt 5160
 gggtatagt cgtgtttac cgggttgga tcaagacat agttacgga taaggcgag 5220
 cgttcgggt gaacgggggg tctgtgaca ccgcacgct tggagcgaac gactacac 5280
 gaactgagat acctacagc tgaactatga gaagcgcca cgttccga aggagaaaag 5340
 gcgcagagt atccgttaag ccgacgggt ggaacaggag agcgcacag gcagcttca 5400
 gggggaacg cctgtatct ttatagctct gtcgggttc gccacotctg acttgagct 5460
 cgattttgt gatctctc agggggcg agcctatga aaaaacgag caacgggccc 5520
 ttttaaggt tctggcctt ttgtggctt ttgtctaaa tgttcttcc tgcgttatcc 5580
 cctgattctg tggataacg tattaccgac tttagtgag ctgatauog tccgcagag 5640

```

cgaacgacgg agcgacgaga gtccgtgagc gaggagcggg aagagcgccc aatacgcaaa 5700
ccgcctctcc ccgcgcggtg gcgcgttcac taatgcagct ggcgcgccac tgcgcagatc 5760
cagttacac tgcattaatt tggcgccaat aatttaaaa aatcgtgta tataatatta 5820
tatgtattat atatatatat catgatgata ctgcacgtca tgtcccatgt ctaaatagac 5880
agactccatc tgcgcctccc aactgatgtt ctaaatattt aagggtcatc ctgcatttgt 5940
ttaataata acagatccca totscgcctt ccaaatgatg ttctaaaaat atattgtatg 6000
aacttatttt tattacttag tattattaga caacttaott gotttargaa aaacacttcc 6060
tatttaggaa acatttata atggcagthc gttcatttaa caatttatgt agataaatg 6120
ttataaatgc gtatggcaaa tottaatat ggtatgata aatgatctat gatttgcta 6180
attcgaaatc aacgcacgcg aaaaaatcc ctgtacaaac ataaatagtc atcgagaaat 6240
atcaactatc aaagacacgc tatccacgcg ttactattga gatttattt ggcgcgaaat 6300
cacacactca actgtctttc tctctcttag aaatacaggt acagtatgtt actattctca 6360
ttgttcatac ttctagtcat ttctccccc atattccttg gatttctct caatgaatga 6420
caittctatc tgcacattca acattatata taagatatac caaagtagcg gtatagtggc 6480
aatcaaaaag ctctctgtgt gtgtctctcg tatitatttt tatitctaat atccatttaa 6540
ggatataat tatttctgt tatataatcc ttitgtttat taactgggct ggtatcatas 6600
aggtattttg atttaatttt ttguttaaat tcaatccccc ctgcttcagt gtcaactgta 6660
atggtaggaa attacactac ttttgaaaga gcaaaaaaaa tgaagaaaaa aaaaantcgt 6720
atttcaggt tagaagttcc gaaatatata gaatgggta tgggtacat ttgttttga 6780
acgtaaaagt tgcgtccctt gagatattgt acatttttgc tttaacagtt acaagtcacat 6840
cgtacaacta tgtactactg ttgtgcata ccaaacagtt ttgtttgttt ttitttgttt 6900
tttttttttc taatgatcca ttacogctat gttacactac ttgtacttgt agtaagocgg 6960
gttatggcg ttcaatlaat catagactta tgaatctgca cgggtgctgc tgcagttcac 7020
tttttagcta tgcattgac ttgggtgtaa tatgggatac tghttgaaaa tcaacggatg 7080
ctcctaggtt cccaaaatcc cagtgcttcc tctccagtcg caaagtaag taccacacag 7140
gttttcgccc gaacattcca cgtgcagcaa cgtcgtgtgc cgtgttaaaa tctggggggg 7200
gggaacacgg acanaggtct ctgtggggg ccgaatgaga gcaaacagcg ggcggtgtgt 7260
ataagggcat ttttgccat ttccctctct cctgtctctc cgaaggtgat ggcgttgtgc 7320
gtcctctatt tctttttatt tcttttgttt ttattctctt gactaacgat ttggtttgat 7380
ttcctcaacc ccaacaaatc aagctcgggc cgaagaaat atatatatac ggcacacgtc 7440
gccttgtaga caaacgtcca ctactctac gatcacccc gtacgtatgt tagtagacaa 7500
caatgatag ttggagcaag ggaagaaagt agagtgtaaa agactcacta tgttcgggcg 7560
ttatctgac caatagccaa agtctggagt ttctgagaga aaagggcaag atacgtatgt 7620

aacaagoga cgcattgtae aataatacgg gaggcatgta tcaatagagag ttatgggttc 7680
gatgatgcca ctggtgcctg gtatgacttt atacggctga ctacatattt gtccctagac 7740
atacaattac agtcaagcac ttacccttgg acatctgtag gtaccccccg gccaaagaca 7800
tctcagcgtg tegtattgtc gattggcgta gtcctctgcg tegtcaattg gctcccatct 7860
actttctctt gcttggtac acaacagctg tctgtatggt ctggttttgc tgccttatct 7920
atcctccacg tattacaaac tcaaatgac atgaigtgat tgggtctaca cttoatatac 7980
agagataagg agtagcacg ttgcataaaa agcccaactc taatcagttt cttoctttct 8040
tgtactagt acaaggtga ttgcgaatat ctggaagctt agttgvcctt aaaaaatca 8100
aaaaagcaa aaaaagaaa acgaacaaac acagttttga gaacagggag gtaacgaagg 8160
atcgtakata tatatatata tatatatata caggatccc gacgcggccc ttgtattctt 8220
cctcaaaccc aacacttctc acaacctaa ttcaaccca tgtccgttgc atccaaagctc 8280
gtctctacg tccgcgcgcg catcgccgtg gtactctttg ccgcctgtgc caactacgcg 8340
gtcctggcgt ccaccattct caccgcata cgaacagag gcttgcccca atgacagtt 8400
gccagagcct tctactactc gttgcgcata ttctgggta tcaagctcaa gctgcragc 8460
cggcaggtga ccgaacocg cgtctggat gctcccaaga tcaagctgcg caaacacacc 8520
aagcccatgg acgacatcac caaacacctg cccgcacatc gattctgat ttcaaaccaa 8580
cagaacgaaa tggacattct gttgctcgtt cgcactctcc cccagtaactg ctccgtcaacc 8640
gccaaaagg cctcaagtg gtacccctct ctgggcagct tcatggcgtt gtcgcgcacc 8700
atctctctgg acgaaagaa ccgaaccaa tccgtgcaga cctcggcgcg cgcctgcaag 8760
accatccaga cgggcaacgg aggcgaaggg cagagcgtct tcatgttccc cgaaggaaac 8820
cgatctact ccaaggacgt cggcatcatg ccttcaaga aggctgtttt ccaactggcg 8880
gtccagtccg gcgtcccat tgtccccgtg gtggccaga acacctcccg aatgtttct 8940
ttcgcccgag gcaagctgga cgcgcgagag atccttgtcg acgtcccgag ccccatgag 9000
accagggtc tggagccag caactgcac gctctcatgg ccaacctta taaggccatg 9060
tgcgagctg ccgaacagat tggctacgtt ggcagagaa ctatgtaggc ggcgcagta 9120
gaagataat atataaatc attgagatat taatgcgct agattagaga gctcataact 9180
gctcggagag aagccaagac gactactcaa aggggattac accatccata tccacagaca 9240
caagctggcg aaagtttcta tataactttt ccgaataaco gtatgttcog atgttatcaa 9300
tggggcgacg caggaattca ggcacttcgg tgtctcgggg tgaatggcg ttcttggcct 9360
ccatcaagtc gtaccatgtc ttaatttgcg tgcacagta aaacagaagc agatgaagaa 9420
tgaacttga gtgaaggaa ttaattgtaa cgaactgaa atttgacag atattgtctc 9480
cgcgtggag ctccagtttt ttctcccttt agtggaggtt aatttcagc ttgcgtaat 9540

catgtcata gctgttctc gtgtgaaatt gttatccgt ccaacgcttc cacacacc 9598

```

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- [US6187366A R0011](#)
- [US6187366B R0011 R03271](#)
- [US6187366B R0011 R03271](#)
- [US2008002343A1 R0031 R0041 R0081 R0251 R01121 R01121 R01141 R01141 R01421 R01431 R01441 R01451 R01481 R01491 R01521 R01581 R01591 R01641 R01681 R02011 R02041 R02121 R02321 R02391 R0243](#)

• US7228452B (0004) (0049) (0153) (0171) (0175) (0178)
 • US2006015681A1 (0004) (0005) (0029) (0033) (0050) (0133) (0295) (0334) (0367) (0387) (0387)
 • US2006026419A1 (0007)
 • US2006009492A1 (0008)
 • US2006002220A1 (0009)
 • US2006016697A1 (0023) (0023) (0023) (0128) (0339) (0387) (0387) (0387) (0387) (0387)
 • US2010007536A1 (0023) (0023) (0023) (0113) (0113) (0113) (0220) (0288) (0288) (0288) (0288) (0288) (0288) (0288) (0306)
 • WC2005047480A (0053)
 • US2006005419A1 (0054) (0055) (0115) (0116) (0250) (0272) (0281) (0367) (0387) (0387) (0387) (0387) (0387) (0387) (0387) (0387) (0387) (0387) (0387) (0387)
 • US7645604A (0055) (0055) (0113) (0113) (0222) (0222) (0249) (0305)
 • US7709278A (0055) (0113) (0113) (0222) (0249) (0285) (0369)
 • US2006025452A1 (0055) (0113) (0113) (0387) (0387)
 • US2006025452A1 (0055) (0113) (0113) (0252) (0281) (0387) (0387)
 • WC2006027081A1 (0061) (0138) (0334) (0368) (0368) (0368) (0387)
 • WC2006028105A1 (0074) (0113) (0137) (0158) (0179) (0220) (0249) (0256) (0281) (0308) (0387) (0387) (0387) (0387)
 • US7125672B (0069)
 • WC2006012824A1 (0113) (0387) (0387)
 • US7256032B (0113) (0113) (0113) (0222) (0249) (0295) (0305) (0367) (0387)
 • WC2006012419A1 (0113) (0387) (0387)
 • US7765602B (0113) (0328) (0249) (0249) (0250) (0281) (0308)
 • US7765602B (0113)
 • US2006027452A1 (0113) (0113) (0281) (0387) (0387)
 • US7465793A (0113)
 • US7550498A (0113) (0235) (0235)
 • US7504258A (0113) (0113) (0223) (0250) (0272) (0387) (0387)
 • US7470522A (0113) (0222) (0262) (0272) (0387)
 • US2010015958A1 (0118)
 • US837877A (0121) (0282) (0285)
 • WC200604076817A2 (0133) (0133) (0133) (0138)
 • WC200604076817A2 (0133) (0133)
 • WC200604076817A2 (0133) (0133)
 • WC200604076817A2 (0133) (0133) (0341) (0342) (0342)
 • WC2006060638A (0137)
 • WC2006061315A (0137) (0334) (0334) (0387)
 • WC2006061414A (0137)
 • WC2006061424A (0143) (0148) (0223)
 • US8880741A (0157)
 • US907176A (0157)
 • US728255B (0158) (0262) (0288) (0281) (0309)
 • US810141A (0161)
 • US8418280A (0171)
 • US879793B (0178)
 • US848564A (0178)
 • US879793B (0182)
 • US879793B (0187)
 • US20060026419A1 (0211)
 • WC20060026419A1 (0227)
 • US720235B (0226) (0249) (0258) (0272) (0265) (0306)
 • US2006009402A1 (0228) (0249) (0249) (0258) (0262) (0272) (0281) (0285) (0295) (0306) (0306) (0344) (0373) (0374)
 • US745854B (0249) (0272)
 • US2006025655A1 (0265) (0324)
 • US20060145667A (0324)
 • WC2006076377A (0334) (0387)
 • US712856B (0334) (0349)
 • US2006016697A1 (0349) (0387)
 • CA8674 (0387)
 • US618726B (0387)
 • WC20060145667A (0387)
 • WC2006076377A (0387) (0387) (0387) (0387)
 • US2007011719A1 (0387) (0387) (0387) (0387)
 • WC2006012122A (0387) (0387)
 • WC2006012122A (0387) (0387)
 • WC2006076377A (0387)
 • US2006016697A1 (0387)
 • WC2007136871A (0387) (0387)
 • US2007029292A1 (0387) (0387)
 • WC2007136871A (0387) (0387)
 • US2007027163A1 (0387) (0387)
 • WC2006013752A (0387) (0387)
 • WC200712388A (0387) (0387)
 • US2007028028A1 (0387) (0387)
 • WC2006054585A (0387) (0387)
 • US2006012122A1 (0387) (0387)
 • WC2006047480A (0387) (0387)
 • WC2007046217A (0387) (0387)
 • US2007008742A1 (0387)

Non-patent literature cited in the description

- TAMAKI, H. et al.J. Biol. Chem., 2007, vol. 282, 34288-34298 [[060601](#)]
- STAHL, U. et al.FEBS Letters, 2008, vol. 582, 305-309 [[060601](#)]
- CHEN, Q. et al.FEBS Letters, 2007, vol. 581, 5511-5516 [[060601](#)]
- BENGHEZAL, M. et al.J. Biol. Chem., 2007, vol. 282, 30845-30855 [[060601](#)]
- RIEKHOFF et al.J. Biol. Chem., 2007, vol. 282, 28344-28352 [[060601](#)]
- HISHIKAWA et al.Proc. Natl. Acad. Sci. U.S.A., 2008, vol. 105, 2830-2835 [[060601](#)]
- KIM, Y.S.S.K. BANGJ. Biol. Chem., 1985, vol. 260, 5098-5104 [[060601](#)]
- KIM, Y.S.H.Z. CHAEBiochem J., 1991, vol. 273, 511-516 [[060601](#)]
- KIM, Y.S.S.W. KANGBiochem J., 1994, 297, 327-333 [[060601](#)]
- DISTEL et al.J. Cell Biol., 1996, vol. 135, 1-13 [[060601](#)]
- KIEL, J. A. K. W. et al.Traffic, 2006, vol. 7, 1291-1303 [[060601](#)]

- WEETEIn: Fungal Lipid BiochemistryPlenum19800000 [f0084f](#)
- YONGMANITCHAIWARDAppl. Environ. Microbiol., 1991, vol. 57, 419-25 [f0088a](#)
- ALTSCHUL, S. F. et al.J. Mol. Biol., 1993, vol. 215, 403-410 [f00887f](#)
- ALTSCHUL et al.J. Mol. Biol., 1990, vol. 215, 403-410 [f0101f](#)
- Comput. Methods Genome Res.W. R. PEARSONProc. Int. Symp.Plenum19920000111-20 [f0101f](#)
- SAMBROOK, J.FRITSCH, E. F.MANIATIS, T.Molecular Cloning: A Laboratory ManualCold Spring Harbor Laboratory19890000 [f0102f](#) [f0200f](#)
- SILHAVY, T. J.BENNAN, M. L.ENQUIST, L. W.Experiments with Gene FusionsCold Spring Harbor Laboratory19840000 [f0102f](#)
- AUSUBEL, F. M. et al.Current Protocols in Molecular BiologyGreene Publishing Assoc. and Wiley-Interscience19870000 [f0102f](#) [f0200f](#)
- KENNEDYWEISSJ. Biol. Chem., 1956, vol. 222, 193-214 [f0123f](#)
- LANDS, W.E.J. Biol. Chem., 1958, vol. 231, 883-888 [f0126f](#)
- STYMNE S.A.K. STOBARTBiochem J., 1984, vol. 223, 2305-314 [f0127f](#)
- J. Biol. Chem., 2009, vol. 284, 11-5 [f0128f](#)
- J. Lipid Res., 2009, vol. 50, S46-S51 [f0128f](#)
- SHINDOU et al.Biochem. Biophys. Res. Comm., 2009, vol. 383, 320-325 [f0128f](#)
- SIMILARLY, LEWIN, T.W. et al.Biochemistry, 1999, vol. 38, 5764-5771 [f0130f](#)
- YAMASHITA et al.Biochim. Biophys. Acta, 2007, vol. 1771, 1202-1215 [f0130f](#)
- STYMNE S.A.K. STOBARTBiochem. J., 1984, vol. 223, 2305-314 [f0130f](#)
- DOMERGUE, F. et al.J. Bio. Chem., 2003, vol. 278, 35115- [f0133f](#)
- ABBADI, A. et al.The Plant Cell, 2004, vol. 16, 2734-2748 [f0133f](#)
- Methods in Enzymology, 1991, vol. 194, 186-187 [f0137f](#)
- CHEN, D. C. et al.Appl. Microbiol. Biotechnol., 1997, vol. 48, 2232-235 [f0137f](#)
- JURETZEK et al.Yeast, 2001, vol. 18, 97-113 [f0160f](#)
- SOUTHERN.J. Mol. Biol., 1975, vol. 98, 503- [f0160f](#)
- KROCZEKJ. Chromatogr. Biomed. Appl., 1993, vol. 618, 1-2133-145 [f0160f](#)
- NAKAHARA, T. et al.Ind. Appl. Single Cell Oils1992000061-97 [f0173f](#)
- Z. JACOBSCritical Reviews in Biotechnology, 1992, vol. 12, 5/6463-491 [f0177f](#)
- A. SINGHO. WARDAdv. Appl. Microbiol., 1997, vol. 45, 271-312 [f0177f](#)
- The New Science of Omega-3 Fatty Acids: Differential Nutritional PharmacologyGILLIES, P.Texas Human Nutrition ConferenceTexas A&M University20100200 [f0204f](#)
- T. J. SILHAVY, M. L. BENNAN, W. ENQUISTExperiments with Gene FusionsCold Spring Harbor Laboratory19840000 [f0200f](#)
- Manual of Methods for General BacteriologyAmerican Society for Microbiology19940000 [f0207f](#)
- THOMAS D. BROCKBiotechnology: A Textbook of Industrial MicrobiologySinauer Associates19890000 [f0207f](#)
- BLIGH, E. G.DYER, W. J.Can. J. Biochem. Physiol., 1959, vol. 37, 911-917 [f0213f](#)
- ROUGHAN, G.NISHIDA I.Arch Biochem Biophys., 1990, vol. 276, 138-46 [f0213f](#)
- DUJON, B. et al.Nature, 2004, vol. 430, 699535-44 [f0225f](#)
- THOMPSON et al.Nucleic Acids Res., 1994, vol. 22, 4673-4680 [f0236f](#)
- ALTSCHUL et al.Nucleic Acids Res., 1997, vol. 25, 3389-3402 [f0238f](#)
- ALTSCHUL et al.FEBS J., 2005, vol. 272, 5101-5109 [f0238f](#)
- HIDEO SHINDOUMIKI ETOARYO MORIMOTOATAKAO SHIMIZUIdentification of membrane O-acyltransferase family motifsBiochemical and Biophysical Research Communications, 2009, vol. 383, 3320-325 [f0238f](#) [f0238f](#) [f0238f](#)
- TAL M. LEWINPING WANGROSALIND A. COLEMANAnalysis of Amino Acid Motifs Diagnostic for the sn-Glycerol-3-phosphate Acyltransferase ReactionBiochemistry, 1999, vol. 38, 1857645771- [f0238f](#)
- ATSUSHI YAMASHITA HIROKI NAKANISHIAHIROSHI SUZUKIARYO KAMATAA KEN TANAKAAKEIZO WAKUATAKAYUKI SUGIURATopology of acyltransferase motifs and substrate specificity and accessibility in 1-acyl-sn-glycero-3-phosphate acyltransferase 1Biochimica et Biophysica Acta, 2007, vol. 1771, 91202-1215 [f0238f](#) [f0238f](#) [f0238f](#)

FORBEDREDE, OPTIMEREDE STAMMER AF YARROWIA LIPOLYTICA TIL FREMSTILLING AF
HØJKONCENTRERET EICOSAPENTAENSYRE

PATENTKRAV

1. Rekombinant *Yarrowia* sp.-værtscelle, der producerer en olie, og som omfatter:
 - 5 (i) mindst 50 vægtprocent eicosapentaensyre målt som en vægtprocent af de samlede fedtsyrer; og
 - (ii) et forhold på mindst 3,5 mellem eicosapentaensyre, målt som en vægtprocent af de samlede fedtsyrer, og linolensyre, målt som en vægtprocent af de samlede fedtsyrer;

hvor værtscellen omfatter:

 - 10 (a) mindst ét multizym, der omfatter et polypeptid, som har mindst én delta-9-elongase forbundet med mindst én delta-8-desaturase;
 - (b) mindst et peroxisombiogenesefaktorprotein, hvis ekspresion er blevet nedreguleret; og
 - (c) mindst ét rekombinant konstrukt, der omfatter et ikke-nativt kimærisk gen, der omfatter en nukleotidsekvens, som koder for en heterolog malonyl-CoA-syntetase, og mindst ét
 15 rekombinant konstrukt, der omfatter et ikke-nativt kimærisk gen, som omfatter en nukleotidsekvens, der koder for en acyl-CoA-lysophospholipidacyltransferase (LPLAT), således at LPLAT overudtrykkes i værtscellen, hvor de kimæriske gener omfatter regulatoriske og kodende sekvenser, der ikke er fundet sammen i naturen.
 2. Rekombinant værtscelle ifølge krav 1, hvor nukleotidsekvensen, der koder for malonyl-CoA-syntetasen, koder for et polypeptid, der i alt væsentligt består af en aminosyre, der er udvalgt fra
 20 gruppen bestående af SEQ ID NO: 40 og SEQ ID NO: 42.
 3. Rekombinant værtscelle ifølge kravene 1 eller 2, hvor nukleotidsekvensen, der koder for acyl-CoA-lysophospholipidacyltransferasen, koder for et polypeptid, der i alt væsentligt består af en aminosyresekvens, der er udvalgt fra gruppen bestående af SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO:
 25 25, SEQ ID NO: 29, SEQ ID NO: 31 og SEQ ID NO: 32.
 4. Rekombinant værtscelle ifølge et hvilket som helst af kravene 1 til 3, hvor LPLAT er en LPAAT.
 5. Rekombinant værtscelle ifølge krav 4, hvor LPAAT omfatter en sekvens med SEQ ID NO: 29.
 6. Rekombinant værtscelle ifølge et hvilket som helst af kravene 1 til 5, hvor multizymet omfatter en linker mellem delta-9-elongasen og delta-8-desaturasen, hvilken linker er udvalgt fra gruppen bestående af:
 30 SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6 og SEQ ID NO: 7.
 7. Rekombinant værtscelle ifølge et hvilket som helst af kravene 1 til 6, hvor multizymet omfatter en sekvens, der er udvalgt fra gruppen bestående af: SEQ ID NO: 9, SEQ ID NO: 11 og SEQ ID NO: 13.
 8. Rekombinant værtscelle ifølge et hvilket som helst af kravene 1 til 7, hvor forholdet mellem
 35 eicosapentaensyre og linolensyre i olien er mindst 4,5.
 9. Rekombinant værtscelle ifølge et hvilket som helst af kravene 1 til 7, hvor cellen producerer mindst 55 vægtprocent eicosopentaensyre, målt som en vægtprocentdel af de samlede fedtsyrer.
 10. Rekombinant værtscelle ifølge et hvilket som helst af kravene 1 til 8, hvor cellen producerer

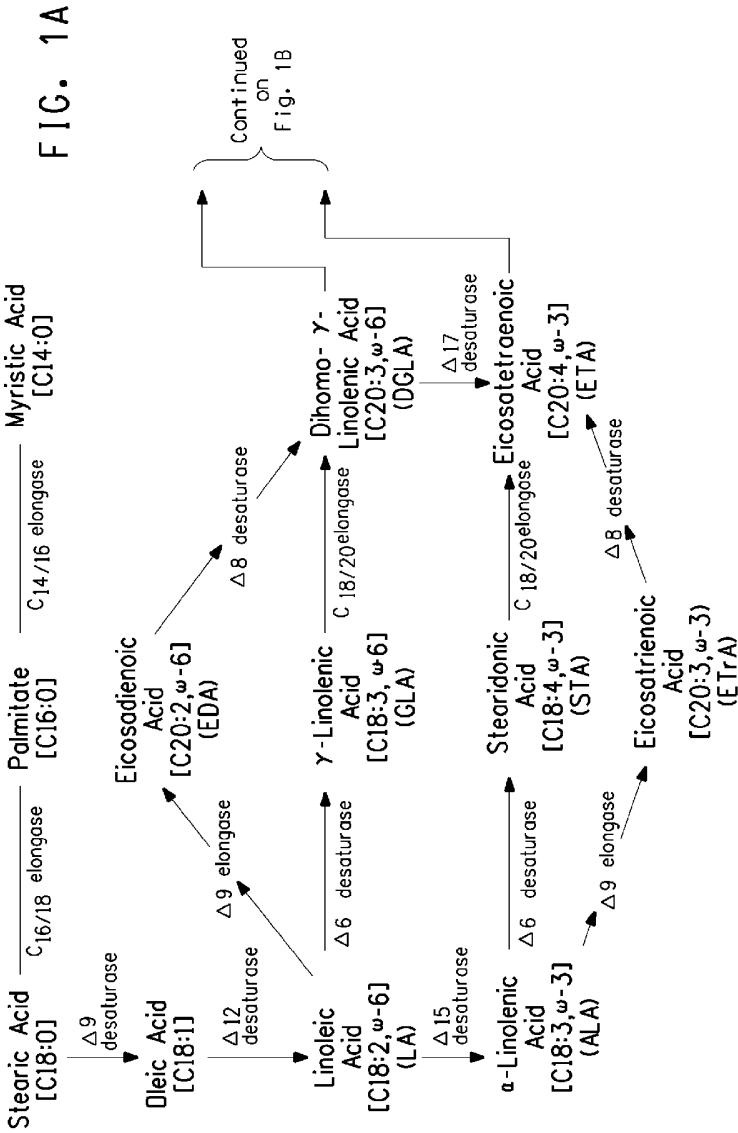
-2-

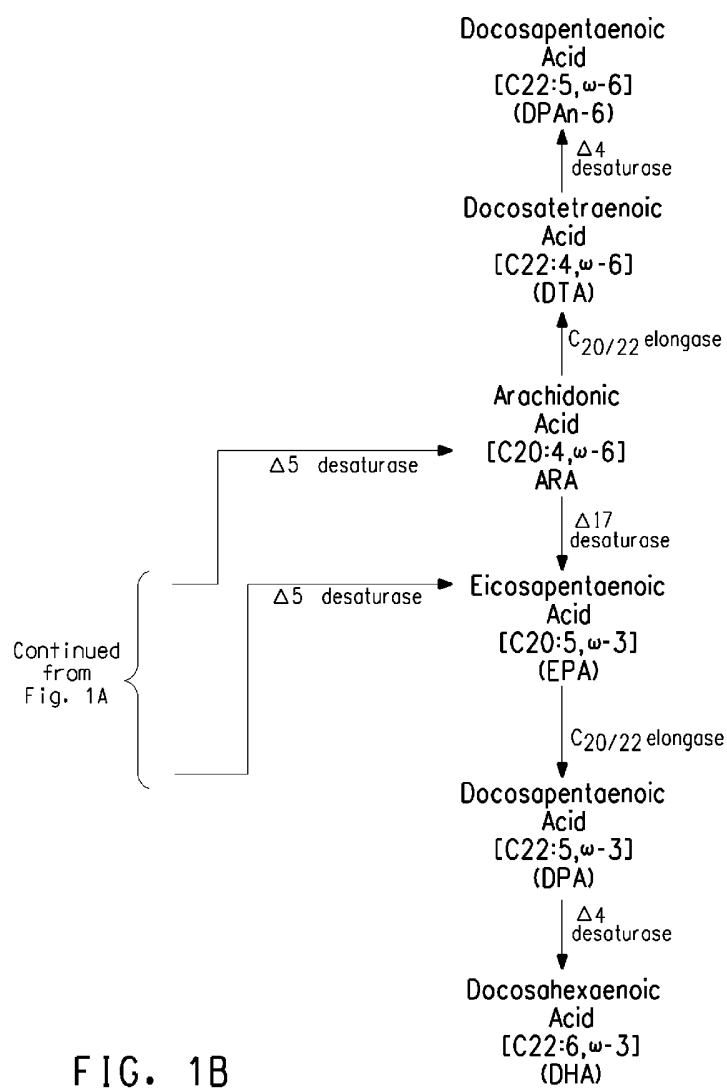
mindst 60 vægtprocent eicosapentaensyre, målt som en vægtprocentdel af de samlede fedtsyrer.

11. Fremgangsmåde til fremstilling af en mikrobiel olie, der omfatter eicosapentaensyre, hvilken fremgangsmåde omfatter:

- a) dyrkning af værtscellen ifølge et hvilket som helst af kravene 1 til 10, hvor der produceres
- 5 en mikrobiel olie, der omfatter eicosapentaensyre; og
- b) eventuelt indvinding af den mikrobielle olie fra trin (a).

DRAWINGS





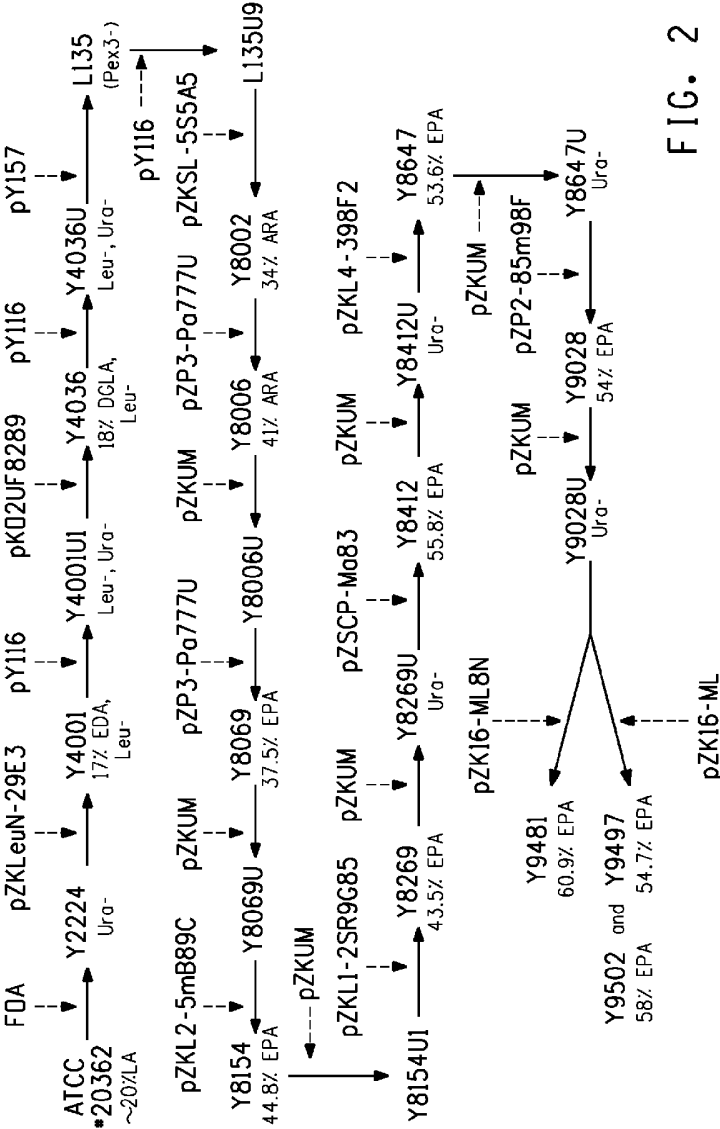


FIG. 2

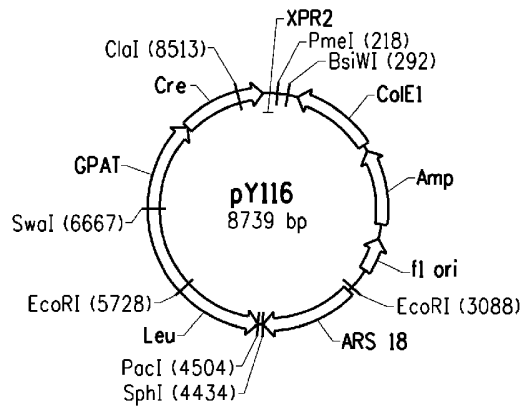


FIG. 3

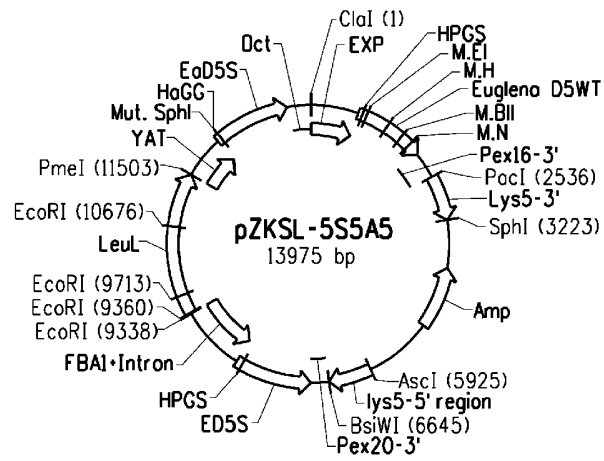


FIG. 4A

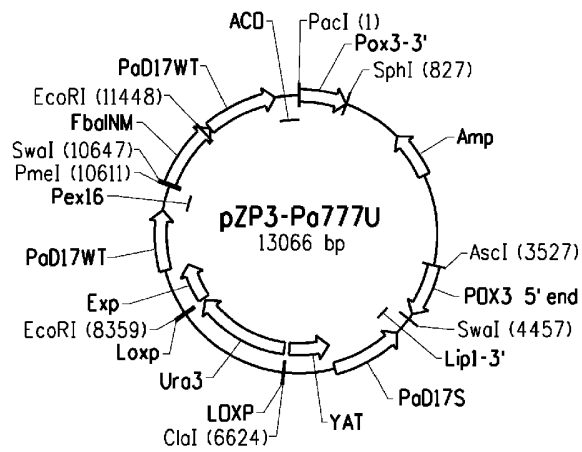


FIG. 4B

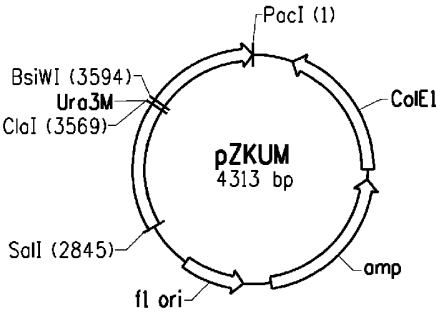


FIG. 5A

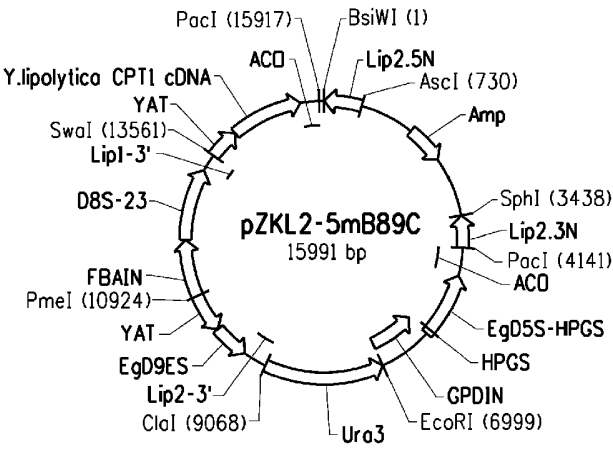


FIG. 5B

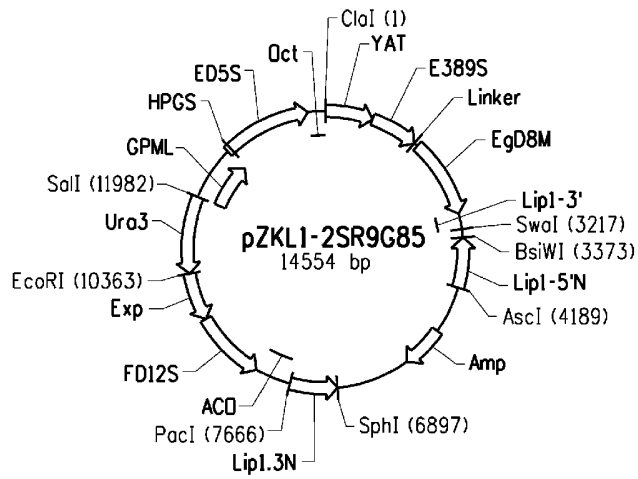


FIG. 6A

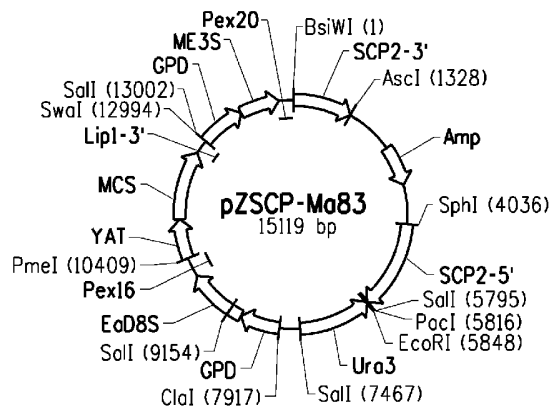


FIG. 6B

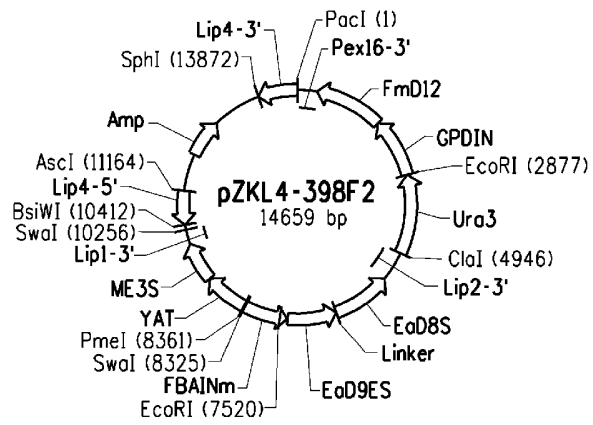


FIG. 7A

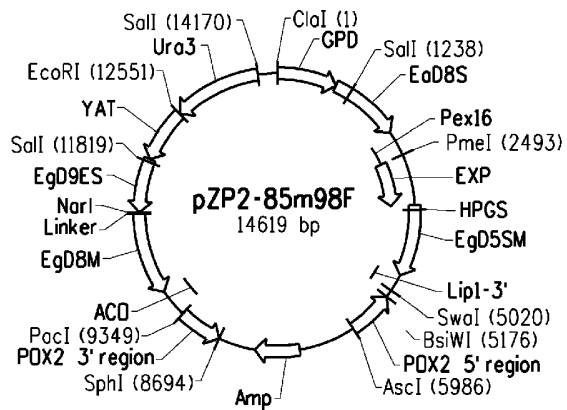


FIG. 7B

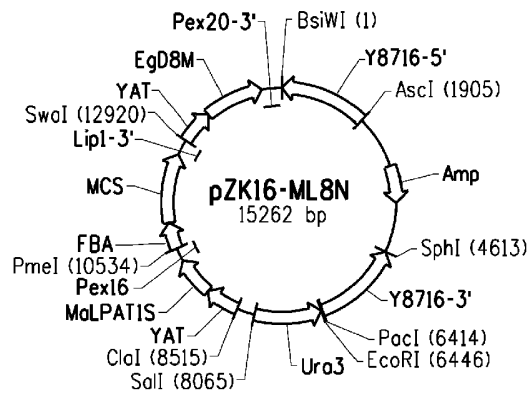


FIG. 8A

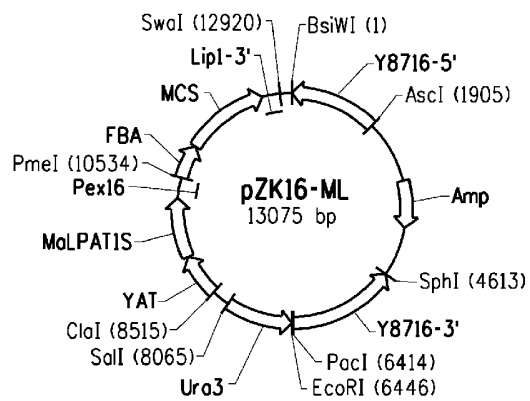


FIG. 8B

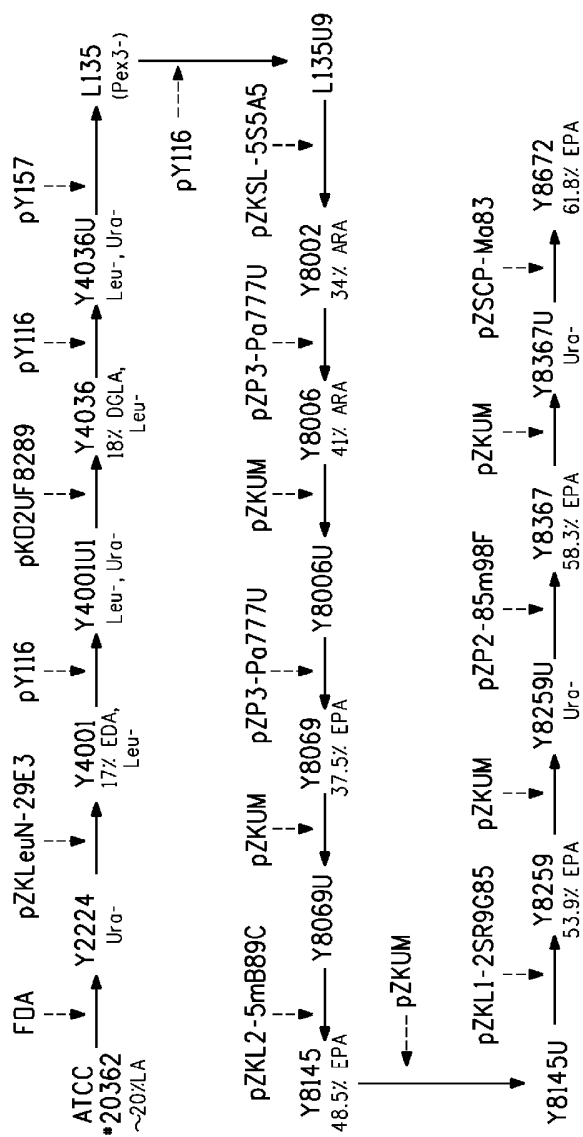


FIG. 9

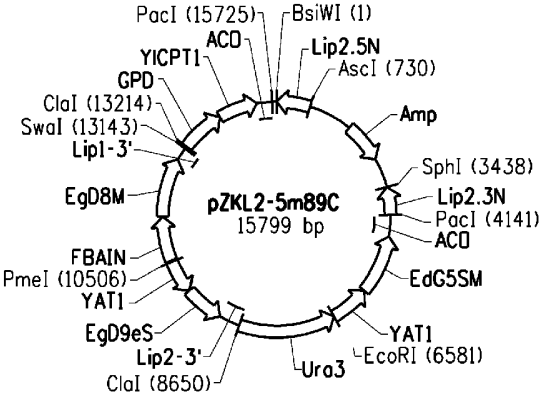


FIG. 10A

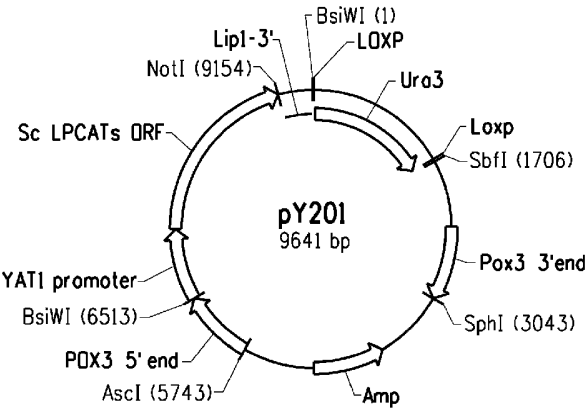


FIG. 10B

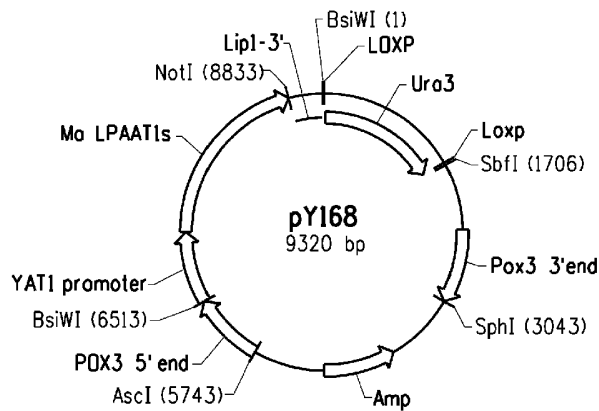


FIG. 11A

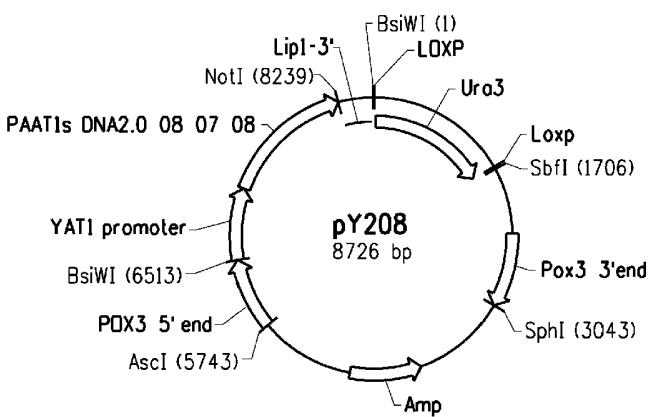


FIG. 11B

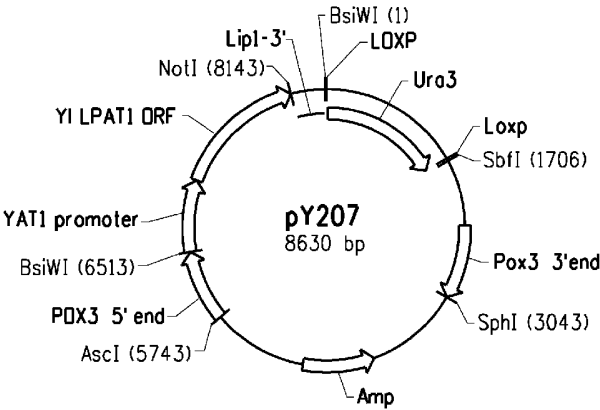


FIG. 12A

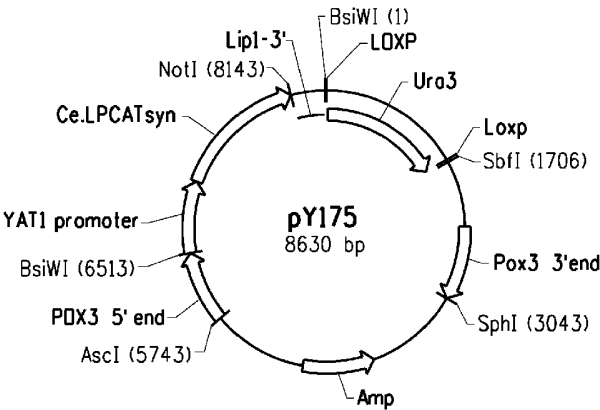


FIG. 12B

FIG. 13

Strain	LA % TFAs	EPA % TFAs	Ratio EPA/LA
Y9481	11.00	60.90	5.54
Y9486	11.90	60.30	5.07
Y9497	11.30	58.70	5.19
Y9502	12.70	57.00	4.49
Y9504	11.30	59.90	5.30
Y9508	13.10	58.70	4.48
Y9510	11.70	58.90	5.03
Y8143	18.10	50.30	2.78
Y8144	18.00	50.60	2.81
Y8145	18.60	48.50	2.61
Y8146	18.70	48.30	2.58
Y8256	18.80	49.90	2.65
Y8259	16.90	53.90	3.19
Y8367	14.20	58.30	4.11
Y8368	17.90	52.50	2.93
Y8369	15.60	55.80	3.58
Y8370	15.70	56.40	3.59
Y8666	15.10	62.20	4.12
Y8669	15.70	61.50	3.92
Y8670	17.00	60.90	3.58
Y8672	16.10	61.80	3.84

Strain	LA % TFAs	EPA % TFAs	Ratio EPA/ LA
Y8404	20.40	51.10	2.50
Y8405	20.50	51.40	2.51
Y8406	20.30	51.20	2.52
Y8407	20.50	51.50	2.51
Y8408	20.20	51.20	2.53
Y8409	20.60	51.00	2.48
Y8410	20.50	50.90	2.48
Y8411	19.90	52.90	2.66
Y8412	19.00	55.80	2.94
Y8413	19.90	52.40	2.63
Y8414	19.90	49.40	2.48
Y8415	16.00	59.00	3.69
Y8647	20.30	53.60	2.64
Y8648	18.60	56.70	3.05
Y8649	18.80	55.80	2.97
Y8650	18.80	56.10	2.98
Y9028	19.80	54.50	2.75
Y9029	19.80	53.80	2.72
Y9031	20.10	52.30	2.60
Y9477	10.00	61.40	6.14

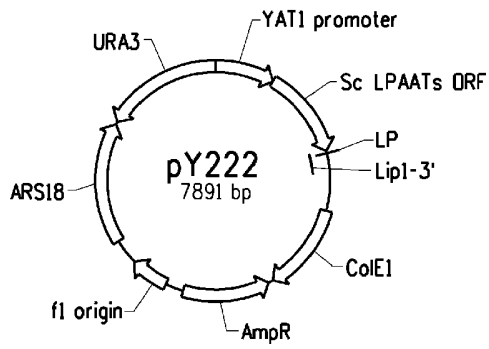


FIG. 14A

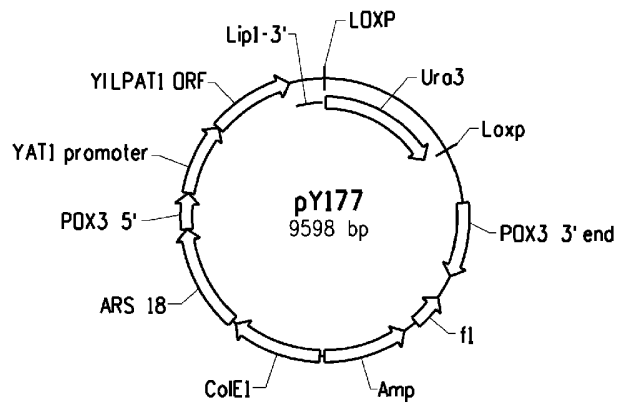


FIG. 14B