(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

PCT

(43) International Publication Date
23 December 2010 (23.12.2010)

(10) International Publication Number
WO 2010/147900 A1

(51) International Patent Classification:
C12P 21/06 (2006.01) *C12P 7/64* (2006.01)

(21) International Application Number:
PCT/US2010/038527

(22) International Filing Date:
14 June 2010 (14.06.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/187,359 16 June 2009 (16.06.2009) US

(71) Applicant *(for all designated States except US)*: **E. I. DU PONT DE NEMOURS AND COMPANY** [US/US]; 1007 Market Street, Wilmington, Delaware 19898 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **YADAV, Narendra, S.** [US/US]; 126 Tuckaway Drive, Wilmington, Delaware 19803 (US). **ZHANG, Hongxiang** [US/US]; 3808 Rotherfield Lane, Chadds Ford, Pennsylvania 19317 (US). **ZHU, Qun** [US/US]; 544 Revere Road, West Chester, Pennsylvania 19382 (US).

(74) Agent: **CHRISTENBURY, Lynne, M.**; E. I. du Pont de Nemours and Company, Legal Patent Records Center, 4417 Lancaster Pike, Wilmington, Delaware 19805 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

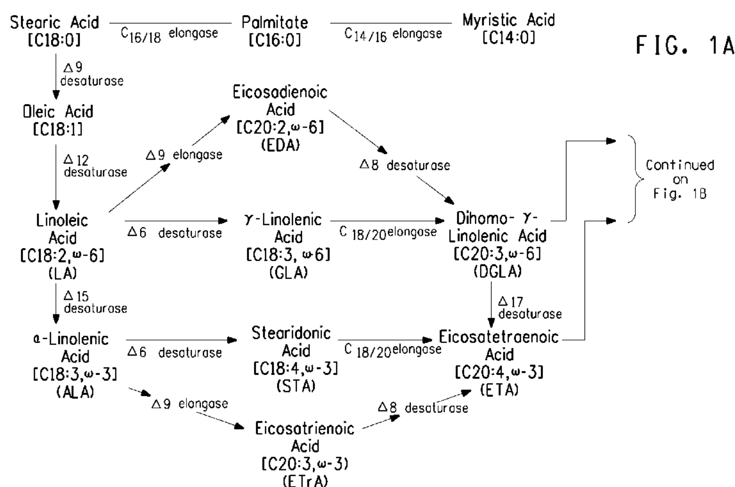
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: IMPROVEMENT OF LONG CHAIN OMEGA-3 AND OMEGA-6 POLYUNSATURATED FATTY ACID BIOSYNTHESIS BY EXPRESSION OF ACYL-CoA LYSOPHOSPHOLIPID ACYLTRANSFERASES



(57) Abstract: Methods for increasing C_{18} to C_{20} elongation conversion efficiency and/or Δ4 desaturation conversion efficiency in long-chain polyunsaturated fatty acid ["LC-PUFA"]-producing recombinant oleaginous microbial host cells are provided herein, based on over-expression of acyl-CoA:lysophospholipid acyltransferases ["LPLATs"] (e.g., Ale1, LPAAT, LPCAT). Production host cells and oils produced by the methods of the invention are also claimed.

TITLE

IMPROVEMENT OF LONG CHAIN OMEGA-3 AND OMEGA-6 POLYUNSATURATED FATTY ACID BIOSYNTHESIS BY EXPRESSION OF ACYL-CoA LYSOPHOSPHOLIPID ACYLTRANSFERASES

This application claims the benefit of U.S. Provisional Applications No. 61/187366, No. 61/187368 and No. 61/187359, each filed June 16, 2009 and each of which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION

This invention is in the field of biotechnology. More specifically, this invention pertains to methods for increasing C_{18} to C_{20} elongation conversion efficiency and/or Δ4 desaturation conversion efficiency in long-chain polyunsaturated fatty acid ["LC-PUFA"]-producing recombinant oleaginous microbial host cells, based on over-expression of genes encoding acyl-CoA:lysophospholipid acyltransferases ["LPLATs"].

BACKGROUND OF THE INVENTION

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. The specific polar head group (e.g., phosphatidic acid, chloline, ethanolamine, glycerol, inositol, serine, cardiolipin) determines the name given to a particular glycerophospholipid, thus resulting in 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.

Glycerophospholipid biosynthesis is complex. Table 1 below summarizes the steps in the *de novo* pathway, originally described by Kennedy and Weiss (*J. Biol. Chem.*, 222:193-214 (1956)):

Table 1: 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 1st 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") **Or** PA → Cytidine Diphosphate Diacylglycerol ("CDP-DG")	Phosphatidic acid phosphatase [E.C. 3.1.3.4] removes a phosphate from PA; DAG can subsequently be converted to PC, 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]) 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 PI, PS, PG or CL

Following their *de novo* synthesis, glycerophospholipids can undergo rapid turnover of the fatty acyl composition at the *sn*-2 position. This "remodeling", or "acyl editing", is important for membrane structure and function, biological response to stress conditions, and manipulation of fatty acid composition and quantity in biotechnological applications. Specifically, the remodeling has been attributed to deacylation of the glycerophospholipid and subsequent reacylation of the resulting lysophospholipid.

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 A_2, 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. LPCAT activity is associated with two structurally distinct protein families, wherein one belongs to the LPAAT family of proteins and the other belongs to the membrane bound *O*-acyltransferase ["MBOAT"] family of proteins.

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)).

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)). Numerous LPLATs have been reported in public and patent literature, based on a variety of conserved motifs.

The effect of LPLATs on polyunsaturated fatty acid ["PUFA"] production has also 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. For example, Intl. App. Pub. No. WO 2004/076617 describes the isolation of an LPCAT from *Caenorhabditis elegans* (clone T06E8.1) and reports increase in the efficiency of $\Delta 6$ desaturation and $\Delta 6$ 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.

Furthermore, Example 16 of Intl. App. Pub. No. WO 2004/087902 describes the isolation of *Mortierella alpina* LPAAT-like proteins (encoded by the proteins of SEQ ID NO:93 and SEQ ID NO:95, 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 in SEQ ID NO:93. Intl. App. Pub. No. WO 2004/087902 also reports expression of one of these proteins using similar methods to those of Intl. App. Pub. No. WO 2004/076617, which results in similar improvements in EDA and ETA biosynthesis.

Both Intl. App. Publications No. WO 2004/076617 and No. WO 2004/087902 teach that the improvements in EDA and ETA biosynthesis are due to reversible LPCAT activity in some LPAAT-like proteins, although not all LPAAT-like proteins have LPCAT activity. They do not teach that LPCAT expression would result in the improvements in strains that do not require exogenous feeding of fatty acid substrates or in microbial species other than *Saccharomyces cerevisiae.* They also do not teach that LPCAT expression in engineered microbes results in increased production of high LC-PUFAs other than EDA and ETA, such as ARA, EPA and DHA, or that LPCAT expression can result in improvement in alternate desaturation reactions, other than Δ6 desaturation. Neither reference teaches the effect of the LPCAT or LPAAT-like proteins on either Δ6 elongation without exogenous feeding of fatty acids or on Δ4 desaturation.

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. Pububblications No. WO 2004/087902, No. WO 2006/069936, No. WO 2006/052870, No. WO 2009/001315, No. WO 2009/014140).

Despite the work describe above, to date no one has studied the effect of LPAATs and LPCATs in an oleaginous organism engineered for high-level

production of LC-PUFAs other than EDA and ETA, such as eicosapentaenoic acid ["EPA"; *cis*-5, 8, 11, 14, 17-eicosapentaenoic acid] and/or docosahexaenoic acid ["DHA"; *cis*-4, 7, 10, 13, 16, 19-docosahexaenoic acid] and for improved C_{18} to C_{20} elongation conversion efficiency, and/or improved Δ4 desaturation conversion efficiency without exogenously feeding fatty acids.

SUMMARY OF THE INVENTION

In one embodiment, the invention concerns a recombinant oleaginous microbial host cell for the improved production of at least one long-chain polyunsaturated fatty acid, said host cell comprising 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:

(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:9 and SEQ ID NO:11;

(ii) a polypeptide having at least one membrane bound *O*-acyltransferase protein family motif selected from the group consisting of: SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:28;

(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:2;

(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:15, SEQ ID NO:17 and SEQ ID NO:18; and,

(v) a polypeptide having at least one 1-acyl-*sn*-glycerol-3-phosphate acyltransferase family motif selected from the group consisting of: SEQ ID NO:19 and SEQ ID NO:20;

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

further wherein the host cell has at least one improvement selected from the group consisting of:

a) an increase in C_{18} to C_{20} elongation conversion efficiency in at least one long-chain polyunsaturated fatty acid-producing oleaginous microbial host cell when compared to a control host cell;

b) an increase in Δ4 desaturation conversion efficiency in at least one long-chain polyunsaturated fatty acid-producing oleaginous microbial host cell when compared to a control host cell.

The recombinant oleaginous microbial host cell can be yeast, preferably, *Yarrowia lipolytica.*

In a second embodiment, the invention concerns a recombinant oleaginous microbial host cell for the improved production of at least one long-chain polyunsaturated fatty acid wherein the long-chain polyunsaturated fatty acid can be selected from the group consisting of: eicosadienoic acid, dihomo-γ–linolenic acid, arachidonic acid, eicosatrienoic acid, eicosatetraenoic acid, eicosapentaenoic acid, docosatetraenoic acid, ω-6 docosapentaenoic acid, ω-3 docosapentaenoic acid and docosahexaenoic acid.

In a third embodiment, the invention concerns a recombinant oleaginous microbial host cell for the improved production of at least one long-chain polyunsaturated fatty acid wherein the polynucleotide encoding a polypeptide having at least acyl-CoA:lysophospholipid acyltransferase activity is stably integrated; and, further wherein the host cell has at least one improvement selected from the group consisting of:

a) an increase in C_{18} to C_{20} elongation conversion efficiency of at least 4% in at least one long-chain polyunsaturated fatty acid-producing oleaginous microbial host cell when compared to a control host cell; and,

b) an increase in Δ4 desaturation conversion efficiency of at least 5% in at least one long-chain polyunsaturated fatty acid-producing oleaginous microbial host cell when compared to a control host cell.

In a fourth embodiment, the improvement in production of at least one long-chain polyunsaturated fatty acid can be selected from the group consisting of:

a) an increase in C_{18} to C_{20} elongation conversion efficiency of at least 13% in an eicosapentaenoic acid-producing host cell when compared to a control host cell;

b) an increase in C_{18} to C_{20} elongation conversion efficiency of at least 4% in a docosahexaenoic acid-producing host cell when compared to a control host cell;

c) an increase in Δ4 desaturation conversion efficiency of at least 18% in a docosahexaenoic acid-producing host cell when compared to a control host cell;

d) an increase of at least 9 weight percent of eicosapentaenoic acid in an eicosapentaenoic acid-producing host cell measured as a weight percent of the total fatty acids when compared to a control host cell;

e) an increase of at least 2 weight percent of eicosapentaenoic acid in a docosahexaenoic acid-producing host cell measured as a weight percent of the total fatty acids when compared to a control host cell; and,

f) an increase of at least 9 weight percent of docosahexaenoic acid in a docosahexaenoic acid-producing host cell measured as a weight percent of the total fatty acids when compared to a control host cell.

In a fifth embodiment, the invention concerns oil comprising eicosapentaenoic acid and/or docosahexaenoic acid obtained from the oleaginous microbial recombinant host cell of the invention.

In a sixth embodiment, the invention concerns a method for making an oil comprising eicosapentaenoic acid and/or docosahexaenoic acid comprising:

a) culturing the oleaginous microbial host cell of Claim 3 wherein an oil comprising eicosapentaenoic acid and/or docosahexaenoic acid is produced; and,

b) optionally recovering the microbial oil of step (a).

In a seventh embodiment, the invention concerns a method for increasing C_{18} to C_{20} elongation conversion efficiency in a long-chain polyunsaturated fatty acid-producing oleaginous microbial recombinant host cell, comprising:

a) introducing into said long-chain polyunsaturated fatty acid-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:

(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:9 and SEQ ID NO:11;

(ii) a polypeptide having at least one membrane bound *O*-acyltransferase protein family motif selected from the group consisting of: SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:28;

(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:2;

(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:15, SEQ ID NO:17 and SEQ ID NO:18; and,

(v) a polypeptide having at least one 1-acyl-*sn*-glycerol-3-phosphate acyltransferase protein family motif selected from the group consisting of: SEQ ID NO:19 and SEQ ID NO:20;

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,

b) growing the oleaginous microbial host cell;

wherein the C_{18} to C_{20} elongation conversion efficiency of the oleaginous microbial host cell is increased relative to the control host cell.

In a eighth embodiment, the invention concerns a method of the invention wherein:

a) the polynucleotide encoding a polypeptide having at least acyl-CoA:lysophospholipid acyltransferase activity is stably integrated; and,

b) the increase in C_{18} to C_{20} elongation conversion efficiency is at least 13% in an eicosapentaenoic acid-producing host cell when compared to the control host cell and/or the increase in C_{18} to C_{20} elongation conversion efficiency is at least 4% in a docosahexaenoic acid-producing host cell when compared to the control host cell.

In an ninth embodiment, the invention concerns a method for increasing Δ4 desaturation conversion efficiency in a long-chain polyunsaturated fatty acid-producing oleaginous microbial recombinant host cell, comprising:

a) introducing into said long-chain polyunsaturated fatty acid-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:

(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:9 and SEQ ID NO:11;

(ii) a polypeptide having at least one membrane bound *O*-acyltransferase protein family motif selected from the group consisting of: SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:28;

(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:2;

(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:15, SEQ ID NO:17 and SEQ ID NO:18; and,

(v) a polypeptide having at least one 1-acyl-*sn*-glycerol-3-phosphate acyltransferase protein family motif selected from the group consisting of: SEQ ID NO:19 and SEQ ID NO:20;

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,

b) growing the oleaginous microbial host cell;

wherein the Δ4 desaturation conversion efficiency of the oleaginous microbial host cell is increased relative to the control host cell.

In a tenth embodiment, the invention concerns a method for increasing Δ4 desaturation conversion efficiency in a long-chain polyunsaturated fatty acid-producing oleaginous microbial recombinant host cell wherein:

a) the polynucleotide encoding a polypeptide having at least acyl-CoA:lysophospholipid acyltransferase activity is stably integrated; and,

b) the increase in Δ4 desaturation conversion efficiency is at least 18% when compared to a control host cell.

BIOLOGICAL DEPOSITS

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
Yarrowia lipolytica Y4128	ATCC PTA-8614	August 23, 2007
Yarrowia lipolytica Y8406	ATCC PTA-10025	May 14, 2009
Yarrowia lipolytica Y8412	ATCC PTA-10026	May 14, 2009

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

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.

FIG. 2 diagrams the development of *Yarrowia lipolytica* strain Y8406, producing greater than 51.2 EPA % TFAs.

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 diagrams the development of *Yarrowia lipolytica* strain Y5037, producing 18.6 EPA % TFAs, 22.8 DPA % TFAs and 9.7 DHA % TFAs.

FIG. 8 provides plasmid maps for the following: (A) pZKL4-220EA41B; and, (B) pZKL3-4GER44.

FIG. 9 provides a plasmid map for pZKLY-G20444.

FIG. 10 provides plasmid maps for the following: (A) pY201, comprising a chimeric YAT1::ScAle1S::Lip1 gene; and, (B) pY168, comprising a chimeric YAT1::YlAle1::Lip1 gene.

FIG. 11 provides plasmid maps for the following: (A) pY208, comprising a chimeric YAT1::MaLPAAT1S::Lip1 gene; and, (B) pY207, comprising a chimeric YAT1::YlLPAAT1::Lip1 gene.

FIG. 12 provides plasmid maps for the following: (A) pY175, comprising a chimeric YAT1::CeLPCATS::Lip1 gene; and, (B) pY153, comprising a chimeric FBAIN::CeLPCATS::YlLPAAT1 gene.

FIG. 13 provides plasmid maps for the following: (A) pY222, comprising a chimeric YAT1::ScLPAATS::Lip1 gene; and (B) pY177, comprising a chimeric YAT1::YlLPAAT1::Lip1 gene.

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

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.

SEQ ID NOs:1-101 are ORFs encoding promoters, genes or proteins (or fragments thereof) or plasmids, as identified in Table 2.

Table 2: Summary of Gene and Protein SEQ ID Numbers

Description	Nucleic acid SEQ ID NO.	Protein SEQ ID NO.
Caenorhabditis elegans LPCAT ("CeLPCAT")	1 (849 bp)	2 (282 AA)
membrane bound *O*-acyltransferase motif M(V/I)LxxKL	--	3
membrane bound *O*-acyltransferase motif RxKYYxxW	--	4
membrane bound *O*-acyltransferase motif SAxWHG	--	5
Synthetic LPCAT derived from *Caenorhabditis elegans*, codon-optimized for expression in *Yarrowia lipolytica* ("CeLPCATS")	6 (859 bp)	7 (282 AA)
Saccharomyces cerevisiae Ale1 ("ScAle1"; also ORF "YOR175C")	8 (1860 bp)	9 (619 AA)
Yarrowia lipolytica Ale1 ("YlAle1")	10 (1539 bp)	11 (512 AA)
Synthetic Ale1 derived from *Saccharomyces cerevisiae*, codon-optimized for expression in *Yarrowia lipolytica* ("ScAle1S")	12 (1870 bp)	13 (619 AA)
Mortierella alpina LPAAT1 ("MaLPAAT1")	14 (945 bp)	15 (314 AA)
Yarrowia lipolytica LPAAT1 ("YlLPAAT1")	16 (1549 bp)	17 (282 AA)
Saccharomyces cerevisiae LPAAT ("ScLPAAT"; also ORF "YDL052C")	--	18 (303 AA)
1-acyl-*sn*-glycerol-3-phosphate acyltransferase motif NHxxxxD	--	19
1-acyl-*sn*-glycerol-3-phosphate acyltransferase motif EGTR	--	20
Synthetic LPAAT1 derived from *Mortierella alpina*, codon-optimized for expression in *Yarrowia lipolytica* ("MaLPAAT1S")	21 (955 bp)	22 (314 AA)
Shindou et al. membrane bound *O*-acyltransferase motif WHGxxxGYxxxF	--	23
Shindou et al. membrane bound *O*-acyltransferase motif YxxxxF	--	24
Shindou et al. membrane bound *O*-acyltransferase motif YxxxYFxxH	--	25
U.S. Pat. Pub. No. 2008-0145867-A1 motif	--	26

M-[V/I]-[L/I]-xxK-[L/V/I]-xxxxxxDG		
U.S. Pat. Pub. No. 2008-0145867-A1 motif RxKYYxxWxxx-[E/D]-[A/G]xxxxGxG-[F/Y]-xG	--	27
U.S. Pat. Pub. No. 2008-0145867-A1 motif $EX_{11}WNX_2$-[T/V]-X_2W	--	28
U.S. Pat. Pub. No. 2008-0145867-A1 motif SAxWHGxxPGYxx-[T/F]-F	--	29
Lewin, T.W. et al. & Yamashita et al. 1-acyl-*sn*-glycerol-3-phosphate acyltransferase motif GxxFI-[D/R]-R	--	30
Lewin, T.W. et al. 1-acyl-*sn*-glycerol-3-phosphate acyltransferase motif [V/I]-[P/X]-[I/V/L]-[I/V]-P-[V/I]	--	31
Yamashita et al. 1-acyl-*sn*-glycerol-3-phosphate acyltransferase motif IVPIVM	--	32
Plasmid pY116	33 (8739 bp)	--
Plasmid pZKSL-5S5A5	34 (13,975 bp)	--
Synthetic mutant Δ5 desaturase ("EgD5SM"), derived from *Euglena gracilis* (“EgD5S”) (U.S. Pat. Pub. No. 2010-0075386-A1)	35 (1350 bp)	36 (449 AA)
Synthetic mutant Δ5 desaturase ("EaD5SM"), derived from *Euglena anabaena* (“EaD5S”) (U.S. Pat. Pub. No. 2010-0075386-A1)	37 (1365 bp)	38 (454 AA)
Plasmid pZP3-Pa777U	39 (13,066 bp)	--
Plasmjd pZKUM	40 (4313 bp)	--
Plasmid pZKL2-5mB89C	41 (15,991 bp)	--
Yarrowia lipolytica diacylglycerol cholinephosphotransferase gene (“YlCPT1”)	42 (1185 bp)	43 (394 AA)
Synthetic mutant Δ8 desaturase ("EgD8M") (U.S. Patent 7,709,239), derived from *Euglena gracilis* (“EgD8S”) (U.S. Patent 7,256,033)	44 (1272 bp)	45 (422 AA)
Synthetic Δ9 elongase derived from *Euglena gracilis*, codon-optimized for expression in *Yarrowia lipolytica* (“EgD9eS”)	46 (777 bp)	47 (258 AA)
Plasmid pZKL1-2SR9G85	48 (14,554 bp)	--
DGLA synthase, comprising E389D9eS/EgD8M gene fusion	49 (2127 bp)	50 (708 AA)
Synthetic Δ12 desaturase derived from *Fusarium moniliforme,* codon-optimized for expression in *Yarrowia lipolytica* (“FmD12S”)	51 (1434 bp)	52 (477 AA)
Plasmid pZSCP-Ma83	53 (15,119 bp)	--
Synthetic $C_{16/18}$ elongase derived from *Mortierella alpina* ELO3, codon-optimized for expression in	54 (828 bp)	55 (275 AA)

Yarrowia lipolytica ("ME3S")		
Synthetic malonyl-CoA synthetase derived from *Rhizobium leguminosarum* bv. viciae 3841 (GenBank Accession No. YP_766603), codon-optimized for expression in *Yarrowia lipolytica* ("MCS")	56 (1518 bp)	57 (505 AA)
Synthetic Δ8 desaturase derived from *Euglena anabaena* UTEX 373, codon-optimized for expression in *Yarrowia lipolytica* ("EaD8S")	58 (1260 bp)	59 (420 AA)
Plasmid pZKL4-220EA41B	60 (16,424 bp)	--
Synthetic C20 elongase derived from *Euglena anabaena*, codon-optimized for expression in *Yarrowia lipolytica* ("EaC20ES")	61 (900 bp)	62 (299 AA)
Synthetic C20 elongase derived from *Euglena gracilis*, codon-optimized for expression in *Yarrowia lipolytica* ("EgC20ES")	63 (912 bp)	64 (303 AA)
Truncated synthetic Δ4 desaturase derived from *Euglena anabaena*, codon-optimized for expression in *Yarrowia lipolytica* ("EaD4S-1")	65 (1644 bp)	66 (547 AA)
Truncated synthetic Δ4 desaturase version B derived from *Euglena anabaena*, codon-optimized for expression in *Yarrowia lipolytica* ("EaD4SB")	67 (1644 bp)	68 (547 AA)
Plasmid pZKL3-4GER44	69 (17,088 bp)	--
Synthetic Δ4 desaturase derived from *Eutreptiella cf_gymnastica* CCMP1594, codon-optimized for expression in *Yarrowia lipolytica* ("E1594D4S")	70 (1548 bp)	71 (515 AA)
Truncated synthetic Δ4 desaturase derived from *Euglena gracilis*, codon-optimized for expression in *Yarrowia lipolytica* ("EgD4S-1")	72 (1542 bp)	73 (513 AA)
Plasmid pZKLY-G20444	74 (15,617 bp)	--
Synthetic DHA synthase derived from *Euglena gracilis*, codon-optimized for expression in *Yarrowia lipolytica* ("EgDHAsyn1S")	75 (2382 bp)	76 (793 AA)
Plasmid pY201	77 (9641 bp)	--
Escherichia coli LoxP recombination site, recognized by a Cre recombinase enzyme	78 (34 bp)	--
Primer 798	79	--
Primer 799	80	--
Primer 800	81	--
Primer 801	82	--
Plasmid pY168	83 (9320 bp)	--
Plasmid pY208	84 (8726 bp)	--
Primer 856	85	--
Primer 857	86	--
Plasmid pY207	87 (8630 bp)	--
Plasmid pY175	88 (8630 bp)	--

Plasmid pY153	89 (8237 bp)	--
Mutant Δ5 desaturase ("EgD5M"), derived from *Euglena gracilis* ("EgD5") (U.S. Pat. Pub. No. 2010-0075386-A1)	90 (1350 bp)	91 (449 AA)
Mortierella alpina LPAAT (corresponding to SEQ ID NOs:16 and 17 within Intl. App. Pub. No. WO 2004/087902)	92 (1254 bp)	93 (417 AA)
Mortierella alpina LPAAT (corresponding to SEQ ID NOs:18 and 19 within Intl. App. Pub. No. WO 2004/087902)	94 (1170 bp)	95 (389 AA)
Synthetic LPAAT derived from *Saccharomyces cerevisiae*, codon-optimized for expression in *Yarrowia lipolytica* ("ScLPAATS")	96 (926 bp)	97 (303 AA)
Primer 869	98	--
Primer 870	99	--
Plasmid pY222	100 (7891 bp)	--
Plasmid pY177	101 (9598 bp)	--

DETAILED DESCRIPTION OF THE INVENTION

Described herein are methods for increasing C_{18} to C_{20} elongation conversion efficiency and/or Δ4 desaturation conversion efficiency in long-chain polyunsaturated fatty acid ["LC-PUFA"]-producing recombinant oleaginous microbial host cells, based on expression of polypeptides (e.g., Ale1, LPAAT, and LPCAT) having LPLAT activity. By increasing the conversion efficiency of C_{18} to C_{20} elongation and/or Δ4 desaturation, the concentration of the LC-PUFAs eicosapentaenoic acid ["EPA"; *cis*-5, 8, 11, 14, 17-eicosapentaenoic acid] and/or docosahexaenoic acid ["DHA"; *cis*-4, 7, 10, 13, 16, 19-docosahexaenoic acid] increased as a weight percent of the total fatty acids. Recombinant host cells are also claimed.

PUFAs, such as EPA and DHA (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 and drink products and may find use as cardiovascular-protective, anti-depression, anti-inflammatory or cholesterol lowering agents. Optionally, the compositions may be used for pharmaceutical use, either human or veterinary.

All patents, patent applications, and publications cited herein are incorporated by reference in their entirety.

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)".

"Diacylglycerol acyltransferase" is abbreviated as "DAG AT" or "DGAT".

"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".

"Long-chain polyunsaturated fatty acid(s)" is abbreviated as "LC-PUFA(s)".

"Acyl-CoA:lysophospholipid acyltransferase(s)" or "lysophospholipid acyltransferase(s)" is abbreviated as "LPLAT(s)".

The term "invention" or "present invention" as used herein is not meant to be limiting to any one specific embodiment of the invention but applies generally to any and all embodiments of the invention as described in the claims and specification.

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 phosphatidylcholines ["PC"], phosphatidylethanolamines ["PE"], phosphatidylglycerols ["PG"], phosphatidylinositols ["PI"], phosphatidylserines ["PS"] and cardiolipins ["CL"].

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

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

The term "acyl-CoA:lysophospholipid acyltransferase" or "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:9), 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)).

More specifically, the term "polypeptide having at least lysophosphtidylcholine acyltransferase ["LPCAT"] activity" will refer to those enzymes capable of catalyzing the reaction: acyl-CoA + 1-acyl-*sn*-glycero-3-phosphocholine = CoA + 1,2-diacyl-*sn*-glycero-3-phosphocholine (EC 2.3.1.23). LPCAT activity has been described in two structurally distinct protein families, i.e., the LPAAT protein family (Hishikawa, 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*).

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:9 (ScAle1) and SEQ ID NO:11 (YlAle1); 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)LxxKL (SEQ ID NO:3), RxKYYxxW (SEQ ID NO:4), SAxWHG (SEQ ID NO:5) and $EX_{11}WNX_2$-[T/V]-X_2W (SEQ ID NO:28). Examples of ALE1 polypeptides include ScAle1 and YlAle1.

The term "ScAle1" refers to a LPCAT (SEQ ID NO:9) isolated from *Saccharomyces cerevisiae* (ORF "YOR175C"), encoded by the nucleotide sequence set forth as SEQ ID NO:8. 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:12 and 13).

The term "YlAle1" refers to a LPCAT (SEQ ID NO:11) isolated from *Yarrowia lipolytica,* encoded by the nucleotide sequence set forth as SEQ ID NO:10.

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:2 (CeLPCAT).

The term "CeLPCAT" refers to a LPCAT enzyme (SEQ ID NO:2) isolated from *Caenorhabditis elegans,* encoded by the nucleotide sequence set forth as SEQ ID NO:1. 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:6 and 7).

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 = CoA + 1,2-diacyl-*sn*-glycerol 3-phosphate (EC 2.3.1.51).

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:15 (MaLPAAT1), SEQ ID NO:17 (YlLPAAT1) and SEQ ID NO:18 (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: NHxxxxD (SEQ ID NO:19) and EGTR (SEQ ID NO:20). Examples of LPAAT polypeptides include ScLPAAT, MaLPAAT1 and YlLPAAT1.

The term "ScLPAAT" refers to a LPAAT (SEQ ID NO:18) isolated from *Saccharomyces cerevisiae* (ORF "YDL052C").

The term "MaLPAAT1" refers to a LPAAT (SEQ ID NO:15) isolated from *Mortierella alpina,* encoded by the nucleotide sequence set forth as SEQ ID NO:14. In contrast, the term "MaLPAAT1S" refers to a synthetic LPAAT

derived from *M. alpina* that is codon-optimized for expression in *Yarrowia lipolytica* (i.e., SEQ ID NOs:21 and 22).

The term "YlLPAAT1" refers to a LPAAT (SEQ ID NO:17) isolated from *Yarrowia lipolytica,* encoded by the nucleotide sequence set forth as SEQ ID NO:16.

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.

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 likely 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.

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.

"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.

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

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 the PC and the PE fractions), but not free fatty acids.

The term “total lipid content” of cells is a measure of TFAs as a percent of the dry cell weight [“DCW”], athough 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.

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).

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.

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.

The term “fatty acids” refers to long chain aliphatic acids (alkanoic acids) of varying chain lengths, from about C_{12} to C_{22}, although both longer and shorter chain-length acids are known. The predominant chain lengths are between C_{16} and C_{22}. 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, which is hereby incorporated herein by reference.

Nomenclature used to describe PUFAs herein is given in Table 3. 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 3: 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	ETrA	*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
Eicosa-tetraenoic	ETA	*cis*-8, 11, 14, 17-eicosatetraenoic	20:4 ω-3
Eicosa-pentaenoic	EPA	*cis*-5, 8, 11, 14, 17-eicosapentaenoic	20:5 ω-3
Docosa-tetraenoic	DTA	*cis*-7, 10, 13, 16-docosatetraenoic	22:4 ω-6
Docosa-pentaenoic	DPAn-6	*cis*-4, 7, 10, 13, 16-docosapentaenoic	22:5 ω-6
Docosa-pentaenoic	DPA	*cis*-7, 10, 13, 16, 19-docosapentaenoic	22:5 ω-3
Docosa-hexaenoic	DHA	*cis*-4, 7, 10, 13, 16, 19-docosahexaenoic	22:6 ω-3

The term “long-chain polyunsaturated fatty acid” [“LC-PUFA”] refers to those PUFAs that have chain lengths of C_{20} or greater. Thus, the term LC-PUFA includes at least EDA, DGLA, ARA, ETrA, ETA, EPA, DTA, DPAn-6, DPA and DHA.

A metabolic pathway, or biosynthetic pathway, in a biochemical sense, can be regarded as a series of chemical reactions occurring in order within a cell, catalyzed by enzymes, to achieve either the formation of a metabolic product to be used or stored by the cell, or the initiation of another metabolic pathway (then called a flux generating step). Many of these pathways are elaborate, and involve a step by step modification of the initial substance to shape it into a product having the exact chemical structure desired.

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, DRA, 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 Intl. App. Pub. No. WO 2006/052870). 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, $C_{14/16}$ elongase, $C_{16/18}$ elongase, $C_{18/20}$ elongase and/or $C_{20/22}$ elongase.

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: Δ8 desaturases; Δ5 desaturases; Δ17 desaturases; Δ12 desaturases; Δ15 desaturases; Δ9 desaturases; Δ6 desaturases; and Δ4 desaturases. Δ17 desaturases, and also Δ15 desaturases, are also occasionally referred to as “omega-3 desaturases”, “w-3 desaturases”, and/or “ω-3 desaturases”, based on their ability to convert ω-6 fatty acids into their ω-3 counterparts.

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., LA, ALA, GLA, STA) and a $C_{20/22}$ elongase (also known as a C20 elongase or Δ5 elongase as the terms can be used interchangeably) 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 Δ6 elongase will catalyze conversion of GLA and STA to DGLA and ETA, respectively, while a Δ9 elongase is able to catalyze the conversion of LA and ALA to EDA and ETrA, respectively.

The terms “conversion efficiency” and “percent substrate conversion” refer to the efficiency by which a particular enzyme, such as a desaturase or elongase, can convert substrate to product. The conversion efficiency is measured according to the following formula: ([product]/[substrate+product])*100, where ‘product’ includes the immediate product and all products in the pathway derived from it.

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 Δ9 elongases or Δ6 elongases.

The terms “Δ9 elongation conversion efficiency” and “Δ9 elongase conversion efficiency” refer to the efficiency by which Δ9 elongase can convert C_{18} substrates (i.e., LA, ALA) to C_{20} products (i.e., EDA, ETrA).

The terms “Δ4 desaturation conversion efficiency” and “Δ4 desaturase conversion efficiency” refer to the efficiency by which Δ4 desaturase can convert substrates (i.e., DTA, DPAn-3) to products (i.e., DPAn-6, DHA).

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). 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. Oleaginous microorganisms include various bacteria, algae, euglenoids, moss, fungi (e.g., *Mortierella*), yeast and stramenopiles (e.g., *Schizochytrium*).

The term “oleaginous yeast” refers to those microorganisms classified as yeasts that can make oil. Examples of oleaginous yeast include, but are no means limited to, the following genera: *Yarrowia, Candida, Rhodotorula, Rhodosporidium*, *Cryptococcus, Trichosporon* and *Lipomyces*.

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, glycerol, monoglycerides, diglycerides, triglycerides, carbon dioxide, methanol, formaldehyde, formate and carbon-containing amines.

As used herein the term “biomass” refers specifically to spent or used cellular material from the fermentation of a recombinant production host producing PUFAs in commercially significant amounts, wherein the preferred production host is a recombinant strain of an oleaginous yeast of the genus *Yarrowia*. 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).

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.

As used herein, a nucleic acid fragment is "hybridizable" to another nucleic acid fragment, such as a cDNA, genomic DNA, or RNA molecule, when a single-stranded form of the nucleic acid fragment can anneal to the other nucleic acid fragment under the appropriate conditions of temperature and solution ionic strength. Hybridization and washing conditions are well known and exemplified in Sambrook, J., Fritsch, E. F. and Maniatis, T. Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989), which is hereby incorporated herein by reference, particularly Chapter 11 and Table 11.1.

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.

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.

As used herein, the terms "homology" and "homologous" are used interchangeably. They refer to nucleic acid fragments wherein changes in one or more nucleotide bases do not affect the ability of the nucleic acid fragment to mediate gene expression or produce a certain phenotype. These terms also refer to modifications of the nucleic acid fragments such as deletion or insertion of one or more nucleotides that do not substantially alter the functional properties of the resulting nucleic acid fragment relative to the initial, unmodified fragment.

Moreover, the skilled artisan recognizes that homologous nucleic acid sequences are also defined by their ability to hybridize, under moderately stringent conditions, e.g., 0.5X SSC, 0.1% SDS, 60 °C, with the sequences exemplified herein, or to any portion of the nucleotide sequences disclosed herein and which are functionally equivalent thereto. Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly related organisms, to highly similar fragments, such as genes that duplicate functional enzymes from closely related organisms. Post-hybridization washes determine stringency conditions. An extensive guide to the hybridization of nucleic acids is found in

Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes,* Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993); and *Current Protocols in Molecular Biology*, Chapter 2, Ausubel et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995).

As used herein, the term “percent identity” refers to a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. “Identity” also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the percentage of match between compared sequences. “Percent identity” and “percent similarity” can be readily calculated by known methods, including but not limited to those described in: 1) Computational Molecular Biology (Lesk, A. M., Ed.) Oxford University: NY (1988); 2) Biocomputing: Informatics and Genome Projects (Smith, D. W., Ed.) Academic: NY (1993); 3) Computer Analysis of Sequence Data, Part I (Griffin, A. M., and Griffin, H. G., Eds.) Humania: NJ (1994); 4) Sequence Analysis in Molecular Biology (von Heinje, G., Ed.) Academic (1987); and, 5) Sequence Analysis Primer (Gribskov, M. and Devereux, J., Eds.) Stockton: NY (1991).

Preferred methods to determine percent identity are designed to give the best match between the sequences tested. Methods to determine percent identity and percent similarity are codified in publicly available computer programs. Sequence alignments and percent identity calculations may be performed using the MegAlign™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences is performed using the “Clustal method of alignment” which encompasses several varieties of the algorithm including the “Clustal V method of alignment” and the “Clustal W method of alignment” (described by Higgins and Sharp, *CABIOS*, 5:151-153 (1989); Higgins, D.G.

et al., *Comput. Appl. Biosci.,* 8:189-191(1992)) and found in the MegAlign™ (version 8.0.2) program of the LASERGENE bioinformatics computing suite (DNASTAR Inc.). 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. After alignment of the sequences using either Clustal program, it is possible to obtain a “percent identity” by viewing the “sequence distances” table in the program.

It is well understood by one skilled in the art that many levels of sequence identity are useful in identifying polypeptides, from other species, wherein such polypeptides have the same or similar function or activity. Useful examples of percent identities include any integer percentage from 34% to 100%, such as 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%. Also, of interest is any full-length or partial complement of this isolated nucleotide fragment. Suitable nucleic acid fragments not only have the above homologies but typically encode a polypeptide having at least 50 amino acids, preferably at least 100 amino acids, more preferably at least 150 amino acids, still more preferably at least 200 amino acids, and most preferably at least 250 amino acids.

“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.

"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.

"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.

"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.

"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.

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.

"RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from post-transcriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA" or "mRNA" refers to the RNA that is without introns and which can be translated into protein by the cell. "cDNA" refers to a double-stranded DNA that is complementary to, and derived from, mRNA.

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.

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.

"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" or "transformant" organisms.

"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.

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.

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 ["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.

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.

As previously described, genes encoding LPLATs are found in all eukaryotic cells, based on their intimate role in *de novo* synthesis and remodeling of glycerophospholipids, wherein LPLATs remove acyl-CoA fatty acids from the cellular acyl-CoA pool and acylate various lysophospholipid substrates at the *sn*-2 position in the phospholipid pool. Publicly available sequences encoding LPLATs include ScAle1 (SEQ ID NO:9), ScLPAAT (SEQ ID NO:18), MaLPAAT1 (SEQ ID NO:15) and CeLPCAT (SEQ ID NO:2). The ScAle1 (SEQ ID NO:9) and ScLPAAT (SEQ ID NO:18) protein sequences were used as a query to identify orthologs from the public *Y. lipolytica* protein database (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 YlAle1 (SEQ ID NO:11) and YlLPAAT1 (SEQ ID NO:17), respectively (see Example 5, *infra*).

When the sequence of a particular LPLAT gene or protein within a preferred host organism is not known, the LPLAT sequences set forth herein as SEQ ID NOs:2, 9, 11, 15, 17 and 18, or portions of them, may be used to search for LPLAT homologs in the same or other algal, fungal, oomycete, euglenoid, stramenopiles, yeast or plant species using sequence analysis software. In general, such computer software matches similar sequences by assigning degrees of homology to various substitutions, deletions, and other

modifications. Use of software algorithms, such as the BLASTP method of alignment with a low complexity filter and the following parameters: Expect value = 10, matrix = Blosum 62 (Altschul, et al., *Nucleic Acids Res.,* 25:3389-3402 (1997)), is well-known for comparing any LPLAT protein against a database of nucleic or protein sequences and thereby identifying similar known sequences within a preferred host organism.

Use of a software algorithm to comb through databases of known sequences is particularly suitable for the isolation of homologs having a relatively low percent identity to publicly available LPLAT sequences, such as those described in SEQ ID NOs:2, 9, 11, 15, 17 and 18. It is predictable that isolation would be relatively easier for LPLAT homologs of at least about 70%-85% identity to publicly available LPLAT sequences. Further, those sequences that are at least about 85%-90% identical would be particularly suitable for isolation and those sequences that are at least about 90%-95% identical would be the most facilely isolated.

LPLAT homologs can also be identified by the use of motifs unique to the LPLAT enzymes. These motifs likely represent regions of the LPLAT protein that are important to the structure, stability or activity of the protein and these motifs are useful as diagnostic tools for the rapid identification of novel LPLAT genes.

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, WHGxxxGYxxxF (SEQ ID NO:23), YxxxxF (SEQ ID NO:24) and YxxxYFxxH (SEQ ID NO:25). Of these, the WD, WHGxxxGYxxxF and YxxxxF motifs are present in ScAle and YlAle1, but the YxxxYFxxH motif is not. Alternate non-plant motifs for Ale1 homologs are also described in U.S. Pat. Pub. No. 2008-0145867-A1; specifically, these

include: M-[V/I]-[L/I]-xxK-[L/V/I]-xxxxxxDG (SEQ ID NO:26), RxKYYxxWxxx-[E/D]-[A/G]xxxxGxG-[F/Y]-xG (SEQ ID NO:27), $EX_{11}WNX_2$-[T/V]-X_2W (SEQ ID NO:28) and SAxWHGxxPGYxx-[T/F]-F (SEQ ID NO:29).

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: NHxxxxD (SEQ ID NO:19), GxxFI-[D/R]-R (SEQ ID NO:30), EGTR (SEQ ID NO:20) and either [V/I]-[P/X]-[I/V/L]-[I/V]-P-[V/I] (SEQ ID NO:31) or IVPIVM (SEQ ID NO:32). The NHxxxxD and EGTR motifs are present in MaLPAAT1, YlLPAAT1 and CeLPCAT, but the other motifs are not.

Based on publicly available Ale1, LPCAT and LPAAT protein sequences, including those described herein, the instant invention concerns the following MBOAT family motifs: M(V/I)LxxKL (SEQ ID NO:3), RxKYYxxW (SEQ ID NO:4), SAxWHG (SEQ ID NO:5) and $EX_{11}WNX_2$-[T/V]-X_2W (SEQ ID NO:28). Similarly, 1-acyl-*sn*-glycerol-3-phosphate acyltransferase family motifs are those set forth as: NHxxxxD (SEQ ID NO:19) and EGTR (SEQ ID NO:20).

Alternatively, publicly available LPLAT sequences or their motifs may be hybridization reagents for the identification of homologs. Hybridization methods are well known to those of ordinary skill in the art as noted above.

Any of the LPLAT nucleic acid fragments or any identified homologs may be used to isolate genes encoding homologous proteins from the same or other algal, fungal, oomycete, euglenoid, stramenopiles, yeast or plant species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to: 1) methods of nucleic acid hybridization; 2) methods of DNA and RNA amplification, as exemplified by various uses of nucleic acid amplification technologies, such as polymerase chain reaction ["PCR"] (U.S. Patent 4,683,202); ligase chain reaction ["LCR"]

(Tabor, S. et al., *Proc. Natl. Acad. Sci. U.S.A.,* 82:1074 (1985)); or strand displacement amplification ["SDA"] (Walker, et al., *Proc. Natl. Acad. Sci. U.S.A.*, 89:392 (1992)); and, 3) methods of library construction and screening by complementation.

For example, genes encoding proteins or polypeptides similar to publicly available LPLAT genes or their motifs could be isolated directly by using all or a portion of those publicly available nucleic acid fragments as DNA hybridization probes to screen libraries from any desired organism using well known methods. Specific oligonucleotide probes based upon the publicly available nucleic acid sequences can be designed and synthesized by methods known in the art (Maniatis, *supra*). Moreover, the entire sequences can be used directly to synthesize DNA probes by methods known to the skilled artisan, such as random primers DNA labeling, nick translation or end-labeling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers can be designed and used to amplify a part or the full length of the publicly available sequences or their motifs. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as probes to isolate full-length DNA fragments under conditions of appropriate stringency.

Typically, in PCR-type amplification techniques, the primers have different sequences and are not complementary to each other. Depending on the desired test conditions, the sequences of the primers should be designed to provide for both efficient and faithful replication of the target nucleic acid. Methods of PCR primer design are common and well known (Thein and Wallace, "The use of oligonucleotides as specific hybridization probes in the Diagnosis of Genetic Disorders", in *Human Genetic Diseases: A Practical Approach*, K. E. Davis Ed., (1986) pp 33-50, IRL: Herndon, VA; Rychlik, W., In Methods in Molecular Biology, White, B. A. Ed., (1993) Vol. 15, pp 31-39, PCR Protocols: Current Methods and Applications. Humania: Totowa, NJ).

Generally two short segments of available LPLAT sequences may be used in PCR protocols to amplify longer nucleic acid fragments encoding

homologous genes from DNA or RNA. PCR may also be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the available nucleic acid fragments or their motifs. The sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding genes.

Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, the skilled artisan can follow the RACE protocol (Frohman et al., *Proc. Natl. Acad. Sci. U.S.A.,* 85:8998 (1988)) to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Primers oriented in the 3' and 5' directions can be designed from the available sequences. Using commercially available 3' RACE or 5' RACE systems (e.g., BRL, Gaithersburg, MD), specific 3' or 5' cDNA fragments can be isolated (Ohara et al., *Proc. Natl. Acad. Sci. U.S.A.,* 86:5673 (1989); Loh et al., *Science,* 243:217 (1989)).

Based on any of these well-known methods just discussed, it would be possible to identify and/or isolate LPLAT gene homologs in any preferred eukaryotic organism of choice. The activity of any putative LPLAT gene can readily be confirmed by expression of the gene within a LC-PUFA-producing host organism, since the C_{18} to C_{20} elongation and/or Δ4 desaturation are increased relative to those within an organism lacking the LPLAT transgene (*supra*).

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. 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 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 criticial 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) Intl. App. Pub. No. WO 2004/076617 A2 (Renz, A. et al.), who provided a gene encoding LPCAT from *Caenorhabditis elegans* (T06E8.1) ["CeLPCAT"] that substantially improved the efficiency of elongation in a genetically introduced Δ6 desaturase/Δ6 elongase pathway in *S. cerevisiae* fed with exogenous fatty acid substrates suitable for Δ6 elongation. Renz 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). However, Intl. App. Pub. No. WO 2004/076617 did not teach the effect of CeLPCAT on Δ6 elongation conversion efficiency in host cells that were not exogenously fed fatty acids, Δ5 elongation conversion efficiency, or Δ4 desaturation conversion efficiency.

Herein, it is demonstrated that LPAAT and LPCAT are indeed important in the accumulation of EPA and DHA in the TAG fraction of *Yarrowia lipolytica.* However, unexpectedly, it was found that over-expression of LPLATs can result in an improvement in the Δ9 elongase conversion efficiency and/or Δ4 desaturase conversion efficiency. As previously defined, conversion efficiency is a term that refers to the efficiency by which a particular enzyme, such as a Δ4 desaturase or Δ9 elongase, can convert substrate to product. Thus, in a strain engineered to produce EPA, improvement in Δ9 elongase conversion efficiency was demonstrated to result in increased EPA % TFAs or EPA % DCW. Similarly, improvement in Δ9 elongase and/or Δ4 desaturase conversion efficiency in a strain

engineered to produce DHA was demonstrated to result in increased DHA % TFAs or DHA % DCW.

PUFA desaturations occur on phospholipids, while fatty acid elongations occur on acyl-CoAs. Based on previous studies, it was therefore expected that LPLAT over-expression would result in improved desaturations due to improved substrate availability in phospholipids, while expression of LPLATs was not expected to result in improved elongations that require improved substrate availability in the CoA pool.

Despite these assumptions, Example 5 demonstrates that LPLAT expression did not improve the conversion efficiency of all desaturations in strains of *Yarrowia* producing DHA, in a comparable manner. Specifically, the conversion efficiency of Δ4 desaturase was selectively improved, while similar improvements were not found in Δ12, Δ8, Δ5 or Δ17 desaturations. It is hypothesized that Δ4 desaturase was therefore limiting as a result of limited availability of the DPA substrate in phospholipids.

Additionally, Examples 4 and 5 demonstrate that LPLAT expression, based on at least one stably integrated polynucleotide encoding the LPLAT polypeptide, significantly improved the Δ9 elongase conversion efficiency in strains of *Yarrowia* producing EPA and DHA, respectively. Surprisingly, however, the LPLATs did not also result in a comparable improvement in the efficiency of the $C_{20/22}$ elongation of EPA to DPA in DHA strains. Generally, there was no significant change in the total lipid content in strains over-expressing LPLATs versus those that were not.

Clearly, broad generalizations are difficult concerning the effect of LPLAT over-expression in host cells producing PUFAs. Instead, the effect of LPLAT activity must be considered based on subsets of desaturases and elongases having specific activity (i.e., Δ12 desaturase, Δ8 desaturase, Δ5 desaturase, Δ17 desaturase, Δ4 desaturase, Δ9 elongase, $C_{14/16}$ elongase, $C_{16/18}$ elongase, $C_{18/20}$ elongase ["also Δ6 elongase"], $C_{20/22}$ elongase ["also Δ5 elongase"]).

On the basis of the above discussion, in one embodiment herein, methods for improving C_{18} to C_{20} elongation conversion efficiency in a LC-PUFA-producing recombinant oleaginous microbial host cell are provided, wherein said method comprises:

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:

(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:9 (ScAle1) and SEQ ID NO:11 (YlAle1);

(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:3), RxKYYxxW (SEQ ID NO:4), SAxWHG (SEQ ID NO:5) and $EX_{11}WNX_2$-[T/V]-X_2W (SEQ ID NO:28);

(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:2 (CeLPCAT);

(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:15 (MaLPAAT1), SEQ ID NO:17 (YlLPAAT1) and SEQ ID NO:18 (ScLPAAT1); and,

(v) a polypeptide having at least one 1-acyl-*sn*-glycerol-3-phosphate acyltransferase protein family motif selected from the group consisting of: NHxxxxD (SEQ ID NO:19) and EGTR (SEQ ID NO:20);

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,

b) growing the oleaginous microbial host cell;

wherein the C_{18} to C_{20} elongation conversion efficiency of the oleaginous microbial host cell is increased relative to the control host cell.

In preferred embodiments, the increase in C_{18} to C_{20} elongation conversion efficiency is at least 4% in at least one LC-PUFA-producing oleaginous microbial host cell, based on at least one stably integrated polynucleotide encoding the LPLAT polypeptide,,when compared to the control host cell, although any increase in C_{18} to C_{20} elongation conversion efficiency greater than 4% is especially preferred, including increases of 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.

For example, in one method demonstrated herein, the increase in C_{18} to C_{20} elongation conversion efficiency is at least 13% in an EPA-producing host cell when compared to the control host cell and the increase in C_{18} to C_{20} elongation conversion efficiency is at least 4% in a DHA-producing host cell when compared to the control host cell.

Similarly, methods are also described herein for increasing Δ4 desaturation conversion efficiency in a LC-PUFA-producing oleaginous microbial recombinant host cell, wherein said method comprises:

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:

(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:9 (ScAle1) and SEQ ID NO:11 (YlAle1);

(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:3), RxKYYxxW (SEQ ID NO:4), SAxWHG (SEQ ID NO:5) and $EX_{11}WNX_2$-[T/V]-X_2W (SEQ ID NO:28);

(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:2 (CeLPCAT);

(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:15 (MaLPAAT1), SEQ ID NO:17 (YlLPAAT1) and SEQ ID NO:18 (ScLPAAT1); and,

(v) a polypeptide having at least one 1-acyl-*sn*-glycerol-3-phosphate acyltransferase protein family motif selected from the group consisting of: NHxxxxD (SEQ ID NO:19) and EGTR (SEQ ID NO:20);

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,

b) growing the oleaginous microbial host cell;

wherein the Δ4 desaturation conversion efficiency of the oleaginous microbial host cell is increased relative to the control host cell.

In preferred embodiments, the increase in Δ4 desaturation conversion efficiency is at least 5% in at least one LC-PUFA-producing oleaginous microbial host cell, based on at least one stably integrated polynucleotide encoding the LPLAT polypeptide, when compared to the control host cell, although any increase in Δ4 desaturation conversion efficiency greater than 5% is especially preferred, including increases of at least about 5-10%, more preferred at least about 10-20%, more preferred at least about 20-40%, and most preferred at least about 40-60% or greater.

For example, in one method demonstrated herein, the increase in Δ4 desaturation conversion efficiency in a DHA-producing host was at least 18% when compared to the control host cell.

Recombinant host cells are also described herein, in addition to the methods set forth above. Specifically, these recombinant host cells comprise 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:

(a) 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:9 (ScAle1) and SEQ ID NO:11 (YlAle1);

(b) 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:3), RxKYYxxW (SEQ ID NO:4), SAxWHG (SEQ ID NO:5) and $EX_{11}WNX_2$-[T/V]-X_2W (SEQ ID NO:28);

(c) 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:2 (CeLPCAT);

(d) 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:15 (MaLPAAT1), SEQ ID NO:17 (YlLPAAT1) and SEQ ID NO:18 (ScLPAAT1); and,

(e) a polypeptide having at least one 1-acyl-*sn*-glycerol-3-phosphate acyltransferase family motif selected from the group consisting of: NHxxxxD (SEQ ID NO:19) and EGTR (SEQ ID NO:20);

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 the recombinant host cells further have at least one improvement selected from the group consisting of:

a) an increase in C_{18} to C_{20} elongation conversion efficiency in at least one LC PUFA-producing oleaginous microbial host cell when compared to the control host cell;

b) an increase in Δ4 desaturation conversion efficiency in at least one LC PUFA-producing oleaginous microbial host cell when compared to the control host cell.

In preferred host cells, the polynucleotide encoding the polypeptide having at least acyl-CoA:lysophospholipid acyltransferase activity is stably integrated; and, further wherein the host cell has at least one improvement selected from the group consisting of:

a) an increase in C_{18} to C_{20} elongation conversion efficiency of at least 4% in at least one long-chain polyunsaturated fatty acid-producing oleaginous microbial host cell when compared to a control host cell; and,

b) an increase in Δ4 desaturation conversion efficiency of at least 5% in at least one long-chain polyunsaturated fatty acid-producing oleaginous microbial host cell when compared to a control host cell.

In more preferred host cells, having at least one stably integrated polynucleotide encoding the LPLAT polypeptide, the at least one improvement is selected from the group consisting of:

a) an increase in C_{18} to C_{20} elongation conversion efficiency of at least 13% in an EPA-producing host cell when compared to the control host cell;

b) an increase of at least 9% EPA of TFAs in an EPA-producing host cell when compared to the control host cell;

c) an increase in C_{18} to C_{20} elongation conversion efficiency of at least of at least 4% in a DHA-producing host cell when compared to the control host cell;

d) an increase of at least 2% EPA of TFAs in a DHA-producing host cell when compared to the control host cell;

e) an increase in Δ4 desaturation conversion efficiency of at least 18% in a DHA-producing host cell when compared to the control host cell; and,

f) an increase of at least 9% DHA of TFAs in a DHA-producing host cell when compared to the control host cell.

Of course, one of skill in the art should understand that the improvements described above should be considered as exemplary, but not limiting to the invention herein.

Based on the above improvements, one of skill in the art will appreciate the value of expressing a LPLAT in a recombinant host cell that is producing long-chain PUFAs, such EDA, DGLA, ARA, DTA, DPAn-6, ETrA, ETA, EPA, DPA and DHA, if it is desirable to optimize the production of these fatty acids.

Standard resource materials that are useful to make recombinant constructs describe, *inter alia*: 1) specific conditions and procedures for construction, manipulation and isolation of macromolecules, such as DNA molecules, plasmids, etc.; 2) generation of recombinant DNA fragments and recombinant expression constructs; and, 3) screening and isolation of clones. See, 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., Bennan, 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).

In general, the choice of sequences included in the construct depends on the desired expression products, the nature of the host cell and the proposed means of separating transformed cells versus non-transformed cells. The skilled artisan is aware of the genetic elements that must be present on the plasmid vector to successfully transform, select and propagate host cells containing the chimeric gene. Typically, however, the vector or cassette contains sequences directing transcription and translation of the

relevant gene(s), a selectable marker and sequences allowing autonomous replication or chromosomal integration. Suitable vectors comprise a region 5' of the gene that controls transcriptional initiation, i.e., 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.

Transcription initiation regions or promoters useful for driving expression of heterologous genes or portions of them in the desired host cell are numerous and well known. These control regions may comprise a promoter, enhancer, silencer, intron sequences, 3’ UTR and/or 5' UTR regions, and protein and/or RNA stabilizing elements. Such elements may vary in their strength and specificity. Virtually any promoter, i.e., native, synthetic, or chimeric, capable of directing expression of these genes in the selected host cell is suitable, although transcriptional and translational regions from the host species are particularly useful. Expression in a host cell can occur in an induced or constitutive fashion. Induced expression occurs by inducing the activity of a regulatable promoter operably linked to the LPLAT gene of interest, while constitutive expression occurs by the use of a constituitive promoter.

3' non-coding sequences encoding transcription termination regions may be provided in a recombinant construct and may be 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 in both the same and different genera and species from which they were derived. Termination regions may also be derived from various genes native to the preferred hosts. The termination region is usually selected more for convenience rather than for any particular property.

Particularly useful termination regions for use in yeast are derived from a yeast gene, particularly *Saccharomyces, Schizosaccharomyces, Candida*,

Yarrowia or *Kluyveromyces*. The 3'-regions of mammalian genes encoding γ-interferon and α-2 interferon are also known to function in yeast. 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 region may be unnecessary, but is highly preferred.

The vector may also comprise a selectable and/or scorable marker, in addition to the regulatory elements described above. Preferably, the marker gene is an antibiotic resistance gene such that treating cells with the antibiotic results in growth inhibition, or death, of untransformed cells and uninhibited growth of transformed cells. For selection of yeast transformants, any marker that functions in yeast is useful with resistance to kanamycin, hygromycin and the amino glycoside G418 and the ability to grow on media lacking uracil, lysine, histine or leucine being particularly useful.

Merely inserting a gene (e.g., encoding a LPLAT) into a cloning vector does not ensure its expression at the desired rate, concentration, amount, etc. In response to the need for a high expression rate, many specialized expression vectors have been created by manipulating a number of different genetic elements that control transcription, RNA stability, translation, protein stability and location, oxygen limitation, and secretion from the host cell. Some of the manipulated features include: the nature of the relevant transcriptional promoter and terminator sequences, the number of copies of the cloned gene and whether the gene is plasmid-borne or integrated into the genome of the host cell, the final cellular location of the synthesized protein, the efficiency of translation and correct folding of the protein in the host organism, the intrinsic stability of the mRNA and protein of the cloned gene 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. Each of these may be used in the methods and host cells described herein to further optimize expression of LPLAT genes.

For example, LPLAT expression 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). Alternately, additional copies of the LPLAT genes may be introduced into the recombinant host cells to thereby increase EPA and/or DHA production and accumulation, either by cloning additional copies of genes within a single expression construct or by introducing additional copies into the host cell by increasing the plasmid copy number or by multiple integration of the cloned gene into the genome.

After a recombinant construct is created comprising at least one chimeric gene comprising a promoter, a LPLAT open reading frame ["ORF"] and a terminator, it is placed in a plasmid vector capable of autonomous replication in the host cell or is 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. Where constructs are targeted to an endogenous locus, all or some of the transcriptional and translational regulatory regions can be provided by the endogenous locus.

When two or more genes are expressed from separate replicating vectors, each vector may have a different means of selection and should lack homology to the other construct(s) 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(s) can be experimentally determined so that all introduced genes are expressed at the necessary levels to provide for synthesis of the desired products.

Constructs comprising the gene(s) of interest may be introduced into a host cell by any standard technique. These techniques include

transformation, e.g., lithium acetate transformation (*Methods in Enzymology*, 194:186-187 (1991)), biolistic impact, electroporation, microinjection, vacuum filtration or any other method that introduces the gene of interest into the host cell.

For convenience, a host cell that has been manipulated by any method to take up a DNA sequence, for example, in an expression cassette, is referred to herein as "transformed" or "recombinant" or “transformant”. The transformed host will have at least one copy of the expression construct and may have two or more, depending upon whether the gene is integrated into the genome, amplified, or is present on an extrachromosomal element having multiple copy numbers.

The transformed host cell can be identified by selection for a marker contained on the introduced construct. Alternatively, a separate marker construct may be co-transformed with the desired construct, as many transformation techniques introduce many DNA molecules into host cells.

Typically, transformed hosts are selected for their ability to grow on selective media, which may incorporate an antibiotic or lack a factor necessary for growth of the untransformed host, such as a nutrient or growth factor. An introduced marker gene may confer antibiotic resistance, or encode an essential growth factor or enzyme, thereby permitting growth on selective media when expressed in the transformed host. Selection of a transformed host can also occur when the expressed marker protein can be detected, either directly or indirectly. Additional selection techniques are described in U.S. Patent 7,238,482 and U.S. Patent 7,259,255.

Regardless of the selected host or expression construct, multiple transformants must be screened to obtain a strain displaying the desired expression level and pattern. For example, 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. Such screening may be accomplished by Southern analysis of DNA blots (Southern, *J. Mol. Biol.,* 98:503 (1975)),

Northern analysis of mRNA expression (Kroczek, *J. Chromatogr. Biomed. Appl.*, 618(1-2):133-145 (1993)), Western analysis of protein expression, phenotypic analysis or GC analysis of the PUFA products.

The metabolic process wherein oleic acid is converted to LC-PUFAs 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 LC-PUFA production.

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 Δ12 desaturase. Then, using the "Δ9 elongase/ Δ8 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 Δ9 elongase; 2) EDA is converted to dihomo-γ-linolenic acid ["DGLA"] by a Δ8 desaturase; 3) DGLA is converted to arachidonic acid ["ARA"] by a Δ5 desaturase; 4) ARA is converted to docosatetraenoic acid ["DTA"] by a $C_{20/22}$ elongase; and, 5) DTA is converted to docosapentaenoic acid ["DPAn-6"] by a Δ4 desaturase.

The "Δ9 elongase/ Δ8 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 Δ15 desaturase; 2) ALA is converted to eicosatrienoic acid ["ETrA"] by a Δ9 elongase; 3) ETrA is converted to eicosatetraenoic acid ["ETA"] by a Δ8 desaturase; 4) ETA is converted to eicosapentaenoic acid ["EPA"] by a Δ5 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 Δ4 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 Δ17 desaturase activity. Advantageously for the purposes herein, the Δ9

elongase/ Δ8 desaturase pathway enables production of an EPA oil that lacks significant amounts of γ-linolenic acid ["GLA"].

Alternate pathways for the biosynthesis of ω-3/ω-6 fatty acids utilize a Δ6 desaturase and $C_{18/20}$ elongase, that is, the "Δ6 desaturase/ Δ6 elongase pathway". More specifically, LA and ALA may be converted to to GLA and stearidonic acid ["STA"], respectively, by a Δ6 desaturase; then, a $C_{18/20}$ elongase converts GLA to DGLA and/or STA to ETA.

A LC-PUFA-producing recombinant host cell will possess at least one of the biosynthetic pathways described above, whether this pathway is native to the host cell or is genetically engineered. Preferably, the host cell will be capable of producing at least about 2-5% LC-PUFAs in the total lipids of the recombinant host cell, more preferably at least about 5-15% LC-PUFAs in the total lipids, more preferably at least about 15-35% LC-PUFAs in the total lipids, more preferably at least about 35-50% LC-PUFAs in the total lipids, more preferably at least about 50-65% LC-PUFAs in the total lipids and most preferably at least about 65-75% LC-PUFAs in the total lipids. The structural form of the LC-PUFAs is not limiting; thus, for example, the EPA or DHA may exist in the total lipids as free fatty acids or in esterified forms such as acylglycerols, phospholipids, sulfolipids or glycolipids.

A variety of eukaryotic microbial organisms, including bacteria, yeast, algae, stramenopile, oomycete, euglenoid and/or fungus, can produce (or can be engineered to produce) LC-PUFAs. These may include hosts that grow on a variety of feedstocks, including simple or complex carbohydrates, fatty acids, organic acids, oils, glycerols and alcohols, and/or hydrocarbons over a wide range of temperature and pH values.

Preferred microbial hosts are oleaginous organisms. These oleaginous organisms are naturally capable of oil synthesis and accumulation, wherein the total oil content can comprise greater than about 25% of the dry cell weight, more preferably greater than about 30% of the dry cell weight, and most preferably greater than about 40% of the dry cell

weight. Various bacteria, algae, euglenoids, moss, fungi, yeast and stramenopiles are naturally classified as oleaginous. Within this broad group of hosts, of particular interest are those organisms that naturally produce ω-3/ω-6 fatty acids. For example, ARA, EPA and/or DHA is produced via *Cyclotella* sp., *Crypthecodinium* sp., *Mortierella* sp., *Nitzschia* sp., *Pythium, Thraustochytrium* sp. and *Schizochytrium* sp. Thus, for example, transformation of *Mortierella alpina,* which is commercially used for production of ARA, with any of the present LPLAT genes under the control of inducible or regulated promoters could yield a transformant organism capable of synthesizing increased quantities of ARA. The method of transformation of *M. alpina* is described by Mackenzie et al. (*Appl. Environ. Microbiol.,* 66:4655 (2000)). Similarly, methods for transformation of Thraustochytriales microorganisms (e.g., *Thraustochytrium*, *Schizochytrium*) are disclosed in U.S. Patent 7,001,772. In alternate embodiments, a non-oleaginous organism can be genetically modified to become oleaginous, e.g., yeast such as *Saccharomyces cerevisiae* (U.S. Pat. Pub. No. 2007/0015237-A1).

In more preferred embodiments, the microbial host cells are oleaginous yeast. Genera typically identified as oleaginous yeast include, but are not limited to: *Yarrowia, Candida, Rhodotorula, Rhodosporidium*, *Cryptococcus, Trichosporon* and *Lipomyces*. More specifically, illustrative oil-synthesizing yeast include: *Rhodosporidium toruloides, Lipomyces starkeyii, L. lipoferus*, *Candida revkaufi, C. pulcherrima, C. tropicalis, C. utilis, Trichosporon pullans, T. cutaneum, Rhodotorula glutinus, R. graminis* and *Yarrowia lipolytica* (formerly classified as *Candida lipolytica*). Most preferred is the oleaginous yeast *Yarrowia lipolytica;* and, in a further embodiment, most preferred are the *Y. lipolytica* strains designated as ATCC #76982, ATCC #20362, ATCC #8862, ATCC #18944 and/or LGAM S(7)1 (Papanikolaou S., and Aggelis G., *Bioresour. Technol.*, 82(1):43-9 (2002)).

Specific teachings applicable for engineering ARA, EPA and DHA production in *Y. lipolytica* are provided in U.S. Pat. Pub. No. 2006-0094092-

A1, U.S. Pat. Pub. No. 2006-0115881-A1, U.S. Pat. Pub. No. 2009-0093543-A1 and U.S. Pat. Pub. No. 2006-0110806-A1, respectively. These references also describe the preferred method of expressing genes in *Yarrowia lipolytica* by integration of a linear DNA fragment into the genome of the host, preferred promoters, termination regions, integration loci and disruptions, and preferred selection methods when using this particular host species.

One of skill in the art would be able to use the cited teachings in U.S. Pat. Pub. No. 2006-0094092-A1, U.S. Pat. Pub. No. 2006-0115881-A1, U.S. Pat. Pub. No. 2009-0093543-A1 and U.S. Pat. Pub. No. 2006-0110806-A1 to recombinantly engineer other host cells for PUFA production.

The transformed recombinant host cell is grown under conditions that optimize expression of chimeric genes (e.g., encoding desaturases, elongases, LPLATs, etc.) and produce the greatest and the most economical yield of LC-PUFA(s). 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)).

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 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.

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.

Nitrogen may be supplied from an inorganic (e.g., $(NH_4)_2SO_4$) 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- and/or DHA-producing host cells and the promotion of the enzymatic pathways for EPA and/or DHA 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)).

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.

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 and/or DHA 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.

In some aspects, the primary product is oleaginous microbial biomass. As such, isolation and purification of the LC-PUFA-containing oils from the biomass may not be necessary (i.e., wherein the whole cell biomass is the product).

However, certain end uses and/or product forms may require partial and/or complete isolation/purification of the LC-PUFA-containing oil from the biomass, to result in partially purified biomass, purified oil, and/or purified LC-PUFAs. Fatty acids, including PUFAs, may be found in the host microorganisms as free fatty acids or in esterified forms such as acylglycerols, phospholipids, sulfolipids or glycolipids. These fatty acids may be extracted from the host cells 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)).

In general, means for the purification of fatty acids (including LC-PUFAs) 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, or combinations thereof. See U.S. Patent 7,238,482.

Many food and feed products incorporate ω-3 and/or ω-6 fatty acids, particularly ALA, GLA, ARA, EPA, DPA and DHA. It is contemplated that oleaginous yeast biomass comprising LC-PUFAs, partially purified biomass

comprising LC-PUFAs, purified oil comprising LC-PUFAs, and/or purified LC-PUFAs made by the methods and host cells described herein impart the health benefits, upon ingestion of foods or feed improved by their addition. These oils can be added to food analogs, drinks, meat products, cereal products, baked foods, snack foods and dairy products, to name a few. See U.S. Pat. Appl. Pub. No. 2006-0094092.

These compositions may impart health benefits by being added to medical foods including medical nutritionals, dietary supplements, infant formula and pharmaceuticals. The skilled artisan will appreciate the amount of the oils to be added to food, feed, dietary supplements, nutriceuticals, pharmaceuticals, and other ingestible products as to impart health benefits. Health benefits from ingestion of these oils are described in the art, known to the skilled artisan and continuously investigated. Such an amount is referred to herein as an “effective” amount and depends on, among other things, the nature of the ingested products containing these oils and the physical conditions they are intended to address.

EXAMPLES

The present invention is further described in the following Examples, which illustrate reductions to practice of the invention but do not completely define all of its possible variations.

GENERAL METHODS

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. Bennan, 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).

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. Costilow, 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.

General molecular cloning was performed according to standard methods (Sambrook et al., *supra*). DNA sequence was generated on an ABI Automatic sequencer using dye terminator technology (U.S. Patent 5,366,860; EP 272,007) using a combination of vector and insert-specific primers. Sequence editing was performed in Sequencher (Gene Codes Corporation, Ann Arbor, MI). All sequences represent coverage at least two times in both directions.

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

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*

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 (e.g., YPD agar medium, Basic Minimal Media ["MM"], Minimal Media + Uracil ["MMU"], Minimal Media + Leucine + Lysine ["MMLeuLys"], Minimal Media + 5-Fluoroorotic Acid ["MM + 5-FOA"], High Glucose Media ["HGM"] and Fermentation medium ["FM"]), as described in U.S. Pat. Appl. Pub. No. 2009-0093543-A1.

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

Fatty Acid Analysis Of *Yarrowia lipolytica*

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 Nishida 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.

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.

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.

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.

For quantitating the amount of an individual fatty acid or the total fatty acids as a weight percent of the dry cell weight ["% DCW"], cells from 10 mL of the culture were collected by centrifugation, washed once with 10 mL water and collected by centrifugation again. Cells were resuspended in 1-2 mL water, poured into a pre-weighed aluminium weighing pan, and rinsed with 1-2 mL water that was also added to the same weighing pan. The pan was placed under vacuum at 80 °C overnight. The pan was weighed and the DCW calculated by subtracting the weight of the empty pan. Determination of the fatty acid as a % DCW can then be calculated based on either μg FAME or μg FA as a fraction of the μg DCW (for example, FAME % DCW was calculated as μg FAME/μg DCW*100).

EXAMPLE 1

Generation Of *Yarrowia lipolytica* Strain Y8406 To Produce About 51% EPA Of Total Fatty Acids

The present Example describes the construction of strain Y8406, derived from *Yarrowia lipolytica* ATCC #20362, capable of producing about 51% EPA relative to the total lipids via expression of a Δ9 elongase/Δ8 desaturase pathway. This strain was used as the EPA-producing host cell in Example 4.

The development of strain Y8406 (FIG. 2) required the construction of strains Y2224, Y4001, Y4001U, Y4036, Y4036U, L135, L135U9, Y8002, Y8006U6, Y8069, Y8069U, Y8154, Y8154U, Y8269 and Y8269U.

Generation Of Y4036U Strain

Briefly, strain Y8406 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 Y4001U1 (*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, Y4001U, Y4036 and Y4036U are described in the General Methods of U.S. Pat. App. Pub. No. 2008-0254191, hereby incorporated herein by reference.

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::EgD8M::Pex20, EXP1::EgD8M::Pex16, GPD::FmD12::Pex20, YAT1::FmD12::OCT (wherein FmD12 is a *Fusarium moniliforme* Δ12 desaturase gene [U.S. Patent 7,504,259]; ME3S is a codon-optimized $C_{16/18}$ elongase gene, derived from *Mortierella alpina* [U.S. Patent 7,470,532]; EgD9e is a *Euglena gracilis* Δ9 elongase gene [U.S. Patent 7,645,604]; EgD9eS is a codon-optimized Δ9 elongase gene, derived from *Euglena gracilis* [U.S. Patent 7,645,604]; EgD8M is a synthetic mutant Δ8 desaturase

[U.S. Patent 7,709,239], derived from *Euglena gracilis* [U.S. Patent 7,256,033]).

Generation Of L135 Strain (*Ura3+, Leu-*, Δpex3) With Chromosomal Deletion Of Pex3

Construction of strain L135 is described in Example 12 of Intl. App. Pub. No. WO 2009/046248, hereby incorporated herein by reference. 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 Δ9 elongase percent conversion efficiency was increased from ca. 48% in strain Y4036 to 83% in strain L135.

The final genotype of strain L135 with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was *Ura3+*, *Leu-, Pex3-, unknown1-,* YAT1::ME3S::Pex16, EXP1::EgD9eS::Lip1, FBAINm::EgD9eS::Lip2, GPAT::EgD9e::Lip2, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, GPD::FmD12::Pex20, YAT1::FmD12::OCT.

Generation of L135U9 (*Leu-, Ura3-*) Strain

Strain L135U was created *via* temporary expression of the *Cre* recombinase enzyme in plasmid pY116 (FIG. 3; SEQ ID NO:33; described in Example 7 of Intl. App. Pub. No. WO 2008/073367, hereby incorporated herein by reference) 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 plates 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

Construct pZKSL-5S5A5 (FIG. 4A; SEQ ID NO:34) 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 4: Description of Plasmid pZKSL-5S5A5 (SEQ ID NO:34)

RE Sites And Nucleotides Within SEQ ID NO:34	Description Of Fragment And Chimeric Gene Components
*Asc*I/*Bsi*WI (5925-6645)	720 bp 5' portion of *Yarrowia Lys5* gene (GenBank Accession No. M34929; labeled as "lys5 5' region" in Figure)
*Pac*I/*Sph*I (2536-3225)	689 bp 3' portion of *Yarrowia Lys5* gene (GenBank Accession No. M34929; labeled as "Lys5-3'" in Figure)
*Eco*RI/ *Bsi*WI (9338-6645)	FBAIN::EgD5SM::Pex20, comprising: • FBAIN: *Yarrowia lipolytica* FBAIN promoter (U.S. Patent 7,202,356); • EgD5SM: Synthetic mutant Δ5 desaturase (SEQ ID NO:35; U.S. Pat. Pub. No. 2010-0075386-A1), derived from *Euglena gracilis* (U.S. Patent 7,678,560) (labeled as "ED5S" in Figure); • Pex20: Pex20 terminator sequence from *Yarrowia Pex20* gene (GenBank Accession No. AF054613)
*Pme*I /*Cla*I (11503-1)	YAT1::EaD5SM::OCT, comprising: • YAT1: *Yarrowia lipolytica* YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • EaD5SM: Synthetic, mutant Δ5 desaturase (SEQ ID NO:37; U.S. Pat. Pub. No. 2010-0075386-A1), derived from *Euglena anabaena* (U.S. Pat. Appl. Pub. No. 2008-0274521-A1) (labeled as "EaD5S" in Figure); • OCT: OCT terminator sequence of *Yarrowia OCT* gene (GenBank Accession No. X69988)

*Cla*I/*Pac*I (1-2536)	EXP1::EgD5M::Pex16, comprising: • EXP1: *Yarrowia lipolytica* export protein (EXP1) promoter (labeled as “EXP” in Figure; Intl. App. Pub. No. WO 2006/052870); • EgD5M: Mutant Δ5 desaturase (SEQ ID NO:90; U.S. Pat. Pub. No. 2010-0075386-A1) with elimination of internal *EcoR*I, *Bgl*II, *Hind*III and *Nco*I restriction enzyme sites, derived from *Euglena gracilis* (U.S. Patent 7,678,560) (labeled as “Euglena D5WT” in Figure); • Pex16: Pex16 terminator sequence from *Yarrowia Pex16* gene (GenBank Accession No. U75433)
*Eco*RI/*Pme*I (9360-11503)	*Yarrowia Leu2* gene (GenBank Accession No. M37309)

The pZKSL-5S5A5 plasmid was digested with *Asc*I/*Sph*I, 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.

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.

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::EgD8M::Pex16, FBAIN::EgD5SM::Pex20, EXP1::EgD5M::Pex16, YAT1::EaD5SM::Oct.

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

Construct pZP3-Pa777U (FIG. 4B; SEQ ID NO:39; described in Table 9 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1, hereby incorporated herein by reference) was generated to integrate three Δ17 desaturase genes into the *Pox3* loci (GenBank Accession No. AJ001301) of strain Y8002.

The pZP3-Pa777U plasmid was digested with *Asc*I/*Sph*I, 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.

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::EgD8M::Pex16, FBAIN::EgD5SM::Pex20, EXP1::EgD5M::Pex16, YAT1::EaD5SM::Oct.

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::EgD8M::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].

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.

<u>Generation Of Strain Y8006U6 (*Ura3-*)</u>

To disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:40; described in Table 15 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1, hereby incorporated herein by reference) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y8006.

Plasmid pZKUM was digested with *SalI/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.

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.

GC analyses showed the presence of 22.9%, 25.5%, 23.6% 21.6%, 21.6% and 25% ARA of TFAs 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).

<u>Generation Of Y8069 Strain To Produce About 37.5% EPA Of TFAs</u>

Construct pZP3-Pa777U (FIG. 4B; SEQ ID NO:39; described in Table 9 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1, hereby incorporated herein

by reference) was used to integrate three Δ17 desaturase genes into the *Pox3* loci (GenBank Accession No. AJ001301) of strain Y8006U6.

The pZP3-Pa777U plasmid was digested with *Asc*I/*Sph*I, 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.

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.

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-, Leu+, 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::EaD5SM::Oct, YAT1::PaD17S::Lip1, EXP1::PaD17::Pex16, FBAINm::PaD17::Aco.

<u>Generation Of Strain Y8069U (*Ura3-*)</u>

To disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:40; 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.

GC analyses showed the presence of 22.4%, 21.9% and 21.5% EPA of TFAs 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

Construct pZKL2-5mB89C (FIG. 5B; SEQ ID NO:41) was generated to integrate one Δ5 desaturase gene, one Δ9 elongase gene, one Δ8 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 5: Description of Plasmid pZKL2-5mB89C (SEQ ID NO:41)

RE Sites And Nucleotides Within SEQ ID NO:41	Description Of Fragment And Chimeric Gene Components
*Asc*I/*Bsi*WI (730-1)	722 bp 5' portion of *Yarrowia Lip2* gene (labeled as "Lip2.5N" in Figure; GenBank Accession No. AJ012632)
*Pac*I/*Sph*I (4141-3438)	697 bp 3' portion of *Yarrowia Lip2* gene (labeled as "Lip2.3N" in Figure; GenBank Accession No. AJ012632)
*Swa*I/*Bsi*WI (13561-1)	YAT1::YlCPT1::Aco, comprising: • YAT1: *Yarrowia lipolytica* YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • YlCPT1: *Yarrowia lipolytica* diacylglycerol cholinephosphotransferase gene (SEQ ID NO:42) (labeled as "Y. lipolytica CPT1 cDNA" in Figure; Intl. App. Pub. No. WO 2006/052870); • Aco: Aco terminator sequence from *Yarrowia Aco* gene (GenBank Accession No. AJ001300)
*Pme*I/*Swa*I (10924-13561)	FBAIN::EgD8M::Lip1 comprising: • FBAIN: *Yarrowia lipolytica* FBAIN promoter (U.S. Patent 7,202,356); • EgD8M: Synthetic mutant Δ8 desaturase (SEQ ID NO:44; U.S. Patent 7,709,239), derived from *Euglena gracilis* ("EgD8S"; U.S. Patent 7,256,033) (labeled as "D8S-23" in

	Figure); • Lip1: Lip1 terminator sequence from *Yarrowia Lip1* gene (GenBank Accession No. Z50020)
*Pme*I/*Cla*I (10924-9068)	YAT1::EgD9eS::Lip2, comprising: • YAT1: *Yarrowia lipolytica* YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • EgD9eS: codon-optimized Δ9 elongase (SEQ ID NO:46), derived from *Euglena gracilis* (U.S. Patent 7,645,604); • Lip2: Lip2 terminator sequence from *Yarrowia Lip2* gene (GenBank Accession No. AJ012632)
*Cla*I/*Eco*RI (9068-6999)	*Yarrowia* Ura3 gene (GenBank Accession No. AJ306421)
*Eco*RI/*Pac*I (6999-4141)	GPDIN::EgD5SM::ACO, comprising: • GPDIN: *Yarrowia lipolytica* GPDIN promoter (U.S. Patent 7,459,546); • EgD5SM: Synthetic mutant Δ5 desaturase (SEQ ID NO:35; U.S. Pat. Pub. No. 2010-0075386-A1), derived from *Euglena gracilis* (U.S. Patent 7,678,560) (labeled as "EgD5S-HPGS" in Figure); • Aco: Aco terminator sequence from *Yarrowia Aco* gene (GenBank Accession No. AJ001300)

The pZKL2-5mB89C plasmid was digested with *Asc*I/*Sph*I, 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.

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.

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::EgD9e::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::YlCPT::Aco.

Generation Of Strain Y8154U1 (*Ura3-*)

To disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:40; 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.

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

Construct pZKL1-2SR9G85 (FIG. 6A; SEQ ID NO:48) was generated to integrate one DGLA synthase, one Δ12 desaturase gene and one Δ5 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 Δ9 elongase linked to a Δ8 desaturase (U.S. Pat. Appl. Pub. No. 2008-0254191-A1).

The pZKL1-2SR9G85 plasmid contained the following components:

Table 6: Description of Plasmid pZKL1-2SR9G85 (SEQ ID NO:48)

RE Sites And Nucleotides Within SEQ ID NO:48	Description Of Fragment And Chimeric Gene Components

*Asc*I/*Bsi*WI (4189-3373)	809 bp 5' portion of *Yarrowia Lip1* gene (labeled as "Lip1-5'N" in Figure; GenBank Accession No. Z50020)
*Pac*I/*Sph*I (7666-6879)	763 bp 3' portion of *Yarrowia Lip1* gene (labeled as "Lip1.3N" in Figure; GenBank Accession No. Z50020)
*Cla*I/*Swa*I (1-3217)	YAT1::E389D9eS/EgD8M::Lip1, comprising: • YAT1: *Yarrowia lipolytica* YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006-0094102-A1); • E389D9eS/EgD8M: gene fusion comprising a codon-optimized Δ9 elongase derived from *Eutreptiella* sp. CCMP389 ("E389D9eS"), a linker, and the synthetic mutant Δ8 desaturase derived from *Euglena gracilis* ("EgD8M") (SEQ ID NO:49) (labeled individually as "E389S", "Linker" and "EgD8M" in Figure; U.S. Pat. Appl. Pub. No. 2008-0254191-A1); • Lip1: Lip1 terminator sequence from *Yarrowia Lip1* gene (GenBank Accession No. Z50020)
*Sal*I/*Cla*I (11982-1)	GPM::EgD5SM::Oct comprising: • GPM: *Yarrowia lipolytica* GPM promoter (labeled as "GPML" in Figure; U.S. Patent 7,202,356); • EgD5SM: Synthetic mutant Δ5 desaturase (SEQ ID NO:35; U.S. Pat. Pub. No. 2010-0075386-A1), derived from *Euglena gracilis* (U.S. Patent 7,678,560) (labeled as "ED5S" in Figure); • OCT: OCT terminator sequence of *Yarrowia OCT* gene (GenBank Accession No. X69988)
*Sal*I/*Eco*RI (11982-10363)	*Yarrowia* Ura3 gene (GenBank Accession No. AJ306421)
*EcoR*I/*Pac*I (10363-7666)	EXP1::FmD12S::ACO, comprising: • EXP1: *Yarrowia lipolytica* export protein (EXP1) promoter (labeled as "Exp" in Figure; Intl. App. Pub. No. WO 2006/052870); • FmD12S: codon-optimized Δ12 elongase (SEQ ID NO:51), derived from *Fusarium moniliforme* (labeled as "FD12S" in Figure; U.S. Patent 7,504,259); • Aco: Aco terminator sequence from *Yarrowia Aco* gene (GenBank Accession No. AJ001300)

The pZKL1-2SR9G85 plasmid was digested with *Asc*I/*Sph*I, 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.

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.

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-, unknown6-,* YAT1::ME3S::Pex16, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9e::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::YlCPT::Aco.

Generation Of Strain Y8269U (*Ura3-*)

To disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:40; 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.

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

Construct pZSCP-Ma83 (FIG. 6B; SEQ ID NO:53) was generated to integrate one Δ8 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 7: Description of Plasmid pZSCP-Ma83 (SEQ ID NO:53)

RE Sites And Nucleotides Within SEQ ID NO:53	Description Of Fragment And Chimeric Gene Components
*Bsi*WI/*Asc*I (1-1328)	1327 bp 3' portion of *Yarrowia SCP2* gene (labeled as "SCP2-3'" in Figure; GenBank Accession No. XM_503410)
*Sph*I/*Pac*I (4036-5816)	1780 bp 5' portion of *Yarrowia SCP2* gene (labeled as "SCP2-5'" in Figure; GenBank Accession No. XM_503410)
*Swa*I/*Bsi*WI (12994-1)	GPD::ME3S::Pex20, comprising: • GPD: *Yarrowia lipolytica* GPD promoter (U.S. Patent 7,259,255); • ME3S: codon-optimized $C_{16/18}$ elongase gene (SEQ ID NO:54), derived from *M. alpina* (U.S. Patent 7,470,532); • Pex20: Pex20 terminator sequence from *Yarrowia Pex20* gene (GenBank Accession No. AF054613)
*Pme*I/*Swa*I (10409-12994)	YAT1::MCS::Lip1 comprising: • YAT1: *Yarrowia lipolytica* YAT1 promoter (labeled as "YAT" in Figure; U.S. Pat. Appl. Pub. No. 2006/0094102-A1); • MCS: codon-optimized malonyl-CoA synthetase gene (SEQ ID NO:56), derived from *Rhizobium leguminosarum* bv. viciae 3841 (U.S. Patent Application No. 12/637877); • Lip1: Lip1 terminator sequence from *Yarrowia Lip1* gene (GenBank Accession No. Z50020)
*Cla*I/*Pme*I (7917-10409)	GPD::EaD8S::Pex16 comprising: • GPD: *Yarrowia lipolytica* GPD promoter (U.S. Patent 7,259,255); • EaD8S: codon-optimized Δ8 desaturase gene (SEQ ID NO:58), derived from *Euglena anabaena* (U.S. Pat. Appl. Pub. No. 2008-0254521-A1); • Pex16: Pex16 terminator sequence from *Yarrowia Pex16* gene (GenBank Accession No. U75433)
*Sal*I/*Eco*RI (7467-5848)	*Yarrowia* Ura3 gene (GenBank Accession No. AJ306421)

The pZSCP-Ma83 plasmid was digested with *Asc*I/*Sph*I, 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.

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.

Knockout of the *SCP2* loci (GenBank Accession No. XM_503410) was not confirmed in any of these EPA strains, produced by transformation with pZSCP-Ma83.

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-, unknown6-, 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::YlCPT::Aco, YAT1::MCS::Lip1.

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

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, as follows.

Specifically, 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 fatty acid analysis (*supra*) and 10 mL dried for dry cell weight ["DCW"] determination.

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.

Data from flask assays are presented as Table 8. The Table presents the total dry cell weight of the cells ["DCW'], the total lipid content of cells

["FAME % 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 FAME % 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.

Table 8: 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)	Total FAME % DCW	% TFAs													EPA FAME % 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

Generation Of Strain Y8406U (*Ura3-*)

To disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:40; 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
5 Y8406 in a manner similar to that described for pZKUM transformation of strain Y8006 (*supra*). Several transformants were grown and identified to possess a *Ura-* phenotype.

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

EXAMPLE 2

Generation Of *Yarrowia lipolytica* Strain Y5037 To Produce About 18.6% EPA, 22.8% DPA And 9.7% DHA Of Total Fatty Acids

15 The present Example describes the construction of strain Y5037, derived from *Yarrowia lipolytica* ATCC #20362, capable of producing about 18.6% EPA, 22.8% DPA and 9.7% DHA relative to the total lipids via expression of a Δ9 elongase/Δ8 desaturase pathway. This strain was used as the DHA-producing host cell in Example 5.

20 Briefly, as diagrammed in FIG. 7, strain Y5037 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 Y4001U1 (*Leu-* and *Ura-*), strain
25 Y4036 (producing 18% DGLA with a *Leu-* phenotype), strain Y4036U (*Leu-* and *Ura-*), strain Y4070 (producing 12% ARA with a *Ura-* phenotype), strain Y4086 (producing 14% EPA), strain Y4086U1 (*Ura3-*), strain Y4128 (producing 37% EPA; deposited with the American Type Culture Collection on August 23, 2007, bearing the designation ATCC
30 PTA-8614), strain Y4128U3 (*Ura-*), strain Y4217 (producing 42% EPA), strain Y4217U2 (*Ura-*), strain Y4259 (producing 46.5% EPA), strain Y4259U2 (*Ura-*), strain Y4305 (producing 53.2% EPA), strain Y4305U3 (*Ura-*), strain Y5004 (producing 17% EPA, 18.7% DPA and 6.4% DHA),

strain Y5004U1 (*Ura-*), strain Y5018 (producing 25.4% EPA, 11.4% DPA and 9.4% DHA), strain Y5018U1 (*Ura-*) and strain Y5037 (producing 18.6% EPA, 22.8% DPA and 9.7% DHA relative to the total TFAs). Further details regarding the construction of strains Y2224, Y4001, Y4001U, Y4036, Y4036U, Y4070, Y4086, Y4086U1, Y4128, Y4128U3, Y4217, Y4217U2, Y4259, Y4259U2, Y4305 and Y4305U3 are described in the General Methods of U.S. Pat. App. Pub. No. 2008-0254191-A1 and in Examples 1-3 of U.S. Pat. App. Pub. No. 2009-0093543-A1, hereby incorporated herein by reference.

The complete lipid profile of strain Y4305 was as follows: 16:0 (2.8%), 16:1 (0.7%), 18:0 (1.3%), 18:1 (4.9%), 18:2 (17.6%), ALA (2.3%), EDA (3.4%), DGLA (2.0%), ARA (0.6%), ETA (1.7%), and EPA (53.2%). The total lipid content of cells ["TFAs % DCW"] was 27.5.

The final genotype of strain Y4305 with respect to wild type *Yarrowia lipolytica* ATCC #20362 was *SCP2- (YALI0E01298g), YALI0C18711g-, Pex10-, YALI0F24167g-, unknown 1-, unknown 3-, unknown 8-,* GPD::FmD12::Pex20, YAT1::FmD12::OCT, GPM/FBAIN::FmD12S::OCT, EXP1::FmD12S::Aco, YAT1::FmD12S::Lip2, YAT1::ME3S::Pex16, EXP1::ME3S::Pex20 (3 copies), GPAT::EgD9e::Lip2, EXP1::EgD9eS::Lip1, FBAINm::EgD9eS::Lip2, FBA::EgD9eS::Pex20, GPD::EgD9eS::Lip2, YAT1::EgD9eS::Lip2, YAT1::E389D9eS::OCT, FBAINm::EgD8M::Pex20, FBAIN::EgD8M::Lip1 (2 copies), EXP1::EgD8M::Pex16, GPDIN::EgD8M::Lip1, YAT1::EgD8M::Aco, FBAIN::EgD5::Aco, EXP1::EgD5S::Pex20, YAT1::EgD5S::Aco, EXP1::EgD5S::ACO, YAT1::RD5S::OCT, YAT1::PaD17S::Lip1, EXP1::PaD17::Pex16, FBAINm::PaD17::Aco, YAT1::YlCPT1::ACO, GPD::YlCPT1::ACO (wherein FmD12 is a *Fusarium moniliforme* Δ12 desaturase gene [U.S. Patent 7,504,259]; FmD12S is a codon-optimized Δ12 desaturase gene, derived from *Fusarium moniliforme* [U.S. Patent 7,504,259]; ME3S is a codon-optimized $C_{16/18}$ elongase gene, derived from *Mortierella alpina* [U.S. Patent 7,470,532]; EgD9e is a *Euglena gracilis* Δ9 elongase gene [U.S. Patent 7,645,604]; EgD9eS is a codon-optimized Δ9 elongase gene, derived from *Euglena*

gracilis [U.S. Patent 7,645,604]; E389D9eS is a codon-optimized Δ9 elongase gene, derived from *Eutreptiella* sp. CCMP389 [U.S. Patent 7,645,604]; EgD8M is a synthetic mutant Δ8 desaturase [U.S. Patent 7,709,239], derived from *Euglena gracilis* [U.S. Patent 7,256,033]; EgD5 is
5 a *Euglena gracilis* Δ5 desaturase [U.S. Patent 7,678,560]; EgD5S is a codon-optimized Δ5 desaturase gene, derived from *Euglena gracilis* [U.S. Patent 7,678,560]; RD5S is a codon-optimized Δ5 desaturase, derived from *Peridinium sp.* CCMP626 [U.S. Patent 7,695,950]; PaD17 is a *Pythium aphanidermatum* Δ17 desaturase [U.S. Patent 7,556,949];
10 PaD17S is a codon-optimized Δ17 desaturase, derived from *Pythium aphanidermatum* [U.S. Patent 7,556,949]; and, YlCPT1 is a *Yarrowia lipolytica* diacylglycerol cholinephosphotransferase gene [Intl. App. Pub. No. WO 2006/052870]).

Strain Y4305U (*Ura3-*) was generated via integrating a *Ura3* mutant
15 gene into the *Ura3* gene of strain Y4305.

Generation Of Y5004 Strain To Produce about 17.0% EPA, 18.7% DPA And 6.4% DHA Of TFAs

Construct pZKL4-220EA41B (FIG. 8A; SEQ ID NO:60) was constructed to integrate two $C_{20/22}$ elongase genes and two Δ4 desaturase
20 genes into the lipase 4-like locus (GenBank Accession No. XM_503825) of strain Y4305U3. The pZKL4-220EA41B plasmid contained the following components:

Table 9: Components Of Plasmid pZKL4-220EA41B (SEQ ID NO:60)

RE Sites And Nucleotides Within SEQ ID NO:60	Description Of Fragment And Chimeric Gene Components
Asc I/*BsiW* I (9777-9025)	745 bp 5' portion of the *Yarrowia* Lipase 4-like gene (GenBank Accession No. XM_503825; labeled as "Lip4" in Figure)
PacI/SphI (13273-12485)	782 bp 3' portion of *Yarrowia* Lipase 4 like gene (GenBank Accession No. XM_503825; labeled as "Lip4-3'" in Figure)

SwaI/BsiW I (6882-9025)	FBAINm::EaC20ES::Pex20, comprising: • FBAINm: *Yarrowia lipolytica* FBAINm promoter (U.S. Patent 7,202,356) • EaC20ES: codon-optimized C20 elongase gene (SEQ ID NO:61), derived from *Euglena anabaena* (U.S. Pat. Appl. Pub. No. 2008/0254191-A1); • Pex20: Pex20 terminator sequence from *Yarrowia Pex20* gene (GenBank Accession No. AF054613)
*Pme*I/*Swa*I (4903-6882)	YAT1::EgC20ES::Lip1, comprising: • YAT1: *Yarrowia lipolytica* YAT1 promoter (U.S. Pat. Appl. Pub. No. 2006/0094102-A1); • EgC20ES: codon-optimized C20 elongase gene (SEQ ID NO:63), derived from *Euglena gracilis* (U.S. Pat. Appl. Pub. No. 2008/0254191-A1); • Lip1: Lip1 terminator sequence from *Yarrowia Lip1* gene (GenBank Accession No. Z50020)
*Pme*I/*Cla*I (4903-2070)	EXP1::EaD4S-1::Lip2, comprising: • EXP1: *Yarrowia lipolytica* export protein (EXP1) promoter (Intl. App. Pub. No. WO 2006/052870); • EaD4S-1: codon-optimized truncated Δ4 desaturase (SEQ ID NO:65), derived from *Euglena anabaena* (U.S. Pat. Appl. Pub. No. 2008/0254191-A1); • Lip2: Lip2 terminator sequence from *Yarrowia Lip2* gene (GenBank Accession No. AJ012632)
*SalI/EcoR*I (1620-1)	*Yarrowia* Ura3 gene (GenBank Accession No. AJ306421)
*EcoR*I/*Pac*I (1-14039)	GPDIN::EaD4SB::Aco, comprising: • GPDIN: *Yarrowia lipolytica* GPDIN promoter (U.S. Patent 7,459,546); • EaD4SB: codon-optimized truncated Δ4 desaturase version B (SEQ ID NO:67), derived from *Euglena anabaena* (U.S. Pat. Appl. Pub. No. 2008/0254191-A1); • Aco: Aco terminator sequence from *Yarrowia Aco* gene (GenBank Accession No. AJ001300)

The pZKL4-220EA41B plasmid was digested with *Asc*I/*Sph*I, and then used for transformation of strain Y4305U3 (*supra*), according to the General Methods. The transformants were selected on MM plates. After 5 days growth at 30 °C, 72 transformants grown on the MM plates were picked and re-streaked onto fresh MM plates. Once grown, these strains were individually inoculated into 3 mL 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.

GC analyses showed the presence of DHA in the transformants with pZKL4-220EA41B, but not in the parent Y4305U strain. Most of the selected 72 strains produced about 22% EPA, 18% DPA and 5% DHA of TFAs. Strain #2 produced 17% EPA, 18.7% DPA and 6.4% DHA, while strain #33 produced 21.5% EPA, 21% DPA and 5.5% DHA. These two strains were designated as Y5004 and Y5005, respectively.

Knockout of the lipase 4-like locus (GenBank Accession No. XM_503825) was not confirmed in either strain Y5004 or Y5005.

Generation Of Strain Y5004U (*Ura3-*)

To disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:40; described in Table 15 of U.S. Pat. App. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y5004, in a manner similar to that described for pZKUM transformation of strain Y8006 (Example 1).

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.

GC analyses showed the presence of 14.8% EPA, 17.4% DPA and 0.4% DHA of TFAs in transformant #5 and 15.3% EPA, 17.2% DPA and 0.4% DHA of TFAs in transformant #8. These two strains were designated as strains Y5004U1 and Y5004U2, respectively (collectively, Y5004U).

Generation Of Strain Y5018 To Produce About 25.4% EPA, 11.4% DPA And 9.4% DHA Of TFAs

Construct pZKL3-4GER44 (FIG. 8B; SEQ ID NO:69) was constructed to integrate one $C_{20/22}$ elongase gene and three $\Delta 4$ desaturase genes into the lipase 3-like locus (GenBank Accession No. XP_506121) of strain Y5004U1. The pZKL3-4GER44 plasmid contained the following components:

Table 10: Components Of Plasmid pZKL3-4GER44 (SEQ ID NO:69)

RE Sites And Nucleotides Within SEQ ID NO:69	Description Of Fragment And Chimeric Gene Components
Asc I/*BsiW* I (10527-9640)	887 bp 5' portion of the *Yarrowia* Lipase 3-like gene (GenBank Accession No. XP_506121)
Pac I/Sph I (14039-13235)	804 bp 3' portion of *Yarrowia* Lipase 3-like gene (GenBank Accession No. XP_506121)
Swa I/BsiW I (7485-9640)	FBAINm::EgC20ES::Pex20, comprising: • FBAINm: *Yarrowia lipolytica* FBAINm promoter (U.S. Patent 7,202,356); • EgC20ES: codon-optimized C20 elongase gene (SEQ ID NO:63), derived from *Euglena gracilis* (U.S. Pat. Appl. Pub. No. 2008/0254191-A1); • Pex20: Pex20 terminator sequence from *Yarrowia Pex20* gene (GenBank Accession No. AF054613)
*Pme*I/*Swa*I (4774-7485)	YAT1::EaD4S-1::Lip1, comprising: • YAT1: *Yarrowia lipolytica* YAT1 promoter (U.S. Pat. Appl. Pub. No. 2006/0094102-A1); • EaD4S-1: codon-optimized truncated Δ4 desaturase (SEQ ID NO:65), derived from *Euglena anabaena* (U.S. Pat. Appl. Pub. No. 2008/0254191-A1); • Lip1: Lip1 terminator sequence from *Yarrowia Lip1* gene (GenBank Accession No. Z50020)
*Cla*I/*Pme*I (2070-4774)	EXP1::E1594D4S::Oct, comprising: • EXP1: *Yarrowia lipolytica* export protein promoter (Intl. App. Pub. No. WO 2006/052870); • E1594D4S: codon-optimized Δ4 desaturase (SEQ ID NO:70), derived from *Eutreptiella cf_gymnastica* CCMP1594 (U.S. Pat. Appl. Pub. No. 2009/0253188-A1) (labeled as "D4S-1594" in Figure); • OCT: OCT terminator sequence of *Yarrowia OCT* gene (GenBank Accession No. X69988)
*SalI/EcoR*I (1620-1)	*Yarrowia* Ura3 gene (GenBank Accession No. AJ306421)
*EcoR*I/ *Pac*I (1-14039)	GPDIN::EgD4S-1::Aco, comprising: • GPDIN: *Yarrowia lipolytica* GPDIN promoter (U.S. Patent 7,459,546); • EgD4S-1: codon-optimized truncated Δ4 desaturase (SEQ ID NO:72), derived from *Euglena gracilis* (U.S. Pat. Appl. Pub. No. 2008/0254191-A1); • Aco: Aco terminator sequence from *Yarrowia Aco* gene (GenBank Accession No. AJ001300)

The pZKL3-4GER44 plasmid was digested with *Asc*I/*Sph*I, and then used for transformation of strain Y5004U1, according to the General Methods. The transformants were selected on MM plates. After 5 days growth at 30 °C, 96 transformants grown on the MM plates were picked

and re-streaked onto fresh MM plates. Once grown, these strains were individually inoculated into 3 mL 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
5 cells were subjected to fatty acid analysis, according to the General Methods.

GC analyses showed that most of the selected 96 strains produced about 19% EPA, 22% DPA and 7% DHA of TFAs. Strain #1 produced 23.3% EPA, 13.7% DPA and 8.9% DHA, while strain #49 produced 25.2%
10 EPA, 11.4% DPA and 9.4% DHA. These two strains were designated as Y5011 and Y5018, respectively.

Knockout of the lipase 3-like locus (GenBank Accession No. XP_506121) was not confirmed in strains Y5011 and Y5018.

Generation Of Strain Y5018U (*Ura3-*)

15 To disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:40; described in Table 15 of U.S. Pat. App. Pub. No. 2009-0093543-A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y5018, in a manner similar to that described for pZKUM transformation of strain Y8006 (Example 1). A total of 18 transformants were grown and
20 identified to possess a *Ura-* phenotype.

GC analyses showed the presence of 16.6% EPA, 10.4% DPA and 0.0% DHA of FAMEs in pZKUM-transformant strain #2 and 17.0% EPA, 10.8% DPA and 0.0% DHA of FAMEs in pZKUM-transformant strain #4. These two strains were designated as strains Y5018U1 and Y5018U2,
25 respectively (collectively, Y5018U).

Generation Of Strain Y5037 To Produce About 18.6% EPA, 22.8% DPA And 9.7% DHA Of TFAs

Construct pZKLY-G20444 (FIG. 9; SEQ ID NO:74) was constructed to integrate one DHA synthase and two Δ4 desaturase genes into the
30 lipase 7-like locus (GenBank Accession No. AJ549519) of strain Y5018U1. A DHA synthase is a multizyme comprising a C20 elongase linked to a Δ4 desaturase. The pZKLY-G20444 plasmid contained the following components:

Table 11: Components Of Plasmid pZKLY-G20444 (SEQ ID NO:74)

RE Sites And Nucleotides Within SEQ ID NO:74	Description Of Fragment And Chimeric Gene Components
*Asc*I/*BsiW*I (9370-8476)	887 bp 5' portion of the *Yarrowia* Lipase 7-like gene (labeled as "LipY-5'" in Figure; GenBank Accession No. AJ549519)
PacI/SphI (12840-12078)	756 bp 3' portion of *Yarrowia* Lipase 7-like gene (labeled as "LipY-3'" in Figure; GenBank Accession No. AJ549519)
*PmeI/Swa*I (4871-8320)	YAT1::EgDHAsyn1S::Lip1, comprising: • YAT1: *Yarrowia lipolytica* YAT1 promoter (U.S. Pat. Appl. Pub. No. 2006/0094102-A1); • EgDHAsyn1S: codon-optimized DHA synthase (SEQ ID NO:75), derived from *Euglena gracilis* (labeled as "EgDHAase" in Figure; U.S. Pat. Appl. Pub. No. 2008/0254191-A1); • Lip1: Lip1 terminator sequence from *Yarrowia Lip1* gene (GenBank Accession No. Z50020)
*Cla*I/ *Pme*I (2070-4871)	EXP1::EaD4S-1::Pex16, comprising: • EXP1: *Yarrowia lipolytica* export protein (EXP1) promoter (Intl. App. Pub. No. WO 2006/052870); • EaD4S-1: codon-optimized truncated Δ4 desaturase (SEQ ID NO:65), derived from *Euglena anabaena* (U.S. Pat. Appl. Pub. No. 2008/0254191-A1); • Pex16: Pex16 terminator sequence from *Yarrowia Pex16* gene (GenBank Accession No. U75433)
*SalI/EcoR*I (1620-1)	*Yarrowia* Ura3 gene (GenBank Accession No. AJ306421)
*EcoR*I/*Pme*I (1-12871)	FBAINm::E1594D4S::Pex16, comprising: • FBAINm: *Yarrowia lipolytica* FBAINm promoter (U.S. Patent 7,202,356); • E1594D4S: codon-optimized Δ4 desaturase (SEQ ID NO:70), derived from *Eutreptiella cf_gymnastica* CCMP1594 (U.S. Pat. Appl. Pub. No. 2009/0253188-A1) (labeled as "D4S-1594" in Figure); • Pex16: Pex16 terminator sequence from *Yarrowia Pex16* gene (GenBank Accession No. U75433)

The pZKLY-G20444 plasmid was digested with *Asc*I/*Sph*I, and then used for transformation of strain Y5018U1, according to the General Methods. The transformants were selected on MM plates. After 5 days growth at 30 °C, 96 transformants grown on the MM plates were picked and re-streaked onto fresh MM plates. Once grown, these strains were individually inoculated into 3 mL 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.

GC analyses showed that most of the selected 96 strains produced about 19% EPA, 22% DPA and 9% DHA of TFAs. Strain #3 produced 18.6% EPA, 22.8% DPA and 9.7% DHA; strain #9 produced 18.4% EPA, 21% DPA and 9.6% DHA; strain #27 produced 17.8% EPA, 20.6% DPA and 10% DHA; and strain #40 produced 18.8% EPA, 21.2% DPA and 9.6% DHA. These four strains were designated as Y5037, Y5038, Y5039 and Y5040, respectively.

Knockout of the lipase 7-like locus (GenBank Accession No, AJ549519) was not confirmed in these knocked out strains.

The final genotype of strains Y5037, Y5038, Y5039 and Y5040 with respect to wild type *Yarrowia lipolytica* ATCC #20362 was *SCP2- (YALI0E01298g), YALI0C18711g-, Pex10-, YALI0F24167g-, unknown 1-, unknown 3-, unknown 8-, unknown 9-, unknown10-, unknown 11-,* GPD::FmD12::Pex20, YAT1::FmD12::OCT, GPM/FBAIN::FmD12S::OCT, EXP1::FmD12S::Aco, YAT1::FmD12S::Lip2, YAT1::ME3S::Pex16, EXP1::ME3S::Pex20 (3 copies), GPAT::EgD9e::Lip2, EXP1::EgD9eS::Lip1, FBAINm::EgD9eS::Lip2, FBA::EgD9eS::Pex20, GPD::EgD9eS::Lip2, YAT1::EgD9eS::Lip2, YAT1::E389D9eS::OCT, FBAINm::EgD8M::Pex20, FBAIN::EgD8M::Lip1 (2 copies), EXP1::EgD8M::Pex16, GPDIN::EgD8M::Lip1, YAT1::EgD8M::Aco, FBAIN::EgD5::Aco, EXP1::EgD5S::Pex20, YAT1::EgD5S::Aco, EXP1::EgD5S::ACO, YAT1::RD5S::OCT, YAT1::PaD17S::Lip1, EXP1::PaD17::Pex16, FBAINm::PaD17::Aco, YAT1::YlCPT1::ACO, GPD::YlCPT1::ACO, FBAINm::EaC20ES::Pex20, YAT1::EgC20ES::Lip1, FBAINm::EgC20ES::Pex20, EXP1::EaD4S-1::Lip2, EXP1::EaD4S-1::Pex16, YAT1::EaD4S-1::Lip1, GPDIN::EaD4SB::Aco, EXP1::E1594D4S::Oct, FBAINm::E1594D4S::Pex16, GPDIN::EgD4S-1::Aco, YAT1::EgDHAsyn1S::Lip1.

<u>Generation Of Strain Y5037U (*Ura3-*)</u>

To disrupt the *Ura3* gene, construct pZKUM (FIG. 5A; SEQ ID NO:40; described in Table 15 of U.S. Pat. App. Pub. No. 2009-0093543-

A1) was used to integrate a *Ura3* mutant gene into the *Ura3* gene of strain Y5037, in a manner similar to that described for pZKUM transformation of strain Y5004 (*supra*). A total of 12 transformants were grown and identified to possess a *Ura-* phenotype.

5 GC analyses showed the presence of 12.1% EPA, 10.2% DPA and 3.3% DHA in pZKUM-transformant strain #4 and 12.4% EPA, 10.3% DPA and 3.5% DHA in pZKUM-transformant strain #11. These two strains were designated as strains Y5037U1 and Y5037U2, respectively (collectively, Y5037U).

10 EXAMPLE 3

Construction Of Various Expression Vectors Comprising Different LPLAT ORFs

The present example describes the construction of a series of vectors, each comprising a LPLAT ORF, suitable for expression in 15 *Yarrowia lipolytica.* LPLAT ORFs included the *Saccharomyces cerevisiae* Ale1, *Yarrowia lipolytica* Ale1, *Mortierella alpina* LPAAT1, *Yarrowia lipolytica* LPAAT1 and *Caenorhabditis elegans* LPCAT. Examples 4, 5 and 6 describe the results obtained following transformation of these vectors into *Yarrowia lipolytica.*

20 Origin Of LPLATs

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 12 summarizes publicly available LPLATs (i.e., ScAle1, ScLPAAT, MaLPAAT1 and CeLPCAT) and LPLAT orthologs 25 identified herein (i.e., YlAle1 and YlLPAAT1) that are utilized in the Examples, following codon-optimization of heterologous genes for expression in *Yarrowia lipolytica* (*infra*).

Table 12: LPLATs Functionally Characterized

LPLAT	Organism	ORF Designation	References	SEQ ID NO
Ale1	*Saccharomyces cerevisiae* *	ORF "YOR175C" or "ScAle1"	GenBank Accession No. NP_014818; U.S. Pat. Appl. Pub. No. 20080145867 (and	8, 9

			corresponding to Intl. App. Pub. No. WO 2008/076377); Intl. App. Pub. No. WO 2009/001315	
	Yarrowia lipolytica	"YALI0F19514p" or "YlAle1"	GenBank Accession No. XP_505624; Intl. App. Pub. No. WO 2009/001315	10, 11
LPAAT	*Saccharo-myces cerevisiae*	ORF "YDL052C" or "ScLPAAT"	GenBank Accession No. NP_010231	18
	Mortierella alpina	"MaLPAAT1"	U.S. Pat. Appl. Pub. No. 2006-0115881-A1; U.S. Pat. Appl. Pub. No. 2009-0325265-A1	14, 15
	Yarrowia lipolytica	"YALI0E18964g" or "YlLPAAT1"	GenBank Accession No. XP_504127; U.S. Patent 7,189,559	16, 17
LPCAT	*Caenor-habditis elegans**	"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)	1, 2

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

More specifically, the ScLPAAT (SEQ ID NO:18) and ScAle1 (SEQ
5 ID NO:9) 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)) using the Washington University in St. Louis School of Medicine BLAST 2.0 (WU-
10 BLAST; http://blast.wustl.edu). Based on analysis of the best hits, the Ale1 and LPAAT orthologs from *Yarrowia lipolytica* are identified herein as YlAle1 (SEQ ID NO:11) and YlLPAAT (SEQ ID NO:17), respectively. The identiy of YlAle1 and YlLPAAT1 as orthologs of ScAle1 and ScLPAAT, respectively, was further confirmed by doing a reciprocal BLAST, i.e.,
15 using SEQ ID NOs:11 and 17 as a query against the *Saccharomyces cerevisiae* public protein database to find ScAle1 and ScLPAAT, repectively, as the best hits.

The LPLAT proteins identified above as ScAle1 (SEQ ID NO:9), YlAle1 (SEQ ID NO:11), ScLPAAT (SEQ ID NO:18), MaLPAAT1 (SEQ ID

NO:15), YlLPAAT1 (SEQ ID NO:17) and CeLPCAT (SEQ ID NO:2) were aligned using the method of Clustal W (slow, accurate, Gonnet option; Thompson et al., *Nucleic Acids Res.*, 22:4673-4680 (1994)) of the MegAlign™ program (version 8.0.2) of the LASERGENE bioinformatics
5 computing suite (DNASTAR, Inc., Madison, WI). This resulted in creation of Table 13, where percent similarity is shown in the upper triangle of the Table while percent divergence is shown in the lower triangle.

Table 13: Percent Identity And Divergence Among Various LPLATs

YlLPAAT1	CeLPCAT	MaLPAAT1	ScAle1	ScLPAAT	YlAle1	
--	26.6	34.0	9.6	43.9	11.7	YlLPAAT1
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	--	YlAle1

10

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
15 34.0% identity (MaLPAAT1 and YlLPAAT1) to 43.9% identity (ScLPAAT and YlLPAAT1), while identity between the ScAle1 and YlAle1 polypeptides was 45%.

Membrane Bound *O*-Acyltransferase ["MBOAT"] Family Motifs: Orthologs of the ScAle1 protein sequence (SEQ ID NO:9) were identified
20 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:9) as the query sequence against fungal proteins in the "nr" protein database
25 (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:

5 Table 14: Fungal Orthologs Of ScAle1 (SEQ ID NO:9) Based On BLAST Analysis

Gen Bank Acession No.	Species
NP_014818.1	*Saccharomyces cerevisiae*
XP_001643411.1	*Vanderwaltozyma polyspora* DSM 70294
XP_448977.1	*Candida glabrata*
XP_455985.1	*Kluyveromyces lactis*
NP_986937.1	*Ashbya gossypii* ATCC 10895
XP_001385654.2	*Pichia stipitis* CBS 6054
XP_001487052.1	*Pichia guilliermondii* ATCC 6260
EDK36331.2	*Pichia guilliermondii* ATCC 6260
XP_001525914.1	*Lodderomyces elongisporus* NRRL YB-4239
XP_461358.1	*Debaryomyces hansenii* CBS767
XP_713184.1	*Candida albicans* SC5314
XP_001645053.1	*Vanderwaltozyma polyspora* DSM 70294
XP_505624.1	*Yarrowia lipolytica*
XP_001805526.1	*Phaeosphaeria nodorum* SN15
XP_001598340.1	*Sclerotinia sclerotiorum* 1980
XP_001907785.1	*Podospora anserine*
XP_001931658.1	*Pyrenophora tritici-repentis* Pt-1C-BFP
XP_001560657.1	*Botryotinia fuckeliana* B05.*10*
XP_963006.1	*Neurospora crassa* OR74A
XP_364011.2	*Magnaporthe grisea* 70-15
XP_001209647.1	*Aspergillus terreus* NIH2624
XP_001822945.1	*Aspergillus oryzae* RIB40
XP_001257694.1	*Neosartorya fischeri* NRRL 181
XP_747591.2	*Aspergillus fumigatus* Af293
XP_001270060.1	*Aspergillus clavatus* NRRL 1
NP_596779.1	*Schizosaccharomyces pombe*
XP_001396584.1	*Aspergillus niger*
XP_001229385.1	*Chaetomium globosum* CBS 148.51
XP_001248887.1	*Coccidioides immitis* RS
XP_664134.1	*Aspergillus nidulans* FGSC A4
XP_566668.1	*Cryptococcus neoformans var. neoformans* JEC21
XP_001839338.1	*Coprinopsis cinerea okayama* 7#130
XP_757554.1	*Ustilago maydis* 521

The yeast and fungal protein sequences of Table 14 were aligned using DNASTAR. Multiple sequence alignments and percent identity
10 calculations were performed using the Clustal W method of alignment (*supra*).

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. Pub. No. 2008-0145867-A1. Specifically, these include: M-[V/I]-[L/I]-xxK-[L/V/I]-xxxxxxDG (SEQ ID NO:26), RxKYYxxWxxx-[E/D]-[A/G]xxxxGxG-[F/Y]-xG (SEQ ID NO:27), $EX_{11}WNX_2$-[T/V]-X_2W (SEQ ID NO:28) and SAxWHGxxPGYxx-[T/F]-F (SEQ ID NO:29) , wherein X encodes any amino acid residue. The His residue in SEQ ID NO:29 has been reported to be a likely active site residue within the protein.

Only one motif, i.e., $EX_{11}WNX_2$-[T/V]-X_2W (SEQ ID NO:28), was completely conserved in all 33 of the organisms aligned. The remaining M-[V/I]-[L/I]-xxK-[L/V/I]-xxxxxxDG (SEQ ID NO:26), RxKYYxxWxxx-[E/D]-[A/G]xxxxGxG-[F/Y]-xG (SEQ ID NO:27) and SAxWHGxxPGYxx-[T/F]-F (SEQ ID NO:29) motifs were only partially conserved. Thus, these motifs were appropriately truncated to fit with 0 mismatch (i.e., SAxWHG [SEQ ID NO:5]), 1 mismatch (i.e., RxKYYxxW [SEQ ID NO:4]), or 2 mismatches (i.e., M(V/I)(L/I)xxK(LVI) [SEQ ID NO:3]) for the purposes of the present methodologies.

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 YlLPAAT1 (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 Ma LPAAT-like proteins disclosed in Intl. App. Pub. No. WO 2004/087902 (i.e., SEQ ID NOs:93 and 95).

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 YlLPAAT1 (SEQ ID NO:17) is
5 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.

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

The *Saccharomyces cerevisiae* ORF designated as “ScAle1” (SEQ ID NO:8) was optimized for expression in *Yarrowia lipolytica*, by DNA 2.0 (Menlo Park, CA). In addition to codon optimization, 5’ *Pci*1 and 3’ *Not*1
15 cloning sites were introduced within the synthetic gene (i.e., ScAle1S; SEQ ID NO:12). 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:13] is identical to that of the wildtype protein sequence [i.e., SEQ ID NO:9]).
20 ScAle1S was cloned into pJ201 (DNA 2.0) to result in pJ201:ScAle1S.

A 1863 bp *Pci*1/*Not*1 fragment comprising ScAle1S was excised from pJ201:ScAle1S and used to create pY201 (SEQ ID NO:77; Table 15; FIG. 10A). In addition to comprising a chimeric YAT1::ScAle1S::Lip1 gene, pY201 also contains a *Yarrowia lipolytica* URA3 selection marker
25 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 *Yarrowia lipolytica* Pox3 gene to facilitate integration by double homologous recombination, although
30 integration into *Yarrowia lipolytica* is known to usually occur without homologous recombination. Thus, construct pY201 thereby contained the following components:

Table 15: Description of Plasmid pY201 (SEQ ID NO:77)

RE Sites And Nucleotides Within SEQ ID NO:77	Description Of Fragment And Chimeric Gene Components
*Bsi*W1/*Sbf*1 (1-1706 bp)	LoxP::*Ura3*::LoxP, comprising: • LoxP sequence (SEQ ID NO:78) • *Yarrowia lipolytica* Ura3 gene (GenBank Accession No. AJ306421); • LoxP sequence (SEQ ID NO:78)
*Sbf*1/*Sph*1 (1706-3043 bp)	3' portion of *Yarrowia lipolytica* POX3 Acyl-CoA oxidase 3 (GenBank Accession No. YALI0D24750g) (i.e., bp 2215-3038 in pY201)
*Sph*1/*Asc*1 (3043-5743 bp)	• *ColE1* plasmid origin of replication; • Ampicillin-resistance gene (Amp^R) for selection in *E. coli* (i.e., bp 3598-4758 [complementary] in pY201); • *E. coli* f1 origin of replication
*Asc*I/*Bsi*WI (5743-6513 bp)	5' portion of *Yarrowia lipolytica* POX3 Acyl-CoA oxidase 3 (GenBank Accession No. YALI0D24750g) (i.e., bp 5743-6512 in pY201)
*Bsi*WI/ *Bsi*WI (6514-1 bp) [a *Not*1 site, located between ScAle1S and Lip1 is present at bp 9154 bp]	YAT1::ScAle1S::Lip1, comprising: • YAT1: *Yarrowia lipolytica* YAT1 promoter (U.S. Pat. Appl. Pub. No. 2006/0094102-A1) (i.e., bp 6514-7291 in pY201) • ScAle1S: codon-optimized Ale1 (SEQ ID NO:12) derived from *Saccharomyces cerevisiae* YOR175C (i.e., bp 7292-9151 in pY201; labeled as "Sc LPCATs ORF" in Figure); • Lip1: Lip1 terminator sequence from *Yarrowia Lip1* 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

The *Yarrowia lipolytica* ORF designated as "YlAle1" (GenBank
5 Accession No. XP_505624; SEQ ID NO:10) was amplified by PCR from
Yarrowia lipolytica ATCC #20362 cDNA library using PCR primers 798
and 799 (SEQ ID NOs:79 and 80, respectively). Additionally, the YAT
promoter was amplified by PCR primers 800 and 801 (SEQ ID NOs:81
and 82, respectively) from pY201 (SEQ ID NO:77). Since the primer pairs
10 were designed to create two PCR products having some overlap with one
another, a YAT1::YlAle1 fusion fragment was then amplified by
overlapping PCR using primers 798 and 801 (SEQ ID NOs:79 and 82,
respectively) and the two PCR fragments as template. The PCR was
carried out in a RoboCycler Gradient 40 PCR machine (Stratagene) using
15 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
5 was carried out, followed by reaction termination at 4 °C.

The PCR product comprising the YAT1::Yl Ale1 fusion fragment was gel purified and digested with *Cla*I/*Not*I. This *Cla*1-*Not*1 fragment was ligated into pY201 that had been similarly digested (thereby removing the YAT1::ScAle1S fragment) to create pY168 (SEQ ID NO:83), comprising a
10 chimeric YAT1::YlAle1::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:83) are identical to those present in pY201, with the exception of the YAT1::YlAle1::Lip1 gene in pY168, instead of the YAT1::ScAle1S::Lip1 gene in pY201 (FIG. 10A). Note that
15 YlAle1 is labeled as "Yl LPCAT" in FIG. 10B.

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

The *Mortierella alpina* ORF designated as "MaLPAAT1" (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*1 and 3' *Not*1
20 cloning sites were introduced within the synthetic gene (i.e., MaLPAAT1S; SEQ ID NO:21). 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:22] is identical to that of the wildtype protein sequence [i.e., SEQ ID NO:15]).
25 MaLPAAT1S was cloned into pJ201 (DNA 2.0) to result in pJ201:MaLPAAT1S.

A 945 bp *Pci*1/*Not*1 fragment comprising MaLPAAT1S was excised from pJ201:MaLPAAT1S and used to create pY208 (SEQ ID NO:84), in a 3-way ligation with two fragments of pY201 (SEQ ID NO:77). Specifically,
30 the MaLPAAT1 fragment was ligated with a 3530 bp *Sph*-*Not*I pY201 fragment and a 4248 bp *Nco*I-*Sph*I pY201 fragment to result in pY208. The components present in pY208 (FIG. 11A; SEQ ID NO:84) are identical to those present in pY201, with the exception of the

YAT1::MaLPAAT1S::Lip1 gene in pY208, instead of the YAT1::Sc Ale1S::Lip1 gene in pY201 (FIG. 10A).

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

A putative LPAAT1 from *Yarrowia lipolytica* (designated herein as "YlLPAAT1"; SEQ ID NO:17) 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".

The YlLPAAT1 ORF (SEQ ID NO:16) was amplified by PCR using *Yarrowia lipolytica* ATCC #20362 cDNA library as a template and PCR primers 856 and 857 (SEQ ID NOs:85 and 86, respectively). The PCR was conducted using the same components and conditions as described above for amplification of the YAT1::Yl Ale1 fusion fragment, prior to synthesis of pY168.

The PCR product comprising YlLPAAT1 ORF was digested with *Pci*I and *Not*I and then utilized in a 3-way ligation with two fragments from pY168. Specifically, the YlLPAAT1 fragment was ligated with a 3530 bp *Sph-Not*I pY168 fragment and a 4248 bp *Nco*I-*Sph*I pY168 fragment, to produce pY207, comprising a chimeric YAT1::YlLPAAT1::Lip1 gene. The *Y. lipolytica* LPAAT1 ORF was confirmed by DNA sequencing. The components present in pY207 (FIG. 11B; SEQ ID NO:87) are identical to those present in pY201, with the exception of the chimeric YAT1::Yl LPAAT1::Lip1 gene in pY207, instead of the YAT1::ScAle1S::Lip1 gene in pY201 (FIG. 10A). Note that YlLPAAT1 is labeled as "Yl LPAT1 ORF" in FIG. 11B.

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

The *Caenorhabditis elegans* ORF designated as "CeLPCAT" (SEQ ID NO:1) was optimized for expression in *Yarrowia lipolytica*, by GenScript Corporation (Piscataway, NJ). In addition to codon optimization, 5' *Nco*1 and 3' *Not*1 cloning sites were introduced within the synthetic gene (i.e., CeLPCATS; SEQ ID NO:6). 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:7] is identical to that of the wildtype protein sequence [i.e., SEQ ID NO:2]).

A *Nco*1-*Not*1 fragment comprising CeLPCATS was used to create pY175 (SEQ ID NO:88), in a 3-way ligation with two fragments from pY168 (SEQ ID NO:83). Specifically, the *Nco*1-*Not*1 fragment comprising CeLPCATS was ligated with a 3530 bp *Sph*-*Not*I pY168 fragment and a 4248 bp *Nco*I-*Sph*I pY168 fragment to result in pY175. The components present in pY175 (FIG. 12A; SEQ ID NO:88) are identical to those present in pY201, with the exception of the YAT1::CeLPCATS::Lip1 gene in pY175, instead of the YAT1::ScAle1S::Lip1 gene in pY201 (FIG. 10A). Note that CeLPCATS is labeled as “Ce.LPCATsyn” in FIG. 12A.

Construction Of pY153, Comprising A *Caenorhabditis elegans* LPCAT Gene

The *Nco*1-*Not*1 fragment comprising CeLPCATS, *supra, was* used to create pY153 (SEQ ID NO:89; FIG. 12B). In addition to comprising a chimeric FBAIN::CeLPCATS::3’ Yl LPAAT1 gene, pY153 also contains a *Yarrowia lipolytica* URA3 selection marker. Both the chimeric FBAIN::CeLPCATS::3’ Yl LPAAT1 gene and the URA3 gene were flanked by fragments having homology to 5’ and 3’ regions of the *Yarrowia lipolytica* LPAAT1 gene to facilitate integration by double homologous recombination, although integration into *Yarrowia lipolytica* is known to usually occur without homologous recombination. Thus, construct pY153 thereby contained the following components:

Table 16: Description of Plasmid pY153 (SEQ ID NO:89)

RE Sites And Nucleotides Within SEQ ID NO:89	Description Of Fragment And Chimeric Gene Components
*Cla*1/*Sap*1 (1-1398 bp)	5’ portion of *Yarrowia lipolytica* gene encoding LPAAT1 (GenBank Accession No. XP_504127) (i.e., bp 1-1112 [complementary] in pY153);
*Sap*1/*Xba*1 (1398-3993 bp)	Vector backbone including: • *ColE1* plasmid origin of replication (i.e., bp 1380-2260 in pY153); • Ampicillin-resistance gene (Amp^R) for selection in *E. coli* (i.e., bp 2330-3190 [complementary] in

	pY153); • *E. coli* f1 origin of replication (i.e., bp 3370-3770 in pY153)
*Xba*1/ *Pme*1 (3993-6719 bp) [a *Nco*1 site, located between CeLPCATS and FBAIN is present at bp 5756; a *Not*1 site, located between CeLPCATS and YlLPAAT1 is present at bp 4904]	FBAIN::CeLPCATS::3' Yl LPAAT1, comprising: • FBAINm: *Yarrowia lipolytica* FBAIN promoter (U.S. Patent 7,202,356) (i.e., bp 5756-6719 [complementary] in pY153); • CeLPCATS: codon-optimized LPCAT (SEQ ID NO:6) derived from *Caenorhabditis elegans* T06E8.1 (GenBank Accession No. CAA98276) (i.e., bp 4910-5758 [complementary] in pY153; labeled as "Ce.LPCATsyn" in Figure); • 3' Yl LPAAT1: 3' untranslated portion of *Yarrowia lipolytica* gene encoding LPAAT1 (GenBank Accession No. XP_504127) (i.e., bp 3987-4905 [complementary] in pY153)
*Pme*1-*Cla*I (6719-1 bp)	*Yarrowia lipolytica* URA3 gene (GenBank Accession No. AJ306421) (i.e., bp 6729-1 [complementary] in pY153)

EXAMPLE 4

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

5 *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
10 chromosome. This was in spite of the host containing its native LPLATs, i.e., Ale1 and LPAAT1.

Transformation And Growth

Yarrowia lipolytica strain Y8406U (Example 1) was individually transformed with linear *Sph*I-*Asc*I fragments of the integrating vectors
15 described in Example 3, wherein each LPLAT was under the control of the *Yarrowia* YAT promoter. Specifically, vectors pY201 (YAT1::ScAle1S::Lip1), pY168 (YAT1::YlAle1::Lip1), pY208 (YAT1::MaLPAAT1S::Lip1), pY207 (YAT1::YlLPAAT1::Lip1) and pY175 (YAT1::CeLPCATS::Lip1) were transformed according to the General
20 Methods.

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
5 Yeast Extract, 6 g KH_2PO_4, 2 g K_2HPO_4, 1.5 g $MgSO_4 \cdot 7H_20$, 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%
10 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
15 prepared by trans-esterification, and subsequently analyzed with a Hewlett-Packard 6890 GC (General Methods).

Based on the fatty acid composition of the 3 mL cultures, selected transformants were further characterized by flask assay. Specifically, clones #5 and #11 of strain Y8406U transformed with expression vector
20 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 YlAle1) was selected and designated as "Y8406U::YlAle1"; clone #8 of strain Y8406U transformed with expression vector pY208 (comprising
25 MaLPAAT1S) was selected and designated as "Y8406U::MaLPAAT1S"; clone #21 of strain Y8406U transformed with expression vector pY207 (comprising YlLPAAT1) was selected and designated as "Y8406U::YlLPAAT1"; and clone #23 of strain Y8406U transformed with expression vector pY175 (comprising CeLPCATS) was selected and
30 designated as "Y8406U::CeLPCATS". Additionally, strain Y8406 (a Ura+ strain that was parent to strain Y8406U (Ura-)) was used as a control.

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
5 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.

For DCW determination, 10 mL culture was harvested by
10 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
15 weight of the cells was determined.

<u>Lipid Content, Fatty Acid Composition And Conversion Efficiencies</u>

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
20 strain Y8406U::YlAle1. Experiment 3 compared control strain Y8406 versus strain Y8406U::YlAle1, strain Y8406U::ScAle1S-11, and strain Y8406U::MaLPAAT1S. Experiment 4 compared control strain Y8406 versus strain Y8406U::MaLPAAT1S, strain Y8406U::YlLPAAT1 and strain Y8406U::CeLPCATS.

25 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 17 below summarizes the total lipid
30 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.

Table 18 summarizes the conversion efficiency of each desaturase and the Δ9 elongase functioning in the PUFA biosynthetic pathway and
5 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
10 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(s)/(product(s)+substrate)*100, where product includes both product and product derivatives.

Table 17: Lipid Content And Composition In LPCAT Transformant Strains Of *Yarrowia lipolytica* Y8406

Expt.	Strain	Replicates	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:: ScAle1S-5	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
		% Ctrl	104	111	100	106	89	67	24	118	168	160	95	115	116	121
2	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
	Y8406U:: YlAle1	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
		% Ctrl	96	109	117	88	61	57	15	115	181	160	100	120	122	117
3	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:: YlAle1	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
		% Ctrl	89	122	100	86	55	59	20	100	152	129	0	na	122	109
	Y8406U:: ScAle1S-11	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
		% Ctrl	107	111	100	107	85	64	24	109	165	129	0	na	115	123
	Y8406U:: MaLPAAT1S	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
		% Ctrl	91	163	114	150	102	55	28	84	161	129	0	na	112	102
4	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:: MaLPAAT1S	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
		% Ctrl	90	171	133	155	111	56	28	86	150	117	0	115	112	101
	Y8406U:: YlLPAAT1	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
		% Ctrl	90	132	117	135	100	63	41	79	130	100	0	110	115	104
	Y8406U:: CeLPCATS	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
		% Ctrl	89	136	100	148	115	67	55	79	130	100	57	110	109	96

Table 18: Desaturase And Elongase 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::YlAle1	AVG.3	96	85	94	91	90
		% Ctrl	103	121	101	98	98
3	Y8406	1	93	72	93	96	89
	Y8406U::YlAle1	AVG.2	96	85	96	92	89
		% Ctrl	104	119	103	96	100
	Y8406U::ScAle1S-11	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
4	Y8406	AVG.2	93	71	94	93	91
	Y8406U::MaLPAAT1S	AVG.2	92	84	96	91	90
		% Ctrl	99	118	102	99	100
	Y8406U::YlLPAAT1	AVG.2	93	82	96	92	92
		% Ctrl	100	115	103	100	101
	Y8406U::CeLPCATS	AVG.2	92	80	96	92	91
		% Ctrl	99	113	102	99	100

Based on the data concerning Experiments 1, 2 and 3 in Table 17 and Table 18, overexpression of LPLAT in EPA strains Y8406U::ScAle1S-5, Y8406U::ScAle1S-11, Y8406U::YlAle1 and Y8406U::MaLPAAT1S 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 Δ9 elongase. Compared to Y8406U::ScAle1S-5 and Y8406U::ScAle1S-11, Y8406U::YlAle1 has lower LA % TFAs, higher EPA % TFAs, better Δ9 elongation conversion efficiency, and slightly lower TFAs % DCW and EPA % DCW. Y8406U::Yl Ale1 and Y8406U::MaLPAAT1S are similar except overexpression of MaLPAAT1S resulted in lower LA % TFAs, EPA % TFAs, and EPA % DCW.

Experiment 4 shows that overexpression of LPLAT in EPA strains Y8406U::YlLPAAT1, Y8406U::MaLPAAT1S 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 Δ9 elongase. Compared to Y8406U::CeLPCATS, Y8406U::YlLPAAT1 and Y8406U::MaLPAAT1S both have lower LA % TFAs, higher EPA % TFAs, higher EPA % DCW, and slightly better TFAs % DCW. Y8406U::YlLPAAT1 and Y8406U::MaLPAAT1S are similar except overexpression of MaLPAAT1S results in lower LA % TFAs, slightly lower EPA % TFAs and EPA % DCW, and slightly better Δ9 elongase conversion efficiency.

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, ScAle1S, YlAle1, MaLPAAT1S and YlLPAAT1 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 ScAle1S, YlAle1, MaLPAAT1S, and YlLPAAT1 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 ScAle1S, YlAle1, MaLPAAT1S, and YlLPAAT1 significantly improved the Δ9 elongase conversion efficiency in strains of *Yarrowia* producing EPA but did not improve the desaturations (measured as Δ12 desaturase conversion efficiency, Δ8 desaturase conversion efficiency, Δ5 desaturase conversion efficiency or Δ17 desaturase conversion efficiency).

CeLPCAT was previously shown to improve Δ6 elongation conversion efficiency in *Saccharomyces cerevisiae* fed LA or GLA (Intl. 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 Δ9 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 Intl. App. Pub. No. WO 2004/076617.

Futhermore, expression of ScAle1S, YlAle1, MaLPAAT1S, YlLPAAT1 and CeLPCATS did not significantly alter either the level of PUFAs accumulated or the total lipid content in strains of *Yarrowia* producing EPA.

Previous studies have shown that both Δ6 elongation and Δ9 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 Δ9 elongase conversion efficiency resulting from over-expression of LPLATs, demonstrated above, it is anticipated that the LPLATs described herein and their orthologs, such as Sc LPAAT, will also improve Δ6 elongation conversion efficiency.

EXAMPLE 5

Functional Characterization Of Different LPLATs In DHA-Producing *Y. lipolytica* Strain Y5037

Yarrowia lipolytica strain Y5037U, producing DHA, was used to functionally characterize the effects of overexpression of the *Saccharomyces*

cerevisiae Ale1, *Mortierella alpina* 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

Yarrowia lipolytica strain Y5037U (Example 2) was individually transformed with linear *Sph*I-*Asc*I fragments of the integrating vectors described in Example 3, wherein ScAle1S and MaLPAAT1S were under the control of the *Yarrowia* YAT promoter, while CeLPCATS was under the control of the *Yarrowia* FBAIN promoter. Specifically, vectors pY201 (YAT1::ScAle1S::Lip1), pY208 (YAT1::MaLPAAT1S::Lip1) and pY153 (FBAIN::CeLPCATS::YlLPAAT1) were transformed according to the General Methods.

Each transformation mix was plated on MM agar plates. Selected transformants were further characterized, as detailed below. More specifically, clone #7 of strain Y5037U, transformed with expression vector pY153 (comprising CeLPCATS) was selected and designated as "Y5037U::FBAIN-CeLPCATS"; clone #18 of strain Y5037U, transformed with expression vector pY201 (comprising ScAle1S) was selected and designated as "Y5037U::ScAle1S"; and clone #6 of strain Y5037U, transformed with expression vector pY208 (comprising MaLPAAT1S) was selected and designated as "Y5037U::MaLPAAT1S". Additionally, strain Y5037 (a Ura+ strain that was parent to strain Y5037 (Ura-)) was used as a control.

A total of four separate experiments were conducted in 3 mL culture based on variable culturing conditions and strains, to examine the effect of LPLAT overexpression on lipid content, fatty acid composition and conversion efficiencies. Experiment 1 compared control strain Y5037 versus strains Y5037U::FBAIN-CeLPCATS and Y5037U::ScAleIS after 2 days of growth in MM medium on a shaker at 200 rpm and 30 °C, followed by 3 days of incubation in 3 mL HGM medium. MM medium (Cat. No. CML-MM, Biomyx Technology), pH 6.1, contains per L: 1.7 g yeast nitrogen base ["YNB"]

without amino acids and NH_4SO_4, 1 g proline, 0.1 g adenine, 0.1 g lysine, and 20 g glucose.

Experiment 2 compared control strain Y5037 versus strain Y5037U::ScAleIS after 2 days of growth in CSM-U medium on a shaker at 200 rpm and 30 ^{o}C, followed by 3 days of incubation in 3 mL HGM medium. CSM-U medium (Cat. No C8140, Teknova Inc,, Hollister, CA) contains: 0.13% amino acid dropout powder minus uracil, 0.17% yeast nitrogen base, 0.5% $(NH_4)_2SO_4$, and 2.0% glucose.

Experiment 3 compared control strain Y5037 versus strains Y5037U::FBAIN-CeLPCATS and Y5037U::ScAleIS after 2 days of growth in MM medium on a shaker at 200 rpm and 30 ^{o}C, followed by 5 days of incubation in 3 mL HGM medium.

Experiment 5 compared control strain Y5037 versus strain Y5037U::MaLPAAT1S after 2 days of growth in FM medium on a shaker at 200 rpm and 30 ^{o}C, followed by 3 days of incubation in 3 mL HGM medium. The composition of FM medium is described in Example 4.

Following growth for 3 days (Experiments 1, 2, and 5) or 5 days (Experiment 3) in HGM, 1 mL aliquots of the cultures were harvested by centrifugation and analyzed by GC, as described in Example 4.

Experiment 4 compared control strain Y5037 versus strains Y5037U::FBAIN-CeLPCATS and Y5037U::ScAleIS after 2 days of growth in 25 mL FM medium followed by 5 days of incubation in HGM medium as described above. Specifically, one loop of freshly streaked cells from MM agar plates was inoculated into 3 mL FM medium and grown overnight at 250 rpm and 30 ^{o}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 ^{o}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 ^{o}C, a 1 mL aliquot was used for GC analysis and 10 mL dried for dry cell weight ["DCW"] determination (*supra,* Example 4).

Lipid Content, Fatty Acid Composition And Conversion Efficiencies

In each experiment, the lipid content and fatty acid composition are quantified for 1, 2, 3 or 4 replicate cultures ["Replicates"] of the control Y5037 strain and the transformant Y5037U strain(s). Additionally, data for each Y5037U transformant is presented as a % of the Y5037 control. Table 19 below summarizes the concentration of each fatty acid as a weight percent of TFAs ["% TFAs"]. 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, DPA, DHA and EDD (corresponding to the sum of EPA plus DPA plus DHA). Additionally, the ratio of DHA % TFAs/ DPA % TFAs is provided.

Table 20 summarizes the total DCW (mg/mL), the total lipid content of cells ["TFAs % DCW"], and the conversion efficiency of each desaturase and elongase functioning in the PUFA biosynthetic pathway and which are required for DHA 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"], Δ4 desaturase conversion efficiency ["Δ4 CE"], Δ9 elongation conversion efficiency ["Δ9e CE"] and Δ5 elongation conversion efficiency ["Δ5e CE"] are provided for each control Y5037 strain and the transformant Y5037U strain(s); data for each Y5037U transformant is presented as a % of the Y5037 control. Conversion efficiency was calculated according to the formula: product(s)/(product(s)+substrate)*100, where product includes both product and product derivatives.

Table 19: Lipid Content and Composition In LPCAT Transformant Strains Of *Yarrowia lipolytica* Y5037

Expt.	Strain	Replicates	% TFAs															
			16:0	16:1	18:0	18:1	18:2	ALA	EDA	DGLA	ARA	ETrA	ETA	EPA	DPA	DHA	EDD	DHA / DPA
1	Y5037	AVG.4	4.1	1.1	3.2	5.4	21.8	0.5	2.9	1.7	0.7	1.3	1.8	18.1	20.6	6.5	45.2	0.3
	Y5037U:: FBAIN-CeLPCATS	1	5.2	1.2	2.7	8.5	11.1	0.3	2.5	3.6	1.1	1.4	2.7	31.7	9.6	11.0	52.4	1.1
		% Ctrl	127	109	84	157	51	60	86	212	157	108	150	175	47	169	116	367
	Y5037U:: ScAleIS	1	4.4	1.4	2.0	4.4	15.7	0.5	3.5	2.7	1.0	1.2	2.2	22.0	16.8	14.4	53.3	0.9
		% Ctrl	107	127	63	81	72	100	121	159	143	92	122	122	82	222	118	300
2	Y5037	AVG.2	4.4	1.1	3.9	5.4	21.8	0.5	3.4	1.6	0.8	1.1	1.7	17.0	21.0	6.7	44.7	0.3
	Y5037U:: ScAleIS	1	4.5	1.5	2.5	4.7	16.6	0.4	4.1	2.6	1.1	1.1	2.1	21.2	17.1	13.2	51.5	0.8
		% Ctrl	102	136	64	87	76	80	121	163	138	100	124	125	81	197	115	267
3	Y5037	AVG.3	3.9	1.1	1.6	4.7	20.7	0.5	3.3	1.8	1.3	1.5	3.9	19.3	20.8	7.9	47.9	0.4
	Y5037U:: FBAIN-CeLPCATS	1	5.8	1.1	2.6	8.0	10.0	0.3	3.0	3.6	1.9	2.2	2.7	31.0	9.9	11.8	52.7	1.2
		% Ctrl	149	100	163	170	48	60	91	200	146	147	69	161	48	149	110	300
	Y5037U:: ScAleIS	1	4.6	1.3	1.8	5.9	18.1	0.3	4.4	2.4	1.3	1.8	4.0	22.1	15.1	11.7	48.9	0.8
		% Ctrl	118	118	113	126	87	60	133	133	100	120	103	115	73	148	102	200
5	Y5037	1	5.1	1.3	1.6	4.7	22.5	2.7	3.9	1.9	1.4	1.3	1.7	20.4	20.7	8.9	50.1	0.4
	Y5037U:: MaLPAT1	1	6.1	1.5	1.8	4.5	21.1	2.2	4.0	2.1	1.5	1.2	1.7	23.4	19.5	10.7	53.7	0.6
		% Ctrl	120	115	113	96	94	81	103	111	107	92	100	115	94	120	107	150
4	Y5037	AVG.3	3.9	1.2	1.3	5.9	22.4	3.9	1.7	1.8	0.8	1.0	1.6	20.0	26.2	6.7	52.9	0.3
	Y5037U::	AVG.3	6.1	1.3	3.4	8.8	10.1	0.7	1.6	3.5	0.7	1.3	2.3	33.9	12.5	10.6	57.0	0.9

	FBAIN-CeLPCATS	% Ctrl	156	108	262	149	45	18	94	194	88	130	144	170	48	158	108	300
	Y5037U:: ScAleIS	AVG.3	5.4	1.4	2.7	8.7	21.1	1.7	5.4	2.5	0.6	1.2	1.4	20.4	19.6	7.3	47.3	0.4
		% Ctrl	138	117	208	147	94	44	318	139	75	120	88	102	75	109	89	133

Table 20: Desaturase And Elongase Conversion Efficiency In LPCAT Transformant Strains Of *Yarrowia lipolytica* Y5037

Expt.	Strain	Replicates	DCW mg/mL	TFA % DCW	Δ12 CE	Δ9e CE	Δ8 CE	Δ5 CE	Δ17 CE	Δ5e CE	Δ4 CE
1	Y5037	AVG.4	nd	nd	93	71	92	93	90	60	24
	Y5037U:: FBAIN-CeLPCATS	1	nd	nd	90	85	94	89	89	39	53
		% Ctrl	nd	nd	96	120	102	96	98	66	221
	Y5037U:: ScAleIS	1	nd	nd	95	80	93	92	89	59	46
		% Ctrl	nd	nd	102	113	101	99	98	98	191
2	Y5037	AVG.2	nd	nd	93	71	91	93	89	62	24
	Y5037U:: ScAleIS	1	nd	nd	94	79	92	92	88	59	44
		% Ctrl	nd	nd	101	111	100	99	98	95	180
3	Y5037	AVG.3	nd	nd	94	74	92	90	89	60	27
	Y5037U:: FBAIN-CeLPCATS	1	nd	nd	91	86	92	90	87	41	54
		% Ctrl	nd	nd	96	117	100	100	97	69	198
	Y5037U:: ScAleIS	1	nd	nd	93	77	90	89	87	55	44
		% Ctrl	nd	nd	99	105	98	99	97	92	160
5	Y5037	1	nd	nd	95	70	91	93	88	59	30
	Y5037U:: MaLPAT1	1	nd	nd	95	73	92	93	88	56	36
		% Ctrl	nd	nd	100	104	101	100	100	95	118
4	Y5037	AVG.3	3.7	19.7	nd	69	95	94	93	62	20
	Y5037U:: FBAIN-CeLPCATS	AVG.3	3.0	14.0	nd	86	96	91	91	40	46
		% Ctrl	82	71	nd	124	100	97	98	65	226
	Y5037U:: ScAleIS	AVG.3	3.7	31.6	nd	72	89	92	85	57	27
		% Ctrl	101	157	nd	104	93	98	92	92	133

Based on the data in Table 19 and Table 20, overexpression of LPLAT in DHA strains Y5037U::CeLPCATS, Y5037U::ScAleIS and Y5037U::MaLPAAT1S results in reduction of the concentration of LA as a weight % of TFAs ["LA % TFAs"], an increase in the concentration of EPA as a weight % of TFAs ["EPA % TFAs"], an increase in the concentration of DHA as a weight % of TFAs ["DHA % TFAs"], an increase in the concentration of EPA + DPA + DHA as a weight % of TFAs ["EDD % TFAs"] (with the exception of strain Y5037U::ScAleIS in Experiment 4), an increase in the ratio of DHA % TFAs to DPA % TFAs ["DHA/DPA"], an increase in the conversion efficiency of the Δ9 elongase and an increase in the conversion efficiency of the Δ4 desaturase.

More specifically, depending on the culture conditions, CeLPCATS overexpression in Y5037U::CeLPCATS can reduce LA % TFAs to 45%, increase EPA % TFAs to 175%, increase DHA % TFAs to 169%, increase Δ9 elongation CE to 124%, and increase Δ4 desaturation CE to 226%, as compared to the control. Similarly, depending on the culture conditions, ScAle1S overexpression in Y5037U::ScAleIS can reduce LA % TFAs to 72%, increase EPA % TFAs to 125%, increase DHA % TFAs to 222%, increase Δ9 elongation CE to 113%, and increase Δ4 desaturation CE to 191%, as compared to the control. Finally, overexpression of MaLPAAT1 in Y5037U::MaLPAAT1S can reduce LA % TFAs to 94%, increase EPA % TFAs to 115%, increase DHA % TFAs to 120%, increase Δ9 elongation CE to 104%, and increase Δ4 desaturation CE to 118%, as compared to the control.

Although Y5037U::CeLPCATS possessed a significantly lower total lipid content ["TFAs % DCW"] in Experiment 4, the total lipid content was significantly improved in strain Y5037U::ScAleIS. This increase in lipid content is a likely explanation for the lower EDD % TFAs in strain Y5037U::ScAleIS.

DHA biosynthesis via EPA involves two steps: elongation of EPA to DPA by $C_{20/22}$ elongase (also known as either a "C20" elongase or a Δ5 elongase) and desaturation of DPA to DHA by Δ4 desaturase. An important

bottleneck in the production of DHA from EPA has been the Δ4 desaturation step, evident by the build up of DPA, although the mechanistic details for this limitation were unknown. The results above show that expression of ScAle1S, YlAle1, YlLPAAT1, MaLPAAT1S, and CeLPCATS proteins significantly improved Δ4 desaturation. Thus, Δ4 desaturation was not limiting because of Δ4 desaturase activity *per se*. Instead, Δ4 desaturation was limiting because of limited availability of the DPA substrate at the *sn*-2 position of phospholipids. The results showed unexpectedly that (unlike other desaturation substrates), limited DPA incorporation into phospholipid can be overcome by overexpression of Ale1, LPAAT and LPCAT proteins.

Previously, Intl. App. Pub. No. WO 2004/076617 showed that expression of CeLPCAT (SEQ ID NO:2) in *Saccharomyces cerevisiae* improved Δ6 elongation of exogenously provided GLA to DGLA. It hypothesized that CeLPCAT removed an acyl chain from the *sn*-2 position of phospholipids, thereby making the removed acyl group available for elongation in the CoA pool. It was shown in the present studies that the expression of the codon-optimized CeLPCATS, under control of the YAT1 promoter, in strains of *Yarrowia lipolytica* engineered to produce high levels of EPA (Example 4) and DHA (Example 5), respectively, improves Δ9 elongation of endogenously produced LA to EDA. However, expression of CeLPCATS in DHA-producing strain Y5037U::CeLPCATS unexpectedly did not result in improved Δ5 elongation of EPA to DPA. In contrast, expression of CeLPCATS in DHA-producing strain Y5037U::CeLPCATS very significantly improved Δ4 desaturation of DPA to DHA (*supra*). This is especially unexpected since desaturations occur mainly at the *sn*-2 position of phospholipids and elongation occurs in the CoA pool.

Based on the improved Δ4 desaturation conversion efficiency resulting from over-expression of LPLATs, demonstrated above, it is anticipated that the LPLATs described herein and their orthologs, such as ScLPAAT, will also improve Δ4 desaturation conversion efficiency.

EXAMPLE 6

Functional Characterization Of Different LPLATs In ARA-Producing *Y. lipolytica* Strain Y8006U

Yarrowia lipolytica strain Y8006U, producing ARA, is used to
5 functionally characterize the effects of overexpression of the *Saccharomyces cerevisiae* Ale1, *Mortierella alpina* LPAAT1 and *Caenorhabditis elegans* LPCAT, following their 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

10 *Yarrowia lipolytica* strain Y8006U (Example 1) will be individually transformed with linear *Sph*I-*Asc*I fragments of the integrating vectors described in Example 3, in a manner comparable to that utilized in Example 4. URA+ transformants will be selected, grown for 2 days in FM medium and 5 days in HGM medium and then 1 mL aliquots of the cultures will be
15 harvested by centrifugation and analyzed by GC (Example 4). Based on the fatty acid composition of the 3 mL cultures, selected transformants will be further characterized using strain Y8006 (a Ura+ strain that was parent to strain Y8006U (Ura-)) as a control.

Each selected transformant and the control will be re-grown in FM and
20 HGM medium, as described in Example 4, and then subjected to GC analysis and DCW determination.

The lipid content, fatty acid composition and ARA as a percent of the DCW will be quantified for the control Y8006 strain and the transformant Y8006U strain(s). Additionally, data for each Y8006U transformant will be
25 determined as a % of the Y8006 control. The conversion efficiency of each desaturase and the Δ9 elongase functioning in the PUFA biosynthetic pathway and which is required for ARA production will also be determined and compared to the control, in a manner similar to that in Examples 4 and 5.

It is hypothesized that overexpression of the ScAle1S, YlAle1,
30 MaLPAAT1S, YlLPAAT1 and CeLPCATS LPLATs in the ARA strains will result in a reduction of the concentration of LA (18:2) as a weight % of TFAs

["LA % TFAs"], an increase in the concentration of ARA as a weight % of TFAs ["ARA % TFAs"], and an increase in the conversion efficiency of the Δ9 elongase.

EXAMPLE 7

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

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 10 included the *Saccharomyces cerevisiae* LPAAT encoding SEQ ID NO:18 and the *Yarrowia lipolytica* LPAAT1 encoding SEQ ID NO:17. Example 8 describes the results obtained following transformation of these vectors into *Y. lipolytica.*

Construction Of pY222, Comprising A Codon-Optimized *Saccharomyces* 15 *cerevisiae* LPAAT Gene

The *Saccharomyces cerevisiae* ORF designated as "ScLPAAT" (SEQ ID NO:18) was optimized for expression in *Yarrowia lipolytica*, by DNA 2.0 (Menlo Park, CA). In addition to codon optimization, 5' *Pci*1 and 3' *Not*1 cloning sites were introduced within the synthetic gene (i.e., ScLPAATS; SEQ 20 ID NO:96). 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:97] is identical to that of the wildtype protein sequence [i.e., SEQ ID NO:18]). ScLPAATS was cloned into pJ201 (DNA 2.0) to result in pJ201:ScLPAATS.

25 A 926 bp *Pci*1/*Not*1 fragment comprising ScLPAATS was excised from pJ201:ScLPAATS and cloned into *Nco*I-*Not*1 cut pYAT-DG2-1 to create pY222 (SEQ ID NO:100; Table 21; FIG. 13A). Thus, pY222 contained the following components:

Table 21: Description of Plasmid pY222 (SEQ ID NO:100)

RE Sites And Nucleotides Within SEQ ID NO:100	Description Of Fragment And Chimeric Gene Components
*Sal*1/*Swa*I (1-2032)	YAT1::ScLPAATS::Lip1, comprising: • YAT1: *Yarrowia lipolytica* YAT1 promoter (U.S. Pat. Appl. Pub. No. 2006/0094102-A1); • ScLPAATS: codon-optimized ScLPAATS (SEQ ID NO:96) (labeled as "Sc LPAATs ORF" in Figure); • Lip1: Lip1 terminator sequence from *Yarrowia Lip1* gene (GenBank Accession No. Z50020) (labeled as "Lip1-3'" in Figure)
*Swa*I/*Ava*I (2032-4946)	• *ColE1* plasmid origin of replication; • Ampicillin-resistance gene (Amp^R) for selection in *E. coli*; • *E. coli* f1 origin of replication
*Ava*I-*Sph*I (4946-6330)	*Yarrowia lipolytica* centromere and autonomously replicating sequence ["ARS"] 18 locus
*Sph*I-*Sal*I (6330-1)	*Yarrowia lipolytica* URA3 gene (GenBank Accession No. AJ306421)

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

The *Yarrowia lipolytica* centromere and autonomously replicating
5 sequence ["ARS"] was amplified by standard PCR using primer 869 (SEQ ID NO:98) and primer 870 (SEQ ID NO:99), with plasmid pYAT-DG2-1 as template. The PCR product was digested with *Asc*I/*Avr*II and cloned into *Asc*I-*Avr*II digested pY207 (SEQ ID NO:87; Example 3) to create pY177 (SEQ ID NO:101; Table 22; FIG. 13B). Thus, the components present in
10 pY177 are identical to those in pY207 (FIG. 11B), except for the replacement of the 373 bp pY207 sequence between *Asc*I and *Avr*II with the 1341 bp sequence containg ARS. More specifically, pY177 contained the following components:

Table 22: Description of Plasmid pY177 (SEQ ID NO:101)

RE Sites And Nucleotides Within SEQ ID NO:101	Description Of Fragment And Chimeric Gene Components
*Bsi*W1/*Sbf*1 (1-1706 bp)	LoxP::*Ura3*::LoxP, comprising: • LoxP sequence (SEQ ID NO:78) • *Yarrowia lipolytica* Ura3 gene (GenBank Accession No. AJ306421); • LoxP sequence (SEQ ID NO:78)
*Sbf*1/*Sph*1 (1706-3043 bp)	3' portion of *Yarrowia lipolytica* POX3 Acyl-CoA oxidase 3 (GenBank Accession No. YALI0D24750g)
*Sph*I/*Asc*I (3043-5743 bp)	• *ColE1* plasmid origin of replication; • Ampicillin-resistance gene (Amp^R) for selection in *E. coli;* • *E. coli* f1 origin of replication
*Asc*I/*Bsi*WI (5743-6513 bp)	5' portion of *Yarrowia lipolytica* POX3 Acyl-CoA oxidase 3 (GenBank Accession No. YALI0D24750g)
*Asc*I/*Avr*II (5743-7084 bp)	*Yarrowia lipolytica* centromere and autonomously replicating sequence ["ARS"] 18 locus
*Avr*II/*Bsi*WI (7084-7481 bp)	5' portion of *Yarrowia lipolytica* POX3 Acyl-CoA oxidase 3 (GenBank Accession No. YALI0D24750g)
*Bsi*WI/ *Bsi*WI (7481-1 bp)	YAT1::YlLPAAT1::Lip1, comprising: • YAT1: *Yarrowia lipolytica* YAT1 promoter (U.S. Pat. Appl. Pub. No. 2006/0094102-A1); • YlLPAAT1: *Yarrowia lipolytica* LPAAT1 ("YALI0E18964g"; GenBank Accession No. XP_504127) (SEQ ID NO:16) (labeled as "Yl LPAT1 ORF" in Figure); • Lip1: Lip1 terminator sequence from *Yarrowia Lip1* gene (GenBank Accession No. Z50020) (labeled as "Lip1-3'" in Figure)

EXAMPLE 8

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

Yarrowia lipolytica strain Y8406U, producing EPA, was used to functionally characterize the effects of expression of the *Saccharomyces cerevisiae* LPAATS (SEQ ID NO:96) and *Yarrowia lipolytica* LPAAT1 (SEQ 10 ID NO:16) without integration on self-replicating plasmids. This was in spite of the host containing its native LPAATs.

Transformation And Growth

Yarrowia lipolytica strain Y8406U (Example 1) was individually transformed with uncut plasmids from Example 7. Specifically, vectors pY177 (YAT1::YlLPAAT1::Lip1) [SEQ ID NO:101] and pY222 (YAT1::ScLPAATS::Lip1) [SEQ ID NO:100] were transformed according to
5 the General Methods.

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
10 0.13% amino acid dropout powder minus uracil, 0.17% yeast nitrogen base, 0.5% $(NH_4)_2SO_4$, 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
15 analyzed by GC, as described in Example 4.

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
20 “Y8406U::ScLPAATS-5” and “Y8406U::ScLPAATS-6”, respectively; clone #1 of strain Y8406U transformed with expression vector pY177 (comprising YlLPAAT1) was selected and designated as “Y8406U::YlLPAAT1”. Additionally, strain Y8406 (a Ura+ strain that was parent to strain Y8406U (Ura-)) was used as a control.

25 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
30 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 4.

Lipid Content, Fatty Acid Composition And Conversion Efficiencies

5 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 23 below summarizes the total lipid content of cells ["TFAs % DCW"], the 10 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.

15 Table 24 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 4.

Table 23: Lipid Content And Composition In ScLPAATS and YlLPAAT1 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:: YlLPAAT1	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 24: Desaturase And Elongase Conversion Efficiency In ScLPAATS and YlLPAAT1 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::YlLPAAT1	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

Based on the data in Table 23 and Table 24 above, overexpression of both ScLPAATS and YlLPAAT1 in EPA strains Y8406U::YlLPAAT1, 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 Δ9 elongase. ScLPAATS and YlLPAAT1 have a similar effect on lipid profile.

The results obtained above were then compared to those obtained in Example 4, although different means were utilized to characterize the LPLATs. Specifically, in Example 4, 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.

In Example 8, the functional characterization of YlLPAAT1 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.

These differences described above can contribute to differences in lipid profile and content, as illustrated by the expression of YlLPAAT1 in Examples 4 and 8. The change over control in LA % TFAs, EPA % TFAs, and Δ9 elongase conversion efficiency were 63%, 115%, and 115%, respectively, upon expression of YlLPAAT in Example 4, whereas the change over control in LA % TFAs, EPA % TFAs, and Δ9 elongase conversion efficiency were were 76%, 101%, and 107%, respectively, upon expression of YlLPAAT in Example 8. Thus, the improvements in Δ9 elongation and LC-PUFA biosynthesis in Example 8 are minimized when compared to those

observed in Example 4. These differences can be attributed to the “position effects” of chromosomal integration and/or different growth conditions.

Since the improvements in LC-PUFA biosynthesis (measured as reduction in LA % TFAs, increase in EPA % TFAs and increase in Δ9
5 elongase conversion efficiency) are similar for both ScLPAATS and YlLPAAT 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 Examples 4 and 5,
10 when YlLPAAT1 was stably integrated into the host chromosome.

15

CLAIMS

What is claimed is:

1. A recombinant oleaginous microbial host cell for the improved
5 production of at least one long-chain polyunsaturated fatty acid, said host cell comprising 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:

a) a polypeptide having at least 45% amino acid identity, based on
10 the Clustal W method of alignment, when compared to an amino acid sequence selected from the group consisting of SEQ ID NO:9 and SEQ ID NO:11;

b) a polypeptide having at least one membrane bound *O*-acyltransferase protein family motif selected from the group
15 consisting of: SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:28;

c) 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:2;

20 d) 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:15, SEQ ID NO:17 and SEQ ID NO:18; and,

e) a polypeptide having at least one 1-acyl-*sn*-glycerol-3-phosphate
25 acyltransferase family motif selected from the group consisting of: SEQ ID NO:19 and SEQ ID NO:20;

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
30 sequence being the same or different; and,

further wherein the host cell has at least one improvement selected from the group consisting of:

(i) an increase in C_{18} to C_{20} elongation conversion efficiency in at least one long-chain polyunsaturated fatty acid-producing oleaginous microbial host cell when compared to a control host cell;

(ii) an increase in Δ4 desaturation conversion efficiency in at least one long-chain polyunsaturated fatty acid-producing oleaginous microbial host cell when compared to a control host cell.

2. The recombinant host cell of claim 1 wherein the at least one long-chain polyunsaturated fatty acid is selected from the group consisting of: eicosadienoic acid, dihomo-γ–linolenic acid, arachidonic acid, eicosatrienoic acid, eicosatetraenoic acid, eicosapentaenoic acid, docosatetraenoic acid, ω-6 docosapentaenoic acid, ω-3 docosapentaenoic acid and docosahexaenoic acid.

3. The recombinant host cell of claim 1 wherein the polynucleotide encoding a polypeptide having at least acyl-CoA:lysophospholipid acyltransferase activity is stably integrated; and,

further wherein the host cell has at least one improvement selected from the group consisting of:

a) an increase in C_{18} to C_{20} elongation conversion efficiency of at least 4% in at least one long-chain polyunsaturated fatty acid-producing oleaginous microbial host cell when compared to a control host cell; and,

b) an increase in Δ4 desaturation conversion efficiency of at least 5% in at least one long-chain polyunsaturated fatty acid-producing oleaginous microbial host cell when compared to a control host cell.

4. The recombinant host cell of claim 3 wherein the improvement is selected from the group consisting of:

a) an increase in C_{18} to C_{20} elongation conversion efficiency of at least 13% in an eicosapentaenoic acid-producing host cell when compared to a control host cell;

b) an increase in C_{18} to C_{20} elongation conversion efficiency of at least 4% in a docosahexaenoic acid-producing host cell when compared to a control host cell;

c) an increase in Δ4 desaturation conversion efficiency of at least 18% in a docosahexaenoic acid-producing host cell when compared to a control host cell;

d) an increase of at least 9 weight percent of eicosapentaenoic acid in an eicosapentaenoic acid-producing host cell measured as a weight percent of the total fatty acids when compared to a control host cell;

e) an increase at least 2 weight percent of eicosapentaenoic acid in an docosahexaenoic acid-producing host cell measured as a weight percent of the total fatty acids when compared to a control host cell; and,

f) an increase of at least 9 weight percent of docosahexaenoic acid in a docosahexaenoic acid-producing host cell measured as a weight percent of the total fatty acids when compared to a control host cell.

5. The recombinant host cell of any of claims 1-4 wherein the microbe is yeast.

6. The recombinant host cell of claim 5 wherein the yeast is *Yarrowia lipolytica*.

7. Oil comprising eicosapentaenoic acid and/or docosahexaenoic acid obtained from the oleaginous microbial recombinant host cell of claim 4.

8. A method for making an oil comprising eicosapentaenoic acid
5 and/or docosahexaenoic acid comprising:

a) culturing the oleaginous microbial host cell of claim 4 wherein an oil comprising eicosapentaenoic acid and/or docosahexaenoic acid is produced; and,

b) optionally recovering the microbial oil of step (a).

10

9. The method of claim 8 wherein the recovered oil of step (b) is further processed.

10. The method of claim 8 wherein the host cell is oleaginous yeast.

15

11. The method of claim 10 wherein the oleaginous yeast is *Yarrowia lipolytica*.

12. A method for increasing C_{18} to C_{20} elongation conversion
20 efficiency in a long-chain polyunsaturated fatty acid-producing oleaginous microbial recombinant host cell, comprising:

a) introducing into said long-chain polyunsaturated fatty acid-producing recombinant host cell at least one isolated polynucleotide encoding a polypeptide having at least acyl-CoA:lysophospholipid acyltransferase
25 activity wherein the polypeptide is selected from the group consisting of:

(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:9 and SEQ ID NO:11;

30 (ii) a polypeptide having at least one membrane bound *O*-acyltransferase protein family motif selected from the group

consisting of: SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:28;

(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:2;

(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:15, SEQ ID NO:17 and SEQ ID NO:18; and,

(v) a polypeptide having at least one 1-acyl-*sn*-glycerol-3-phosphate acyltransferase protein family motif selected from the group consisting of: SEQ ID NO:19 and SEQ ID NO:20;

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,

b) growing the oleaginous microbial host cell;

wherein the C_{18} to C_{20} elongation conversion efficiency of the oleaginous microbial host cell is increased relative to the control host cell.

13. The method of claim 12 wherein:

a) the polynucleotide encoding a polypeptide having at least acyl-CoA:lysophospholipid acyltransferase activity is stably integrated; and,

b) the increase in C_{18} to C_{20} elongation conversion efficiency is at least 13% in an eicosapentaenoic acid-producing host cell when compared to the control host cell.

14. The method of claim 12 wherein:

a) the polynucleotide encoding a polypeptide having at least acyl-CoA:lysophospholipid acyltransferase activity is stably integrated; and,

b) the increase in C_{18} to C_{20} elongation conversion efficiency is at least 4% in a docosahexaenoic acid-producing host cell when compared to the control host cell.

15. The method of claim 12 wherein the host cell is oleaginous yeast.

16. The method of claim 15 wherein the oleaginous yeast is *Yarrowia lipolytica*.

17. A method for increasing Δ4 desaturation conversion efficiency in a long-chain polyunsaturated fatty acid-producing oleaginous microbial recombinant host cell, comprising:

a) introducing into said long-chain polyunsaturated fatty acid-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:

(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:9 and SEQ ID NO:11;

(ii) a polypeptide having at least one membrane bound *O*-acyltransferase protein family motif selected from the group consisting of: SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 and SEQ ID NO:28;

(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:2;

(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:15, SEQ ID NO:17 and SEQ ID NO:18; 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: SEQ ID NO:19 and SEQ ID NO:20;

wherein the at least one isolated polynucleotide encoding a polypeptide having at least acyl-CoA:lysophospholipid acyltransferase activity 10 is operably linked to at least one regulatory sequence, said regulatory sequence being the same or different; and,

b) growing the oleaginous microbial host cell;

wherein the Δ4 desaturation conversion efficiency of the oleaginous microbial host cell is increased relative to a control host cell.

15

18. The method of claim 17 wherein:

a) the polynucleotide encoding a polypeptide having at least acyl-CoA:lysophospholipid acyltransferase activity is stably integrated; and,

20 b) the increase in Δ4 desaturation conversion efficiency is at least 18% when compared to the control host cell.

19. The method of claim 17 wherein the host cell is oleaginous yeast.

25 20. The method of claim 19 wherein the oleaginous yeast is *Yarrowia lipolytica*.

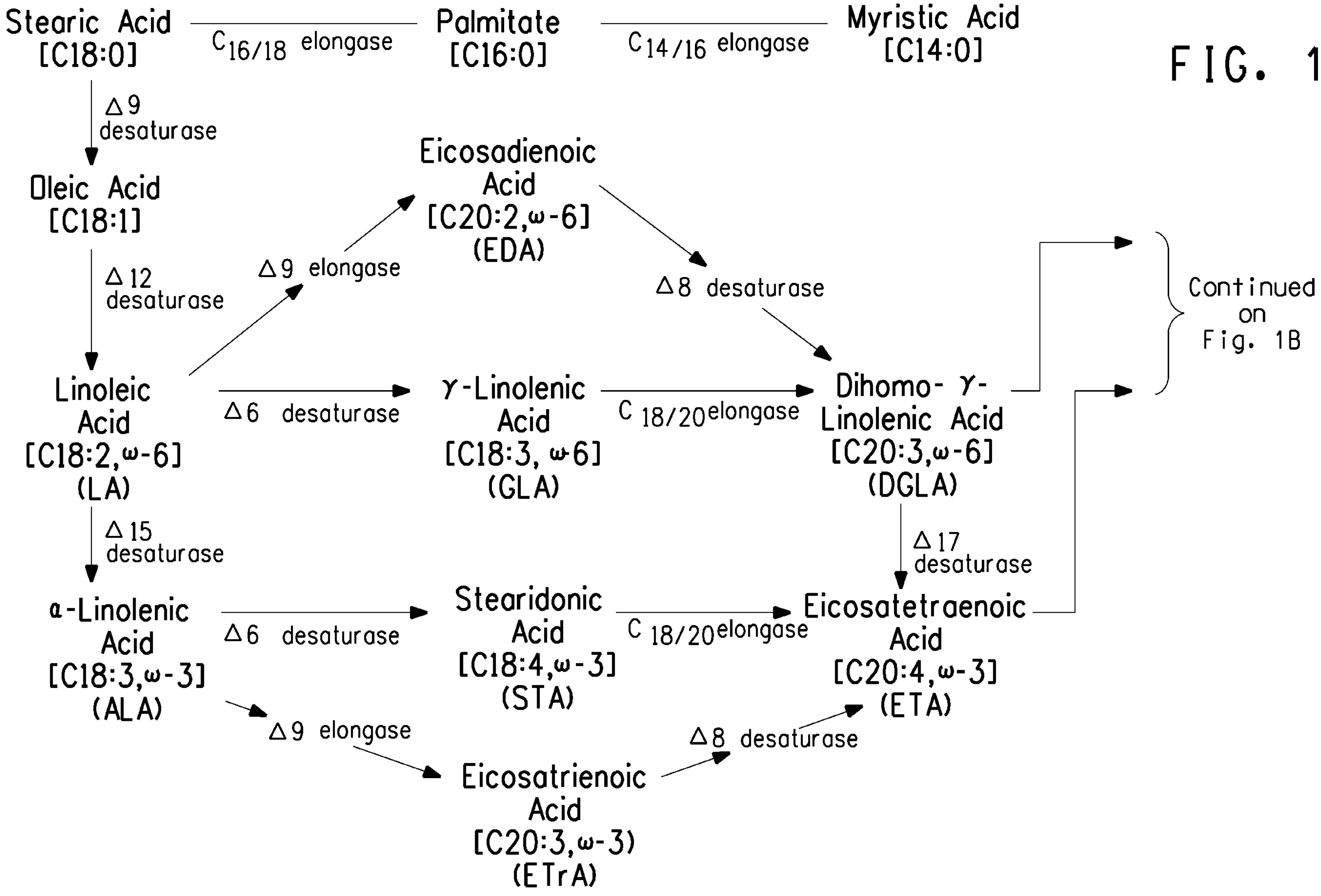


FIG. 1A

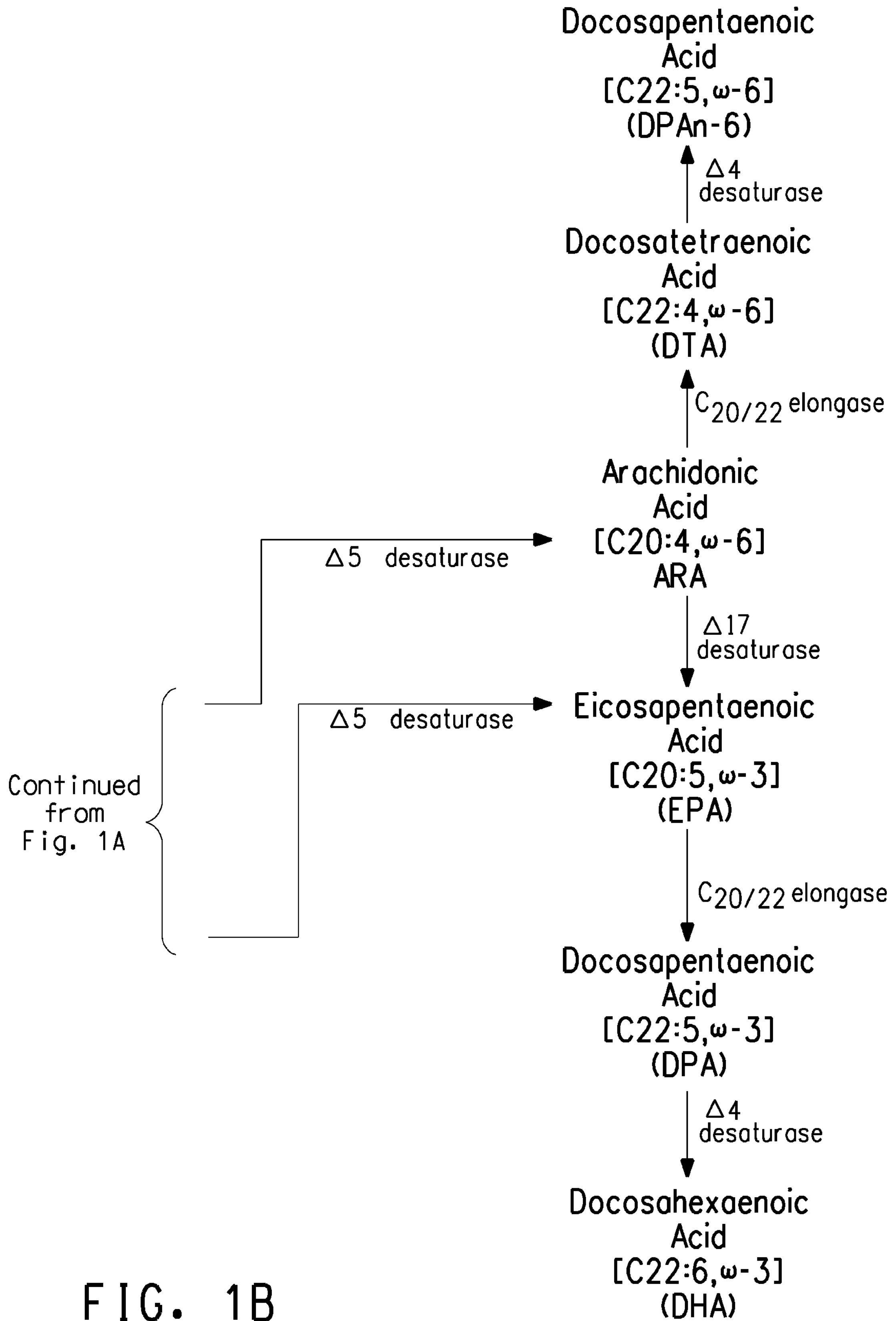


FIG. 1B

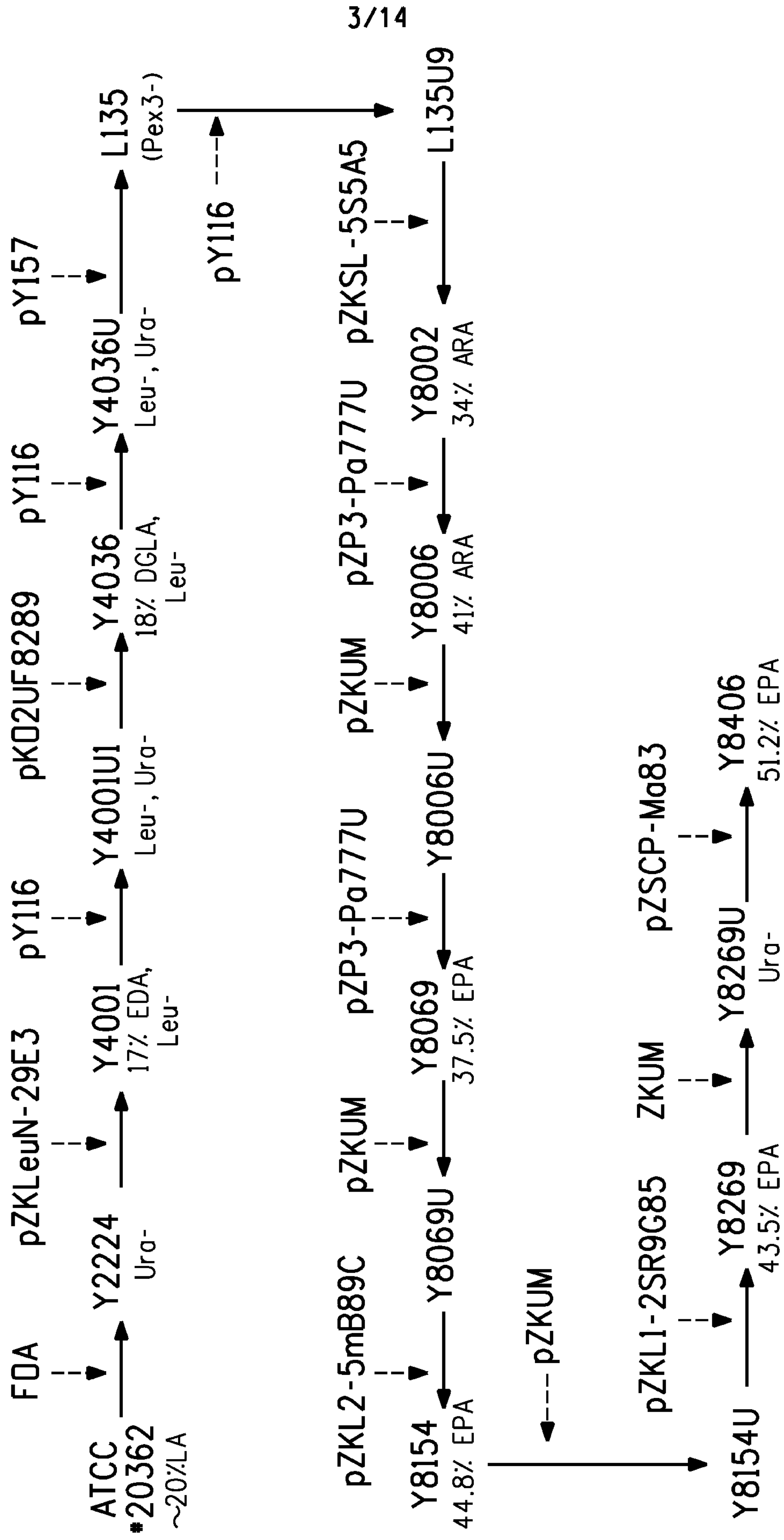


FIG. 2

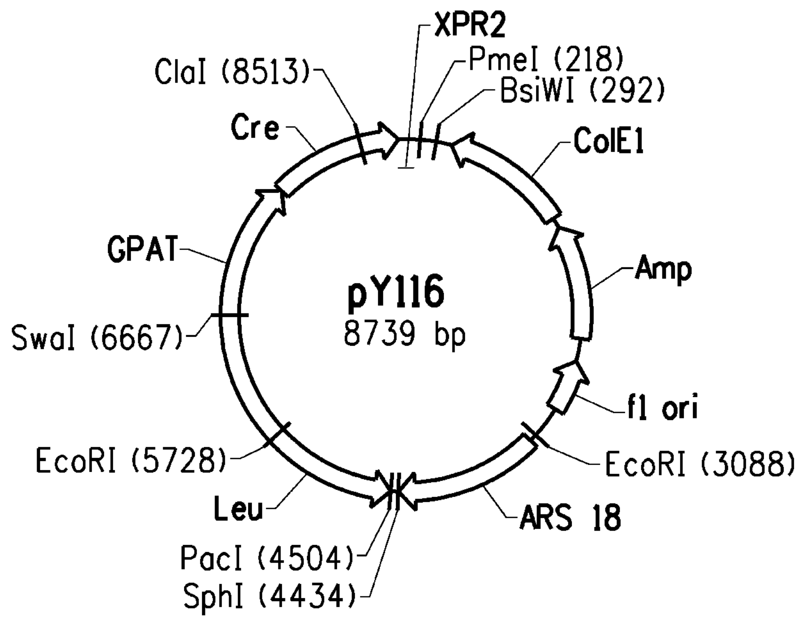


FIG. 3

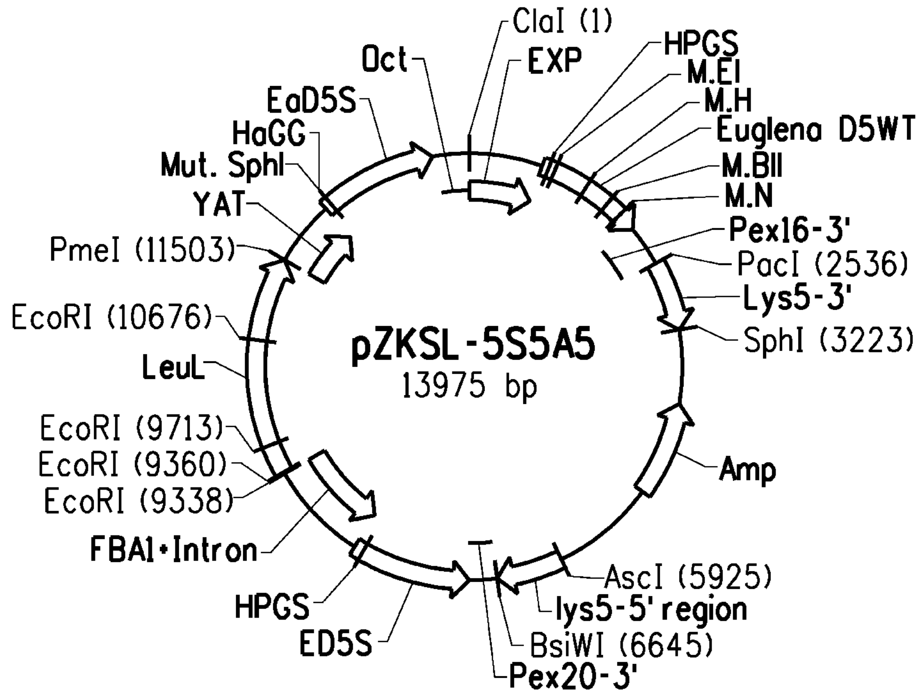


FIG. 4A

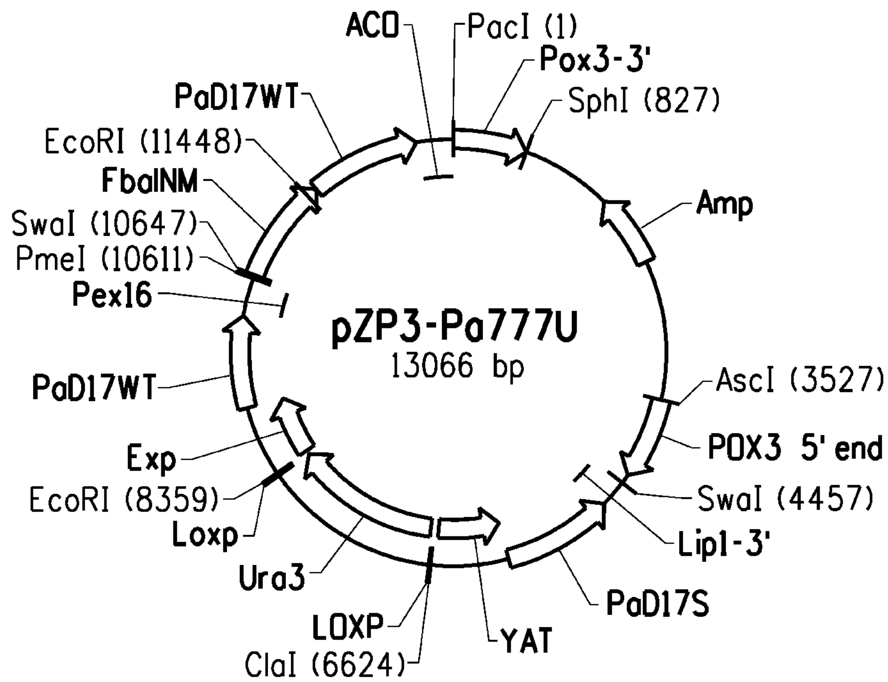


FIG. 4B

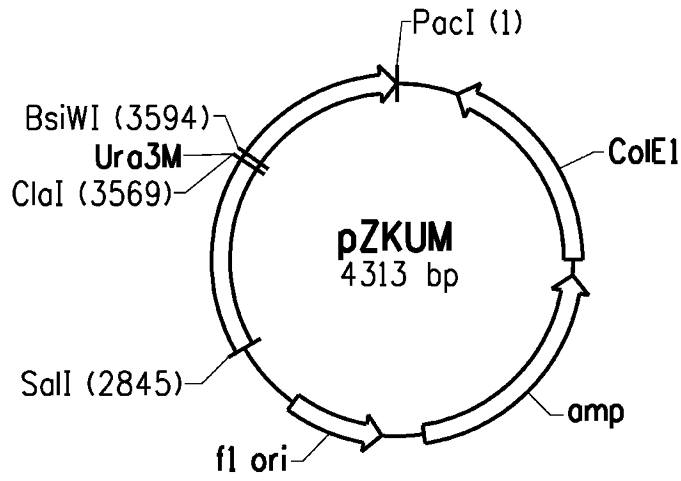


FIG. 5A

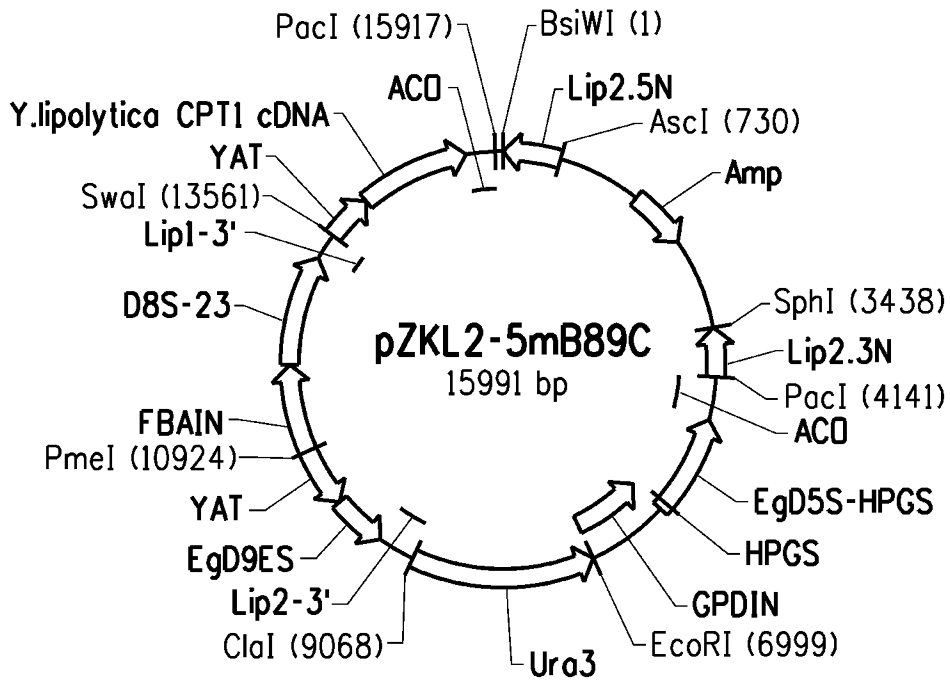


FIG. 5B

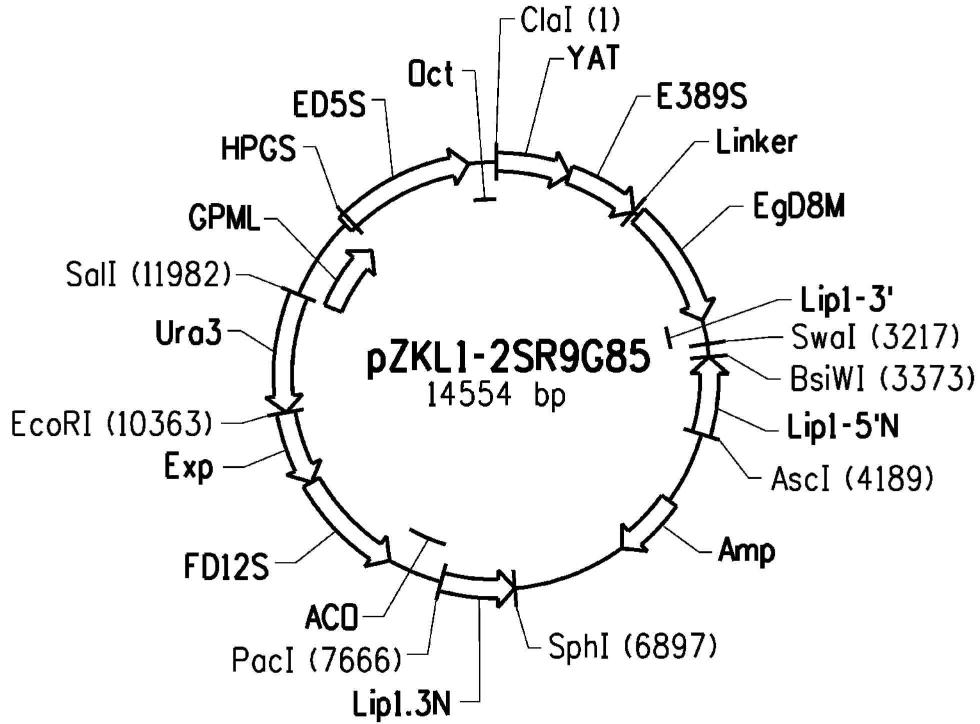


FIG. 6A

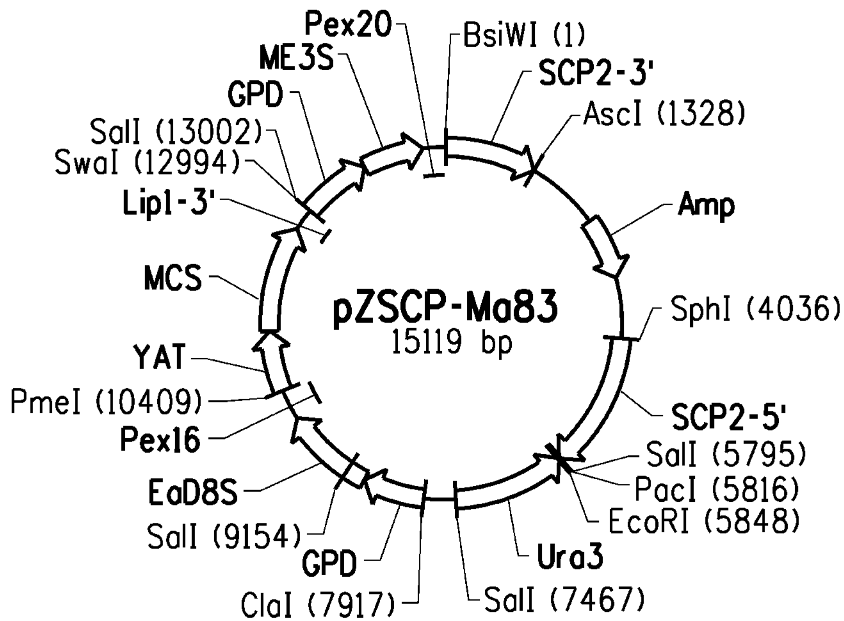


FIG. 6B

ATCC
*20362
~20%LA
FOA
Y2224
Ura-
Chr. E
pZKLeuN-29E3
Y4001
17% EDA,
Leu-, Chr. C
pY116
Y4001U1
Leu-, Ura-
pKO2UF8289
Y4036
18% DGLA,
Leu-
pY116
Y4036U
Leu-, Ura-
pZKSL-555R
Y4069,
Y4070
12% ARA,
Ura-
pZP3-Pa777U
Y4086
14% EPA
pY117
Y4086U1
Ura-
pZP2-2988
Y4128
37.6% EPA
pZKUE3S
Y4128U3
Ura-
pZKL2-5U89GC
Y4217
42% EPA,
49% Ω3
pZKUE3S
Y4217U2
pZKL1-2SP98C
Y4259
46.5% EPA,
52% Ω3
pZKUM
Y4259U2
Ura-
pZKD2-5U89A2
Y4305
53.2% EPA,
58% Ω3
pZKUM
Y4305U3
Ura-
pZKL4-220EA41B
Y5004
pZKUM
Y5004U1
Ura-
pZKL3-4GER44
Y5018
pZKUM
Y5018U1
Ura-
pZKLY-G20444
Y5037

FIG. 7

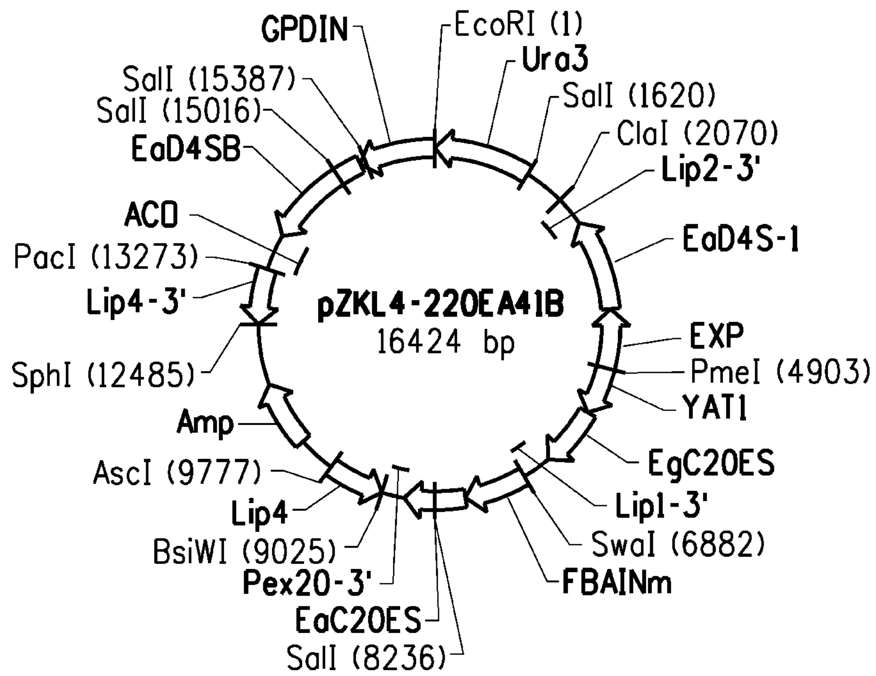


FIG. 8A

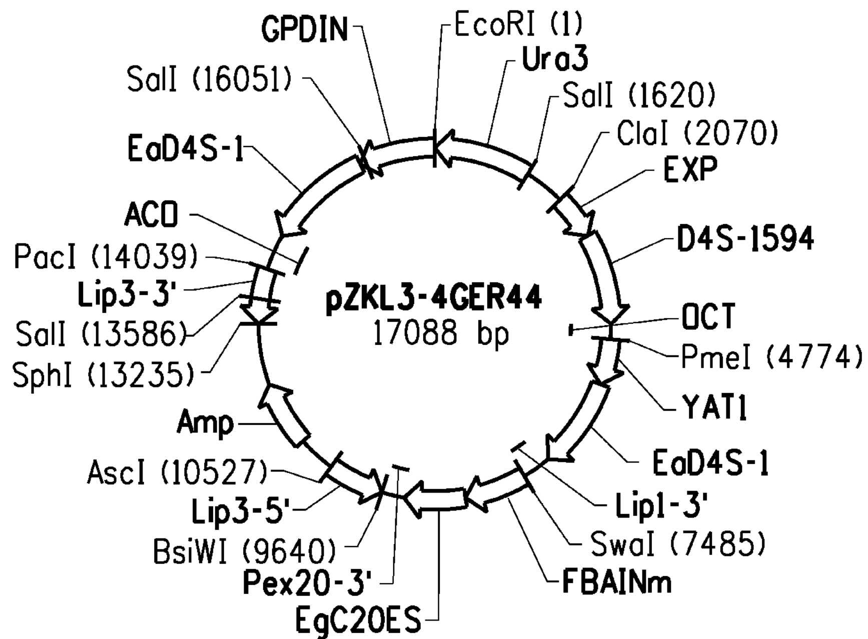


FIG. 8B

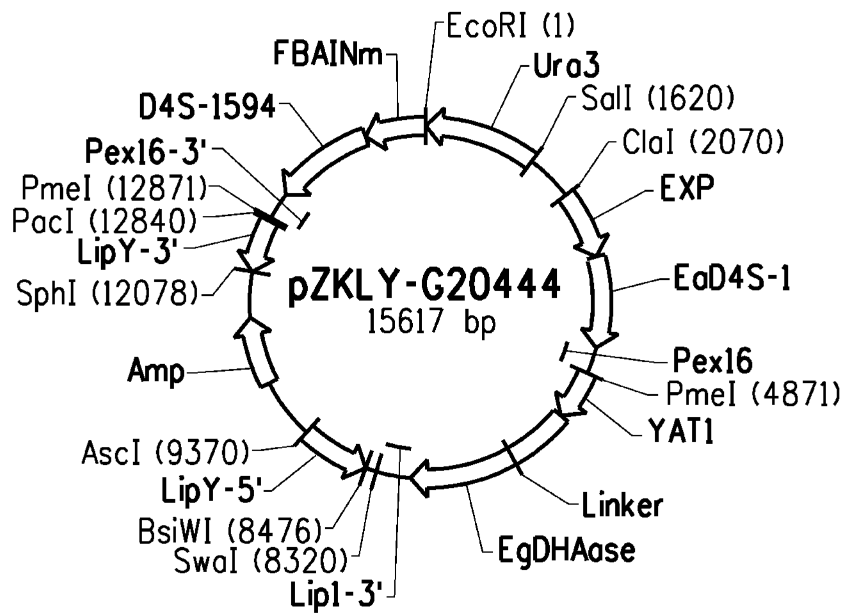


FIG. 9

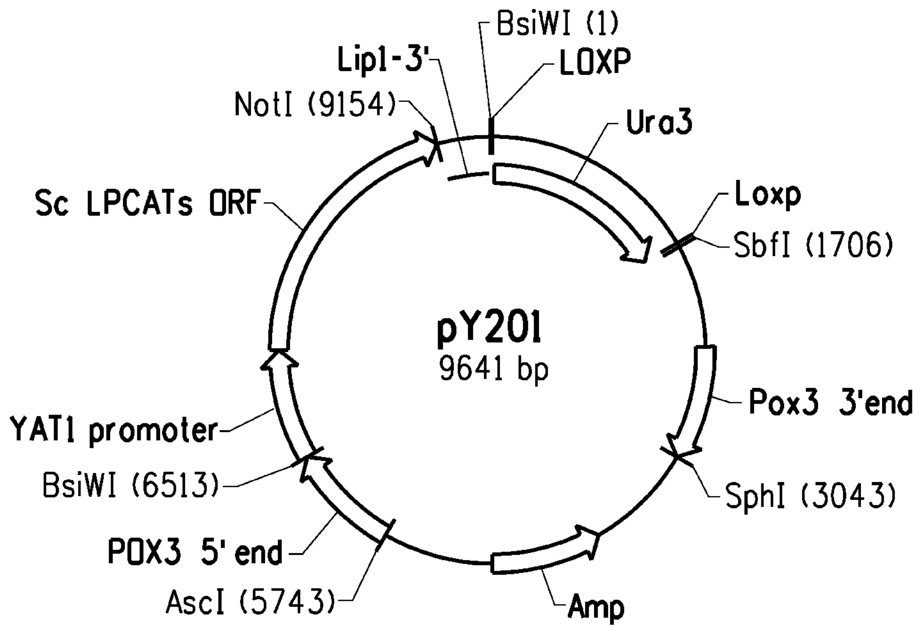


FIG. 10A

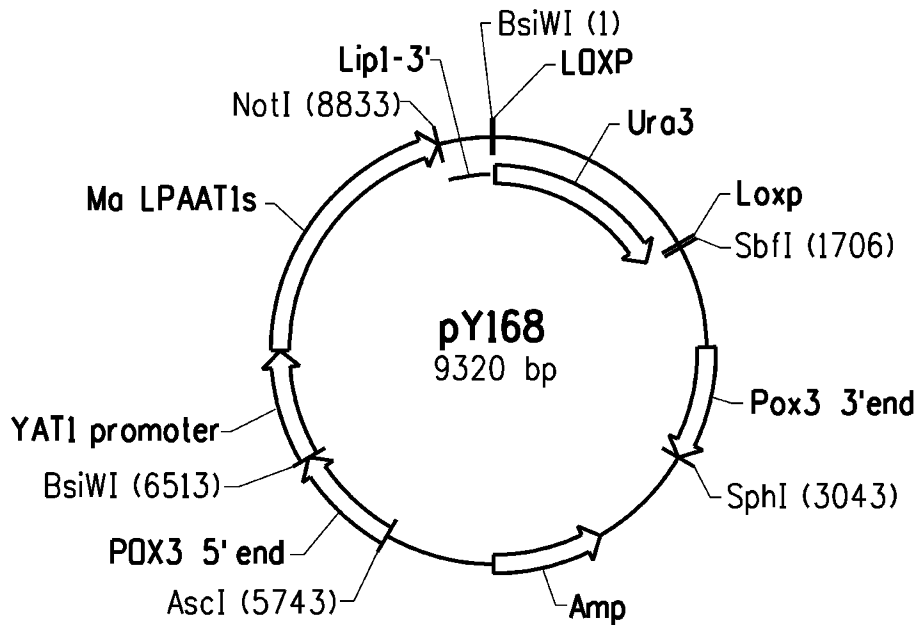


FIG. 10B

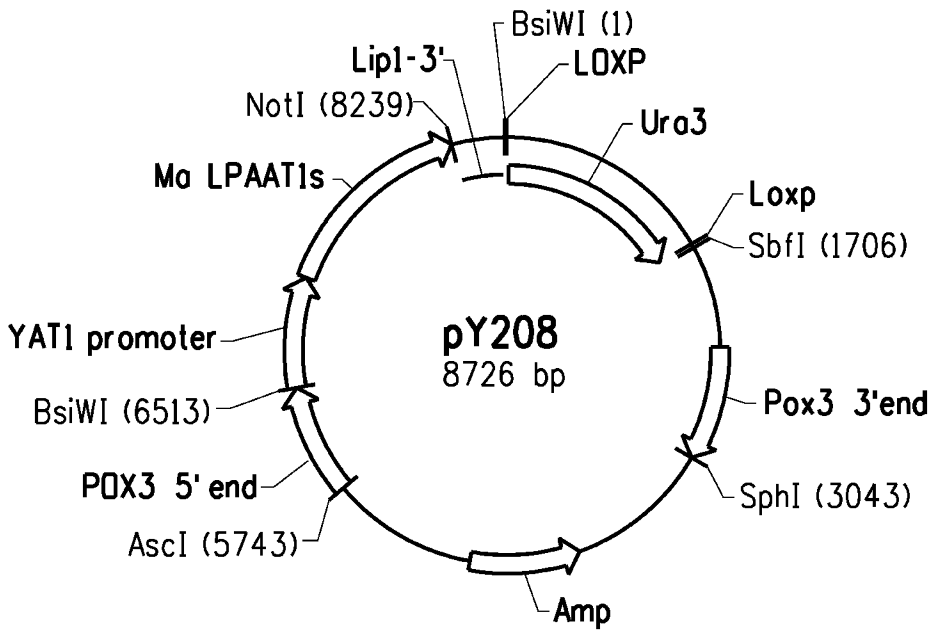


FIG. 11A

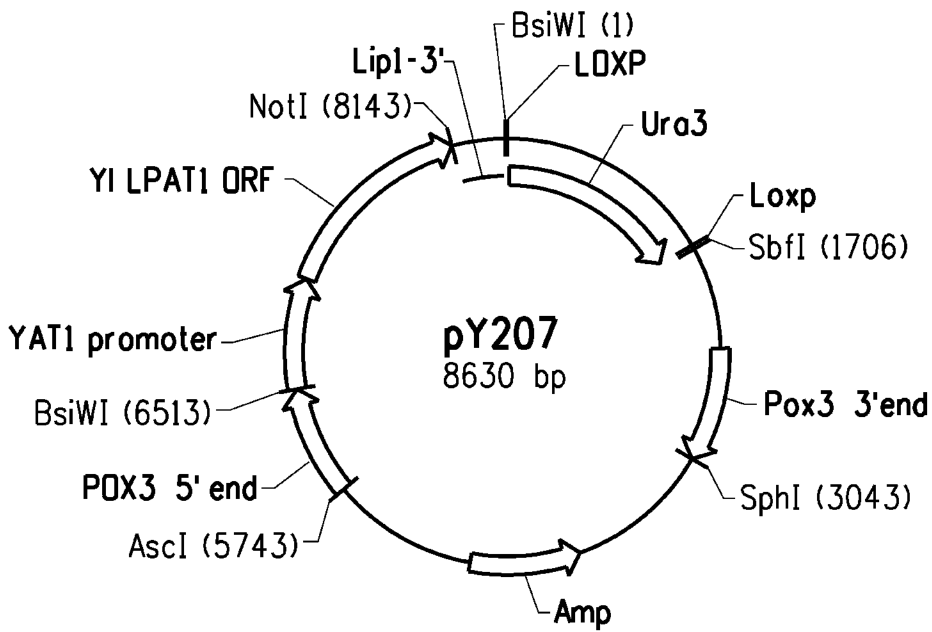


FIG. 11B

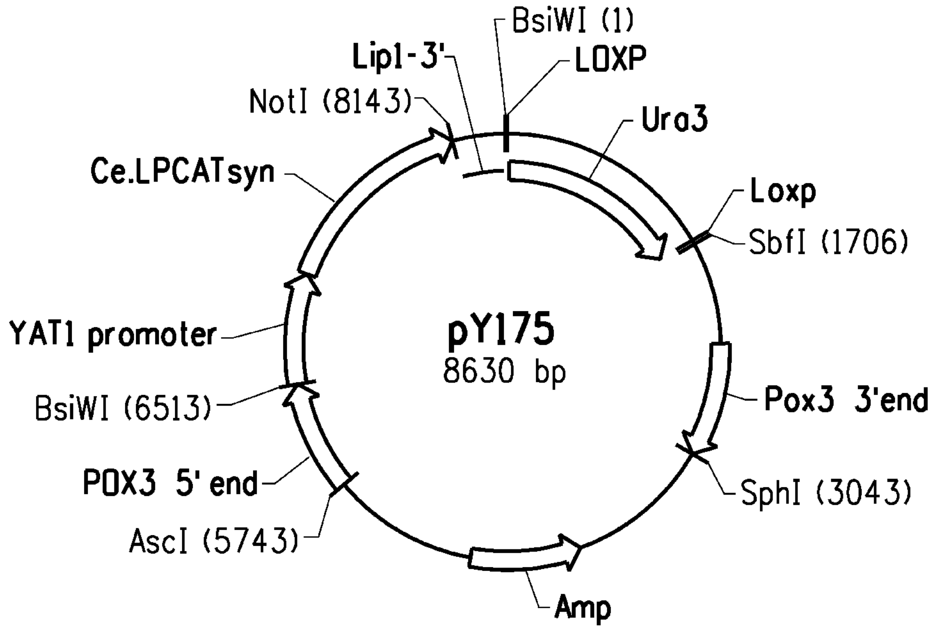


FIG. 12A

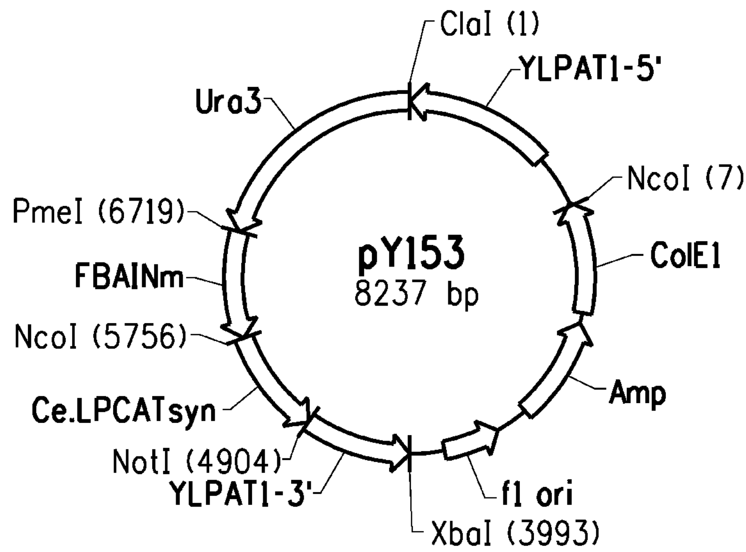


FIG. 12B

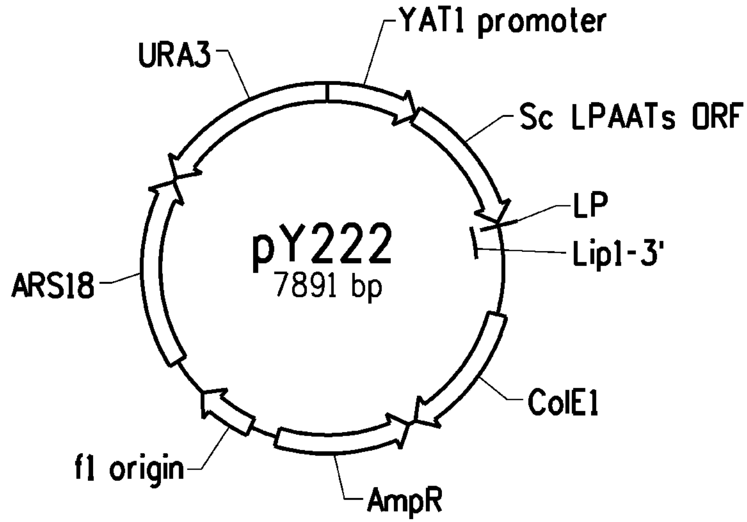


FIG. 13A

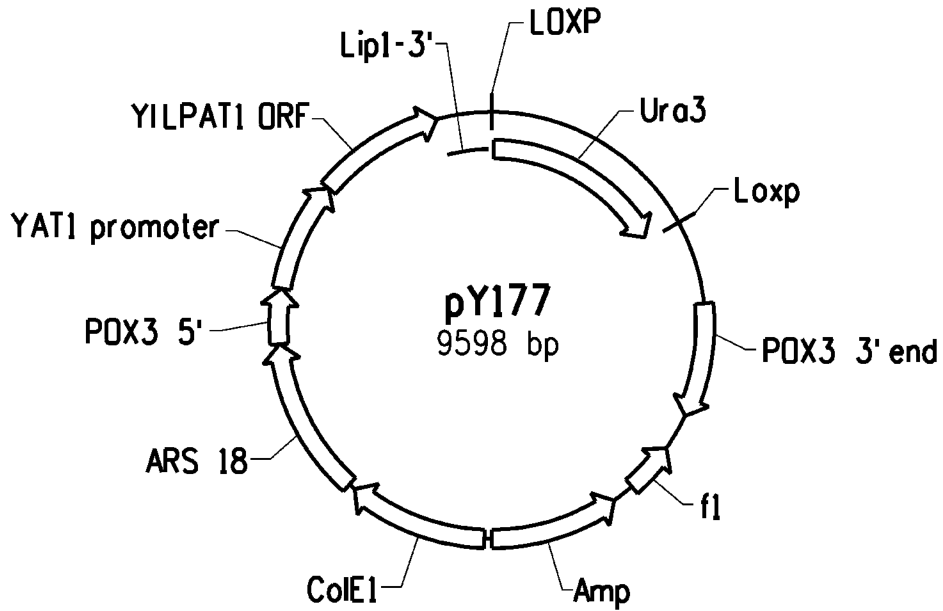


FIG. 13B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 10/38527

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C12P 21/06; C12P 7/64 (2010.01)
USPC - USPC: 435/69.1; 435/134
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8): C12P 21/06; C12P 7/64 (2010.01)
USPC: 435/69.1; 435/134

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
USPC: 435/471

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PubWEST (PGPB,USPT,EPAB,JPAB): yeast, oil, polyunsaturated, fatty acid, acyltransferase, LPLAT, Yarrowia, delta, elongase, Clustal, Clustal W, EDA, ETA, DGLA, eicosapentaenoic, delta 4
--- Please see continuation on attached additional sheet ---

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2006/0168687 A1 (RENZ et al.) 27 July 2006 (27.07.2006), abstract; para [0017], [0027], [0028], [0031], [0033], [0035], [0036], [0051], [0058], [0061], [0064], [0066], [0092], [0100], [0225], [0247].	1-20
Y	WO 2008/034648 A1 (PUZIO et al.) 27 March 2008 (27.03.2008), para [0002.1.28.28], [0122.1.10.10, [0160.1.1.1], [0264.1.6.6]; SEQ ID NO: 56383. This document can be viewed at the following url: http://v3.espacenet.com/publicationDetails/originalDocument?CC=WO&NR=2008034648A1&KC=A1&FT=D&date=20080327&DB=EPODOC&locale=en_EP	1-20
Y	US 2006/0115881 A (DAMUDE et al.) 1 June 2006 (01.06.2006), abstract; para [0204], [0257], [0264]).	1-20

☐ Further documents are listed in the continuation of Box C. ☐

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 August 2010 (10.08.2010)	01 OCT 2010
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 10/38527

Continuation of (B): Fields Searched - Electronic data base consulted during the international search (name of data base and, where practicable, search terms used):

Google Scholar: conversion efficiency yeast ?4 desaturase polyunsaturated fatty acid

esp@cenet: Narendra Yadav, omega, delta