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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 June 2008 (19.06.2008)

PCT

(10) International Publication Number
WO 2008/073271 A1

(51) International Patent Classification:
C12N 9/12 (2006.01)

(21) International Application Number:
PCT/US2007/025001

(22) International Filing Date:
5 December 2007 (05.12.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/635,258 7 December 2006 (07.12.2006) 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)*: **DAMUDE, Howard, Glenn** [CA/US]; 4 Kenwick Road, Hockessin, Delaware 19707 (US). **HE, Hongxian** [CN/US]; 2 Bloomfield Avenue, #25, Drexel Hill, Pennsylvania 19026 (US). **LIAO, Der-ing** [US/US]; 169 East Green Valley Circle, Newark, Delaware 19711 (US). **ZHU, Quinn, Qun** [US/US]; 544 Revere Road, West Chester, Pennsylvania 19382 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, 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, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

WO 2008/073271 A1

(54) **Title:** MUTANT DELTA8 DESATURASE GENES ENGINEERED BY TARGETED MUTAGENESIS AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

(57) **Abstract:** The present invention relates to mutant Δ8 desaturase genes, which have the ability to convert eicosadienoic acid [20:2 ω-6, EDA] to dihomo-γ-linolenic acid [20:3, DGLA] and/or eicosatrienoic acid [20:3 ω-3, ETrA] to eicosatetraenoic acid [20:3 ω-3, ETA]. Isolated nucleic acid fragments and recombinant constructs comprising such fragments encoding Δ8 desaturase along with methods of making long-chain polyunsaturated fatty acids (PUFAs) using these mutant Δ8 desaturases in plants and oleaginous yeast are disclosed.

TITLE

MUTANT Δ8 DESATURASE GENES ENGINEERED BY TARGETED MUTAGENESIS AND THEIR USE IN MAKING POLYUNSATURATED FATTY ACIDS

FIELD OF THE INVENTION

This invention is in the field of biotechnology. More specifically, this invention pertains to the creation of nucleic acid fragments encoding mutant Δ8 fatty acid desaturase enzymes and the use of these desaturases in making long-chain polyunsaturated fatty acids (PUFAs).

BACKGROUND OF THE INVENTION

The importance of PUFAs is undisputed. For example, certain PUFAs are important biological components of healthy cells and are considered "essential" fatty acids that cannot be synthesized *de novo* in mammals and instead must be obtained either in the diet or derived by further desaturation and elongation of linoleic acid (LA;18:2 ω-6) or α-linolenic acid (ALA; 18:3 ω-3). Additionally PUFA's are constituents of plasma membranes of cells, where they may be found in such forms as phospholipids or triacylglycerols. PUFA's are necessary for proper development (particularly in the developing infant brain) and for tissue formation and repair and, are precursors to several biologically active eicosanoids of importance in mammals (e.g., prostacyclins, eicosanoids, leukotrienes, prostaglandins). Studies have shown that a high intake of long-chain ω-3 PUFAs produces cardiovascular protective effects (Dyerberg, J. et al., *Amer. J. Clin. Nutr.*, 28:958-966 (1975); Dyerberg, J. et al., *Lancet,* 2(8081):117-119 (July 15, 1978); Shimokawa, H., *World Rev. Nutr. Diet*, 88:100-108 (2001); von Schacky, C. and Dyerberg, J., *World Rev. Nutr. Diet*, 88:90-99 (2001)). The literature reports wide-ranging health benefits conferred by administration of ω-3 and/or ω-6 PUFAs against a variety of symptoms and diseases (e.g., asthma, psoriasis, eczema, diabetes, cancer).

A variety of different hosts including plants, algae, fungi and yeast are being investigated as means for commercial PUFA production. Genetic engineering has demonstrated that the natural abilities of some hosts can be substantially altered to produce various long-chain ω-3/ω-6 PUFAs. For example, production of arachidonic acid (ARA; 20:4 ω-6), eicosapentaenoic acid (EPA; 20:5 ω-3) and docosahexaenoic acid (DHA; 22:6 ω-3) all require expression of either the Δ9 elongase/Δ8 desaturase pathway or the Δ6 desaturase/Δ6 elongase pathway. The Δ9 elongase/Δ8 desaturase pathway is present for example in euglenoid species and is characterized by the production of eicosadienoic acid ["EDA"; 20:2 ω-6] and/or eicosatrienoic acid ["ETrA"; 20:3 ω-3]. (Figure 1). The Δ6 desaturase/Δ6 elongase pathway is predominantly found in algae, mosses, fungi, nematodes and humans and is characterized by the production of γ–linoleic acid ["GLA"; 18:3 ω-6] and/or stearidonic acid ["STA"; 18:4 ω-3]) (Figure 1).

For some applications the Δ9 elongase/Δ8 desaturase pathway is favored. However Δ8 desaturase enzymes are not well known in the art leaving the construction of a recombinant Δ9 elongase/Δ8 desaturase pathway with limited options . The few Δ8 desaturase enzymes identified thus far have the ability to convert both EDA to dihomo-γ–linolenic acid [20:3, DGLA] and ETrA to eicosatetraenoic acid [20:4, ETA] (wherein ARA are EPA are subsequently synthesized from DGLA and ETA, respectively, following reaction with a Δ5 desaturase, while DHA synthesis requires subsequent expression of an additional $C_{20/22}$ elongase and a Δ4 desaturase).

Several Δ8 desaturase enzymes are known and have been partially characterized (see for example Δ8 desaturases from *Euglena gracilis* Wallis et al., *Arch. Biochem. and Biophys.*, 365(2):307-316 (May 1999); WO 2000/34439; U.S. Patent No. 6,825,017; WO 2004/057001; WO 2006/012325; WO 2006/012326). Additionally WO 2005/103253 (published April 22, 2005) discloses amino acid and nucleic acid

sequences for a Δ8 desaturase enzyme from *Pavlova salina* (see also U.S. Publication No. 2005/0273885). Sayanova et al. (*FEBS Lett.*, 580:1946-1952 (2006)) describes the isolation and characterization of a cDNA from the free living soil amoeba *Acanthamoeba castellanii* that, when expressed in *Arabidopsis*, encodes a C_{20} Δ8 desaturase. Furthermore, commonly owned and co-pending United States Provisional Application No. 60/795810 filed April 28, 2006 discloses amino acid and nucleic acid sequences for a Δ8 desaturase enzyme from *Pavlova lutheri* (CCMP459).

A need remains therefore for additional Δ8 desaturase enzymes to be used in recombinant pathways for the production of PUFA's. Applicants have solved the stated need by developing a synthetically engineered mutant *Euglena gracilis* Δ8 desaturase.

SUMMARY OF THE INVENTION

The present invention relates to new recombinant constructs encoding mutant polypeptides having Δ8 desaturase activity, and their use in plants and yeast for the production of PUFAs and particularly ω-3 and/or ω-6 fatty acids.

Accordingly the invention provides, an isolated polynucleotide comprising:(a) a nucleotide sequence encoding a mutant polypeptide having Δ8 desaturase activity having an amino acid sequence as set forth in SEQ ID NO:2 and wherein SEQ ID NO:2 is not identical to SEQ ID NO:10; or, (b) a complement of the nucleotide sequence of part (a), wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

In an alternate embodiment the invention provides an isolated polynucleotide comprising: (a) a nucleotide sequence encoding a mutant polypeptide having Δ8 desaturase activity, having an amino acid sequence as set forth in SEQ ID NO: 198and wherein SEQ ID NO:198 is not identical to SEQ ID NO:10; or, (b) a complement of the nucleotide sequence of part (a), wherein the complement and the nucleotide

sequence consist of the same number of nucleotides and are 100% complementary.

It is one aspect of the invention to provide polypeptides encoded by the polynucleotides of the invention as well as genetic chimera and host cells transformed and expressing the same.

In another aspect the invention provides a method for making long-chain polyunsaturated fatty acids in a yeast cell comprising: (a) providing a yeast cell of the invention ; and (b) growing the yeast cell of (a) under conditions wherein long-chain polyunsaturated fatty acids are produced.

In another aspect of the invention provides microbial oil obtained from the yeast of the invention.

In an alternate embodiment the invention provides an oleaginous yeast producing at least about 25% of its dry cell weight as oil comprising:
a) a first recombinant DNA construct comprising an isolated polynucleotide encoding a Δ8 desaturase polypeptide of the invention operably linked to at least one regulatory sequence; and,
b) at least one second recombinant DNA construct comprising an isolated polynucleotide operably linked to at least one regulatory sequence, the construct encoding a polypeptide selected from the group consisting of: a Δ4 desaturase, a Δ5 desaturase, Δ6 desaturase, a Δ9 desaturase, a Δ12 desaturase, a Δ15 desaturase, a Δ17 desaturase, a Δ9 elongase, a $C_{14/16}$ elongase, a $C_{16/18}$ elongase, a $C_{18/20}$ elongase and a $C_{20/22}$ elongase.

In another aspect the invention provides a food or feed product comprising the microbial oil of the invention.

In another embodiment the invention provides a method for producing dihomo–γ–linoleic acid comprising:

a) providing an oleaginous yeast comprising:

(i) a recombinant construct encoding a Δ8 desaturase polypeptide having an amino acid sequence as set forth in SEQ ID NO:2, wherein SEQ ID NO:2 is not identical to SEQ ID NO:10; and,

(ii) a source of eicosadienoic acid;

b) growing the yeast of step (a) under conditions wherein the recombinant construct encoding a Δ8 desaturase polypeptide is expressed and eicosadienoic acid is converted to dihomo-γ-linoleic acid, and;

c) optionally recovering the dihomo-γ-linoleic acid of step (b).

In an alternate embodiment the invention provides a method for producing eicosatetraenoic acid comprising:

a) providing an oleaginous yeast comprising:

(i) a recombinant construct encoding a Δ8 desaturase polypeptide having an amino acid sequence as set forth in SEQ ID NO:2, wherein SEQ ID NO:2 is not identical to SEQ ID NO:10; and,

(ii) a source of eicosatrienoic acid;

b) growing the yeast of step (a) under conditions wherein the recombinant construct encoding a Δ8 desaturase polypeptide is expressed and eicosatrienoic acid is converted to eicosatetraenoic acid, and;

c) optionally recovering the eicosatetraenoic acid of step (b).

In another embodiment the invention provides a method for the production of dihomo–γ–linoleic acid comprising:

a) providing a yeast cell comprising:

i) a first recombinant DNA construct comprising the isolated polynucleotide of the invention operably linked to at least one regulatory sequence, and;

ii) at least one second recombinant DNA construct comprising an isolated polynucleotide encoding a Δ9 elongase polypeptide, operably linked to at least one regulatory sequence;

b) providing the yeast cell of (a) with a source of linolenic acid, and;

c) growing the yeast cell of (b) under conditions where dihomo–γ–linoleic acid is formed.

BIOLOGICAL DEPOSITS

The following plasmid has been deposited with the American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, and bears the following designation, Accession Number and date of deposit (Table 1).

TABLE 1

ATCC Deposits

Plasmid	Accession Number	Date of Deposit
pKR72	PTA-6019	May 28, 2004

BRIEF DESCRIPTION OF THE DRAWINGS AND SEQUENCE LISTINGS

Figure 1 is a representative PUFA biosynthetic pathway.

Figure 2 shows a topological model of EgD8S.

Figure 3 shows an alignment of EgD8S (SEQ ID NO:10), a Δ6 desaturase of *Amylomyces rouxii* (SEQ ID NO:13), a Δ6 desaturase of *Rhizopus orizae* (SEQ ID NO:14), a Δ8 fatty acid desaturase-like protein of *Leishmania major* (GenBank Accession No. CAJ09677; SEQ ID NO:15), and a Δ6 desaturase of *Mortierella isabellina* (GenBank Accession No. AAG38104; SEQ ID NO:16). The method of alignment used corresponds to the "Clustal W method of alignment".

Figure 4 shows an alignment of EgD8S (SEQ ID NO:10), the cytochrome b_5 of *Saccharomyces cerevisiae* (GenBank Accession No. P40312; SEQ ID NO:178) and a probable cytochrome b_5 1 of *Schizosaccharomyces pombe* (GenBank Accession No. O94391; SEQ ID NO:179). The method of alignment used corresponds to the "Clustal W method of alignment".

Figure 5 provides plasmid maps for the following: (A) pZKLeuN-29E3; and, (B) pY116.

Figure 6 provides plasmid maps for the following: (A) pKUNFmkF2; (B) pDMW287F; (C) pDMW214; and, (D) pFmD8S.

Figure 7 diagrams the synthesis of the Mutant EgD8S-5B, by ligation of fragments from Mutant EgD8S-1 and Mutant EgD8S-2B.

Figure 8A diagrams the synthesis of Mutant EgD8S-008, by ligation of fragments from Mutant EgD8S-001 and Mutant EgD8S-003. Similarly, Figure 8B diagrams the synthesis of Mutant EgD8S-009, by ligation of fragments from Mutant EgD8S-001 and Mutant EgD8S-004.

Figure 9A diagrams the synthesis of Mutant EgD8S-013, by ligation of fragments from Mutant EgD8S-009 and Mutant EgD8S-23. Similarly, Figure 9B diagrams the synthesis of Mutant EgD8S-015, by ligation of fragments from Mutant EgD8S-008 and Mutant EgD8S-28.

Figure 10 shows an alignment of EgD8S (SEQ ID NO:10), Mutant EgD8S-23 (SEQ ID NO:4), Mutant EgD8S-013 (SEQ ID NO:6) and Mutant EgD8S-015 (SEQ ID NO:8). The method of alignment used corresponds to the "Clustal W method of alignment".

Figure 11 provides plasmid maps for the following: (A) pKo2UFm8; and, (B) pKO2UF8289.

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.

A Sequence Listing is provided herewith on Compact Disk. The contents of the Compact Disk containing the Sequence Listing are hereby

incorporated by reference in compliance with 37 CFR 1.52(e). The Compact Disks are submitted in triplicate and are identical to one another. The disks are labeled "Copy 1 – Sequence Listing", "Copy 2 - Sequence Listing", and CRF. The disks contain the following file: CL3495 Seq Listing_11.27.06_ST25 having the following size: 293,000 bytes and which was created December 6, 2006.

SEQ ID NOs:1-17, 19-23, 165 and 172-177 are ORFs encoding genes or proteins (or portions thereof) or plasmids, as identified in Table 2.

Table 2

Summary Of Nucleic Acid And Protein SEQ ID Numbers

Description and Abbreviation	**Nucleic acid SEQ ID NO.**	**Protein SEQ ID NO.**
Synthetic mutant Δ8 desaturase, derived from *Euglena gracilis* ("EgD8S-consensus") optionally comprising M1, M2, M3, M8, M12, M15, M16, M18, M21, M26, M38, M45, M46, M51, M63, M68, M69 and M70 mutation sites	1 (1272 bp)	2 (422 AA)
Synthetic mutant Δ8 desaturase, derived from *Euglena gracilis* ("EgD8S-consensus") optionally comprising M1, M2, M3, M6, M8, M12, M14, M15, M16, M18, M19, M21, M22, M26, M38, M39, M40, M41, M45, M46, M49, M50, M51, M53, M54, M58, M63, M68, M69 and M70 mutation sites	197 (1272 bp)	198 (422 AA)
Synthetic mutant Δ8 desaturase, derived from *Euglena gracilis* ("Mutant EgD8S-23")	3 (1272 bp)	4 (422 AA)
Synthetic mutant Δ8 desaturase, derived from *Euglena gracilis* ("Mutant EgD8S-013")	5 (1272 bp)	6 (422 AA)
Synthetic mutant Δ8 desaturase, derived from *Euglena gracilis* ("Mutant EgD8S-015")	7 (1272 bp)	8 (422 AA)
Synthetic Δ8 desaturase, derived from *Euglena gracilis,* codon-optimized for expression in *Yarrowia lipolytica* ("EgD8S")	9 (1272 bp)	10 (422 AA)
Euglena gracilis Δ8 desaturase (full-length gene is nucleotides 4-1269 (Stop)) ("EgD8")	11 (1271 bp)	12 (421 AA)
Amylomyces rouxii Δ6 desaturase (GenBank Accession No. AAR27297)	--	13 (467 AA)
Rhizopus orizae Δ6 desaturase (GenBank Accession No. AAS93682)	--	14 (445 AA)

Leishmania major Δ8 fatty acid desaturase-like protein (GenBank Accession No. CAJ09677)	--	15 (382 AA)
Mortierella isabellina Δ6 desaturase (GenBank Accession No. AAG38104)	--	16 (439 AA)
Saccharomyces cerevisiae cytochrome b_5 (GenBank Accession No. P40312)	--	178 (120 AA)
Schizosaccharomyces pombe probable cytochrome b_5 1 (GenBank Accession No. O94391)	--	179 (124 AA)
Plasmid pZKLeuN-29E3	17 (14,655 bp)	--
Synthetic $C_{16/18}$ elongase gene derived from *Mortierella alpina* ELO3, codon-optimized for expression in *Yarrowia lipolytica*	19 (828 bp)	--
Plasmid pFmD8S	20 (8,910 bp)	--
Plasmid pKUNFmkF2	21 (7,145 bp)	--
Plasmid pDMW287F	22 (5,473 bp)	--
Plasmid pDMW214	23 (9,513 bp)	--
Plasmid pKO2UFkF2	165 (8,560 bp)	--
Isochrysis galbana Δ9 elongase (GenBank Accession No. AF390174)	172 (1064 bp)	173 (263 AA)
Synthetic Δ9 elongase gene derived from *Isochrysis galbana*, codon-optimized for expression in *Yarrowia lipolytica*	174 (792 bp)	173 (263 AA)
Euglena gracilis Δ9 elongase	175 (777 bp)	176 (258 AA)
Synthetic Δ9 elongase gene derived from *Euglena gracilis*, codon-optimized for expression in *Yarrowia lipolytica*	177 (777 bp)	176 (258 AA)
Plasmid pY116	180 (8739 bp)	--
Plasmid pKO2UF8289	181 (15,304 bp)	--
Synthetic mutant Δ8 desaturase, derived from *Euglena gracilis* ("modified Mutant EgD8S-23"), comprising a 5' *Not*1 site	182 (1288 bp)	--
Plasmid pKR457	183 (5252 bp)	--
Plasmid pKR1058	184 (6532 bp)	--

Plasmid pKR607	185 (7887 bp)	--
Plasmid pKR1060	186 (11,766 bp)	--
Plasmid pKR906	189 (4311 bp)	--
Plasmid pKR72	190 (7085 bp)	--
Plasmid pKR1010	191 (7873 bp)	--
Plasmid pKR1059	192 (11752 bp)	--
Euglena gracilis Δ9 elongase -- 5' sequence of the cDNA insert from clone eeg1c.pk001.n5.f.	194 (757 bp)	--
Euglena gracilis Δ9 elongase -- 3' sequence of the cDNA insert from clone eeg1c.pk001.n5.f.	195 (774 bp)	--
Euglena gracilis Δ9 elongase --sequence aligned from SEQ ID NO:1 and SEQ ID NO:2 (full cDNA sequence excluding polyA tail)	196 (1201 bp)	--

SEQ ID NO:18 corresponds to a LoxP recombination site that is recognized by a Cre recombinase enzyme.

SEQ ID NOs:24-164 correspond to 70 pairs of nucleotide primers (i.e., 1A, 1B, 2A, 2B, 3A, 3B, etc. up to 69A, 69B, 70A and 70B, respectively), used to create specific targeted mutations at mutation sites M1, M2, M3, etc. up to M70.

SEQ ID NOs:166-171 correspond to His-rich motifs that are featured in membrane-bound fatty acid desaturases belonging to a super-family of membrane di-iron proteins.

SEQ ID NOs:187 and 188 correspond to primers oEugEL1-1 and oEugEL1-2, respectively, used to amplify a *Euglena gracilis* Δ9 elongase.

SEQ ID NO:193 is the M13F universal primer.

DETAILED DESCRIPTION OF THE INVENTION

As used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a plant" includes a plurality of

such plants, reference to "a cell" includes one or more cells and equivalents thereof known to those skilled in the art, and so forth.

All patents, patent applications, and publications cited herein are incorporated by reference in their entirety. This specifically includes the following commonly owned and copending applications: U.S. Patent Applications No. 10/840478, No. 10/840579 and No. 10/840325 (filed May 6, 2004), U.S. Patent Application No. 10/869630 (filed June 16, 2004), U.S. Patent Application No. 10/882760 (filed July 1, 2004), U.S. Patent Applications No. 10/985254 and No. 10/985691 (filed November 10, 2004), U.S. Patent Application No. 10/987548 (filed November 12, 2004), U.S. Patent Applications No. 11/024545 and No. 11/024544 (filed December 29, 2004), U.S. Patent Application No. 11/166993 (filed June 24, 2005), U.S. Patent Application No. 11/183664 (filed July 18, 2005), U.S. Patent Application No. 11/185301 (filed July 20, 2005), U.S. Patent Application No. 11/190750 (filed July 27, 2005), U.S. Patent Application No. 11/198975 (filed August 8, 2005), U.S. Patent Application No. 11/225354 (filed September 13, 2005), U.S. Patent Application No. 11/251466 (filed October 14, 2005), U.S. Patent Applications No. 11/254173 and No. 11/253882 (filed October 19, 2005), U.S. Patent Applications No. 11/264784 and No. 11/264737 (filed November 1, 2005), U.S. Patent Application No. 11/265761 (filed November 2, 2005), U.S. Patent Application No. 60/739989 (filed November 23, 2005), U.S. Patent Application No. 60/793575 (filed April 20, 2006), U.S. Patent Application No. 60/795810 (filed April 28, 2006), U.S. Patent Application No. 60/796637 (filed May 1, 2006) and U.S. Patent Applications No. 60/801172 and No. 60/801119 (filed May 17, 2006).

Additionally, commonly owned U.S. Patent Application No. 10/776311, (published August 26, 2004) relating to the production of PUFAs in plants, and U.S. Patent Application No. 10/776889 (published August 26, 2004) relating to annexin promoters and their use in expression of transgenes in plants, are incorporated by reference in their entirety.

The present invention provides mutant Δ8 desaturase enzymes and genes encoding the same, that may be used for the manipulation of biochemical pathways for the production of healthful PUFAs.

PUFAs, or derivatives thereof, made by the methodology disclosed herein can be 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. PUFAs may also be used as anti-inflammatory or cholesterol lowering agents as components of pharmaceutical or veterinary compositions.

Definitions

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

"Open reading frame" is abbreviated ORF.

"Polymerase chain reaction" is abbreviated PCR.

"American Type Culture Collection" is abbreviated ATCC.

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

"Triacylglycerols" are abbreviated TAGs.

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}. Additional details concerning the differentiation between "saturated fatty acids" versus "unsaturated fatty acids", "monounsaturated fatty acids" versus "polyunsaturated fatty acids" (or "PUFAs"), and "omega-6 fatty acids" (ω-6 or *n*-6) versus "omega-3 fatty acids" (ω-3 or *n*-3) are provided in WO2004/101757.

Fatty acids are described herein by a simple notation system of "X:Y", wherein X is the number of carbon (C) atoms in the particular fatty acid and Y is the number of double bonds. The number following the fatty acid designation indicates the position of the double bond from the

carboxyl end of the fatty acid with the "c" affix for the *cis*-configuration of the double bond [e.g., palmitic acid (16:0), stearic acid (18:0), oleic acid (18:1, 9c), petroselinic acid (18:1, 6c), LA (18:2, 9c,12c), GLA (18:3, 6c,9c,12c) and ALA (18:3, 9c,12c,15c)]. Unless otherwise specified, 18:1, 18:2 and 18:3 refer to oleic, LA and ALA fatty acids. If not specifically written as otherwise, double bonds are assumed to be of the *cis* configuration. For instance, the double bonds in 18:2 (9,12) would be assumed to be in the *cis* configuration.

Nomenclature used to describe PUFAs in the present disclosure is shown below 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 each compounds' chemical name.

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–Linoleic	GLA	*cis*-6, 9, 12-octadecatrienoic	18:3 ω-6
Eicosadienoic	EDA	*cis*-11, 14- eicosadienoic	20:2 ω-6
Dihomo-γ–Linoleic	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
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-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 terms "triacylglycerol", "oil" and "TAGs" refer to neutral lipids composed of three fatty acyl residues esterified to a glycerol molecule (and such terms will be used interchangeably throughout the present disclosure herein). Such oils can contain long-chain PUFAs, as well as shorter saturated and unsaturated fatty acids and longer chain saturated fatty acids. Thus, "oil biosynthesis" generically refers to the synthesis of TAGs in the cell.

"Percent (%) PUFAs in the total lipid and oil fractions" refers to the percent of PUFAs relative to the total fatty acids in those fractions. The term "total lipid fraction" or "lipid fraction" both refer to the sum of all lipids (i.e., neutral and polar) within an oleaginous organism, thus including those lipids that are located in the phosphatidylcholine (PC) fraction, phosphatidyletanolamine (PE) fraction and triacylglycerol (TAG or oil) fraction. However, the terms "lipid" and "oil" will be used interchangeably throughout the specification.

The term "PUFA biosynthetic pathway" refers to a metabolic process that converts oleic acid to LA, EDA, GLA, DGLA, ARA, ALA, STA, ETrA, ETA, EPA, DPA and DHA. This process is well described in the literature (e.g., see 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 desaturation and elongation enzymes (i.e., "PUFA biosynthetic pathway enzymes") present in the endoplasmic reticulim membrane. More specifically, "PUFA biosynthetic pathway enzyme"

refers to any of the following enzymes (and genes which encode said enzymes) associated with the biosynthesis of a PUFA, including: a Δ4 desaturase, a Δ5 desaturase, a Δ6 desaturase, a Δ12 desaturase, a Δ15 desaturase, a Δ17 desaturase, a Δ9 desaturase, a Δ8 desaturase, a $C_{14/16}$ elongase, a Δ9 elongase, a $C_{16/18}$ elongase, a $C_{18/20}$ elongase and/or a $C_{20/22}$ elongase.

The term "ω-3/ω-6 fatty acid biosynthetic pathway" refers to a set of genes which, when expressed under the appropriate conditions encode enzymes that catalyze the production of either or both ω-3 and ω-6 fatty acids. Typically the genes involved in the ω-3/ω-6 fatty acid biosynthetic pathway encode some or all of the following enzymes: Δ12 desaturase, Δ6 desaturase, $C_{18/20}$ elongase, $C_{20/22}$ elongase, Δ5 desaturase, Δ17 desaturase, Δ15 desaturase, Δ9 desaturase, Δ8 desaturase, a Δ9 elongase and Δ4 desaturase. A representative pathway is illustrated in Figure 1, providing for the conversion of oleic acid through various intermediates to DHA, which demonstrates how both ω-3 and ω-6 fatty acids may be produced from a common source. The pathway is naturally divided into two portions where one portion will generate ω-3 fatty acids and the other portion, only ω-6 fatty acids. That portion that only generates ω-3 fatty acids will be referred to herein as the ω-3 fatty acid biosynthetic pathway, whereas that portion that generates only ω-6 fatty acids will be referred to herein as the ω-6 fatty acid biosynthetic pathway.

The term "functional" as used herein in context with the ω-3/ω-6 fatty acid biosynthetic pathway means that some (or all of) the genes in the pathway express active enzymes, resulting in *in vivo* catalysis or substrate conversion. It should be understood that "ω-3/ω-6 fatty acid biosynthetic pathway" or "functional ω-3/ω-6 fatty acid biosynthetic pathway" does not imply that all the genes listed in the above paragraph are required, as a number of fatty acid products will only require the expression of a subset of the genes of this pathway.

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 that will desaturate a fatty acid between the 8th and 9th carbon atom numbered from the carboxyl-terminal end of the molecule and that can, for example, catalyze the conversion of EDA to DGLA and/or ETrA to ETA. Other desaturases include: 1.) Δ5 desaturases that catalyze the conversion of DGLA to ARA and/or ETA to EPA; 2.) Δ6 desaturases that catalyze the conversion of LA to GLA and/or ALA to STA; 3.) Δ4 desaturases that catalyze the conversion of DPA to DHA; 4.) Δ12 desaturases that catalyze the conversion of oleic acid to LA; 5.) Δ15 desaturases that catalyze the conversion of LA to ALA and/or GLA to STA; 6.) Δ17 desaturases that catalyze the conversion of ARA to EPA and/or DGLA to ETA; and 7.) Δ9 desaturases that catalyze the conversion of palmitate to palmitoleic acid (16:1) and/or stearate to oleic acid (18:1). In the art, Δ15 and Δ17 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 (e.g., conversion of LA into ALA and ARA into EPA, respectively). In some embodiments, it is most desirable to empirically determine the specificity of a particular fatty acid desaturase by transforming a suitable host with the gene for the fatty acid desaturase and determining its effect on the fatty acid profile of the host.

For the purposes herein, the term "EgD8" refers to a Δ8 desaturase enzyme (SEQ ID NO:12) isolated from *Euglena gracilis,* encoded by SEQ ID NO:11 herein. EgD8 is 100% identical and functionally equivalent to "Eg5", as described in WO 2006/012325 and WO 2006/012326 [US2005-0287652-A1].

Similarly, the term "EgD8S" refers to a synthetic Δ8 desaturase derived from *Euglena gracilis* that is codon-optimized for expression in *Yarrowia lipolytica* herein (i.e., SEQ ID NOs:9 and 10). EgD8S is 100% identical and functionally equivalent to "D8SF", as described in WO 2006/012325 and WO 2006/012326.

The term ""mutant EgD8S" refers to a Δ8 desaturase of the present invention that has at least one mutation with respect to the synthetic Δ8 desaturase derived from *Euglena gracilis* that is codon-optimized for expression in *Yarrowia lipolytica* (i.e., EgD8S). Although "mutations" may include any deletions, insertions and point mutations (or combinations thereof), in preferred embodiments the mutant EgD8S is set forth in SEQ ID NO:2, wherein: (i) SEQ ID NO:2 comprises at least one mutation selected from the group consisting of: 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 54A to G, 55F to Y, 66P to Q, 67S to A, 108S to L, 117G to A, 118Y to F, 120L to M, 121M to L, 125Q to H, 126M to L, 132V to L, 133L to V, 162L to V, 163V to L, 170G to A, 171L to V, 279T to L, 280L to T, 293L to M, 346I to V, 347I to L, 348T to S, 407A to S, 408V to Q, 418A to G, 419G to A and 422L to Q, wherein the mutations are defined with respect to the synthetic codon-optimized sequence of EgD8S (i.e., SEQ ID NO:10); and (ii) SEQ ID NO:2 is not 100% identical to SEQ ID NO:10. In more preferred embodiments, the mutant EgD8S has at least about 10-18 mutations with respect to the synthetic codon-optimized sequence of EgD8S, more preferably at least about 19-25 mutations, and most preferably at least about 26-33 mutations with respect to synthetic codon-optimized sequence of EgD8S (i.e., SEQ ID NO:10). In another embodiment, the Δ8 desaturase activity of the mutant EgD8S is at least about functionally equivalent to the Δ8 desaturase activity of the synthetic codon-optimized EgD8S (SEQ ID NO:10).

A mutant EgD8S is "at least about functionally equivalent" to EgD8S when enzymatic activity and specific selectivity of the mutant EgD8S sequence are comparable to that of EgD8S, despite differing polypeptide sequences. Thus, a functionally equivalent mutant EgD8S

sequence will possess Δ8 desaturase activity that is not substantially reduced with respect to that of EgD8S when the "conversion efficiency" of each enzyme is compared (i.e., a mutant EgD8S will have at least about 50%, preferably at least about 75%, more preferably at least about 85%, and most preferably at least about 95% of the enzymatic activity of EgD8S). In more preferred embodiments, the mutant EgD8S will have increased enzymatic activity and specific selectivity when compared to that of EgD8S (i.e., at least about 105%, more preferably at least about 115% and most preferably at least about 125% of the enzymatic activity of EgD8S).

The terms "conversion efficiency" and "percent substrate conversion" refer to the efficiency by which a particular enzyme (e.g., a desaturase) 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 "elongase system" refers to a suite of four enzymes that are responsible for elongation of a fatty acid carbon chain to produce a fatty acid that is 2 carbons longer than the fatty acid substrate that the elongase system acts upon. More specifically, the process of elongation occurs in association with fatty acid synthase, whereby CoA is the acyl carrier (Lassner et al., *The Plant Cell,* 8:281-292 (1996)). In the first step, which has been found to be both substrate-specific and also rate-limiting, malonyl-CoA is condensed with a long-chain acyl-CoA to yield CO_2 and a β-ketoacyl-CoA (where the acyl moiety has been elongated by two carbon atoms). Subsequent reactions include reduction to β-hydroxyacyl-CoA, dehydration to an enoyl-CoA and a second reduction to yield the elongated acyl-CoA. Examples of reactions catalyzed by elongase systems are the conversion of GLA to DGLA, STA to ETA and EPA to DPA.

For the purposes herein, an enzyme catalyzing the first condensation reaction (i.e., conversion of malonyl-CoA to β-ketoacyl-CoA)

will be referred to generically as an "elongase". In general, the substrate selectivity of elongases is somewhat broad but segregated by both chain length and the degree of unsaturation. Accordingly, elongases can have different specificities. For example, a $C_{14/16}$ elongase will utilize a C_{14} substrate (e.g., myristic acid), a $C_{16/18}$ elongase will utilize a C_{16} substrate (e.g., palmitate), a $C_{18/20}$ elongase will utilize a C_{18} substrate (e.g., GLA, STA) and a $C_{20/22}$ elongase will utilize a C_{20} substrate (e.g., EPA). In like manner, a Δ9 elongase is able to catalyze the conversion of LA and ALA to EDA and ETrA, respectively (e.g., WO 2002/077213). It is important to note that some elongases have broad specificity and thus a single enzyme may be capable of catalyzing several elongase reactions (e.g., thereby acting as both a $C_{16/18}$ elongase and a $C_{18/20}$ elongase).

The term "Δ9 elongase/ Δ8 desaturase pathway" refers to a biosynthetic pathway for production of long-chain PUFAs, said pathway minimally comprising a Δ9 elongase and a Δ8 desaturase and thereby enabling biosynthesis of DGLA and/or ETA from LA and ALA, respectively. This pathway may be advantageous in some embodiments, as the biosynthesis of GLA and/or STA is excluded.

The term "amino acid" will refer to the basic chemical structural unit of a protein or polypeptide. The amino acids are identified by either the one-letter code or the three-letter codes for amino acids, in conformity with the IUPAC-IYUB standards described in *Nucleic Acids Research,* 13:3021-3030 (1985) and in the *Biochemical Journal,* 219 (2):345-373 (1984), which are herein incorporated by reference.

The term "conservative amino acid substitution" refers to a substitution of an amino acid residue in a given protein with another amino acid, without altering the chemical or functional nature of that protein. For example, it is well known in the art that alterations in a gene that result in the production of a chemically equivalent amino acid at a given site (but do not affect the structural and functional properties of the encoded, folded protein) are common. For the purposes of the present invention,

"conservative amino acid substitutions" are defined as exchanges within one of the following five groups:

1. Small aliphatic, nonpolar or slightly polar residues: Ala [A], Ser [S], Thr [T] (Pro [P], Gly [G]);
2. Polar, negatively charged residues and their amides: Asp [D], Asn [N], Glu [E], Gln [Q];
3. Polar, positively charged residues: His [H], Arg [R], Lys [K];
4. Large aliphatic, nonpolar residues: Met [M], Leu [L], Ile [I], Val [V] (Cys [C]); and
5. Large aromatic residues: Phe [F], Tyr [Y], Trp [W].

Thus, Ala, a slightly hydrophobic amino acid, may be substituted by another less hydrophobic residue (e.g., Gly). Similarly, changes which result in substitution of one negatively charged residue for another (e.g., Asp for Glu) or one positively charged residue for another (e.g., Lys for Arg) can also be expected to produce a functionally equivalent product. As such, conservative amino acid substitutions generally maintain: 1.) the structure of the polypeptide backbone in the area of the substitution; 2.) the charge or hydrophobicity of the molecule at the target site; or 3.) the bulk of the side chain. Additionally, in many cases, alterations of the N-terminal and C-terminal portions of the protein molecule would also not be expected to alter the activity of the protein.

The term "non-conservative amino acid substitution" refers to an amino acid substitution that is generally expected to produce the greatest change in protein properties. Thus, for example, a non-conservative amino acid substitution would be one whereby: 1.) a hydrophilic residue is substituted for/by a hydrophobic residue (e.g., Ser or Thr for/by Leu, Ile, Val); 2.) a Cys or Pro is substituted for/by any other residue; 3.) a residue having an electropositive side chain is substituted for/by an electronegative residue (e.g., Lys, Arg or His for/by Asp or Glu); or, 4.) a residue having a bulky side chain is substituted for/by one not having a side chain (e.g., Phe for/by Gly). Sometimes, non-conservative amino

acid substitutions between two of the five groups will not affect the activity of the encoded protein.

The terms "polynucleotide", "polynucleotide sequence", "nucleic acid sequence", "nucleic acid fragment" and "isolated nucleic acid fragment" are used interchangeably herein. 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.

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), particularly Chapter 11 and Table 11.1 therein (entirely incorporated herein by reference). The conditions of temperature and ionic strength determine the "stringency" of the hybridization. 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. One set of preferred conditions uses a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45 °C for 30 min, and then repeated twice with 0.2X SSC, 0.5% SDS at 50 °C for 30 min. A more preferred set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2X SSC, 0.5% SDS was increased to 60 °C. Another preferred set of highly

stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65 °C. An additional set of stringent conditions include hybridization at 0.1X SSC, 0.1% SDS, 65 °C and washed with 2X SSC, 0.1% SDS followed by 0.1X SSC, 0.1% SDS, for example.

Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of T*m* for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher T*m*) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating T*m* have been derived (see Sambrook et al., *supra*, 9.50-9.51). For hybridizations with shorter nucleic acids, i.e., oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (see Sambrook et al., *supra*, 11.7-11.8). In one embodiment the length for a hybridizable nucleic acid is at least about 10 nucleotides. Preferably a minimum length for a hybridizable nucleic acid is at least about 15 nucleotides; more preferably at least about 20 nucleotides; and most preferably the length is at least about 30 nucleotides. Furthermore, the skilled artisan will recognize that the temperature and wash solution salt concentration may be adjusted as necessary according to factors such as length of the probe.

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 putatively identify 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 (e.g., *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 instant specification teaches the complete amino acid and nucleotide sequence encoding encoding one or more particular Δ8 desaturase proteins. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

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. Accordingly, the invention herein also includes isolated nucleic acid fragments that are complementary to the complete sequences as reported in the accompanying Sequence Listing, as well as those substantially similar nucleic acid sequences.

The terms "homology", "homologous", "substantially similar" and "corresponding substantially" are used interchangeably herein. 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 of the instant invention 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. It is therefore understood, as those skilled in the art will appreciate, that the invention encompasses more than the specific exemplary sequences.

Moreover, the skilled artisan recognizes that substantially similar nucleic acid sequences encompassed by this invention 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 to any of the nucleic acid sequences disclosed herein.

"Codon degeneracy" refers to the nature in the genetic code permitting variation of the nucleotide sequence without effecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleic acid fragment that encodes all or a substantial portion of the amino acid sequence encoding the instant Δ8 desaturase polypeptides as set forth in SEQ ID NOs:2, 10 and 12. 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.

"Chemically synthesized", as related to a sequence of DNA, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of DNA may be accomplished using well-established procedures or, automated chemical synthesis can be performed using one of a number of commercially available machines. "Synthetic genes" can be assembled from oligonucleotide building blocks that are chemically

synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed 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.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, and that 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 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 not normally found in the host organism, but 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.

An "allele" is one of several alternative forms of a gene occupying a given locus on a chromosome. When all the alleles present at a given locus on a chromosome are the same that plant is homozygous at that

locus. If the alleles present at a given locus on a chromosome differ, that plant is heterozygous at that locus.

"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, but are not limited to: promoters, translation leader sequences, 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. The promoter sequence may consist of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a DNA sequence that can stimulate promoter activity, and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. 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. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of some variation may have identical promoter activity. Promoters that cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found

in the compilation by Okamuro, J. K., and Goldberg, R. B. *Biochemistry of Plants,* 15:1-82 (1989).

"Translation leader sequence" refers to a polynucleotide sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner, R. and Foster, G. D., *Mol. Biotechnol.*, 3:225-236 (1995)).

The terms "3' non-coding sequences", "transcription terminator" and "termination sequences" 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. The use of different 3' non-coding sequences is exemplified by Ingelbrecht, I. L., et al. (*Plant Cell,* 1:671-680 (1989)).

"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. A RNA transcript is referred to as the mature RNA when it is a RNA sequence derived from post-transcriptional processing of the primary transcript. "Messenger RNA" or "mRNA" refers to the RNA that is without introns and that can be translated into protein by the cell. "cDNA" refers to a DNA that is complementary to, and synthesized from, a mRNA template using the enzyme reverse transcriptase. The cDNA can be single-stranded or converted into double-stranded form using the Klenow fragment of DNA polymerase I. "Sense" RNA refers to RNA transcript that includes the mRNA and can be

translated into protein within a cell or *in vitro*. "Antisense RNA" refers to a RNA transcript that is complementary to all or part of a target primary transcript or mRNA, and that blocks the expression of a target gene (U.S. 5,107,065; WO 99/28508). The complementarity of an antisense RNA may be with any part of the specific gene transcript, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. "Functional RNA" refers to antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes. The terms "complement" and "reverse complement" are used interchangeably herein with respect to mRNA transcripts, and are meant to define the antisense RNA of the message.

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 regulated 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 (i.e., the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in a sense or antisense orientation.

The terms "plasmid", "vector" and "cassette" 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 which is capable of introducing a promoter fragment and DNA sequence for a selected gene product along with appropriate 3' untranslated sequence into a cell. "Expression cassette" refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that allow for enhanced expression of that gene in a foreign host.

The terms "recombinant construct", "expression construct", "chimeric construct", "construct", and "recombinant DNA construct" are used interchangeably herein. A recombinant construct comprises an artificial combination of nucleic acid fragments, e.g., regulatory and coding sequences that are not found together in nature. For example, a chimeric construct 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. Such a construct may be used by itself or may be used in conjunction with a vector. If a vector is used, then the choice of vector is dependent upon the method that will be used to transform host cells as is well known to those skilled in the art. For example, a plasmid vector can be used. The skilled artisan is well aware of the genetic elements that must be present on the vector in order to successfully transform, select and propagate host cells comprising any of the isolated nucleic acid fragments of the invention. The skilled artisan will also recognize that different independent transformation events will result in different levels and patterns of expression (Jones et al., *EMBO J.* 4:2411-2418 (1985); De Almeida et al., *Mol. Gen. Genetics* 218:78-86 (1989)), and thus that multiple events must be screened in order to obtain strains or lines displaying the desired expression level and pattern. Such screening may be accomplished by Southern analysis of DNA blots (Southern, *J. Mol. Biol.,* 98:503 (1975)), Northern analysis of mRNA expression (Kroczek, *J. Chromatogr. Biomed. Appl.*, 618(1-2):133-145 (1993)), Western and/or Elisa analyses of protein expression, phenotypic analysis or GC analysis of the PUFA products, among others.

The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragments of the invention. Expression may also refer to translation of mRNA into a polypeptide.

"Mature" protein refers to a post-translationally processed polypeptide (i.e., one from which any pre- or propeptides present in the

primary translation product have been removed). "Precursor" protein refers to the primary product of translation of mRNA (i.e., with pre- and propeptides still present). Pre- and propeptides may be, but are not limited to, intracellular localization signals.

"Transformation" refers to the transfer of a nucleic acid fragment into a genome of a host organism, including both nuclear and organellar genomes, resulting in genetically stable inheritance. 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. Thus, the nucleic acid molecule used for transformation may be a plasmid that replicates autonomously, for example, or, it may integrate into the genome of the host organism. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" or "recombinant" or "transformed" organisms.

"Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of the target protein. "Co-suppression" refers to the production of sense RNA transcripts capable of suppressing the expression of identical or substantially similar foreign or endogenous genes (U.S. 5,231,020). Co-suppression constructs in plants previously have been designed by focusing on overexpression of a nucleic acid sequence having homology to an endogenous mRNA, in the sense orientation, which results in the reduction of all RNA having homology to the overexpressed sequence (Vaucheret et al., *Plant J.*, 16:651-659 (1998); Gura, *Nature,* 404:804-808 (2000)). The overall efficiency of this phenomenon is low, and the extent of the RNA reduction is widely variable. Subsequent work has described the use of "hairpin" structures that incorporate all, or part, of a mRNA encoding sequence in a complementary orientation that results in a potential "stem-loop" structure for the expressed RNA (WO 99/53050; WO 02/00904). This increases the frequency of co-suppression in the recovered transgenic plants. Another variation describes the use of plant

viral sequences to direct the suppression, or "silencing", of proximal mRNA encoding sequences (WO 98/36083). Both of these co-suppressing phenomena have not been elucidated mechanistically, although genetic evidence has begun to unravel this complex situation (Elmayan et al., *Plant Cell,* 10:1747-1757 (1998)).

The term "oleaginous" refers to those organisms that tend to store their energy source in the form of lipid (Weete, In: Fungal Lipid Biochemistry, 2nd Ed., Plenum, 1980). Generally, the cellular oil or TAG content of these 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)).

The term "oleaginous yeast" refers to those microorganisms classified as yeasts that can make oil. It is not uncommon for oleaginous microorganisms to accumulate in excess of about 25% of their dry cell weight as oil. Examples of oleaginous yeast include, but are no means limited to, the following genera: *Yarrowia, Candida, Rhodotorula, Rhodosporidium, Cryptococcus, Trichosporon* and *Lipomyces*.

The term "percent identity", as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "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 identity are designed to give the best match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs.

The "Clustal V method of alignment" corresponds to the alignment method labeled Clustal V (described by Higgins and Sharp, *CABIOS.* 5:151-153 (1989); Higgins, D.G. et al. (1992) *Comput. Appl. Biosci. 8*:189-191) and found in the MegAlign™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Sequence alignments and percent identity calculations may be performed using the MegAlign™ program. Multiple alignment of the sequences is performed using the Clustal method of alignment (Higgins and Sharp, *CABIOS,* 5:151153 (1989); Higgins, D.G. et al., *Comput. Appl. Biosci.*, 8:189-191 (1992)) with default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10), unless otherwise specified. Default parameters for pairwise alignments and calculation of percent identity of protein sequences using the Clustal V method are: KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. After alignment of the sequences using the Clustal program, it is possible to obtain a "percent identity" by viewing the "sequence distances" table in the same program.

The "Clustal W method of alignment" corresponds to the alignment method labeled Clustal W (described by Thompson et al., *Nucleic Acids Res.* 22:4673-4680 (1994)) and found in the MegAlign™ v5.07 program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Default parameters for multiple alignments and calculation of percent identity of protein sequences are GAP PENALTY=10, GAP LENGTH PENALTY=0.2, DELAY DIVERGENCE SEQS(%)=30, DNA TRANSITION WEIGHT=0.50, protein weight matrix=Gonnet series and DNA weight matrix=IUB, unless otherwise specified. Default parameters for pairwise alignments and calculation of percent identity of protein

sequences are GAP PENALTY=10, GAP LENGTH PENALTY=0.1, protein weight matrix=Gonnet 250 and DNA weight matrix=IUB, unless otherwise specified.

"BLASTN method of alignment" is an algorithm provided by the National Center for Biotechnology Information (NCBI) to compare nucleotide sequences using default parameters.

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, but are not limited to, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or any integer percentage from 50% to 100%. Indeed, any integer amino acid identity from 50% to 100% may be useful in describing the present invention, such as 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.

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.

Suitable nucleic acid fragments (isolated polynucleotides of the present invention) encode polypeptides that are at least about 85% identical to the amino acid sequences reported herein. More preferred nucleic acid fragments encode amino acid sequences that are at least about 90% identical to the amino acid sequences reported herein while most preferred are nucleic acid fragments that encode amino acid sequences that are at least about 95% identical. 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; although preferred ranges are described above, any integer percentage from 85% to 100% is useful for the purposes herein.

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.

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully 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) (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 (1987).

An Overview: Biosynthesis Of Omega Fatty Acids And Triacylglycerols

The metabolic process wherein oleic acid is converted to ω-3/ω-6 fatty acids 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 reticulim membrane. However, as seen in Figure 1 and as described below, there are often multiple alternate pathways for production of a specific ω-3/ω-6 fatty acid.

Specifically, all pathways require the initial conversion of oleic acid to LA, the first of the ω-6 fatty acids, by a Δ12 desaturase. Then, using the "Δ9 elongase/Δ8 desaturase pathway", ω-6 fatty acids are formed as follows: (1) LA is converted to EDA by a Δ9 elongase; (2) EDA is converted to DGLA by a Δ8 desaturase; and (3) DGLA is converted to ARA by a Δ5 desaturase. Alternatively, the "Δ9 elongase/Δ8 desaturase pathway" can be utilized for formation of 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 ETrA by a Δ9 elongase; (3) ETrA is converted to ETA by a Δ8 desaturase; (4) ETA is converted to EPA by a Δ5 desaturase; (5) EPA is converted to DPA by a $C_{20/22}$ elongase; and (6) DPA is converted to 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.

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

It is contemplated that the particular functionalities required to be expressed in a specific host organism for production of ω-3/ω-6 fatty acids will depend on the host cell (and its native PUFA profile and/or desaturase/elongase profile), the availability of substrate, and the desired

end product(s). For example, expression of the Δ9 elongase/Δ8 desaturase pathway may be preferred in some embodiments, as opposed to expression of the Δ6 desaturase/Δ6 elongase pathway, since PUFAs produced via the former pathway are devoid of GLA.

One skilled in the art will be able to identify various candidate genes encoding each of the enzymes desired for ω-3/ω-6 fatty acid biosynthesis. Useful desaturase and elongase sequences may be derived from any source, e.g., isolated from a natural source (from bacteria, algae, fungi, plants, animals, etc.), produced via a semi-synthetic route or synthesized *de novo*. Although the particular source of the desaturase and elongase genes introduced into the host is not critical, considerations for choosing a specific polypeptide having desaturase or elongase activity include: 1.) the substrate specificity of the polypeptide; 2.) whether the polypeptide or a component thereof is a rate-limiting enzyme; 3.) whether the desaturase or elongase is essential for synthesis of a desired PUFA; and/or 4.) co-factors required by the polypeptide. The expressed polypeptide preferably has parameters compatible with the biochemical environment of its location in the host cell (see WO 2004/101757).

In additional embodiments, it will also be useful to consider the conversion efficiency of each particular desaturase and/or elongase. More specifically, since each enzyme rarely functions with 100% efficiency to convert substrate to product, the final lipid profile of un-purified oils produced in a host cell will typically be a mixture of various PUFAs consisting of the desired ω-3/ω-6 fatty acid, as well as various upstream intermediary PUFAs. Thus, consideration of each enzyme's conversion efficiency is also an important variable when optimizing biosynthesis of a desired fatty acid that must be considered in light of the final desired lipid profile of the product.

With each of the considerations above in mind, candidate genes having the appropriate desaturase and elongase activities (e.g., Δ6 desaturases, $C_{18/20}$ elongases, Δ5 desaturases, Δ17 desaturases, Δ15 desaturases, Δ9 desaturases, Δ12 desaturases, $C_{14/16}$ elongases, $C_{16/18}$

elongases, Δ9 elongases, Δ8 desaturases, Δ4 desaturases and $C_{20/22}$ elongases) can be identified according to publicly available literature (e.g., GenBank), the patent literature, and experimental analysis of organisms having the ability to produce PUFAs. These genes will be suitable for introduction into a specific host organism, to enable or enhance the organism's synthesis of PUFAs.

Once fatty acids are synthesized within an organism (including saturated and unsaturated fatty acids and short-chain and long-chain fatty acids), they may be incorporated into triacylglycerides (TAGs). TAGs (the primary storage unit for fatty acids, including PUFAs) are formed by a series of reactions that involve: 1.) the esterification of one molecule of acyl-CoA to glycerol-3-phosphate via an acyltransferase to produce lysophosphatidic acid; 2.) the esterification of a second molecule of acyl-CoA via an acyltransferase to yield 1,2-diacylglycerol phosphate (commonly identified as phosphatidic acid); 3.) removal of a phosphate by phosphatidic acid phosphatase to yield 1,2-diacylglycerol (DAG); and 4.) the addition of a third fatty acid by the action of an acyltransferase to form TAG.

Sequence Identification Of A *Euglena gracilis* Δ8 Desaturase

Commonly owned WO 2006/012325 and WO 2006/012326 disclose a *E. gracilis* Δ8 desaturase able to desaturate EDA and EtrA (identified therein as "Eg5 and assigned SEQ ID NO:2). In the present application, the *E. gracilis* Δ8 desaturase described as "EgD8" (SEQ ID NOs:11 and 12 herein) is 100% identical and equivalent to the nucleotide and amino acid sequences of Eg5.

As is well known in the art, codon-optimization can be a useful means to further optimize the expression of an enzyme in an alternate host, since use of host-preferred codons can substantially enhance the expression of the foreign gene encoding the polypeptide. As such, a synthetic Δ8 desaturase derived from *Euglena gracilis* and codon-optimized for expression in *Yarrowia lipolytica* was also disclosed in WO 2006/012325 and WO 2006/012326 as SEQ ID NOs:112 and 113

(designated therein as "D8SF"). Specifically, 207 bp (16.4%) of the 1263 bp coding region were modified, corresponding to codon-optimization of 192 codons. Additionally, "D8SF" had one additional valine amino acid inserted between amino acid residues 1 and 2 of the wildtype Eg5; thus, the total length of the codon-optimized desaturase is 422 amino acids. Expression of the codon-optimized gene (i.e., "D8SF") in *Y. lipolytica* demonstrated more efficient desaturation of EDA to DGLA and/or ETrA to ETA than the wildtype gene (i.e., Eg5). In the present application, the synthetic Δ8 desaturase derived from *E. gracilis* and codon-optimized for expression in *Y. lipolytica* described as "EgD8S" (SEQ ID NOs:9 and 10 herein) is 100% identical and equivalent to the nucleotide and amino acid sequences of D8SF.

Engineering Targeted Mutations Within The Synthetic Δ8 Desaturase, Derived From *Euglena gracilis* And Codon-Optimized For Expression In *Yarrowia lipolytica*

Methods for synthesizing sequences and bringing sequences together are well established in the literature. Many techniques are commonly employed in the literature to obtain mutations of naturally occurring desaturase genes (wherein such mutations may include deletions, insertions and point mutations, or combinations thereof). This would permit production of a polypeptide having desaturase activity, respectively, *in vivo* with more desirable physical and kinetic parameters for function in the host cell such as a longer half-life or a higher rate of production of a desired PUFA. Or, if desired, the regions of a polypeptide of interest (i.e., a desaturase) important for enzymatic activity can be determined through routine mutagenesis, expression of the resulting mutant polypeptides and determination of their activities. All such mutant proteins and nucleotide sequences encoding them that are derived from the wildtype (i.e., SEQ ID NOs:11 and 12) and synthetic codon-optimized (SEQ ID NOs:9 and 10) Δ8 desaturase described *supra* are within the scope of the present invention.

More specifically in the invention herein, mutant sequences encoding Δ8 desaturases were synthetically engineered, by making targeted mutations within the known, functional *Euglena gracilis* Δ8 desaturase that was codon-optimized for expression in *Yarrowia lipolytica* (i.e., "EgD8S", as set forth in SEQ ID NOs:9 and 10). The effect of each mutation on the Δ8 desaturase activity of the resulting mutant EgD8S was screened. Although not to be construed as limiting to the invention herein, a mutant EgD8S enzyme (SEQ ID NO:2) was ultimately created comprising at least one amino acid mutation (and up to about 33 amino acid mutations) with respect to the synthetic codon-optimized EgD8S and having functionally equivalent Δ8 desaturase activity, using the methodology described below.

Creation Of A Topological Model And Identification Of Suitable Amino Acid Sites For Mutation

General characteristics of Δ8 desaturases, based on desaturase evolution, are well-described by P. Sperling et al. (*Prostaglandins Leukot. Essent. Fatty Acids,* 68:73–95 (2003)). Along with Δ6, Δ5 and Δ4 desaturases, Δ8 desaturases are known as long-chain PUFA "front-end" desaturases (wherein desaturation occurs between a pre-existing double bond and the carboxyl terminus of the fatty acid's acyl group, as opposed to methyl-directed desaturation). These desaturases are characterized by three histidine boxes [$H(X)_{3-4}H$ (SEQ ID NOs:166 and 167), $H(X)_{2-3}HH$ (SEQ ID NOs:168 and 169) and $H/Q(X)_{2-3}HH$ (SEQ ID NOs:170 and 171)] and are members of the cytochrome b_5 fusion superfamily, since they possess a fused cytochrome b_5 domain at their N-terminus which serves as an electron donor. The cytochrome b_5 domain also contains an absolutely conserved heme-binding motif (i.e., "HPGG") which has been shown to be necessary for enzyme activity (J.A. Napier, et al., *Prostaglandins Leukot. Essent. Fatty Acids,* 68:135–143 (2003); P. Sperling, et al., *supra*). Finally, although the crystal structure of a "front-end" desaturase is not yet available, hydropathy plots reveal 4–6

membrane spanning helices that account for nearly 30% of the amino acid sequence of these proteins.

Based on the generalizations above, the protein sequence of EgD8S (SEQ ID NO:10) was specifically analyzed to enable creation of a topological model (Figure 2). As expected, EgD8S contained two domains: an N-terminal cytochrome b_5 domain (located between amino acid residues 5 to 71 of SEQ ID NO:10) and a C-terminal desaturase domain (located between amino acid residues 79 to 406 of SEQ ID NO:10). Four membrane-spanning helices were identified at amino acid residues 88-109, 113-132, 266-283 and 287-309 of SEQ ID NO:10 (labeled as regions I, II, III and IV on Figure 2), with both the N- and C-termini located on the cytoplasmic side of the membrane. Two additional hydrophobic regions were located at amino acid residues 157-172 and 223-245. Finally, the three histidine boxes were located between amino acid residues 146-150, 183-187 and 358-362, and the conserved heme-binding motif ("HPGG") was located at amino acid residues 27-30.

Using the topological model, alignment of EgD8S with other front-end desaturases and alignment of EgD8S's cytochrome b_5 domain with other cytochrome b_5 genes, 70 sites within EgD8S were subsequently selected as possibly suitable for mutagenesis (criteria for selection are described in detail in Example 2). These sites corresponded to 126 individual amino acid mutations (i.e., 57.9% conserved amino acid substitutions and 42.1% non-conserved amino acid substitutions), as detailed in Table 7 of Example 2.

Site-Directed Mutagenesis For Creation Of EgD8S Mutants

Although a variety of approaches may be used for mutagenesis of a Δ8 desaturase enzyme, based on the strategies herein it was desirable to create specific point mutations within EgD8S using oligonucleotide-mediated site-directed mutagenesis. Furthermore, although numerous site-directed mutagenesis protocols exist (e.g., Ishii, T. M., et al., *Methods Enzymol.*, 293:53–71 (1998); Ling M. M. and B.H. Robinson, *Anal. Biochem.*, 254:157–178 (1997); Braman J. (ed.) *In Vitro Mutagenesis*

Protocols. 2nd Ed., Humania: Totowa, NJ (2002); Kunkel T. A., et al., *Methods Enzymol.*, 154:367–382 (1987); Sawano A. and Miyawaki, A. *Nucleic Acids Res.*, 28:e78 (2000)), the QuikChange® site-directed mutagenesis kit (Stratagene, La Jolla, CA) was selected for use based on its facile implementation and high efficiency. Specifically, the kit requires no specialized vectors, unique restriction sites, or multiple transformations and allows site-specific mutation in virtually any double-stranded plasmid. The basic procedure utilizes a supercoiled double-stranded DNA vector with an insert of interest and two synthetic oligonucleotide primers containing the desired mutation. The oligonucleotide primers, each complementary to opposite strands of the vector, are extended during temperature cycling by a DNA polymerase. Incorporation of the oligonucleotide primers generates a mutated plasmid containing staggered nicks. Following temperature cycling, the product is treated with *Dpn* I endonuclease (specific for methylated and hemi-methylated DNA) as a means to digest the parental DNA template and to select for newly synthesized mutant DNA. The nicked vector DNA containing the desired mutations is then transformed and propogated in an *Escherichia coli* host.

Using the techniques described above, the feasibility of engineering a synthetic EgD8S (having multiple point mutations with respect to EgD8S but maintaining functional equivalence with respect to the enzyme's Δ8 desaturase activity) was then tested. Specifically, selected individual point mutations were introduced by site-directed mutagenesis into EgD8S (within a plasmid construct comprising a chimeric FBAINm::EgD8S::XPR gene), transformed into *E. coli,* and then screened for Δ8 desaturase activity based on GC analyses.

The skilled person will be able to envision additional screens for the selection of genes encoding proteins having Δ8 desaturase activity. For example, desaturase activity may be demonstrated by assays in which a preparation containing an enzyme is incubated with a suitable form of substrate fatty acid and analyzed for conversion of this substrate to the predicted fatty acid product. Alternatively, a DNA sequence proposed to

encode a desaturase protein may be incorporated into a suitable vector construct and thereby expressed in cells of a type that do not normally have an ability to desaturate a particular fatty acid substrate. Activity of the desaturase enzyme encoded by the DNA sequence can then be demonstrated by supplying a suitable form of substrate fatty acid to cells transformed with a vector containing the desaturase-encoding DNA sequence and to suitable control cells (e.g., transformed with the empty vector alone). In such an experiment, detection of the predicted fatty acid product in cells containing the desaturase-encoding DNA sequence and not in control cells establishes the desaturase activity.

Results from the experiment described above resulted in the identification of some mutations that resulted in completely non-functional mutant Δ8 desaturases having 0% Δ8 desaturase activity (e.g., simultaneous mutation of 48V to F and 49M to L or simultaneous mutation of 304G to F and 305F to G). Despite this, ~75% of the individual mutations tested did not significantly diminish the mutant enzyme's Δ8 desaturase activity as compared to the Δ8 desaturase activity of EgD8S. More specifically, the following mutations were identified as preferred mutations, wherein Δ8 desaturase activity was functionally equivalent (i.e., 100%) or greater than that of EgD8S (i.e., SEQ ID NO:10):

Table 4

Initial Preferred Mutations Within EgD8S

Mutation Site	Sequence Mutations With Respect to EgD8S (SEQ ID NO:10)	% Activity*
M3	16T to K, 17T to V	100%
M8	66P to Q, 67S to A	100%
M12	407A to S, 408V to Q	100%
M14	416Q to V, 417P to Y	100%
M16	108S to L	100%
M19	122V to S	100%
M38	54A to G, 55F to Y	100%
M39	64I to L, 65N to D	100%
M40	69E to D, 70L to V	100%
M41	75A to G, 76V to L	100%

M45	117G to A, 118Y to F	100%
M46	132V to L, 133L to V	100%
M49	297F to V, 298V to L	100%
M50	309I to V, 310V to I	100%
M51	347I to L, 348T to S	100%
M51B	346I to V, 347I to L, 348T to S	100%
M53	9L to V	100%
M54	19D to E, 20V to I	100%
M58	65N to Q	100%
M63	279T to L, 280L to T	100%
M68	162L to V, 163V to L	100%
M69	170G to A, 171L to V	100%
M70	418A to G, 419G to A	100%
M2	12T to V	110%
M22	127Y to Q	110%
M26	293L to M	110%
M6	59K to L	110%
M1	4S to A, 5K to S	115%
M21	125Q to H, 126M to L	120%
M15	422L to Q	125%

* "% Activity" refers to the Δ8 desaturase activity of each mutant EgD8S with respect to the Δ8 desaturase activity of EgD8S, as set forth as SEQ ID N0:10.

It will be appreciated by one of skill in the art that the useful mutant Δ8 desaturases of the present invention are not limited to the 30 mutation combinations described above. For example, although the mutation site described as M3 includes two specific amino acid mutations (i.e., 16T to K and 17T to V), one skilled in the art will expect that a single mutation of either 16T to K or 17T to V will have utility in the design of a mutant Δ8 desaturase whose Δ8 desaturase activity is at least about functionally equivalent to the Δ8 desaturase activity of the synthetic codon-optimized EgD8S. Thus, in actuality, Table 4 presents 52 single amino acid mutations that are useful for the purposes herein, in the design of a mutant Δ8 desaturase having Δ8 desaturase activity that is at least about functionally equivalent to the Δ8 desaturase activity of SEQ ID NO:10.

Based on the results above, experimental work was continued in an effort to "stack" appropriate individual amino acid mutations within the synthetic codon-optimized EgD8S sequence. This resulted in creation of a mutant Δ8 desaturase as set forth in SEQ ID NO:2 having "n" amino acid mutations, wherein "n" is any integer from 1 to 33 inclusive (i.e., 1, 2, 3, 4,

5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33), and having Δ8 desaturase activity comparable to that of EgD8S. Specifically, SEQ ID NO:2 comprises at least one mutation selected from the group consisting of: 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 54A to G, 55F to Y, 66P to Q, 67S to A, 108S to L, 117G to A, 118Y to F, 120L to M, 121M to L, 125Q to H, 126M to L, 132V to L, 133L to V, 162L to V, 163V to L, 170G to A, 171L to V, 279T to L, 280L to T, 293L to M, 346I to V, 347I to L, 348T to S, 407A to S, 408V to Q, 418A to G, 419G to A and 422L to Q, wherein the mutations are defined with respect to the synthetic codon-optimized sequence of EgD8S (i.e., SEQ ID NO:10); wherein SEQ ID NO:2 is not 100% identical to SEQ ID NO:10; and wherein the mutant EgD8S is at least about functionally equivalent to EgD8S (SEQ ID NO:10). It will be appreciated by the skilled person that each of the above mutations can be used in any combination with one another. And, all such mutant proteins and nucleotide sequences encoding them that are derived from EgD8 and/or EgD8S as described herein are within the scope of the present invention. In more preferred embodiments, the mutant EgD8S has at least about 10-18 conservative and non-conservative amino acid substitutions (i.e., mutations) with respect to the synthetic codon-optimized sequence of EgD8S, more preferably at least about 19-25 conservative and non-conservative amino acid substitutions, and most preferably at least about 26-33 conservative and non-conservative amino acid substitutions with respect to the synthetic codon-optimized sequence of EgD8S (i.e., SEQ ID NO:10). Thus, for example, in one preferred embodiment mutant EgD8S-23 (SEQ ID NO:4) comprises the following 24 amino acid mutations with respect to the synthetic codon-optimized EgD8S sequence set forth as SEQ ID NO:10: 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 108S to L, 117G to A, 118Y to F, 120L to M, 121M to L, 125Q to H, 126M to L, 132V to L, 133 L to V, 162L to V, 163V to L, 293L to M, 407A to S, 408V to Q, 418A to G, 419G to A and 422L to Q. Pairwise alignment of the mutant EgD8S-23 amino acid sequence to the synthetic

codon-optimized sequence of SEQ ID NO:10 using default parameters of Vector NTI®'s AlignX program (Invitrogen Corporation, Carlsbad, CA) revealed 94.3% sequence identity and 97.9% consensus between the two proteins over a length of 422 amino acids.

In another preferred embodiment, mutant EgD8S-013 (SEQ ID NO:6) comprises the following 28 amino acid mutations with respect to the synthetic codon-optimized EgD8S sequence set forth as SEQ ID NO:10: 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 54A to G, 55F to Y, 66P to Q, 67S to A, 108S to L, 117G to A, 118Y to F, 120L to M, 121M to L, 125Q to H, 126M to L, 132V to L, 133L to V, 162L to V, 163V to L, 170G to A, 171L to V, 293L to M, 407A to S, 408V to Q, 418A to G, 419G to A and 422L to Q. Pairwise alignment of the mutant EgD8S-013 amino acid sequence to the synthetic codon-optimized sequence of SEQ ID NO:10 using default parameters of Vector NTI®'s AlignX program revealed 93.4% sequence identity and 97.9% consensus between the two proteins over a length of 422 amino acids.

In another preferred embodiment, mutant EgD8S-015 (SEQ ID NO:8) comprises the following 31 amino acid mutations with respect to the synthetic codon-optimized EgD8S sequence set forth as SEQ ID NO:10: 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 54A to G, 55F to Y, 66P to Q, 67S to A, 108S to L, 117G to A, 118Y to F, 120L to M, 121M to L, 125Q to H, 126M to L, 162L to V, 163V to L, 170G to A, 171L to V, 293L to M, 279T to L, 280L to T, 346I to V, 347I to L, 348T to S, 407A to S, 408V to Q, 418A to G, 419G to A and 422L to Q. Pairwise alignment of the mutant EgD8S-015 amino acid sequence to the synthetic codon-optimized sequence of SEQ ID NO:10 using default parameters of Vector NTI®'s AlignX program revealed 92.7% sequence identity and 97.4% consensus between the two proteins over a length of 422 amino acids.

Thus, in one embodiment, the present invention concerns an isolated polynucleotide comprising:

(a) a nucleotide sequence encoding a mutant polypeptide having

Δ8 desaturase activity, wherein the mutant polypeptide has an amino acid sequence as set forth in SEQ ID NO:2, wherein:

(i) SEQ ID NO:2 comprises at least one mutation selected from the group consisting of: 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 54A to G, 55F to Y, 66P to Q, 67S to A, 108S to L, 117G to A, 118Y to F, 120L to M, 121M to L, 125Q to H, 126M to L, 132V to L, 133L to V, 162L to V, 163V to L, 170G to A, 171L to V, 279T to L, 280L to T, 293L to M, 346I to V, 347I to L, 348T to S, 407A to S, 408V to Q, 418A to G, 419G to A and 422L to Q, wherein the mutations are defined with respect to the synthetic codon-optimized sequence of EgD8S (i.e., SEQ ID NO:10); and,

(ii) SEQ ID NO:2 is not identical to SEQ ID NO:10; or,

(b) a complement of the nucleotide sequence, wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

In further preferred embodiments, the Δ8 desaturase activity of SEQ ID NO:2, as described above, is at least about functionally equivalent to the Δ8 desaturase activity of SEQ ID NO:10.

It is contemplated that a mutant Δ8 desaturase having Δ8 desaturase activity at least about functionally equivalent to that of EgD8 (SEQ ID NO:12) or EgD8S (SEQ ID NO:10) may be evolved as set forth in SEQ ID NO:2, wherein: (i) SEQ ID NO:2 comprises at least one mutation selected from the group consisting of: 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 54A to G, 55F to Y, 66P to Q, 67S to A, 108S to L, 117G to A, 118Y to F, 120L to M, 121M to L, 125Q to H, 126M to L, 132V to L, 133L to V, 162L to V, 163V to L, 170G to A, 171L to V, 279T to L, 280L to T, 293L to M, 346I to V, 347I to L, 348T to S, 407A to S, 408V to Q, 418A to G, 419G to A and 422L to Q, wherein the mutations are defined with respect to the synthetic codon-optimized sequence of EgD8S (i.e., SEQ ID

NO:10); and, (ii) SEQ ID NO:2 is not 100% identical to SEQ ID NO:10. Furthermore, it will be appreciated that the invention encompasses not only the specific mutations described above, but also those that allow for the substitution of chemically equivalent amino acids. So, for example, where a substitution of an amino acid with the aliphatic, nonpolar amino acid alanine is made, it will be expected that the same site may be substituted with the chemically equivalent amino acid serine.

In other embodiments, any of the Δ8 desaturase nucleic acid fragments described herein may be used for creation of new and improved fatty acid desaturases by domain swapping, wherein a functional domain from any of the Δ8 desaturase nucleic acid fragments is exchanged with a functional domain in an alternate desaturase gene to thereby result in a novel protein.

Methods For Production Of Various ω-3 And/Or ω-6 Fatty Acids

It is expected that introduction of chimeric genes encoding the Δ8 desaturases described herein, under the control of the appropriate promoters will result in increased production of DGLA and/or ETA in the transformed host organism, respectively. As such, the present invention encompasses a method for the direct production of PUFAs comprising exposing a fatty acid substrate (i.e., EDA or ETrA) to the desaturase enzymes described herein (i.e., those mutants derived from EgD8 or EgD8S, or homologs thereof), such that the substrate is converted to the desired fatty acid product (i.e., DGLA and/or ETA).

More specifically, it is an object of the present invention to provide a method for the production of DGLA in a host cell (e.g., oleaginous yeast, soybean), wherein the host cell comprises:

a.) a recombinant construct encoding a Δ8 desaturase polypeptide as set forth in SEQ ID NO:2 wherein SEQ ID NO:2 is not 100% identical to SEQ ID NO:10; and,

b.) a source of EDA;

wherein the host cell is grown under conditions such that the Δ8 desaturase is expressed and the EDA is converted to DGLA, and wherein the DGLA is optionally recovered.

The person of skill in the art will recognize that the broad substrate range of the Δ8 desaturase may additionally allow for the use of the enzyme for the conversion of ETrA to ETA. Accordingly the invention provides a method for the production of ETA, wherein the host cell comprises:

a.) a recombinant construct encoding a Δ8 desaturase polypeptide as set forth in SEQ ID NO:2 wherein SEQ ID NO:2 is not 100% identical to SEQ ID NO:10; and,

b.) a source of ETrA;

wherein the host cell is grown under conditions such that the Δ8 desaturase is expressed and the ETrA is converted to ETA, and wherein the ETA is optionally recovered.

Alternatively, each Δ8 desaturase gene and its corresponding enzyme product described herein can be used indirectly for the production of ω-3 fatty acids (see WO 2004/101757). Indirect production of ω-3/ω-6 PUFAs occurs wherein the fatty acid substrate is converted indirectly into the desired fatty acid product, via means of an intermediate step(s) or pathway intermediate(s). Thus, it is contemplated that the Δ8 desaturases described herein (i.e., those mutants derived from EgD8 or EgD8S, or homologs thereof) may be expressed in conjunction with additional genes encoding enzymes of the PUFA biosynthetic pathway (e.g., Δ6 desaturases, $C_{18/20}$ elongases, Δ17 desaturases, Δ15 desaturases, Δ9 desaturases, Δ12 desaturases, $C_{14/16}$ elongases, $C_{16/18}$ elongases, Δ9 elongases, Δ5 desaturases, Δ4 desaturases, $C_{20/22}$ elongases) to result in higher levels of production of longer-chain ω-3/ω-6 fatty acids (e.g., ARA, EPA, DPA and DHA). In preferred embodiments, the Δ8 desaturases of the present invention will minimally be expressed in conjunction with a Δ9 elongase (e.g., a Δ9 elongase as set forth in SEQ ID NO:173 or SEQ ID

NO:176). However, the particular genes included within a particular expression cassette will depend on the host cell (and its PUFA profile and/or desaturase/elongase profile), the availability of substrate and the desired end product(s).

Microbial Expression Systems, Cassettes & Vectors, And Transformation

The Δ8 desaturase genes and gene products described herein (i.e., those mutants derived from EgD8 or EgD8S, or homologs thereof) may also be produced in heterologous microbial host cells, particularly in the cells of oleaginous yeasts (e.g., *Yarrowia lipolytica*).

Microbial expression systems and expression vectors containing regulatory sequences that direct high level expression of foreign proteins are well known to those skilled in the art. Any of these could be used to construct chimeric genes for production of any of the gene products of the instant sequences. These chimeric genes could then be introduced into appropriate microorganisms via transformation to provide high-level expression of the encoded enzymes.

Vectors or DNA cassettes useful for the transformation of suitable microbial host cells are well known in the art. The specific choice of sequences present in the construct is dependent upon the desired expression products (*supra*), the nature of the host cell and the proposed means of separating transformed cells versus non-transformed cells. 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 (e.g., a promoter) 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 microbial host cell, although it is to be understood that such control regions need not be derived from the genes native to the specific species chosen as a production host.

Initiation control regions or promoters which are useful to drive expression of the instant Δ8 desaturase ORFs in the desired microbial host cell are numerous and familiar to those skilled in the art. Virtually any promoter capable of directing expression of these genes in the selected host cell is suitable for the present invention. Expression in a microbial host cell can be accomplished in a transient or stable fashion. Transient expression can be accomplished by inducing the activity of a regulatable promoter operably linked to the gene of interest. Stable expression can be achieved by the use of a constitutive promoter operably linked to the gene of interest. As an example, when the host cell is yeast, transcriptional and translational regions functional in yeast cells are provided, particularly from the host species (e.g., see WO 2004/101757 and WO 2006/052870 for preferred transcriptional initiation regulatory regions for use in *Yarrowia lipolytica*). Any one of a number of regulatory sequences can be used, depending upon whether constitutive or induced transcription is desired, the efficiency of the promoter in expressing the ORF of interest, the ease of construction and the like.

Nucleotide sequences surrounding the translational initiation codon 'ATG' have been found to affect expression in yeast cells. If the desired polypeptide is poorly expressed in yeast, the nucleotide sequences of exogenous genes can be modified to include an efficient yeast translation initiation sequence to obtain optimal gene expression. For expression in yeast, this can be done by site-directed mutagenesis of an inefficiently expressed gene by fusing it in-frame to an endogenous yeast gene, preferably a highly expressed gene. Alternatively, one can determine the consensus translation initiation sequence in the host and engineer this sequence into heterologous genes for their optimal expression in the host of interest.

The termination region can be derived from the 3' region of the gene from which the initiation region was obtained or from a different gene. A large number of termination regions are known and function satisfactorily in a variety of hosts (when utilized both in the same and

different genera and species from where they were derived). The termination region usually is selected more as a matter of convenience rather than because of any particular property. Preferably, when the microbial host is a yeast cell, the termination region is 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. Termination control regions may also be derived from various genes native to the preferred hosts. Optionally, a termination site may be unnecessary; however, it is most preferred if included. Although not intended to be limiting, termination regions useful in the disclosure herein include: ~100 bp of the 3' region of the *Yarrowia lipolytica* extracellular protease (XPR; GenBank Accession No. M17741); the acyl-coA oxidase (Aco3: GenBank Accession No. AJ001301 and No. CAA04661; Pox3: GenBank Accession No. XP_503244) terminators; the Pex20 (GenBank Accession No. AF054613) terminator; the Pex16 (GenBank Accession No. U75433) terminator; the *Lip1* (GenBank Accession No. Z50020) terminator; the *Lip2* (GenBank Accession No. AJ012632) terminator; and the 3-oxoacyl-coA thiolase (OCT; GenBank Accession No. X69988) terminator.

As one of skill in the art is aware, merely inserting a gene into a cloning vector does not ensure that it will be successfully expressed at the level needed. 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 aspects of transcription, translation, protein stability, oxygen limitation and secretion from the microbial host cell. More specifically, some of the molecular features that have been manipulated to control gene expression include: 1.) the nature of the relevant transcriptional promoter and terminator sequences; 2.) the number of copies of the cloned gene and whether the gene is plasmid-borne or integrated into the genome of the host cell; 3.) the final cellular location of the synthesized foreign protein; 4.) the efficiency of translation

and correct folding of the protein in the host organism; 5.) the intrinsic stability of the mRNA and protein of the cloned gene within the host cell; and 6.) 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 types of modifications are encompassed in the present invention, as means to further optimize expression of the mutant Δ8 desaturases described herein.

Once the DNA encoding a polypeptide suitable for expression in an appropriate microbial host cell (e.g., oleaginous yeast) has been obtained (e.g., a chimeric gene comprising a promoter, ORF and terminator), it is placed in a plasmid vector capable of autonomous replication in a host cell, or it 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 within 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.

In the present invention, the preferred method of expressing genes in *Yarrowia lipolytica* is by integration of linear DNA into the genome of the host; and, integration into multiple locations within the genome can be particularly useful when high level expression of genes are desired [e.g., in the *Ura3* locus (GenBank Accession No. AJ306421), the *Leu2* gene locus (GenBank Accession No. AF260230), the *Lys5* gene (GenBank Accession No. M34929), the *Aco2* gene locus (GenBank Accession No. AJ001300), the *Pox3* gene locus (Pox3: GenBank Accession No. XP_503244; or, Aco3: GenBank Accession No. AJ001301), the Δ12 desaturase gene locus (WO2004/104167), the *Lip1* gene locus (GenBank Accession No. Z50020) and/or the *Lip2* gene locus (GenBank Accession No. AJ012632)].

Advantageously, the *Ura3* gene can be used repeatedly in combination with 5-fluoroorotic acid (5-fluorouracil-6-carboxylic acid monohydrate; "5-FOA") selection (*infra*), to readily permit genetic

modifications to be integrated into the *Yarrowia* genome in a facile manner.

Where two or more genes are expressed from separate replicating vectors, it is desirable that each vector has a different means of selection and should lack homology to the other 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 of interest may be introduced into a microbial host cell by any standard technique. These techniques include transformation (e.g., lithium acetate transformation [*Methods in Enzymology*, 194:186-187 (1991)]), protoplast fusion, bolistic impact, electroporation, microinjection, or any other method that introduces the gene of interest into the host cell. More specific teachings applicable for oleaginous yeasts (i.e., *Yarrowia lipolytica*) include U.S. 4,880,741 and U.S. 5,071,764 and Chen, D. C. et al. (*Appl. Microbiol. Biotechnol.*, 48(2):232-235 (1997)).

For convenience, a host cell that has been manipulated by any method to take up a DNA sequence (e.g., an expression cassette) will be referred to as "transformed" or "recombinant" herein. 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 various selection techniques, as described in WO2004/101757 and WO 2006/052870. Preferred selection methods for use herein are resistance to kanamycin, hygromycin and the amino glycoside G418, as well as ability to grow on media lacking uracil, leucine, lysine, tryptophan or histidine. In alternate embodiments, 5-FOA is used for selection of yeast Ura- mutants. The

compound is toxic to yeast cells that possess a functioning URA3 gene encoding orotidine 5'-monophosphate decarboxylase (OMP decarboxylase); thus, based on this toxicity, 5-FOA is especially useful for the selection and identification of Ura$^-$ mutant yeast strains (Bartel, P.L. and Fields, S., Yeast 2-Hybrid System, Oxford University: New York, v. 7, pp 109-147, 1997). More specifically, one can first knockout the native Ura3 gene to produce a strain having a Ura- phenotype, wherein selection occurs based on 5-FOA resistance. Then, a cluster of multiple chimeric genes and a new Ura3 gene can be integrated into a different locus of the *Yarrowia* genome to thereby produce a new strain having a Ura+ phenotype. Subsequent integration produces a new Ura3- strain (again identified using 5-FOA selection), when the introduced Ura3 gene is knocked out. Thus, the Ura3 gene (in combination with 5-FOA selection) can be used as a selection marker in multiple rounds of transformation.

Following transformation, substrates suitable for the instant Δ8 desaturases (and, optionally other PUFA enzymes that are co-expressed within the host cell) may be produced by the host either naturally or transgenically, or they may be provided exogenously.

Microbial host cells for expression of the instant genes and nucleic acid fragments may include hosts that grow on a variety of feedstocks, including simple or complex carbohydrates, fatty acids, organic acids, oils and alcohols, and/or hydrocarbons over a wide range of temperature and pH values. The genes of the present invention have been isolated for expression in an oleaginous yeast (and in particular *Yarrowia lipolytica*). It is contemplated that because transcription, translation and the protein biosynthetic apparatus is highly conserved, any bacteria, yeast, algae, oomycete and/or fungus will be a suitable microbial host for expression of the present nucleic acid fragments.

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

preferably greater than about 40% of the cellular dry weight. 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 yeasts 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 #20362, ATCC #8862, ATCC #18944, ATCC #76982 and/or LGAM S(7)1 (Papanikolaou S., and Aggelis G., *Bioresour. Technol.* 82(1):43-9 (2002)).

Historically, various strains of *Y. lipolytica* have been used for the manufacture and production of: isocitrate lyase; lipases; polyhydroxyalkanoates; citric acid; erythritol; 2-oxoglutaric acid; γ-decalactone; γ-dodecalatone; and pyruvic acid. Specific teachings applicable for engineering ARA, EPA and DHA production in *Y. lipolytica* are provided in U.S. Patent Applications No. 11/264784 (WO 2006/055322), No. 11/265761 (WO 2006/052870) and No. 11/264737 (WO 2006/052871), respectively.

Other preferred microbial hosts include oleaginous bacteria, algae, Oomycetes and other fungi; and, within this broad group of microbial hosts, of particular interest are microorganisms that synthesize ω-3/ω-6 fatty acids (or those that can be genetically engineered for this purpose [e.g., other yeast such as *Saccharomyces cerevisiae*]). Thus, for example, transformation of *Mortierella alpina* (which is commercially used for production of ARA) with any of the present Δ8 desaturase genes under the control of inducible or regulated promoters could yield a transformant organism capable of synthesizing increased quantities of EDA. 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 are disclosed in U.S. 7,001,772.

Based on the teachings described above, in one embodiment this invention is drawn to a method of producing either DGLA or ETA, respectively, comprising:

a) providing an oleaginous yeast comprising:

(i) a nucleotide sequence encoding a mutant polypeptide having Δ8 desaturase activity, wherein the mutant polypeptide has an amino acid sequence as set forth in SEQ ID NO:2 and wherein SEQ ID NO:2 is not identical to SEQ ID NO:10; and,

(ii) a source of desaturase substrate consisting of either EDA or ETrA, respectively; and,

b) growing the yeast of step (a) in the presence of a suitable fermentable carbon source wherein the gene encoding the Δ8 desaturase polypeptide is expressed and EDA is converted to DGLA or ETrA is converted to ETA, respectively; and,

c) optionally recovering the DGLA or ETA, respectively, of step (b).

Substrate feeding may be required.

Of course, since naturally produced PUFAs in oleaginous yeast are limited to 18:2 fatty acids (i.e., LA), and less commonly, 18:3 fatty acids (i.e., ALA), in more preferred embodiments of the present invention the oleaginous yeast will be genetically engineered to express multiple enzymes necessary for long-chain PUFA biosynthesis (thereby enabling production of e.g., ARA, EPA, DPA and DHA), in addition to the Δ8 desaturase described herein.

Specifically, in one embodiment this invention concerns an oleaginous yeast comprising:

a) a first recombinant DNA construct comprising an isolated

polynucleotide encoding a mutant Δ8 desaturase polypeptide, operably linked to at least one regulatory sequence; and,

b) at least one additional recombinant DNA construct comprising an isolated polynucleotide, operably linked to at least one regulatory sequence, encoding a polypeptide selected from the group consisting of: a Δ4 desaturase, a Δ5 desaturase, Δ6 desaturase, a Δ9 desaturase, a Δ12 desaturase, a Δ15 desaturase, a Δ17 desaturase, a Δ9 elongase, a $C_{14/16}$ elongase, a $C_{16/18}$ elongase, a $C_{18/20}$ elongase and a $C_{20/22}$ elongase.

In particularly preferred embodiments, the at least one additional recombinant DNA construct encodes a polypeptide having Δ9 elongase activity, e.g., the Δ9 elongase isolated and/or derived from *Isochrysis galbana* (GenBank Accession No. AF390174) and set forth in SEQ ID NOs:172-174 or the Δ9 elongase isolated and/or derived from *Euglena gracilis* and set forth in SEQ ID NOs:175-177.

Metabolic Engineering Of Omega-3 And/Or Omega-6 Fatty Acid Biosynthesis In Microbes

Methods for manipulating biochemical pathways are well known to those skilled in the art; and, it is expected that numerous manipulations will be possible to maximize ω-3 and/or ω-6 fatty acid biosynthesis in oleaginous yeasts, and particularly, in *Yarrowia lipolytica*. This may require metabolic engineering directly within the PUFA biosynthetic pathway or additional coordinated manipulation of various other metabolic pathways.

In the case of manipulations within the PUFA biosynthetic pathway, it may be desirable to increase the production of LA to enable increased production of ω-6 and/or ω-3 fatty acids. Introducing and/or amplifying genes encoding Δ9 and/or Δ12 desaturases may accomplish this. To maximize production of ω-6 unsaturated fatty acids, it is well known to one skilled in the art that production is favored in a host microorganism that is

substantially free of ALA; thus, preferably, the host is selected or obtained by removing or inhibiting Δ15 or ω-3 type desaturase activity that permits conversion of LA to ALA. Alternatively, it may be desirable to maximize production of ω-3 fatty acids (and minimize synthesis of ω-6 fatty acids). In this example, one could utilize a host microorganism wherein the Δ12 desaturase activity that permits conversion of oleic acid to LA is removed or inhibited; subsequently, appropriate expression cassettes would be introduced into the host, along with appropriate substrates (e.g., ALA) for conversion to ω-3 fatty acid derivatives of ALA (e.g., STA, ETrA, ETA, EPA, DPA, DHA).

In alternate embodiments, biochemical pathways competing with the ω-3 and/or ω-6 fatty acid biosynthetic pathways for energy or carbon, or native PUFA biosynthetic pathway enzymes that interfere with production of a particular PUFA end-product, may be eliminated by gene disruption or down-regulated by other means (e.g., antisense mRNA).

Detailed discussion of manipulations within the PUFA biosynthetic pathway as a means to increase ARA, EPA or DHA (and associated techniques thereof) are presented in WO 2006/055322, WO 2006/052870 and WO 2006/052871, respectively, as are desirable manipulations in the TAG biosynthetic pathway and the TAG degradation pathway (and associated techniques thereof).

Within the context of the present invention, it may be useful to modulate the expression of the fatty acid biosynthetic pathway by any one of the strategies described above. For example, the present invention provides methods whereby genes encoding key enzymes in the Δ9 elongase/ Δ8 desaturase biosynthetic pathway are introduced into oleaginous yeasts for the production of ω-3 and/or ω-6 fatty acids. It will be particularly useful to express the present the Δ8 desaturase genes in oleaginous yeasts that do not naturally possess ω-3 and/or ω-6 fatty acid biosynthetic pathways and coordinate the expression of these genes, to maximize production of preferred PUFA products using various means for metabolic engineering of the host organism.

Microbial Fermentation Processes For PUFA Production

The transformed microbial host cell is grown under conditions that optimize expression of chimeric desaturase and elongase genes and produce the greatest and the most economical yield of the preferred PUFAs. In general, media conditions that may be optimized include the type and amount of carbon source, the type and amount of nitrogen source, the carbon-to-nitrogen ratio, 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. Microorganisms of interest (i.e., *Yarrowia lipolytica*) are generally grown in complex media (e.g., 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 in the present invention must contain a suitable carbon source. Suitable carbon sources are taught in WO 2004/101757. Although it is contemplated that the source of carbon utilized in the present invention may encompass a wide variety of carbon-containing sources, preferred carbon sources are sugars, glycerol, and/or fatty acids. Most preferred is glucose and/or fatty acids containing between 10-22 carbons.

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 microorganism and promotion of the enzymatic pathways necessary for PUFA production. Particular attention is given to several metal ions (e.g., Mn^{+2}, Co^{+2}, Zn^{+2}, 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 in the present invention 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 the transformant host cells 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 PUFAs in *Yarrowia lipolytica*. This approach is described in WO 2004/101757, as are various suitable fermentation process designs (i.e., batch, fed-batch and continuous) and considerations during growth.

Purification And Processing Of PUFA Oils

PUFAs may be found in the host microorganisms and plants as free fatty acids or in esterified forms such as acylglycerols, phospholipids, sulfolipids or glycolipids, and 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 PUFAs may include extraction with organic solvents, sonication, supercritical fluid extraction (e.g., using carbon dioxide), saponification and physical means such as presses, or combinations thereof. One is referred to the teachings of WO 2004/101757 for additional details.

DESCRIPTION OF PREFERRED EMBODIMENTS

One object of the present invention is the synthesis of suitable Δ8 desaturases that will enable expression of the Δ9 elongase/ Δ8 desaturase pathway in plants and oleaginous yeast.

In commonly owned WO 2006/012325 and WO 2006/012326 [US2005-0287652] applicants describe the isolation of a Δ8 desaturase from *Euglena gracilis* ("Eg5"), and the enzyme's functional characterization upon expression in *Saccharomyces cerevisiae*. The wildtype Eg5 sequence was additionally codon-optimized for expression in *Yarrowia lipolytica,* resulting in the synthesis of a synthetic, functional codon-optimized Δ8 desaturase designated as "D8SF". Upon co-expression of D8SF with a codon-optimized Δ9 elongase (derived from *Isochrysis galbana* (GenBank Accession No. 390174) in *Yarrowia lipolytica,* 6.4% DGLA (with no co-synthesis of GLA) was demonstrated (Example 12 in WO 2006/012325 and WO 2006/012326 [US2005-0287652-A1]).

In the present Application, the synthetic codon-optimized Δ8 desaturase designated as "D8SF" (and designated herein as EgD8S") was subjected to targeted mutations. Ultimately, a mutant EgD8S enzyme (SEQ ID NO:2) was created comprising at least one amino acid mutation (and up to about 33 amino acid mutations) with respect to the synthetic codon-optimized EgD8S, wherein: (i) the at least one mutation is selected from the group consisting of: 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 54A to G, 55F to Y, 66P to Q, 67S to A, 108S to L, 117G to A, 118Y to F, 120L to M, 121M to L, 125Q to H, 126M to L, 132V to L, 133L to V, 162L to V, 163V to L, 170G to A, 171L to V, 279T to L, 280L to T, 293L to M, 346I to V, 347I to L, 348T to S, 407A to S, 408V to Q, 418A to G, 419G to A and 422L to Q, wherein the mutations are defined with respect to the synthetic codon-optimized sequence set forth in SEQ ID NO:10; (ii) SEQ ID NO:2 is not identical to SEQ ID NO:10; and, (iii) SEQ ID NO:2 is at least about functionally equivalent to SEQ ID NO:10.

Given the teaching of the present application the skilled person will recognize the commercial utility of the recombinant genes of the present

invention encoding Δ8 desaturases, to enable production of a variety of ω-3 and/or ω-6 PUFAs via expression of a heterologous Δ9 elongase/Δ8 desaturase pathway.

EXAMPLES

The present invention is further defined in the following Examples. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions. Thus, various modifications of the invention in addition to those shown and described herein will be apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims.

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 (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), GIBCO/BRL (Gaithersburg, MD), or Sigma Chemical Company (St. Louis, MO), unless otherwise specified.

General molecular cloning was performed according to standard methods (Sambrook et al., *supra*). Oligonucleotides were synthesized by Sigma-Genosys (Spring, TX). DNA sequence was generated on an ABI Automatic sequencer using dye terminator technology (U.S. 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. Comparisons of genetic sequences were accomplished using DNASTAR software (DNAStar Inc., Madison, WI). Alternatively, manipulations of genetic sequences were accomplished using the suite of programs available from the Genetics Computer Group Inc. (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI). The GCG program "Pileup" was used with the gap creation default value of 12, and the gap extension default value of 4. The GCG "Gap" or "Bestfit" programs were used with the default gap creation penalty of 50 and the default gap extension penalty of 3. Unless otherwise stated, in all other cases GCG program default parameters were used.

The meaning of abbreviations is as follows: "sec" means second(s), "min" means minute(s), "hr" 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) and "kB" means kilobase(s). Parts and percentages are by weight and degrees are Celsius.

Transformation And Cultivation Of *Yarrowia lipolytica*

Yarrowia lipolytica strains ATCC #20362, #76982 and #90812 were purchased from the American Type Culture Collection (Rockville, MD). *Y. lipolytica* strains were usually grown at 28 °C on YPD agar (1% yeast extract, 2% bactopeptone, 2% glucose, 2% agar).

Transformation of *Yarrowia lipolytica* was performed according to the method of Chen, D. C. et al. (*Appl. Microbiol. Biotechnol.*, 48(2):232-235 (1997)), unless otherwise noted. Briefly, *Yarrowia* was streaked onto a YPD plate and grown at 30 °C for approximately 18 hr. Several large loopfuls of cells were scraped from the plate and resuspended in 1 mL of transformation buffer containing: 2.25 mL of 50% PEG, average MW 3350; 0.125 mL of 2 M Li acetate, pH 6.0; 0.125 mL of 2 M DTT; and 50 μg sheared salmon sperm DNA. Then, approximately 500 ng of linearized plasmid DNA was incubated in 100 μl of resuspended cells, and maintained at 39 °C for 1 hr with vortex mixing at 15 min intervals. The cells were plated onto selection media plates and maintained at 30 °C for 2 to 3 days.

For selection of transformants, minimal medium ("MM") was generally used; the composition of MM is as follows: 0.17% yeast nitrogen base (DIFCO Laboratories, Detroit, MI) without ammonium sulfate or amino acids, 2% glucose, 0.1% proline, pH 6.1). Supplements of uracil or leucine were added as appropriate to a final concentration of 0.01% (thereby producing "MMU" or "MMLeu" selection media, respectively, each prepared with 20 g/L agar).

Alternatively, transformants were selected on 5-fluoroorotic acid ("FOA"; also 5-fluorouracil-6-carboxylic acid monohydrate) selection media, comprising: 0.17% yeast nitrogen base (DIFCO Laboratories, Detroit, MI) without ammonium sulfate or amino acids, 2% glucose, 0.1% proline, 75 mg/L uracil, 75 mg/L uridine, 900 mg/L FOA (Zymo Research Corp., Orange, CA) and 20 g/L agar.

Fatty Acid Analysis Of *Yarrowia lipolytica*

For fatty acid 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 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* culture (3 mL) was harvested, washed once in distilled water, and dried under vacuum in a Speed-Vac for 5-10 min. Sodium methoxide (100 µl of 1 %) was added to the sample, and then the sample was vortexed and rocked for 20 min. 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 as described above.

Example 1

Development Of A Topological Model For EgD8S

BLASTP analysis showed that EgD8S contained two domains: an N-terminal cytochrome b_5 domain (located between amino acid residues 5 to 71 of SEQ ID NO:10) and a C-terminal desaturase domain (located between amino acid residues 79 to 406 of SEQ ID NO:10). In order to mutate the amino acid sequence of EgD8S without negatively affecting the Δ8 desaturase activity, a topological model (Figure 2) was developed based on the logic and analyses below.

First, the TMHMM program ("Prediction of transmembrane helices in proteins"; TMHMM Server v. 2.0, Center for Biological Sequence Analysis, BioCentrum-DTU, Technical University of Denmark, DK-2800 Lyngby, Denmark) predicted that EgD8S had four membrane-spanning helices (amino acid residues 113-132, 223-245, 266-283 and 287-309), with both the N- and C- termini located on the cytoplasmic side of the membrane.

The membrane-bound fatty acid desaturases belong to a super-family of membrane di-iron proteins that feature three His-rich motifs: $HX_{(3-4)}H$ (SEQ ID NOs:166 and 167), $HX_{(2-3)}HH$ (SEQ ID NOs:168 and 169) and $(H/Q)X_{(2-3)}HH$ (SEQ ID NOs:170 and 171). These His-rich residues have been predicted to be located in the cytoplasmic face of the membrane and have been shown to be important for enzyme activity (Shanklin, J. et al., *Biochemistry,* 33:12787-12794 (1994); Shanklin, J., and Cahoon, E. B., *Annu. Rev. Plant Physiol. Plant Mol. Biol.,* 49:611-641 (1998)). Within SEQ ID NO:10, these three histidine boxes were located between amino acid residues 146-150, 183-187 and 358-362; two additional His residues are located at amino acid residues 27 and 50. Each of these His residues are depicted on Figure 2 with a small round circle.

If the model predicted by TMHMM (*supra*) were accepted without alteration, the first two His-rich regions (i.e., the regions spanning between amino acid residues 146-150 and 183-187) would be located in the periplasmic space, thus preventing their participation in the iron-active site.

The conflict noted above was resolved when hydropathy plot analysis (Kyte, J., and Doolittle, R., *J. Mol. Biol.*, 157:105-132 (1982)) predicted one more hydrophobic region located between amino acid residues 88-109 that immediately preceded the first predicted transmembrane segment (i.e., residues 113-132). Since the N-terminal cytochrome-b_5 domain is located in the cytoplasmic space, it was predicted that the hydrophobic region (i.e., residues 88-109) should be the first membrane-spanning segment (i.e., region I, as shown in Figure 2), while the predicted transmembrane segment corresponding to residues 113-132 was designated as the second membrane-spanning segment (i.e., region II, as shown in Figure 2). As a result, the transmembrane segment found between residues 223-245 that was originally predicted by TMHMM to span through the membrane was instead predicted to lie in the cytoplasmic face, such that the first two His-rich motifs (i.e., the regions spanning between amino acid residues 146-150 and 183-187) could be adjusted to be within the cytoplasmic side.

Finally, the hydropathy plot analysis also predicted another hydrophobic region (i.e., residues 157-172) between the first two His-rich motifs. Because the substrate for the desaturase is highly hydrophobic, it was expected to most likely partition into the lipid bilayer of the cytoplasmic membrane. This suggested that the desaturase active site assembled from the His-rich motifs might be at (or very near) the membrane surface. Thus, it was hypothesized that both hydrophobic regions (i.e., residues 157-172 and residues 223-245) lie near the membrane surface to ensure that the active site sits close the membrane.

The last two membrane-spanning helices predicted by TMHMM (i.e., residues 266-283 and 287-309) are included within the final topological model shown in Figure 2 as region III and region IV.

Thus, the topological model depicted in Figure 2 includes four transmembrane regions labeled as regions I, II, III and IV, which correspond to amino acid residues 88-109, 113-132, 266-283 and 287-309, respectively. Two additional hydrophobic regions are located at amino acid residues 157-172 and 223-245. Finally, "IN" corresponds with the cytoplasmic space, while "OUT" corresponds with the periplasmic space.

Example 2

Strategies To Select Amino Acid Residues For Mutation

Close homologs to the EgD8S sequence were determined by conducting BLAST (Basic Local Alignment Search Tool; Altschul, S. F., et al., *J. Mol. Biol.*, 215:403-410 (1993)) searches for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the SWISS-PROT protein sequence database, EMBL and DDBJ databases). Specifically, EgD8S (SEQ ID NO:10) was compared for similarity to all publicly available protein sequences contained in the "nr" database, using the BLASTP algorithm (Altschul et al., *J. Mol. Biol.* 215:403-410 (1990)) provided by the NCBI.

Ignoring all hits to any Δ8 desaturase isolated from *Euglena gracilis*, the BLASTP searches showed that EgD8S was most homologous to the following proteins:

Table 6

Homologous Proteins To EgD8S, Based On BLASTP Analysis

GenBank Accession No.	Protein	Organism	% Identity	% Similarity	E-Value
CAE65324	hypothetical protein CBG10258	*Caenorhabditis briggsae*	38	56	1E-65
AAR27297	Δ6 desaturase	*Amylomyces rouxii*	35	52	3E-65
AAS93682	Δ6 desaturase	*Rhizopus orizae*	32	53	4E-64

* "% Identity" is defined as the percentage of amino acids that are identical between the two proteins.

** "% Similarity" is defined as the percentage of amino acids that are identical or conserved between the two proteins.

*** "Expectation value" estimates the statistical significance of the match, specifying the number of matches, with a given score, that are expected in a search of a database of this size absolutely by chance.

In order to select the amino acid residues that could be mutated within EgD8S without affecting the Δ8 desaturase activity, a set of criteria were developed to identify preferred targets for mutation, as outlined below.

1. Preferred amino acid residue targets of the EgD8S desaturase domain (located between amino acid residues 79 to 406 of SEQ ID NO:10) are not conserved, when compared to the Δ6 desaturase of *A. rouxii* (*supra;* "ArD6"; SEQ ID NO:13), the Δ6 desaturase of *R. orizae* (*supra;* "RoD6"; SEQ ID NO:14) and representatives of other desaturases such as the Δ8 fatty acid desaturase-like protein of *Leishmania major* (GenBank Accession No. CAJ09677; "LmD8L"; SEQ ID NO:15), and the Δ6 desaturase of *Mortierella isabellina* (GenBank Accession No. AAG38104; "MiD6"; SEQ ID NO:16). An alignment of these proteins is shown in Figure 3, using the method of Clustal W (slow, accurate, Gonnet

option; Thompson et al., *Nucleic Acids Res.*, 22:4673-4680 (1994)) of the MegAlign™ program of DNASTAR™ software. It was hypothesized that changes in the non-conserved regions among these 5 different desaturases should not affect the Δ8 desaturase activity of EgD8S.

2. Preferred amino acid residue targets of the cytochrome b_5 domain of EgD8S (located between amino acids 5 to 71 of SEQ ID NO:10) are not conserved, when compared to the cytochrome b_5 genes of *Saccharomyces cerevisiae* (GenBank Accession No. P40312; "SCb5"; SEQ ID NO:178) and *Schizosaccharomyces pombe* (GenBank Accession No. O94391; SPb5; SEQ ID NO:179). An alignment of the N-terminal portion of EgD8S (i.e., amino acid residues 1-136 of SEQ ID NO:10) with SCb5 and SPb5 is shown in Figure 4, using the method of Clustal W (*supra*) of the MegAlign™ program of DNASTAR™ software. It was hypothesized that changes in the non-conserved region among these 3 different proteins should not affect the electron transport function of the cytochrome b_5 domain of EgD8S and thus not affect the Δ8 desaturase activity.

3. Preferred amino acid residue targets are on the transmembrane helices close to the endoplasmic reticulum (ER) side of the membrane or exposed to the ER lumen.

4. Preferred amino acid residue targets are close to the N- terminal or C-terminal ends of the EgD8S enzyme, since non-conserved residues in these regions may tolerate more mutations.

Based on the above criteria, a set of 70 target mutation sites (comprising one, two or three amino acid residues) within EgD8S (i.e., SEQ ID NO:10) were selected for mutation as described below in Table 7. Of the individual 126 amino acid residue mutations, 53 (i.e., 42.1%) were identified as "non-conservative amino acid substitutions" while 73 (i.e., 57.9%) were identified as "conservative amino acid substitutions".

Table 7

Selected Amino Acid Residues Suitable For Targeted Mutation

Mutation Site	Sequence Mutations Within SEQ ID NO:10	Mutation Site	Sequence Mutations Within SEQ ID NO:10
M1	4S to A, 5K to S	M36	37Y to F, 38Q to N
M2	12T to V	M37	51S to T, 52Q to N
M3	16T to K, 17T to V	M38	54A to G, 55F to Y
M4	25N to D, 26F to E	M39	64I to L, 65N to D
M5	31A to D, 32E to S	M40	69E to D, 70L to V
M6	59K to L	M41	75A to G, 76V to L
M7	61M to A, 62P to V	M42	89E to D, 90L to I
M8	66P to Q, 67S to A	M43	97D to E, 98A to V
M9	72P to Q, 73Q to P	M44	110T to S, 111L to V
M10	79A to Q, 80Q to A	M45	117G to A, 118Y to F
M11	87R to A, 88E to I	M46	132V to L, 133L to V
M12	407A to S, 408V to Q	M47	198D to E, 199I to L
M13	412M to S, 413A to Q	M48	231L to V, 232V to L
M14	416Q to V, 417P to Y	M49	297F to V, 298V to L
M15	422L to Q	M50	309I to V, 310V to I
M16	108S to L	M51A	347I to L, 348T to S
M17	110T to A	M51B	346I to V, 347I to L, 348T to S
M18	120L to M, 121M to L	M52	400V to I, 401I to V
M19	122V to S	M53	9L to V
M20	123Q to Y, 124Y to Q	M54	19D to E, 20V to I
M21	125Q to H, 126M to L	M55	33I to L
M22	127Y to Q	M56	45A to G, 46F to Y
M23	288S to N	M57	57K to R, 58L to I
M24	289I to P, 290L to M	M58	65N to Q
M25	291T to V, 292S to V	M59	73Q to N, 74A to G
M26	293L to M	M60	96F to Y
M27	296F to T	M61	239F to I, 240I to F
M28	298V to S	M62	271L to M, 272A to S
M29	392N to T, 393P to T	M63	279T to L, 280L to T
M30	394L to G, 395P to M	M64	130G to A, 131A to G
M31	7Q to L, 8A to S	M65	304G to F, 305F to G
M32	10P to W, 11L to Q	M66	229F to Y, 230Y to F
M33	21S to F, 22A to S	M67	291T to S, 292S to L
M34	46F to S, 47M to L	M68	162L to V, 163V to L
M35	48V to F, 49M to L	M69	170G to A, 171L to V
		M70	418A to G, 419G to A

Example 3

Generation Of *Yarrowia lipolytica* Strains Y4001 And Y4001U To Produce About 17% EDA Of Total Lipids

The present Example describes the construction of strains Y4001 and Y4001U, derived from *Yarrowia lipolytica* ATCC #20362, and each capable of producing 17% EDA (C20:2) relative to the total lipids. Both strains were engineered to test functional expression of EgD8S and mutations thereof. Thus, it was necessary to construct host strains capable of producing the Δ8 desaturase substrate, EDA.

The development of strain Y4001U, having a *Leu-* and *Ura-* phenotype, required the construction of strain Y2224 (a FOA resistant mutant from an autonomous mutation of the *Ura3* gene of wildtype *Yarrowia* strain ATCC #20362) and strain Y4001.

Generation Of Strain Y2224

Strain Y2224 was isolated in the following manner: *Yarrowia lipolytica* ATCC #20362 cells from a YPD agar plate (1% yeast extract, 2% bactopeptone, 2% glucose, 2% agar) were streaked onto a MM plate (75 mg/L each of uracil and uridine, 6.7 g/L YNB with ammonia sulfate, without amino acid, and 20 g/L glucose) containing 250 mg/L 5-FOA (Zymo Research). Plates were incubated at 28 °C and four of the resulting colonies were patched separately onto MM plates containing 200 mg/mL 5-FOA and MM plates lacking uracil and uridine to confirm uracil *Ura3* auxotrophy.

Generation Of Strain Y4001 To Produce 17% EDA Of Total Lipids

Strain Y4001 was created via integration of construct pZKLeuN-29E3 (Figure 5A; comprising four chimeric genes [a Δ12 desaturase, a $C_{16/18}$ elongase and two Δ9 elongases]) into the *Leu2* loci of Y2224 strain to thereby enable production of EDA.

Construct pZKLeuN-29E3 contained the following components:

Table 8

Description of Plasmid pZKLeuN-29E3 (SEQ ID NO:17)

RE Sites And Nucleotides Within SEQ ID NO:17	Description Of Fragment And Chimeric Gene Components
BsiW I/*Asc* I (7797-7002)	795 bp 3' part of *Yarrowia Leu2* gene (GenBank Accession No. AF260230)
Sph I/*Pac* I (4302-3591)	703 bp 5' part of *Yarrowia Leu2* gene (GenBank Accession No. AF260230)
Swa I/*BsiW* I (10500-7797)	GPD::F.D12::Pex20, comprising: • GPD: *Yarrowia lipolytica* GPD promoter (WO 2005/003310) • F.D12: *Fusarium moniliforme* Δ12 desaturase gene (WO 2005/047485) • Pex20: Pex20 terminator sequence from *Yarrowia Pex20* gene (GenBank Accession No. AF054613)
Bgl II/*Swa* I (12526-10500)	Exp pro::EgD9E::Lip1, comprising: • Exp pro: *Yarrowia lipolytica* export protein (EXP1) promoter (WO 2006/052870 and U.S. Patent Application No. 11/265761) • EgD9E: codon-optimized Δ9 elongase gene (SEQ ID NO:177), derived from *Euglena gracilis* (SEQ ID NOs:175 and 176; U.S. Patent Application No. 60/739989; see also Example 12 herein) • Lip1: Lip1 terminator sequence from *Yarrowia Lip1* gene (GenBank Accession No. Z50020)
Pme I/*Cla* I (12544-1)	FBAINm::EgD9S::Lip2, comprising: • FBAINm: *Yarrowia lipolytica* FBAINm promoter (WO 2005/049805) • EgD9S: codon-optimized Δ9 elongase gene (SEQ ID NO:177; *supra*) • Lip2: Lip2 terminator sequence from *Yarrowia Lip2* gene (GenBank Accession No. AJ012632)
Cla I/*EcoR* I (1-1736)	LoxP::*Ura3*::LoxP, comprising: • LoxP sequence (SEQ ID NO:18) • *Yarrowia* Ura3 gene (GenBank Accession No. AJ306421) • LoxP sequence (SEQ ID NO:18)
EcoR I/*Pac* I (1736-3591)	NT::ME3S::Pex16, comprising: • NT: *Yarrowia lipolytica* YAT1 promoter (Patent Publication US 2006/0094102-A1) • ME3S: codon-optimized $C_{16/18}$ elongase gene (SEQ ID NO:19), derived from *M. alpina* (see U.S. Patent Application No. 11/253882 and also WO 2006/052870) • Pex16: Pex16 terminator sequence of *Yarrowia Pex 16*

	gene (GenBank Accession No. U75433)

Plasmid pZKLeuN-29E3 was digested with *Asc* I/*Sph* I, and then used for transformation of *Y. lipolytica* strain Y2224 (i.e., ATCC #20362 Ura3-) according to the General Methods. The transformant cells were plated onto MMLeu media plates and maintained at 30 °C for 2 to 3 days. The colonies were picked and streaked onto MM and MMLeu selection plates. The colonies that could grow on MMLeu plates but not on MM plates were selected as *Leu-* strains. Single colonies of *Leu-* strains were then inoculated into liquid MMLeu at 30 °C and shaken at 250 rpm/min for 2 days. The cells were collected by centrifugation, lipids were extracted, and fatty acid methyl esters were prepared by trans-esterification, and subsequently analyzed with a Hewlett-Packard 6890 GC.

GC analyses showed the presence of EDA in the transformants containing the 4 chimeric genes of pZKLeuN-29E3, but not in the *Yarrowia* Y2224 control strain. Most of the selected 36 *Leu-* strains produced about 12 to 16.9% EDA of total lipids. There were 3 strains (i.e., strains #11, #30 and #34) that produced about 17.4%, 17% and 17.5% EDA of total lipids; they were designated as strains Y4001, Y4002 and Y4003, respectively.

Generation Of Strain Y4001U (*Leu-, Ura-*) To Produce 17% EDA Of Total Lipids

Strain Y4001U was created via temporary expression of the Cre recombinase enzyme in plasmid pY116 (Figure 5B) within strain Y4001 to produce a *Leu-* and *Ura-* phenotype. Construct pY116 contained the following components:

Table 9

Description of Plasmid pY116 (SEQ ID NO:180)

RE Sites And Nucleotides Within SEQ ID NO:180	Description Of Fragment And Chimeric Gene Components
1328-448	ColE1 plasmid origin of replication

2258-1398	Ampicillin-resistance gene (Amp^R)
3157-4461	*Yarrowia* autonomous replication sequence (ARS18; GenBank Accession No. A17608)
PacI/SawI 6667-4504	*Yarrowia* Leu2 gene (GenBank Accession No. AF260230)
6667-180	GPAT::Cre::XPR2, comprising: • GPAT: *Yarrowia lipolytica* GPAT promoter (WO 2006/031937) • Cre: Enterobacteria phage P1 Cre gene for recombinase protein (GenBank Accession No. X03453) • XPR2: ~100 bp of the 3' region of the *Yarrowia Xpr* gene (GenBank Accession No. M17741)

Plasmid pY116 was used for transformation of freshly grown Y4001 cells according to the General Methods. The transformant cells were plated onto MMU plates containing 280 μg/mL sulfonylurea and maintained at 30 °C for 3 to 4 days. Four 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 MMU media, and 100 μL was plated onto new YPD plates and maintained at 30 °C for 2 days. Colonies were picked and streaked onto MMLeu and MMLeu+Ura selection plates. The colonies that could grow on MMLeu+Ura 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, produced about 17% EDA of total lipids and was designated as Y4001U.

Example 4

Generation Of Auto-Replicating Plasmid pFmD8S

The present Example describes the construction of plasmid pFmD8S comprising a chimeric FBAINm::EgD8S::XPR gene. Plasmid pFmD8S (SEQ ID NO:20; Figure 6D) was constructed by three-way ligation using fragments from plasmids pKUNFmkF2, pDMW287F and pDMW214. Plasmid pFmD8S, an auto-replicating plasmid that will reside in *Yarrowia* in 1-3 copies, was utilized to test functional expression of EgD8S (and mutations thereof), as described in Examples 5-10, *infra*.

Plasmid pKUNFmkF2

pKUNFmkF2 (SEQ ID NO:21; Figure 6A; WO 2006/012326) is a construct comprising a chimeric FBAINm::F.D12::Lip2 gene (wherein "FBAINmK" is the *Yarrowia lipolytica* FBAINm promoter [WO 2005/049805], "F.D12" is the *Fusarium moniliforme* Δ12 desaturase [WO 2005/047485], and "Lip2" is the *Yarrowia lipolytica* Lip2 terminator sequence (GenBank Accession No. AJ012632)).

Plasmid pDMW287F

pDMW287F (SEQ ID NO:22; Figure 6B; WO 2006/012326) is a construct comprising the synthetic Δ8 desaturase, derived from wildtype *Euglena gracilis,* and codon-optimized for expression in *Yarrowia lipolytica* (wherein EgD8S is identified as "D8SF" in the Figure). The desaturase gene is flanked by a *Yarrowia lipolytica* FBAIN promoter (WO 2005/049805; identified as "FBA1+intron" in the Figure) and a Pex16 terminator sequence of the *Yarrowia Pex16* gene (GenBank Accession No. U75433).

Plasmid pDMW214

pDMW214 (SEQ ID NO:23; Figure 6C; WO 2005/049805) is a shuttle plasmid that could replicate both in *E. coli* and *Yarrowia lipolytica*. It contained the following components:

Table 10

Description Of Plasmid pDMW214 (SEQ ID NO:23)

RE Sites And Nucleotides Within SEQ ID NO:23	Description Of Fragment And Chimeric Gene Components
1150-270	ColE1 plasmid origin of replication
2080-1220	Ampicillin-resistance gene (Amp^R)
2979-4256	*Yarrowia* autonomous replication sequence (ARS18; GenBank Accession No. A17608)
PmeI/SphI 6501-4256	*Yarrowia* Leu2 gene (GenBank Accession No. AF260230)

6501-1	FBA1+intron::GUS::XPR2, comprising: • FBA1+intron: *Yarrowia lipolytica* FBAIN promoter (WO 2005/049805) • GUS: *E. coli* gene encoding β-glucuronidase (Jefferson, R.A. *Nature.* 342:837-838 (1989) • XPR2: ~100 bp of the 3' region of the *Yarrowia Xpr* gene (GenBank Accession No. M17741)

Final Construction Of Plasmid pFmD8S

The *PmeI*/*NcoI* fragment of plasmid pKUNFmkF2 (Figure 6A; comprising the FBAINm promoter) and the *NcoI*/*NotI* fragment of plasmid pDMW287F (Figure 6B; comprising the synthetic Δ8 desaturase gene EgD8S) were used directionally to replace the *PmeI*/*Not* I fragment of pDMW214 (Figure 6C). This resulted in generation of pFmD8S (SEQ ID NO:20; Figure 6D), comprising a chimeric FBAINm::EgD8S::XPR2 gene. Thus, the components of pFmD8S are as described in Table 11 below.

Table 11

Components Of Plasmid pFmD8S (SEQ ID NO:20)

RE Sites And Nucleotides Within SEQ ID NO:20	Description Of Fragment And Chimeric Gene Components
Swa I/*Sac II* (7988-1461)	FBAINm::EgD8S::XPR2, comprising: • FBAINm: *Yarrowia lipolytica* FBAINm promoter (WO 2005/049805) • EgD8S: codon-optimized Δ8 desaturase gene (SEQ ID NO:9, identified as "D8-corrected" in Figure 6D), derived from *E. gracilis* (SEQ ID NO:11) • XPR2: ~100 bp of the 3' region of the *Yarrowia Xpr* gene (GenBank Accession No. M17741)
2601-1721	ColE1 plasmid origin of replication
3531-2671	Ampicillin-resistance gene (Amp^R) for selection in *E. coli*
4430-5734	*Yarrowia* autonomous replication sequence (ARS18; GenBank Accession No. A17608)
7942-5741	*Yarrowia Leu2* gene (GenBank Accession No. AF260230)

Example 5

Development Of A Quick Screen To Functionally Analyze Δ8 Desaturase Activity In *Yarrowia lipolytica*

A set of 40 mutations was created using pFmD8S (Example 4) as template and 40 pairs of oligonucleotide primers to individually mutate targeted amino acid residues within EgD8S (SEQ ID NO:10) by site-directed mutagenesis (QuikChange® Kit, Stratagene, La Jolla, CA). Specific mutations were selected from those set forth in Table 7 of Example 2 and primer pairs were selected from the oligonucleotides set forth in SEQ ID NOs:24-164, such that creation of the M1 mutation (i.e., 4S to A, 5K to S within SEQ ID NO:10) required primers 1A and 1B (SEQ ID NOs:24 and 25, respectively), etc. Plasmids from each mutation were transformed into *E. coli* XL2Blue cells (Stratagene). Four colonies from each of the 40 transformations were picked and grown individually in liquid media at 37 °C overnight. Plasmids (i.e., 160 total) were isolated from these cultures and sequenced individually to confirm the mutations.

Plasmid pFmD8S and the isolated mutant plasmids were transformed into strain Y4001 (Example 3) individually, as described in the General Methods. The transformants were selected on MM plates. After 2 days growth at 30 °C, transformants were scraped from each plate, lipids were extracted, and fatty acid methyl esters were prepared by trans-esterification, and subsequently analyzed with a Hewlett-Packard 6890 GC.

GC analyses showed that there were about 7% DGLA and 12% EDA of total lipids produced by the transformants with plasmid pFmD8S; the average conversion efficiency whereby EgD8S converted EDA to DGLA in the transformants was 36.8%. The conversion efficiency was 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.

GC analyses of transformants carrying mutations within EgD8S showed that 30 of the 40 mutations did not affect the Δ8 desaturase activity (when compared to the synthetic codon-optimized EgD8S Δ8 desaturase activity in transformants carrying plasmid pFmD8S). These results suggested that the screening procedure described herein (i.e., with pFmD8S as parent plasmid and strain Y4001 as a host) could be used to quickly screen the EgD8S mutants and identify which mutations negatively affected Δ8 desaturase activity.

Based on these results, the remaining 30 mutations set forth in Table 7 of Example 2 were synthesized (although some mutations were introduced in combination for efficiency), using the methodology described above. Table 12 describes the Δ8 desaturase activity attributed to each mutation site (i.e., M1 to M70), as a percent of the Δ8 desaturase activity resulting in each mutant EgD8S with respect to the Δ8 desaturase activity of the synthetic codon-optimized EgD8S (SEQ ID NO:10). As seen in the Table below, Δ8 desaturase activity ranged from 0% up to 125%.

Table 12

Mutant Δ8 Desaturase Activities

Mutations	Primers Used	% Activity
M1	1A, 1B (SEQ ID NOs:24 and 25)	115%
M2	2A, 2B (SEQ ID NOs:26 and 27)	110%
M3	3A, 3B (SEQ ID NOs:28 and 29)	100%
M4	4A, 4B (SEQ ID NOs:30 and 31)	N/A
M5	5A, 5B (SEQ ID NOs:32 and 33)	N/A
M6	6A, 6B (SEQ ID NOs:34 and 35)	110%
M7	7A, 7B (SEQ ID NOs:36 and 37)	30%
M8	8A, 8B (SEQ ID NOs:38 and 39)	100%
M9	9A, 9B (SEQ ID NOs:40 and 41)	N/A
M10	10A, 10B (SEQ ID NOs:42 and 43)	N/A
M11	11A, 11B (SEQ ID NOs:44 and 45)	20%
M12	12A, 12B (SEQ ID NOs:46 and 47)	100%
M13	13A, 13B (SEQ ID NOs:48 and 49)	N/A
M14	14A, 14B (SEQ ID NOs:50 and 51)	100%
M15	15A, 15B (SEQ ID NOs:52 and 53)	125%
M16	16A, 16B (SEQ ID NOs:54 and 55)	100%
M17	17A, 17B (SEQ ID NOs:56 and 57)	50%
M18	18A, 18B (SEQ ID NOs:58 and 59)	N/A
M19	19A, 19B (SEQ ID NOs:60 and 61)	100%
M20	20A, 20B (SEQ ID NOs:62 and 63)	80%
M21	21A, 21B (SEQ ID NOs:64 and 65)	120%
M22	22A, 22B (SEQ ID NOs:66 and 67)	110%
M23	23A, 23B (SEQ ID NOs:68 and 69)	80%
M24	24A, 24B (SEQ ID NOs:70 and 71)	N/A
M25	25A, 25B (SEQ ID NOs:72 and 73)	N/A

Mutations	Primers Used	% Activity
M26	26A, 26B (SEQ ID NOs:74 and 75)	110%
M27	27A, 27B (SEQ ID NOs:76 and 77)	80%
M28	28A, 28B (SEQ ID NOs:78 and 79)	90%
M29	29A, 29B (SEQ ID NOs:80 and 81)	N/A
M30	30A, 30B (SEQ ID NOs:82 and 83)	85%
M31	31A, 31B (SEQ ID NOs:84 and 85)	85%
M32	32A, 32B (SEQ ID NOs:86 and 87)	N/A
M33	33A, 33B (SEQ ID NOs:88 and 89)	N/A
M34	34A, 34B (SEQ ID NOs:90 and 91)	N/A
M35	35A, 35B (SEQ ID NOs: 92 and 93)	0%
M36	36A, 36B (SEQ ID NOs: 94 and 95)	80%
M37	37A, 37B (SEQ ID NOs: 96 and 97)	90%
M38	38A, 38B (SEQ ID NOs: 98 and 99)	100%
M39	39A, 39B (SEQ ID NOs:100 and 101)	100%
M40	40A, 40B (SEQ ID NOs:102 and 103)	100%
M41	41A, 41B (SEQ ID NOs:104 and 105)	100%
M42	42A, 42B (SEQ ID NOs:106 and 107)	0%
M43	43A, 43B (SEQ ID NOs:108 and 109)	90%
M44	44A, 44B (SEQ ID NOs:110 and 111)	N/A
M45	45A, 45B (SEQ ID NOs:112 and 113)	100%
M46	46A, 46B (SEQ ID NOs:114 and 115)	100%
M47	47A, 47B (SEQ ID NOs:116 and 117)	0%
M48	48A, 48B (SEQ ID NOs:118 and 119)	N/A
M49	49A, 49B (SEQ ID NOs:120 and 121)	100%
M50	50A, 50B (SEQ ID NOs:122 and 123)	100%
M51	51A, 51B (SEQ ID NOs:124 and 125)	100%

Mutations	Primers Used	% Activity
M51B	51A, 51B (SEQ ID NOs:126 and 125)	100%
M52	52A, 52B (SEQ ID NOs:127 and 128)	80%
M53	53A, 53B (SEQ ID NOs:129 and 130)	100%
M54	54A, 54B (SEQ ID NOs:131 and 132)	100%
M55	55A, 55B (SEQ ID NOs:133 and 134)	40%
M56	56A, 56B (SEQ ID NOs:135 and 136)	0%
M57	57A, 57B (SEQ ID NOs:137 and 138)	N/A
M58	58A, 58B (SEQ ID NOs:139 and 140)	100%
M59	59A, 59B (SEQ ID NOs:141 and 142)	90%
M60	60A, 60B (SEQ ID NOs:143 and 144)	50%
M61	61A, 61B (SEQ ID NOs:145 and 146)	50%
M62	62A, 62B (SEQ ID NOs:147 and 148)	50%
M63	63A, 63B (SEQ ID NOs:149 and 150)	100%
M64	64A, 64B (SEQ ID NOs:151 and 152)	60%
M65	65A, 65B (SEQ ID NOs:153 and 154)	0%
M66	66A, 66B (SEQ ID NOs:155 and 156)	N/A
M67	67A, 67B (SEQ ID NOs:157 and 158)	N/A
M68	68A, 68B (SEQ ID NOs:159 and 160)	100%
M69	69A, 69B (SEQ ID NOs:161 and 162)	100%
M70	70A, 70B (SEQ ID NOs:163 and 164)	100%

* N/A is reported when the desired mutation was not successfully produced or when GC data was lacking.

Example 6

Generation Of pFmD8S-1, pFmD8S-001, pFmD8S-2A, pFmD8S-2B, pFmD8S-3A, pFmD8S-3B, pFmD8S-003, pFmD8S-4, pFmD8S-004, pFmD8S-005 And pFmD8S-006 Constructs By Site-Directed Mutagenesis Of EgD8S Within Construct pFmD8S

A series of plasmids were generated by consecutive rounds of continued site-directed mutagenesis to introduce multiple select mutations into EgD8S (SEQ ID NOs:9 and 10). pFmD8S (Example 4), comprising the synthetic codon-optimized EgD8S, was used as the starting template in Tables 13, 14, 16 and 17, while a mutant created thereof (i.e., pFmD8S-M45, comprising 117G to A and 118Y to F mutations with respect to SEQ ID NO:10) was used as the starting template in Table 15. The resulting plasmids comprising mutant EgD8S sequences, as well as details concerning the primers used to produce these mutations, are described below in Tables 13, 14, 15, 16 and 17.

The column titled "Mutation Site Introduced" refers to the specific amino acid sites selected for mutation, as listed in Table 7 of Example 2. In the column titled "Total Mutations In Resultant Plasmid With Respect to EgD8S (SEQ ID NO:10)", those amino acid mutations that are highlighted in bold-face text correspond to newly introduced mutations that were not present in the template in the indicated round of site-directed mutagenesis. The number shown in parentheses corresponds with the number of total mutations in the resultant plasmid with respect to EgD8S (SEQ ID NO:10).

Table 13

Generation Of pFmD8S-1 And pFmD8S-001 Constructs

Round	Mutation Site Introduced	Template	Primers	Resultant Plasmid	Total Mutations In Resultant Plasmid With Respect to EgD8S (SEQ ID NO:10)
1	M3	pFmD8S	3A, 3B (SEQ ID NOs:28 and 29)	pFmD8S-M3	**16T to K, 17T to V** (2)
2	M1	pFmD8S-M3	1A, 1B (SEQ ID NOs:24 and 25)	pFmD8S-M3,1	16 T to K, 17T to V, **4S to A, 5K to S** (4)
3	M2	pFmD8S-M3,1	2A, 2B (SEQ ID NOs:26 and 27)	pFmD8S-M3,1,2	16 T to K, 17T to V, 4S to A, 5K to S, **12T to V** (5)
4	M8	pFmD8S-M3,1,2	8A, 8B (SEQ ID NOs:38 and 39)	pFmD8S-1	16 T to K, 17T to V, 4S to A, 5K to S, 12T to V, **66P to Q, 67S to A** (7)
5	M38	pFmD8S-1	38A, 38B (SEQ ID NOs:98 and 99)	pFmD8S-001	16 T to K, 17T to V, 4S to A, 5K to S, 12T to V, 65P to Q, 66S to A, **54A to G, 55F to Y** (9)

Table 14

Generation Of pFmD8S-2A And pFmD8S-003 Constructs

Round	Mutation Site Introduced	Template	Primers	Resultant Plasmid	Total Mutations In Resultant Plasmid With Respect to EgD8S (SEQ ID NO:10)
1	M45	pFmD8S	45A, 45B (SEQ ID NOs:112 and 113)	pFmD8S-M45	**117G to A, 118Y to F** (2)
2	M21	pFmD8S-M45	21A, 21B (SEQ ID NOs:64 and 65)	pFmD8S-M45,21	117G to A, 118Y to F, **125Q to H, 126M to L** (4)

3	M16	pFmD8S-M45,21	16A, 16B (SEQ ID NOs:54 and 55)	pFmD8S-M45,21,16	117G to A, 118Y to F, 125Q to H, 126M to L, **108S to L** (5)
4	M18	pFmD8S-M45,21,16	18A, 18B (SEQ ID NOs:58 and 59)	pFmD8S-2A	117G to A, 118Y to F, 125Q to H, 126M to L, 108S to L, **120L to M, 121M to L** (7)
5	M68, M69	pFmD8S-2A	68A, 68B (SEQ ID NOs:159 and 160) 69A, 69B (SEQ ID NOs:161 and 162)	pFmD8S-003	117G to A, 118Y to F, 125Q to H, 126M to L, 108S to L, 120L to M, 121M to L, **162L to V, 163V to L, 170G to A, 171L to V** (11)

Table 15

Generation Of pFmD8S-2B And pFmD8S-004 Constructs

Round	Mutation Site Introduced	Template	Primers	Resultant Plasmid	Total Mutations In Resultant Plasmid With Respect to EgD8S (SEQ ID NO:10)
1	M46	pFmD8S-M45	46A, 46B (SEQ ID NOs:114 and 115)	pFmD8S-M45,46	117G to A, 118Y to F, **132V to L, 133 L to V** (4)
2	M16, M21	pFmD8S-M45,46	16A,16B (SEQ ID NOs:54 and 55) 21A, 21B (SEQ ID NOs:64 and 65)	pFmD8S-M45,46,16,21	117G to A, 118Y to F, 132V to L, 133 L to V, **108S to L, 125Q to H, 126M to L** (7)
3	M18	pFmD8S-M45,46,16,21	18A, 18B (SEQ ID NOs:58 and 59)	pFmD8S-2B	117G to A, 118Y to F, 132V to L, 133 L to V, 108S to L, 125Q to H, 126M to L, **120L to M, 121M to L** (9)
4	M68, M69	pFmD8S-2B	68A, 68B (SEQ ID NOs:159 and 160) 69A, 69B (SEQ ID NOs:161 and 162)	pFmD8S-004	117G to A, 118Y to F, 132V to L, 133 L to V, 108S to L, 125Q to H, 126M to L, 120L to M, 121M to L, **162L to V, 163V to L, 170G to A, 171L to V** (13)

Table 16

Generation Of pFmD8S-3A, pFmD8S-3B And pFmD8S-005 Constructs

Round	Mutation Site Introduced	Template	Primers	Resultant Plasmid	Total Mutations In Resultant Plasmid With Respect to EgD8S (SEQ ID NO:10)
1	M49	pFmD8S	49A, 49B (SEQ ID NOs:120 and 121)	pFmD8S-M49	**297F to V, 298V to L** (2)
2	M26	pFmD8S-M49	26A, 26B (SEQ ID NOs:74 and 75)	pFmD8S-M49,26	297F to V, 298V to L, **293L to M** (3)
3A	M61	pFmD8S-M49,26	61A, 61B (SEQ ID NOs:145 and 146)	pFmD8S-3A	297F to V, 298V to L, 293L to M, **239F to I, 240I to F** (5)
3B	M62, M63	pFmD8S-M49,26	62A, 62B (SEQ ID NOs:147 and 148) 63A, 63B (SEQ ID NOs:149 and 150)	pFmD8S-3B	297F to V, 298V to L, 293L to M, **271L to M, 272A to S, 279T to L, 280L to T** (7)
4	M63	pFmD8S-3A	63A, 63B (SEQ ID NOs:149 and 150)	pFmD8S-005	297F to V, 298V to L, 293L to M, 239F to I, 240I to F, **279T to L, 280L to T** (7)

Table 17

Generation Of pFmD8S-4 And pFmD8S-006 Constructs

Round	Mutation Site Introduced	Template	Primers	Resultant Plasmid	Total Mutations In Resultant Plasmid With Respect to EgD8S (SEQ ID NO:10)
1	M51B	pFmD8S	51A, 51B (SEQ ID NOs:126 and 125)	pFmD8S-M51	**346I to V, 347I to L, 348T to S** (3)

2	M15	pFmD8S-M51	15A, 15B (SEQ ID NOs:52 and 53)	pFmD8S-M51,15	346I to V, 347I to L, 348T to S, **422L to Q** (4)
3	M14	pFmD8S-M51,15	14A, 14B (SEQ ID NOs:50 and 51)	pFmD8S-M51,15,14	346I to V, 347I to L, 348T to S, 422L to Q, **416Q to V, 417P to Y** (6)
4	M12	pFmD8S-M51,15,14	12A, 12B (SEQ ID NOs:46 and 47)	pFmD8S-4	346I to V, 347I to L, 348T to S, 422L to Q, 416Q to V, 417P to Y, **407A to S, 408V to Q** (8)
5	M70	pFmD8S-4	70, 70B (SEQ ID NOs:163 and 164)	pFmD8S-006	346I to V, 347I to L, 348T to S, 422L to Q, 416Q to V, 417P to Y, 407A to S, 408V to Q, **418A to G, 419G to A** (10)

After each round of mutagenesis, the mutations in each resulting plasmid was confirmed by DNA sequencing. Additionally, the Δ8 desaturase activity of each mutant EgD8S within each mutant plasmid was compared with the Δ8 desaturase activity of the synthetic codon-optimized EgD8S within pFmD8S by transforming the plasmids into strain Y4001 (Example 3) and assaying activity based on the methodology described in Example 5. Based on these functional analyses, it was demonstrated that the mutations in all 24 of the mutant EgD8S genes within the resultant plasmids generated in Tables 11, 12, 13, 14 and 15 did not affect Δ8 desaturase activity (i.e., pFmD8S-M3; pFmD8S-M3,1; pFmD8S-M3,1,2; pFmD8S-1; pFmD8S-001; pFmD8S-M45; pFmD8S-M45,21; pFmD8S-M45,21,16; pFmD8S-2A; pFmD8S-003; pFmD8S-M45,46; pFmD8S-M45,46,16,21; pFmD8S-2B; pFmD8S-004; pFmD8S-M49; pFmD8S-M49,26; pFmD8S-3A; pFmD8S-3B; pFmD8S-005; pFmD8S-M51; pFmD8S-M51,15; pFmD8S-M51,15,14; pFmD8S-4; and pFmD8S-006).

Example 7

Generation Of Complex Construct pFmD8S-5B By Digestion And Ligation Of Multiple Parent Plasmids

Plasmid pFmD8S-5B contained 16 mutant amino acids within the first half of EgD8S. This plasmid was generated by 3-way ligation, wherein the 318 bp *Nco* I/*Bgl* II fragment from pFmD8S-1 (containing 7 amino acid mutations, corresponding to M1, M2, M3 and M8) and the 954 bp *Bgl* II/*Not* I fragment from pFmD8S-2B (containing 9 amino acid mutations, corresponding to M16, M18, M21, M45 and M46) were used to replace the *Nco* I/*Not* I fragment of pFmD8S (Example 4; Figure 6D). DNA sequence confirmed that pFmD8S-5B contained the expected 16 amino acid mutations within EgD8S.

The synthesis of plasmid pFmD8S-5B is schematically diagrammed in Figure 7 (and a similar format is used in Figures 8 and 9). For clarity, the pFmD8S vector backbone in which each mutant EgD8S is contained is not included within the figure; instead, only the 1272 bases corresponding to the mutant EgD8S are shown (wherein the coding sequence for the Δ8 desaturase corresponds to nucleotide bases 2-1270). Thus, the mutant

EgD8S fragment labeled as "Mutant EgD8S-1" in Figure 7 corresponds to the mutant EgD8S found within plasmid pFmD8S-1 and the mutant EgD8S fragment labeled as "Mutant EgD8S-2B" in Figure 7 corresponds to the mutant EgD8S fround within plasmid pFmD8S-2B.

Similarly, the *Nco* I and *Not* I restriction enzyme sites that flank each mutant EgD8S gene are not included in the figure. The *Nco* I nucleotide recognition sequence ("CCATGG") corresponds to the –2 to +4 region of the mutant EgD8S, wherein the 'A' position of the 'ATG' translation initiation codon is designated as +1; the first nucleotide recognized as part of the *Not* I nucleotide recognition sequence is nucleotide +1271 of mutant EgD8S, wherein the 'TAA' STOP codon of mutant EgD8S is located at +1269 to +1270.

Mutation sites are labeled on each mutant EgD8S. Those mutation sites shown with an asterisk correspond to a single amino acid mutation (i.e., M2* corresponds to a mutation of 12T to V), while those lacking an asterisk correspond to two individual amino acid mutations (i.e., M1 corresponds to mutations 4S to A and 5K to S); those mutation sites shown with 2 asterisks correspond to a triple amino acid mutation (i.e., M51** corresponds to mutations 346I to V, 347I to L and 348T to S).

The Δ8 desaturase activity of mutant EgD8S-5B within pFmD8S-5B was compared with the Δ8 desaturase activity of the synthetic codon-optimized EgD8S within pFmD8S by transforming each plasmid into strain Y4001 (Example 3) and assaying the activity based on the methodology described in Example 5. Based on this analysis, it was determined that the 16 amino acid mutations within mutant EgD8S-5B (i.e., 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 108S to L, 120L to M, 121M to L, 125Q to H, 126M to L, 117G to A, 118Y to F, 132V to L and 133 L to V, corresponding to mutation sites M1, M2, M3, M8, M16, M18, M21, M45 and M46) in pFmD8S-5B did not affect the Δ8 desaturase activity.

Example 8

Generation Of pFmD8S-12, pFmD8S-13, pFmD8S-23 And pFmD8S-28 Constructs By Additional Site-Directed Mutagenesis Of Mutant EgD8S-5B Within Construct pFmD8S-5B

An additional series of plasmids were generated by consecutive rounds of continued site-directed mutagenesis to introduce multiple select mutations into mutant EgD8S-5B, using pFmD8S-5B (Example 7) as the starting template. The resulting plasmids comprising mutant EgD8S sequences, as well as details concerning the primers used to produce these mutations, are described below in Table 18. Format and column titles of Table 18 are the same as defined above in Example 6.

Table 18

Generation Of pFmD8S-12, pFmD8S-13, pFmD8S-23 And pFmD8S-28 Constructs

Round	Mutation Site Introduced	Template	Primers	Resultant Plasmid	Total Mutations In Resultant Plasmid With Respect to EgD8S (SEQ ID NO:10)
1A	M12, M15, M26	pFmD8S-5B	12A, 12B (SEQ ID NOs:46 and 47) 15A, 15B (SEQ ID NOs:52 and 53) 26A, 26B (SEQ ID NOs:74 and 75)	pFmD8S-12	4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 108S to L, 120L to M, 121M to L, 125Q to H, 126M to L, 117G to A, 118Y to F, 132V to L, 133 L to V, **407A to S, 408V to Q, 422L to Q, 293L to M** (20)
1B	M12, M15, M26, M51B	pFmD8S-5B	12A, 12B (SEQ ID NOs:46 and 47) 15A, 15B (SEQ ID NOs:52 and 53) 26A, 26B (SEQ ID NOs:74 and 75) 51A, 51B (SEQ ID NOs:126 and 125)	pFmD8S-13	4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 108S to L, 120L to M, 121M to L, 125Q to H, 126M to L, 117G to A, 118Y to F, 132V to L, 133 L to V, **407A to S, 408V to Q, 422L to Q, 293L to M, 346I to V, 347I to L, 348T to S** (23)
2A	M68, M70	pFmD8S-12	68A, 68B (SEQ ID NOs:159 and 160) 70A, 70B (SEQ ID NOs:163 and 164)	pFmD8S-23	4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 108S to L, 120L to M, 121M to L, 125Q to H, 126M to L, 117G to A, 118Y to F, 132V to L, 133 L to V, 407A to S, 408V to Q, 422L to Q, 293L to M, **162L to V, 163V to L, 418A to G, 419G to A** (24)
2B	M38, M63, M68, M69, M70	pFmD8S-13	38A, 38B (SEQ ID NOs:98 and 99) 63A, 63B (SEQ ID NOs:149 and 150)	pFmD8S-28	4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 108S to L, 120L to M, 121M to L,

		68A, 68B (SEQ ID NOs:159 and 160) 69A, 69B (SEQ ID NOs:161 and 162) 70A, 70B (SEQ ID NOs:163 and 164)		125Q to H, 126M to L, 117G to A, 118Y to F, 132V to L, 133 L to V, 407A to S, 408V to Q, 422L to Q, 293L to M, 346I to V, 347I to L, 348T to S, **54A to G, 55F to Y, 279T to L, 280L to T, 162L to V, 163V to L, 170G to A, 171L to V, 418A to G, 419G to A** (33)

After each round of mutagenesis, the mutations in the resulting plasmid were confirmed by DNA sequencing. Additionally, the Δ8 desaturase activity of each mutant EgD8S within each mutated plasmid was compared with the Δ8 desaturase activity of the synthetic codon-optimized EgD8S within pFmD8S by transforming the plasmids into strain Y4001 (Example 3) and assaying activity based on the methodology described in Example 5. Based on these functional analyses, it was demonstrated that the 20 mutations in mutant EgD8S-12 within pFmD8S-12, the 23 mutations in mutant EgD8S-13 within pFmD8S-13, the 24 mutations in mutant EgD8S-23 within pFmD8S-23 and the 33 mutations in mutant EgD8S-28 within pFmD8S-28 did not affect the Δ8 desaturase activity.

Example 9

Generation Of Complex Constructs pFmD8S-008, pFmD8S-009, pFmD8S-013 And pFmD8S-015 By Digestion And Ligation Of Multiple Parent Plasmids

Plasmids pFmD8S-008 and pFmD8S-009 contained 20 and 22 mutant amino acids within the first half of EgD8S, respectively. These plasmids were generated by 3-way ligation, as shown in Figure 8A and 8B, respectively (Figure format is identical to that described for Figure 7 in Example 7). Specifically, the 318 bp *Nco* I/*Bgl* II fragment from pFmD8S-001 (containing 9 amino acid mutations in mutant EgD8S-001 corresponding to M1, M2, M3, M8 and M38) and the 954 bp *Bgl* II/*Not* I fragment from either pFmD8S-003 (containing 11 amino acid mutations in mutant EgD8S-003 corresponding to M16, M18, M21, M45, M68 and M69) or pFmD8S-004 (containing 13 amino acid mutations in mutant EgD8S-004, corresponding to M16, M18, M21, M45, M46, M68 and M69) were used to replace the *Nco* I/*Not* I fragment of pFmD8S (Example 4; Figure 6D) to generate mutant EgD8S-008 within pFmD8S-008 and mutant EgD8S-009 within pFmD8S-009, respectively. DNA sequence confirmed that mutant EgD8S-008 contained 20 amino acid mutations and mutant EgD8S-009 contained 22 amino acid mutations, as expected.

Plasmids pFmD8S-013 and pFmD8S-015, containing 28 and 31 amino acid mutations within mutant EgD8S-013 and mutant EgD8S-015, respectively, were created using a similar 3-way ligation strategy as shown in Figure 9A and 9B (Figure format is identical to that described for Figure 7 in Example 7). The 639 bp *Nco* I/*Xho* I fragment from either pFmD8S-009 (containing 22 amino acid mutations within mutant EgD8S-009) or pFmD8S-008 (containing 20 amino acid mutations within mutant mutant EgD8S-008) and the 633 bp *Xho* I/*Not* I fragment from either pFmD8S-23 (Example 8, containing 6 amino acid mutations within mutant EgD8S-23, corresponding to M12, M15, M26 and M70) or pFmD8S-28 (Example 8, containing 11 amino acid mutations within mutant EgD8S-28, corresponding to M12, M15, M26, M51B, M63 and M70) were used to replace the *Nco* I/*Not* I fragment of pFmD8S (Example 4; Figure 6D) to generate pFmD8S-013 and pFmD8S-015, respectively. DNA sequence confirmed that mutant EgD8S-013 and mutant EgD8S-015 contained 28 amino acid mutations and 31 amino acid mutations, respectively.

The Δ8 desaturase activity of mutant EgD8S-008 within pFmD8S-008, mutant EgD8S-009 within pFmD8S-009, mutant EgD8S-013 within pFmD8S-013 and mutant EgD8S-015 within pFmD8S-015 were compared with the Δ8 desaturase activity of the synthetic codon-optimized EgD8S within pFmD8S by transforming these plasmids into strain Y4001 (Example 3) and assaying activity based on the methodology described in Example 5. Based on these functional analyses, it was demonstrated the Δ8 desaturase activity was not affected by the 20 mutations in mutant EgD8S-008 within pFmD8S-008 (i.e., 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 108S to L, 120L to M, 121M to L, 125Q to H, 126M to L, 54A to G, 55F to Y, 117G to A, 118Y to F, 162L to V, 163V to L, 170G to A and 171L to V, corresponding to mutation sites M1, M2, M3, M8, M16, M18, M21, M38, M45, M68 and M69), the 22 mutations in mutant EgD8S-009 within pFmD8S-009 (i.e., 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 108S to L, 120L to M, 121M to L, 125Q to H, 126M to L, 54A to G, 55F to Y, 117G to A, 118Y to F, 132V to

L, 133L to V, 162L to V, 163V to L, 170G to A and 171L to V, corresponding to mutation sites M1, M2, M3, M8, M16, M18, M21, M38, M45, M46, M68 and M69), the 28 mutations in mutant EgD8S-013 within pFmD8S-013 (i.e., 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 407A to S, 408V to Q, 422L to Q, 108S to L, 120L to M, 121M to L, 125Q to H, 126M to L, 293L to M, 54A to G, 55F to Y, 117G to A, 118Y to F, 132V to L, 133L to V, 162L to V, 163V to L, 170G to A, 171L to V, 418A to G and 419G to A, corresponding to mutation sites M1, M2, M3, M8, M12, M15, M16, M18, M21, M26, M38, M45, M46, M68, M69, M70) or the 31 mutations in mutant EgD8S-015 within pFmD8S-015 (i.e., 4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 407A to S, 408V to Q, 422L to Q, 108S to L, 120L to M, 121M to L, 125Q to H, 126M to L, 293L to M, 54A to G, 55F to Y, 117G to A, 118Y to F, 346I to V, 347I to L, 348T to S, 279T to L, 280L to T, 162L to V, 163V to L, 170G to A, 171L to V, 418A to G and 419G to A, corresponding to mutation sites M1, M2, M3, M8, M12, M15, M16, M18, M21, M26, M38, M45, M51B, M63, M68, M69, M70).

Figure 10 shows an alignment of EgD8S (SEQ ID NO:10), Mutant EgD8S-23 (SEQ ID NO:4; Example 8), Mutant EgD8S-013 (SEQ ID NO:6; Example 9) and Mutant EgD8S-015 (SEQ ID NO:8; Example 9). The method of alignment used corresponds to the "Clustal W method of alignment".

Example 10

Comparison Of Δ8 Desaturase Activities Among The Synthetic Codon-Optimized EgD8S And Its Mutants Upon Integration Into The *Yarrowia lipolytica* Genome

This Example describes quantitative analyses of the Δ8 desaturase activities of EgD8S and mutants thereof included within pFmD8S-23, pFmD8S-013 and pFmD8S-015. This comparison required each of the chimeric genes comprising EgD8S (or a mutant thereof) to be inserted into the pKO2UFkF2 vector backbone. Specifically, pKO2UFkF2 comprised a 5' and 3' portion of the *Yarrowia lipolytica* Δ12 desaturase gene that was

designed to target integration to this locus (although plasmid integration could also occur via random integration into other sites of the genome). Thus, the activities of the chimeric genes containing the synthetic codon-optimized EgD8S, mutant EgD8S-023, mutant EgD8S-013 and mutant EgD8S-015 in the *Yarrowia* genome (i.e., 1 copy) could be more fairly compared upon integration into the *Yarrowia* genome, as opposed to the Δ8 desaturase activity levels that were obtained upon plasmid expression (i.e., via expression in pFmD8S as 1-3 copies) and reported in previous examples.

The components of pKO2UFkF2 are as described in Table 19 below.

Table 19

Components Of Plasmid pKO2UFkF2 (SEQ ID NO:165)

RE Sites And Nucleotides Within SEQ ID NO:165	Description Of Fragment And Chimeric Gene Components
SwaI/BsiWI 7638-1722	FBAINm::F.D12::Pex20, comprising: • FBAINm: *Yarrowia lipolytica* FBAINm promoter (WO 2005/049805) • F.D12: *Fusarium moniliforme* Δ12 desaturase gene (WO 2005/047485) • Pex20: Pex20 terminator sequence of *Yarrowia Pex20* gene (GenBank Accession No. AF054613)
AscI/BsiWI 2459-1722	5' portion of *Yarrowia lipolytica* Δ12 desaturase gene (WO 2004/104167)
EcoRI/SphI 5723-5167	3' portion of *Yarrowia lipolytica* Δ12 desaturase gene (WO 2004/104167)
EcoRI/PacI 5723-7240	*Yarrowia Ura3* gene (GenBank Accession No. AJ306421)

First, the *Swa* I/*Not* I fragment from pFmD8S (comprising the chimeric FBAINm::EgD8S gene, wherein EgD8S is identified as "D8-corrected" in Figure 6D) was used to replace the *SwaI/NotI* fragment of pKO2UFkF2 (comprising the chimeric FBAINm::F.D12 gene) to generate construct pKO2UFm8 (Figure 11A). The same methodology was used to

replace the *SwaI/NotI* fragment of pKO2UFkF2 with the *Swa I/Not I* fragments of pFmD8S-23, pFmD8S-013 and pFmD8S-015, respectively, thereby creating constructs pKO2UFm8-23, pKO2UFm8-013 and pKO2UFm8-015, respectively. As such, the synthetic codon-optimized EgD8S, mutant EgD8S-023, mutant EgD8S-013 and mutant EgD8S-015 were each under the control of the FBAINm promoter and the Pex20 terminator.

Plasmids pKO2UFm8, pKO2UFm8-23, pKO2UFm8-013 and pKO2UFm8-015 were digested with *AscI/SphI*, and then used for transformation of strain Y4001 individually according to the General Methods. Following transformation, cells were plated onto MM plates and maintained at 30 °C for 2 to 3 days.

A total of 6 transformants from each transformation were picked and re-streaked onto fresh MM plates. Once grown, these strains were individually inoculated into liquid MM at 30 °C and grown with shaking at 250 rpm/min for 1 day. The cells were collected by centrifugation, lipids were extracted, and fatty acid methyl esters were prepared by trans-esterification, and subsequently analyzed with a Hewlett-Packard 6890 GC. Delta-8 desaturase activity of each Δ8 desaturase are shown below in Table 20; conversion efficiency was calculatd as described in Example 5.

Table 20

Δ8 Desaturase Activity In EgD8S And Its Mutants

Plasmid	Mutations With Respect To EgD8S (SEQ ID NO:10)	Conversion Efficiency
pKO2UFm8 (comprising wildtype EgD8S [SEQ ID NO:10])	none	37.9% (average)
pKO2UFm8S-23 (comprising mutant EgD8S-23 [SEQ ID NO:4])	4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 108S to L, 120L to M, 121M to L, 125Q to H, 126M to L, 117G to A, 118Y to F, 132V to L, 133 L to V, 407A to S, 408V to Q, 422L to Q, 293L to M, 162L to V, 163V to L, 418A to G,	35%, 35%, 36.1%, 36.2%, 39.8%, 40%

	419G to A (24)	
pKO2UFm8S-013 (comprising mutant EgD8S-013 [SEQ ID NO:6])	4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 407A to S, 408V to Q, 422L to Q, 108S to L, 120L to M, 121M to L, 125Q to H, 126M to L, 293L to M, 54A to G, 55F to Y, 117G to A, 118Y to F, 132V to L, 133L to V, 162L to V, 163V to L, 170G to A, 171L to V, 418A to G and 419G to A (28)	17.8% 18.4%, 18.6%, 24.4%, 34.4% 39.1% 70.8%
pKO2UFm8S-015 (comprising mutant EgD8S-015 [SEQ ID NO:8])	4S to A, 5K to S, 12T to V, 16T to K, 17T to V, 66P to Q, 67S to A, 407A to S, 408V to Q, 422L to Q, 108S to L, 120L to M, 121M to L, 125Q to H, 126M to L, 293L to M, 54A to G, 55F to Y, 117G to A, 118Y to F, 346I to V, 347I to L, 348T to S, 279T to L, 280L to T, 162L to V, 163V to L, 170G to A, 171L to V, 418A to G and 419G to A (31)	17.3%, 19.8%, 20%, 20.1% 29.2%, 33.5% 38.5%

The different conversion efficiencies observed for each specific mutant EgD8S may be attributed to a "position effect" based on the respective locations of each gene's integration within the *Yarrowia* genome. In any case, the results demonstrate that several of the transformants comprising mutant EgD8S-23 (SEQ ID NO:4), mutant EgD8S-013 (SEQ ID NO:6) and mutant EgD8S-015 (SEQ ID NO:8) possessed Δ8 desaturase activity that was at least functionally equivalent (or increased) with respect to that of the synthetic codon-optimized EgD8S (SEQ ID NO:10).

Example 11

Generation Of *Yarrowia lipolytica* Strains Y4031, Y4032 And Y4033 To Produce About 10 -13.6% DGLA Of Total Lipids

The present Example describes the construction of strains Y4031, Y4032 and Y4033, derived from *Yarrowia lipolytica* Y4001U (Example 3), capable of producing 10-13.6% DGLA (C20:3) relative to the total lipids. These strains were engineered to express the Δ9 elongase/Δ8 desaturase pathway, via expression of a mutant Δ8 desaturase of the present invention and a Δ9 elongase.

More specifically, construct pKO2UF8289 (Figure 11B; SEQ ID NO:181) was created to integrate a cluster of four chimeric genes (comprising a Δ12 desaturase, two copies of the mutant EgD8-23 and one Δ9 elongase) into the Δ12 gene locus of *Yarrowia* genome in strain Y4001U. Construct pKO2UF8289 contained the following components:

Table 21

Description of Plasmid pKO2UF8289 (SEQ ID NO:181)

RE Sites And Nucleotides Within SEQ ID NO:181	Description Of Fragment And Chimeric Gene Components
AscI/BsiW I (10304-9567)	5' portion of *Yarrowia lipolytica* Δ12 desaturase gene (WO 2004/104167)
EcoRI/Sph I (13568-13012)	3' portion of *Yarrowia lipolytica* Δ12 desaturase gene (WO 2004/104167)
Cla I/*EcoR* I (1-13568)	LoxP::*Ura3*::LoxP, comprising: • LoxP sequence (SEQ ID NO:18) • *Yarrowia* Ura3 gene (GenBank Accession No. AJ306421) • LoxP sequence (SEQ ID NO:18)
PmeI/ClaI (2038-1)	GPAT::EgD9E::Lip2, comprising: • GPAT: *Yarrowia lipolytica* GPAT promoter (WO 2006/031937) • EgD9E: codon-optimized Δ9 elongase gene (SEQ ID NO:177), derived from *Euglena gracilis* (SEQ ID NOs:175 and 176; U.S. Patent Application No. 60/739989; see also Example 12 herein) • Lip2: Lip2 terminator sequence from *Yarrowia Lip1* gene (GenBank Accession No. AJ012632)
PmeI/PacI (4581-2124)	Exp::D8-23::Pex16, comprising: • Exp: *Yarrowia lipolytica* export protein (EXP1) promoter (WO 2006/052870 and U.S. Patent Application No. 11/265761) • D8-23: mutant EgD8S-23 (Example 8; SEQ ID NO:4) • Pex16: Pex16 terminator sequence of *Yarrowia Pex 16* gene (GenBank Accession No. U75433)
Swa I/Pme I (7055-4581)	YAT::F. D12::Oct, comprising: • YAT: *Yarrowia lipolytica* YAT1 promoter (Patent Publication US 2006/0094102-A1) • F.D12: *Fusarium moniliforme* Δ12 desaturase gene (WO 2005/047485)

	• OCT: OCT terminator sequence of *Yarrowia OCT* gene (GenBank Accession No. X69988)
Swa I/BsiW I (7055-9567)	FBAINm::D8S-23::Pex20, comprising: • FBAINm: *Yarrowia lipolytica* FBAINm promoter (WO 2005/049805) • D8S-23: mutant EgD8S-23 (Example 8; SEQ ID NO:4) • Pex20: Pex20 terminator sequence of *Yarrowia Pex20* gene (GenBank Accession No. AF054613)

Plasmid pKO2UF8289 was digested with *Asc I/Sph* I, and then used for transformation of *Y. lipolytica* strain Y4001U (Example 3) according to the General Methods. The transformant cells were plated onto MMLeu media plates and maintained at 30 °C for 2 to 3 days. The colonies were picked and streaked onto MMLeu selection plates at 30 °C for 2 days. These cells were then inoculated into liquid MMLeu at 30 °C and shaken at 250 rpm/min for 2 days. The cells were collected by centrifugation, lipids were extracted, and fatty acid methyl esters were prepared by trans-esterification, and subsequently analyzed with a Hewlett-Packard 6890 GC.

GC analyses showed the presence of DGLA in the transformants containing the 4 chimeric genes of pKO2UF8289, but not in the *Yarrowia* Y4001U control strain. Most of the selected 12 strains produced about 4 to 8% DGLA of total lipids. There were 3 strains (i.e., strains #7, #8 and #12) that produced about 11.3%, 10% and 13.6% DGLA of total lipids; they were designated as strains Y4031, Y4031 and Y4033, respectively.

Example 12

Identification Of A Δ9 Elongase From *Euglena gracilis*

The present Example, disclosed in U.S. Patent Application No. 60/739989, describes the isolation of a Δ9 elongase from *Euglena gracilis* (SEQ ID NOs:175 and 176).

Euglena gracilis Growth Conditions, Lipid Profile And mRNA Isolation

Euglena gracilis was obtained from Dr. Richard Triemer's lab at Michigan State University (East Lansing, MI). From 10 mL of actively growing culture, a 1 mL aliquot was transferred into 250 mL of *Euglena gracilis* (Eg) Medium in a 500 mL glass bottle. Eg medium was made by

combining 1 g of sodium acetate, 1 g of beef extract (U126-01, Difco Laboratories, Detroit, MI), 2 g of Bacto® tryptone (0123-17-3, Difco Laboratories), 2 g of Bacto® yeast extract (0127-17-9, Difco Laboratories) in 970 mL of water. After filter sterilizing, 30 mL of soil-water supernatant (15-3790, Carolina Biological Supply Company, Burlington, NC) was aseptically added to give the final Eg medium. *Euglena gracilis* cultures were grown at 23 °C with a 16 h light, 8 h dark cycle for 2 weeks with no agitation.

After 2 weeks, 10 mL of culture was removed for lipid analysis and centrifuged at 1,800 x g for 5 min. The pellet was washed once with water and re-centrifuged. The resulting pellet was dried for 5 min under vacuum, resuspended in 100 μL of trimethylsulfonium hydroxide (TMSH) and incubated at room temperature for 15 min with shaking. After this, 0.5 mL of hexane was added and the vials were incubated for 15 min at room temperature with shaking. Fatty acid methyl esters (5 μL injected from hexane layer) were separated and quantified using a Hewlett-Packard 6890 Gas Chromatograph fitted with an Omegawax 320 fused silica capillary column (Supelco Inc., Cat. No. 24152). The oven temperature was programmed to hold at 220 °C for 2.7 min, increase to 240 °C at 20 °C /min and then hold for an additional 2.3 min. Carrier gas was supplied by a Whatman hydrogen generator. Retention times were compared to those for methyl esters of standards commercially available (Nu-Chek Prep, Inc. Cat. No. U-99-A).

The remaining 2 week culture (240 mL) was pelleted by centrifugation at 1,800 x g for 10 min, washed once with water and re-centrifuged. Total RNA was extracted from the resulting pellet using the RNA STAT-60™ reagent (TEL-TEST, Inc., Friendswood, TX) and following the manufacturer's protocol provided (use 5 mL of reagent, dissolved RNA in 0.5 mL of water). In this way, 1 mg of total RNA (2 mg/mL) was obtained from the pellet. The mRNA was isolated from 1 mg of total RNA using the mRNA Purification Kit (Amersham Biosciences, Piscataway, NJ)

following the manufacturer's protocol provided. In this way, 85 μg of mRNA was obtained.

Euglena gracilis cDNA Synthesis, Library Construction And Sequencing

A cDNA library was generated using the Cloneminer™ cDNA Library Construction Kit (Cat. No.18249-029, Invitrogen Corporation, Carlsbad, CA) and following the manufacturer's protocol provided (Version B, 25-0608). Using the non-radiolabeling method, cDNA was synthesized from 3.2 μg of mRNA (described above) using the Biotin-*att*B2-Oligo(dT) primer. After synthesis of the first and second strand, the *att*B1 adapter was added, ligated and the cDNA was size fractionated using column chromatography. DNA from fractions 7 and 8 (size ranging from ~800-1500 bp) were concentrated, recombined into pDONR™222 and transformed into *E. coli* ElectroMAX™ DH10B™ T1 Phage-Resistant cells (Invitrogen Corporation). The *Euglena gracilis* library was named eeg1c.

For sequencing, clones first were recovered from archived glycerol cultures grown/frozen in 384-well freezing media plates, and replicated with a sterile 384 pin replicator (Genetix, Boston, MA) in 384-well microtiter plates containing LB + 75 μg/mL Kanamycin (replicated plates). Plasmids then were isolated, using the Templiphi DNA sequencing template amplification kit method (Amersham Biosciences) following the manufacturer's protocol. Briefly, the Templiphi method uses bacteriophage ϕ29 DNA polymerase to amplify circular single-stranded or double-stranded DNA by isothermal rolling circle amplification (Dean et al., *Genome Res.,* 11:1095-1099 (2001); Nelson et al., *Biotechniques,* 32:S44-S47 (2002)). After growing 20 h at 37 °C, cells from the replicated plate were added to 5 μL of dilution buffer and denatured at 95 °C for 3 min to partially lyse cells and release the denatured template. Templiphi premix (5 μL) was then added to each sample and the resulting reaction mixture was incubated at 30 °C for 16 h, then at 65 °C for 10 min to inactivate the ϕ29 DNA polymerase activity. DNA quantification with the

PicoGreen® dsDNA Quantitation Reagent (Molecular Probes) was performed after diluting the amplified samples 1:3 in distilled water.

The amplified products then were denatured at 95 °C for 10 min and end-sequenced in 384-well plates, using the M13F universal primer (SEQ ID NO:193), and the ABI BigDye version 3.1 Prism Sequencing Kit. For the sequencing reaction, 100-200 ng of templates and 6.4 pmol of primers were used, and the following reaction conditions were repeated 25 times: 96 °C for 10 sec, 50 °C for 5 sec and 60 °C for 4 min. After ethanol-based cleanup, cycle sequencing reaction products were resolved and detected on Perkin-Elmer ABI 3730xl automated sequencers.

Identification Of Long-Chain Polyunsaturated Fatty Acid Elongation Enzyme Homologs From *Euglena gracilis* cDNA Library eeg1c

cDNA clones encoding long-chain polyunsaturated fatty acid elongation enzyme homologs (LC-PUFA ELO homologs or Δ9 elongases) were identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul et al., *J. Mol. Biol.* 215:403-410 (1993)) searches for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL and DDBJ databases). The *Euglena gracilis* cDNA sequences obtained above were analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm provided by the National Center for Biotechnology Information (NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish and States, *Nat. Genet.*, 3:266-272 (1993)) provided by the NCBI. For convenience, the P-value (probability) of observing a match of a cDNA sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which represent the negative of the logarithm of the reported P-value. Accordingly, the greater

the pLog value, the greater the likelihood that the cDNA sequence and the BLAST "hit" represent homologous proteins.

The BLASTX search using the nucleotide sequences from clone eeg1c.pk001.n5.f revealed similarity of the protein encoded by the cDNA to the long-chain PUFA elongation enzyme from *Isochrysis galbana* (SEQ ID NO:173) (GenBank Accession No. AAL37626 (GI 17226123), locus AAL37626, CDS AF390174; Qi et al., *FEBS Lett.* 510(3):159-165 (2002)). The sequence of a portion of the cDNA insert from clone eeg1c.pk001.n5.f is shown in SEQ ID NO:194 (5' end of cDNA insert). Additional sequence was obtained from the 3' end of the cDNA insert of eeg1c.pk001.n5.1 as described above, but using the poly(A) tail-primed WobbleT oligonucleotides. Briefly, the WobbleT primer is an equimolar mix of 21mer poly(T)A, poly(T)C, and poly(T)G, used to sequence the 3' end of cDNA clones.

The 3' end sequence is shown in SEQ ID NO:195. Both the 5' and 3' sequences were aligned using Sequencher™ (Version 4.2, Gene Codes Corporation, Ann Arbor, MI) and the resulting sequence for the cDNA is shown in SEQ ID NO:196 (1201 bp). Sequence for the coding sequence from the cDNA in eeg1c.pk001.n5.f and the corresponding deduced amino acid sequence is shown in SEQ ID NO:175 (777 bp) and SEQ ID NO:176 (258 amino acids), respectively.

The amino acid sequence set forth in SEQ ID NO:176 was evaluated by BLASTP, yielding a pLog value of 38.70 (E value of 2e-39) versus the *Isochrysis galbana* sequence (SEQ ID NO:173). The *Euglena gracilis* Δ9 elongase is 39.4% identical to the *Isochrysis galbana* Δ9 elongase sequence using the Jotun Hein method. Sequence percent identity calculations performed by the Jotun Hein method (Hein, J. J., *Meth. Enz.,* 183:626-645 (1990)) were done using the MegAlign™ v6.1 program of the LASERGENE™ bioinformatics computing suite (DNASTAR™ Inc., Madison, WI) with the default parameters for pairwise alignment (KTUPLE=2). The *Euglena gracilis* Δ9 elongase is 31.8% identical to the *Isochrysis galbana* Δ9 elongase sequence using the

Clustal V method. Sequence percent identity calculations performed by the Clustal V method (Higgins, D.G. and Sharp, P.M., *Comput. Appl. Biosci.,* 5:151-153 (1989); Higgins et al., *Comput. Appl. Biosci.,* 8:189-191 (1992)) were done using the MegAlign™ v6.1 program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI) with the default parameters for pairwise alignment (KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5 and GAP LENGTH PENALTY=10). BLAST scores and probabilities indicate that the nucleic acid fragment described herein as SEQ ID NO:175 encodes an entire *Euglena gracilis* Δ9 elongase.

CLAIMS

What is claimed is:

1. An isolated polynucleotide comprising:
   (a) a nucleotide sequence encoding a mutant polypeptide having Δ8 desaturase activity having an amino acid sequence as set forth in SEQ ID NO:2 and wherein SEQ ID NO:2 is not identical to SEQ ID NO:10; or,
   (b) a complement of the nucleotide sequence of part (a), wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

2. The isolated polynucleotide of Claim 1 wherein the nucleotide sequence comprises SEQ ID NO:1 and wherein SEQ ID NO:1 is not identical to SEQ ID NO:9.

3. An isolated polynucleotide comprising:
   (a) a nucleotide sequence encoding a mutant polypeptide having Δ8 desaturase activity, having an amino acid sequence as set forth in SEQ ID NO: 198and wherein SEQ ID NO:198 is not identical to SEQ ID NO:10; or,
   (b) a complement of the nucleotide sequence of part (a), wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

4. The isolated polynucleotide of Claim 3 wherein the nucleotide sequence comprises SEQ ID NO:197 and wherein SEQ ID NO:197 is not identical to SEQ ID NO:9.

5. A polypeptide encoded by the isolated polynucleotide of claim 1 having Δ8 desaturase activity.

6. A polypeptide encoded by the isolated polynucleotide of claim 3 having Δ8 desaturase activity.

7. The polypeptide of claim 5 wherein the Δ8 desaturase activity is at least about functionally equivalent to the Δ8 desaturase activity of the polypeptide as set forth in SEQ ID NO:10.

8. The polypeptide of claim 6 wherein the Δ8 desaturase activity is at least about functionally equivalent to the Δ8 desaturase activity of the polypeptide as set forth in SEQ ID NO:10.

9. A recombinant construct comprising the isolated polynucleotide of any one of Claims 1 or 3 operably linked to at least one regulatory sequence.

10. A cell comprising the isolated polynucleotide of either claim 1 or claim 3.

11. The cell of Claim 10 wherein said cell is a yeast.

12. The cell of Claim 11 wherein the yeast is an oleaginous yeast producing at least about 25% of its dry cell weight as oil.

13. The cell of claim 12 wherein the oleaginous yeast is selected from the group consisting of: *Yarrowia, Candida, Rhodotorula, Rhodosporidium, Cryptococcus, Trichosporon* and *Lipomyces*.

14. A method for making long-chain polyunsaturated fatty acids in a yeast cell comprising:

(a) providing a yeast cell according to claim 10; and

(b) growing the yeast cell of (a) under conditions wherein long-chain polyunsaturated fatty acids are produced.

15. A method according to claim 14 wherein the yeast is oleaginous yeast producing at least about 25% of its dry cell weight as oil.

16. A method according to claim 15 wherein the yeast is a *Yarrowia sp.*

17. Microbial oil obtained from the yeast of claim 12.

18. An oleaginous yeast producing at least about 25% of its dry cell weight as oil comprising:

a) a first recombinant DNA construct comprising an isolated polynucleotide encoding a Δ8 desaturase polypeptide according to claim 1 or claim 3, operably linked to at least one regulatory sequence; and,

b) at least one second recombinant DNA construct comprising an isolated polynucleotide operably linked to at least one regulatory sequence, the construct encoding a polypeptide selected from the group consisting of: a Δ4 desaturase, a Δ5 desaturase, Δ6 desaturase, a Δ9 desaturase, a Δ12 desaturase, a Δ15 desaturase, a Δ17 desaturase, a Δ9 elongase, a $C_{14/16}$ elongase, a $C_{16/18}$ elongase, a $C_{18/20}$ elongase and a $C_{20/22}$ elongase.

19. The yeast of claim 18 selected from the group consisting of: *Yarrowia, Candida, Rhodotorula, Rhodosporidium, Cryptococcus, Trichosporon* and *Lipomyces*.

20. The yeast of Claim 19 wherein the yeast cell is a *Yarrowia sp.* and the oil comprises a long-chain polyunsaturated fatty acid selected

from the group consisting of: arachidonic acid, eicosadienoic acid, eicosapentaenoic acid, eicosatetraenoic acid, eicosatrienoic acid, dihomo-γ–linolenic acid, docosapentaenoic acid and docosahexaenoic acid.

21. The oleaginous yeast of claim 18 wherein the first recombinant DNA construct comprises a polynucleotide encoding a Δ8 desaturase polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:198.

22. A method for producing dihomo–γ–linoleic acid comprising:

a) providing an oleaginous yeast comprising:

(i) a recombinant construct encoding a Δ8 desaturase polypeptide having an amino acid sequence as set forth in SEQ ID NO:2, wherein SEQ ID NO:2 is not identical to SEQ ID NO:10; and,

(ii) a source of eicosadienoic acid;

b) growing the yeast of step (a) under conditions wherein the recombinant construct encoding a Δ8 desaturase polypeptide is expressed and eicosadienoic acid is converted to dihomo-γ-linoleic acid, and;

c) optionally recovering the dihomo-γ-linoleic acid of step (b).

23. A method for producing eicosatetraenoic acid comprising:

a) providing an oleaginous yeast comprising:

(i) a recombinant construct encoding a Δ8 desaturase polypeptide having an amino acid sequence as set forth in SEQ ID NO:2, wherein SEQ ID NO:2 is not identical to SEQ ID NO:10; and,

(ii) a source of eicosatrienoic acid;

b) growing the yeast of step (a) under conditions wherein the recombinant construct encoding a Δ8 desaturase

polypeptide is expressed and eicosatrienoic acid is converted to eicosatetraenoic acid, and;

c) optionally recovering the eicosatetraenoic acid of step (b).

24. A method for the production of dihomo–γ–linoleic acid comprising:

a) providing a yeast cell comprising:

i) a first recombinant DNA construct comprising the isolated polynucleotide of either Claim 1 or Claim 3 operably linked to at least one regulatory sequence, and;

ii) at least one second recombinant DNA construct comprising an isolated polynucleotide encoding a Δ9 elongase polypeptide, operably linked to at least one regulatory sequence;

b) providing the yeast cell of (a) with a source of linolenic acid, and;

c) growing the yeast cell of (b) under conditions wherein dihomo–γ–linoleic acid is formed.

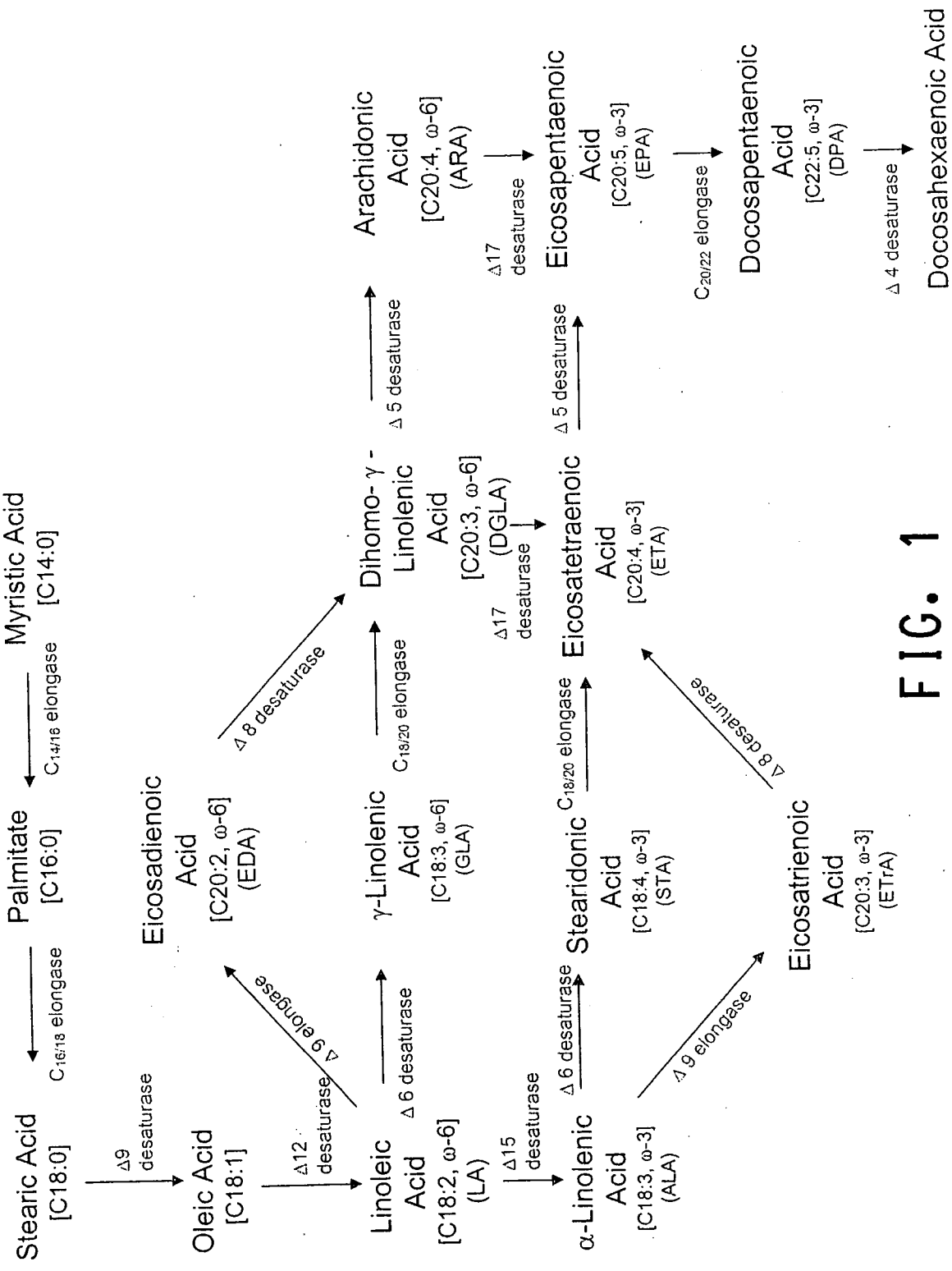
Stearic Acid [C18:0]
Palmitate [C16:0]
Myristic Acid [C14:0]
C16/18 elongase
C14/16 elongase
Δ9 desaturase
Oleic Acid [C18:1]
Eicosadienoic Acid [C20:2, ω-6] (EDA)
Δ12 desaturase
Δ 8 desaturase
Δ 9 elongase
Linoleic Acid [C18:2, ω-6] (LA)
γ-Linolenic Acid [C18:3, ω-6] (GLA)
Dihomo- γ - Linolenic Acid [C20:3, ω-6] (DGLA)
Arachidonic Acid [C20:4, ω-6] (ARA)
Δ 6 desaturase
C18/20 elongase
Δ 5 desaturase
Δ15 desaturase
Δ17 desaturase
Δ17 desaturase
α-Linolenic Acid [C18:3, ω-3] (ALA)
Stearidonic Acid [C18:4, ω-3] (STA)
Eicosatetraenoic Acid [C20:4, ω-3] (ETA)
Eicosapentaenoic Acid [C20:5, ω-3] (EPA)
Δ 6 desaturase
C18/20 elongase
Δ 5 desaturase
Δ 9 elongase
Δ 8 desaturase
C20/22 elongase
Eicosatrienoic Acid [C20:3, ω-3] (ETrA)
Docosapentaenoic Acid [C22:5, ω-3] (DPA)
Δ 4 desaturase
Docosahexaenoic Acid [C22:6, ω-3] (DHA)

FIG. 1

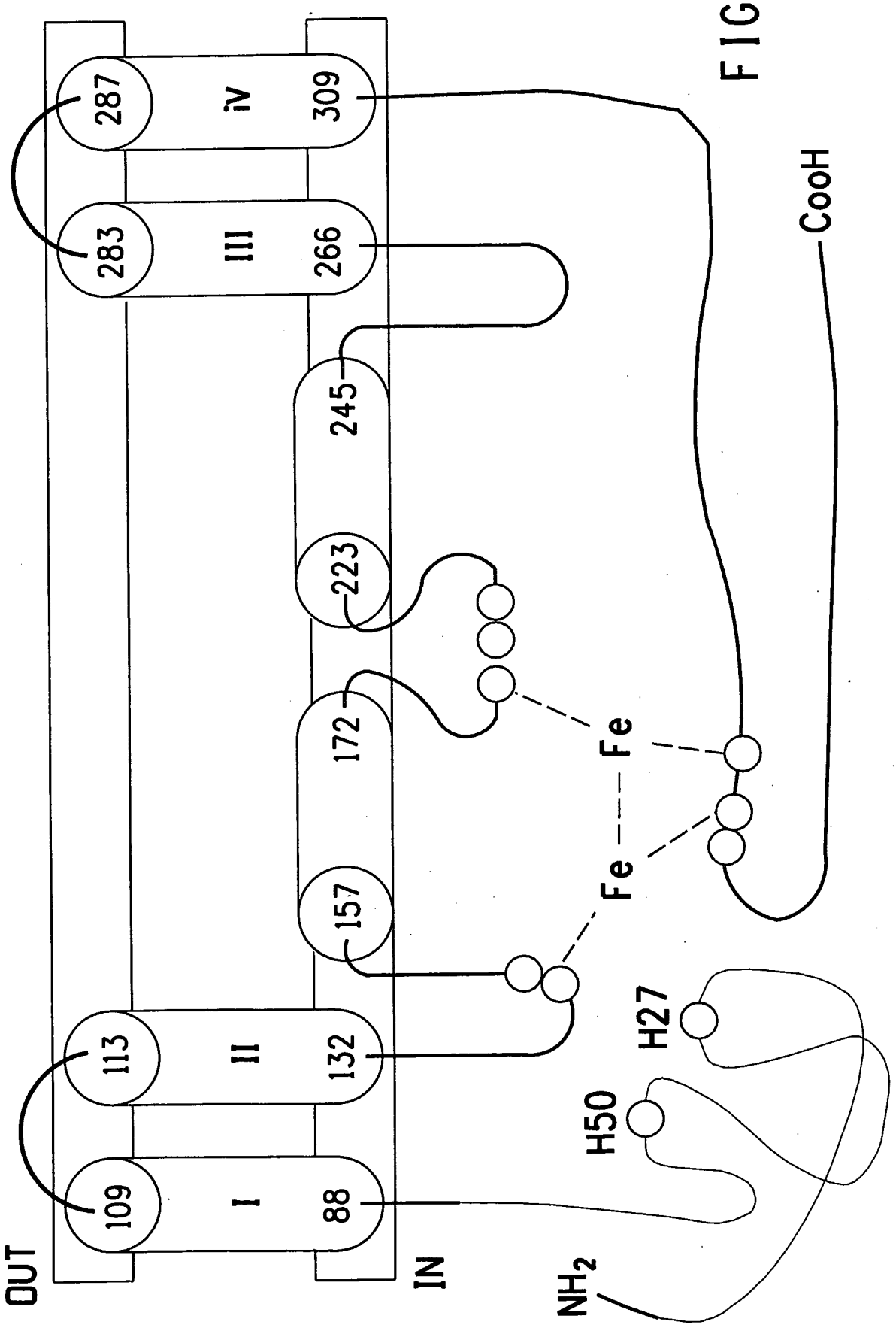
287
iv
309
283
III
266
245
223
172
157
113
II
132
109
I
88
OUT
IN
CooH
Fe
Fe
H27
H50
NH2

FIG. 2

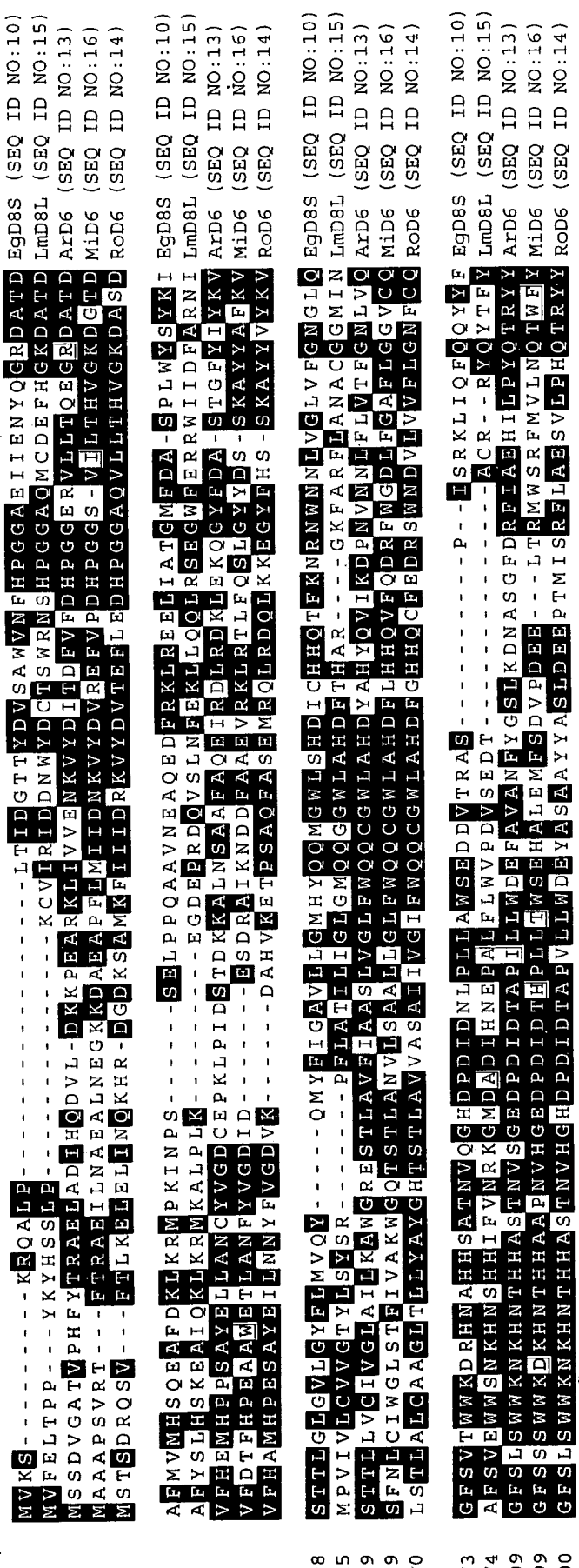
EgD8S (SEQ ID NO:10)
LmD8L (SEQ ID NO:15)
ArD6 (SEQ ID NO:13)
MiD6 (SEQ ID NO:16)
RoD6 (SEQ ID NO:14)

FIG. 3A

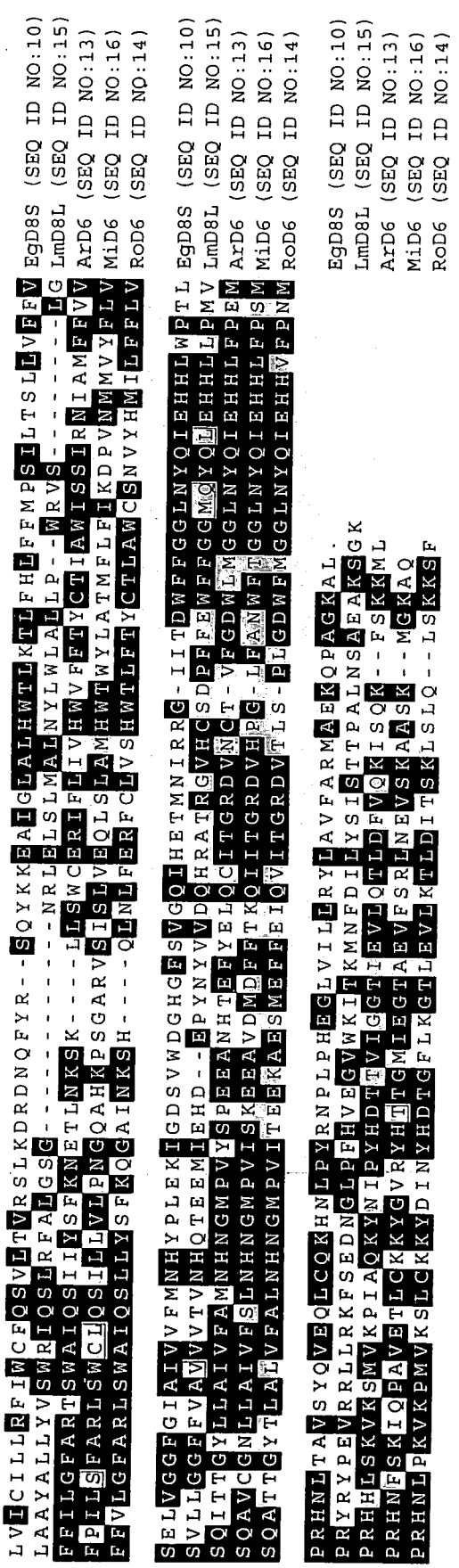
231 LVICILLRFIWCFQSVLTVRSLKDRDNQFYR--SQYKKEAIGLALHWTLKTLFHLFFMPSILTSLLVFFV EgD8S (SEQ ID NO:10)
228 LAAYALLYVSWRIQSLRFALGSG----------NRLELSLMALNYLWLALLP--WRVS-------LG LmD8L (SEQ ID NO:15)
279 FFILGFARTSWAIQSIIYSFKNETLNKSK----LLSWCERIFLIVHWVFFTYCTIAWISSIRNIAMFFVV ArD6 (SEQ ID NO:13)
266 FPILSFARLSWCIQSILLVLPNGQAHKPSGARVSISLVEQLSLAMHWTWYLATMFLFIKDPVNMVYFLV MiD6 (SEQ ID NO:16)
270 FFVLGFARLSWAIQSLLYSFKQGAINKSH----QLNLFERFCLVSHWTLFTYCTLAWCSNVYHMILFFLV RoD6 (SEQ ID NO:14)
299 SELVGGFGIAIVVFMNHYPLEKIGDSVWDGHGFSVGQIHETMNIRRG-IITDWFFGGLNYQIEHHLWPTL EgD8S (SEQ ID NO:10)
276 SVLLGGFFVAVVVTVNHQTEEMIEHD--EPYNYVVDQHRATRGVHCSDPFFEWFFGGMQYQLEHHLLPMV LmD8L (SEQ ID NO:15)
345 SQITTGYLLAIVFAMNHNGMPVYSPEEANHTEEFYELQCITGRDVNCT-VFGDWLMGGLNYQIEHHLFPEM ArD6 (SEQ ID NO:13)
336 SQAVCGNLLAIVFSLNHNGMPVISKEEAVDMDFFTKQIITGRDVHPG-LFANWFTGGLNYQIEHHLFPSM MiD6 (SEQ ID NO:16)
336 SQATTGYTLALVFALNHNGMPVITEEKAESMEFFEIQVITGRDVTLS-PLGDWFMGGLNYQIEHHVFPNM RoD6 (SEQ ID NO:14)
368 PRHNLTAVSYQVEQLCQKHNLPYRNPLPHEGLVILLRYLAVFARMAEKQPAGKAL. EgD8S (SEQ ID NO:10)
344 PRYRYPEVRLLRKFSEDNGLPFHVEGVWKITKMNFDILYSISTTPALNSAEAKSGK LmD8L (SEQ ID NO:15)
414 PRHHLSKVKSMVKPIAQKYNIPYHDTTVIGGTIEVLQTLDFVQKISQK--FSKKML ArD6 (SEQ ID NO:13)
405 PRHNFSKIQPAVETLCKKYGVRYHTTGMIEGTAEVFSRLNEVSKAASK--MGKAQ MiD6 (SEQ ID NO:16)
405 PRHNLPKVKPMVKSLCKKYDINYHDTGFLKGTLEVLKTLDITSKLSLQ--LSKKSF RoD6 (SEQ ID NO:14)

FIG. 3B

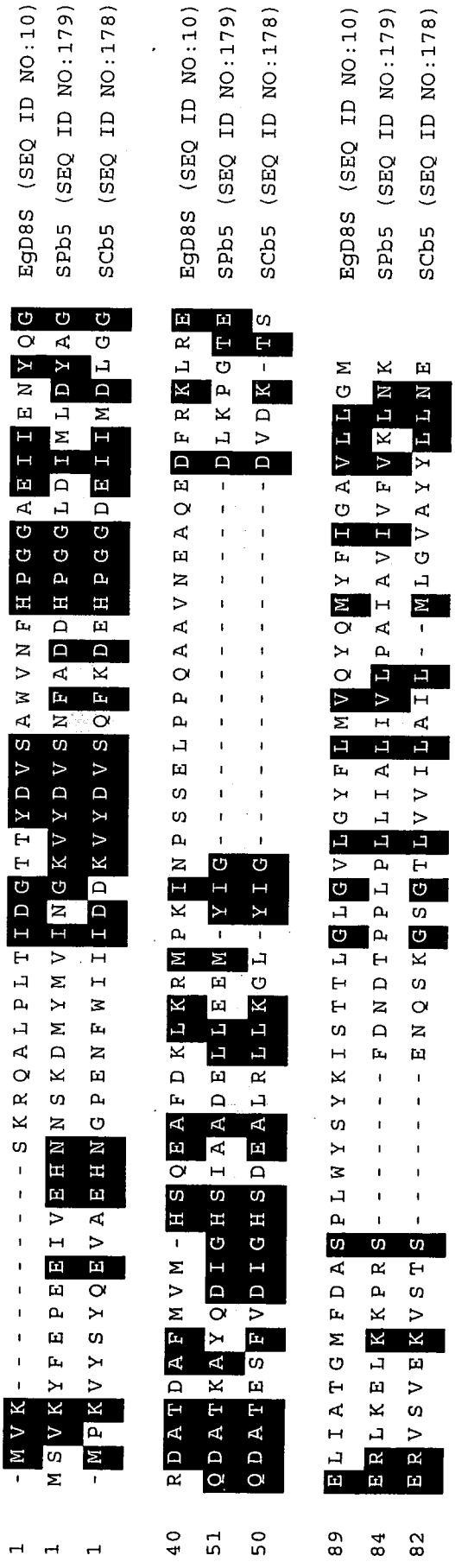
EgD8S (SEQ ID NO:10)
SPb5 (SEQ ID NO:179)
SCb5 (SEQ ID NO:178)
EgD8S (SEQ ID NO:10)
SPb5 (SEQ ID NO:179)
SCb5 (SEQ ID NO:178)
EgD8S (SEQ ID NO:10)
SPb5 (SEQ ID NO:179)
SCb5 (SEQ ID NO:178)

FIG. 4

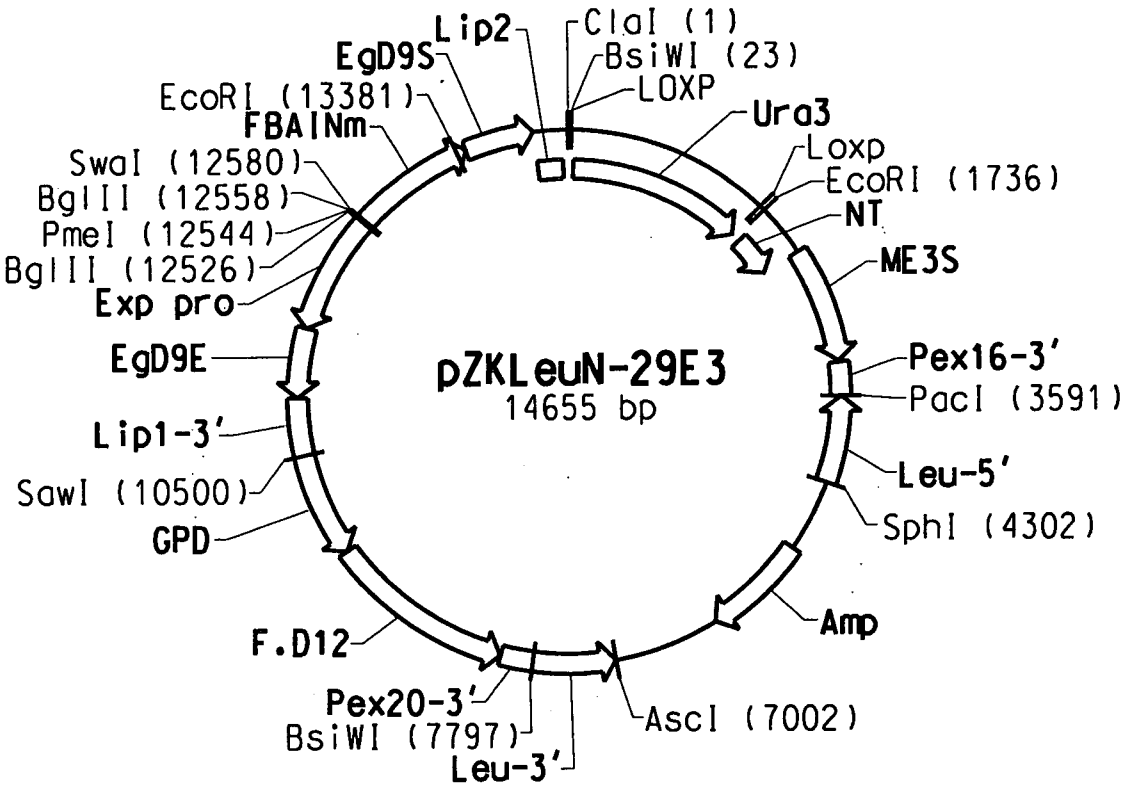
pZKLeuN-29E3
14655 bp
Lip2
ClaI (1)
BsiWI (23)
LOXP
Ura3
Loxp
EcoRI (1736)
NT
ME3S
Pex16-3'
PacI (3591)
Leu-5'
SphI (4302)
Amp
AscI (7002)
Leu-3'
BsiWI (7797)
Pex20-3'
F.D12
GPD
SawI (10500)
Lip1-3'
EgD9E
Exp pro
BglII (12526)
PmeI (12544)
BglII (12558)
SwaI (12580)
FBAINm
EcoRI (13381)
EgD9S

FIG. 5A

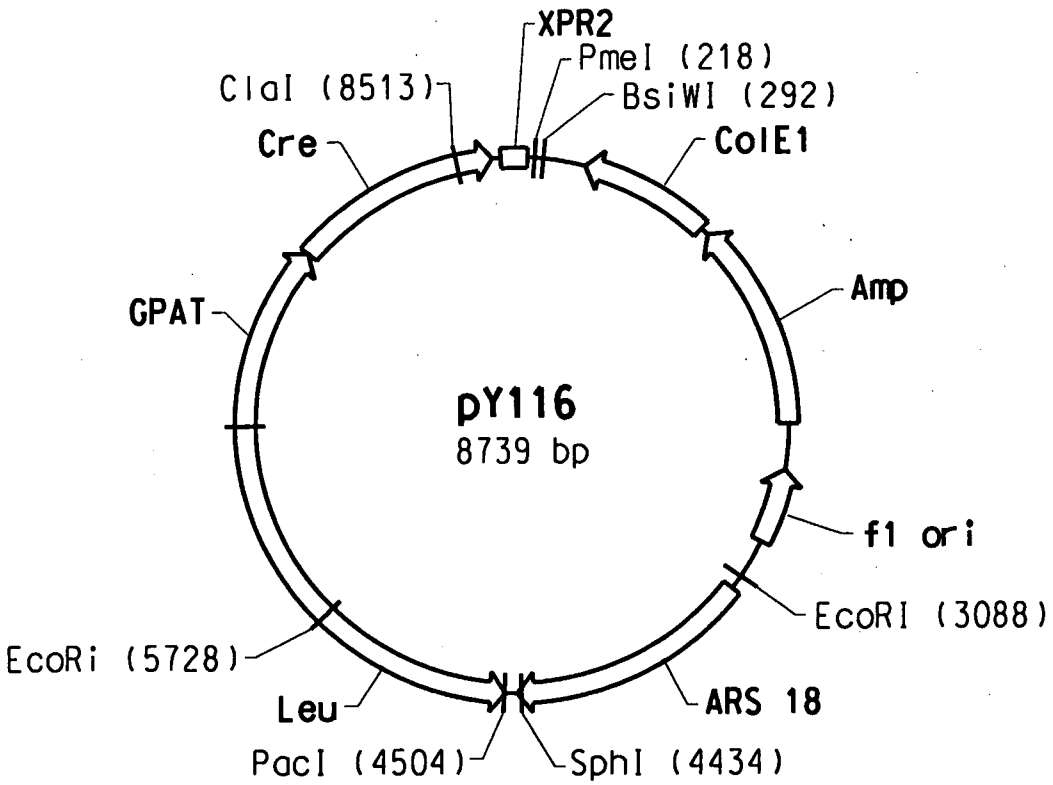
pY116
8739 bp
XPR2
PmeI (218)
BsiWI (292)
ColE1
Amp
f1 ori
EcoRI (3088)
ARS 18
SphI (4434)
PacI (4504)
Leu
EcoRi (5728)
GPAT
Cre
ClaI (8513)

FIG. 5B

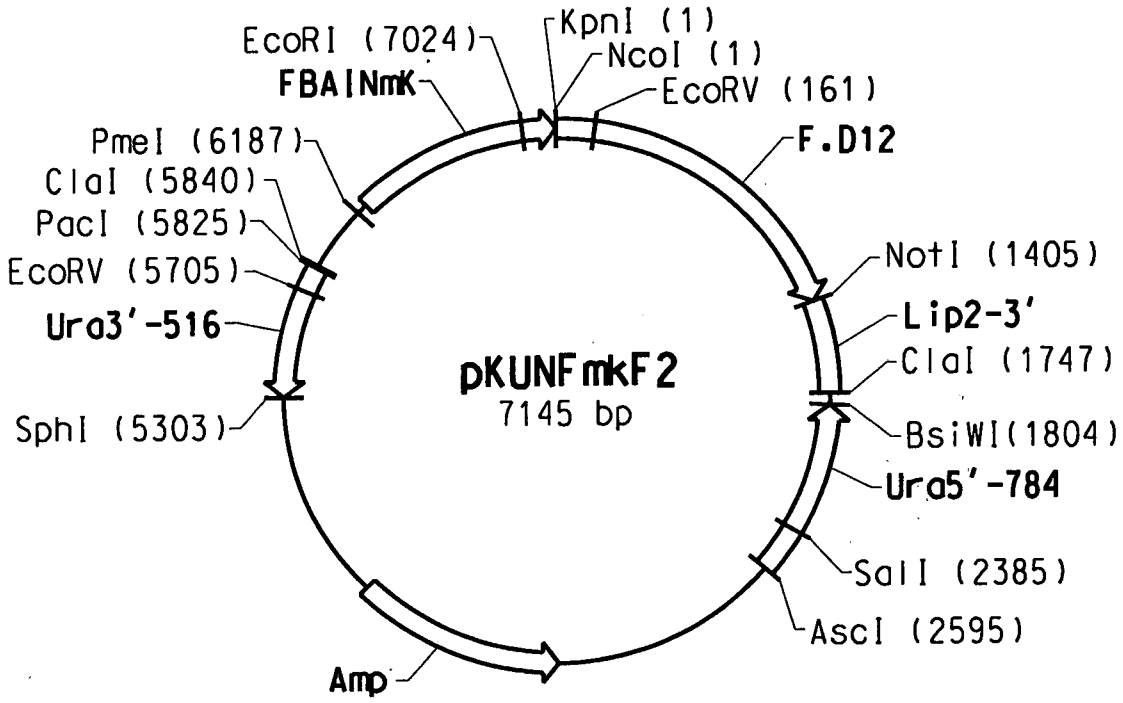
EcoRI (7024)
KpnI (1)
NcoI (1)
EcoRV (161)
FBAINmK
F.D12
PmeI (6187)
ClaI (5840)
PacI (5825)
EcoRV (5705)
Ura3'-516
NotI (1405)
Lip2-3'
ClaI (1747)
pKUNFmkF2
7145 bp
SphI (5303)
BsiWI(1804)
Ura5'-784
SalI (2385)
AscI (2595)
Amp

FIG. 6A

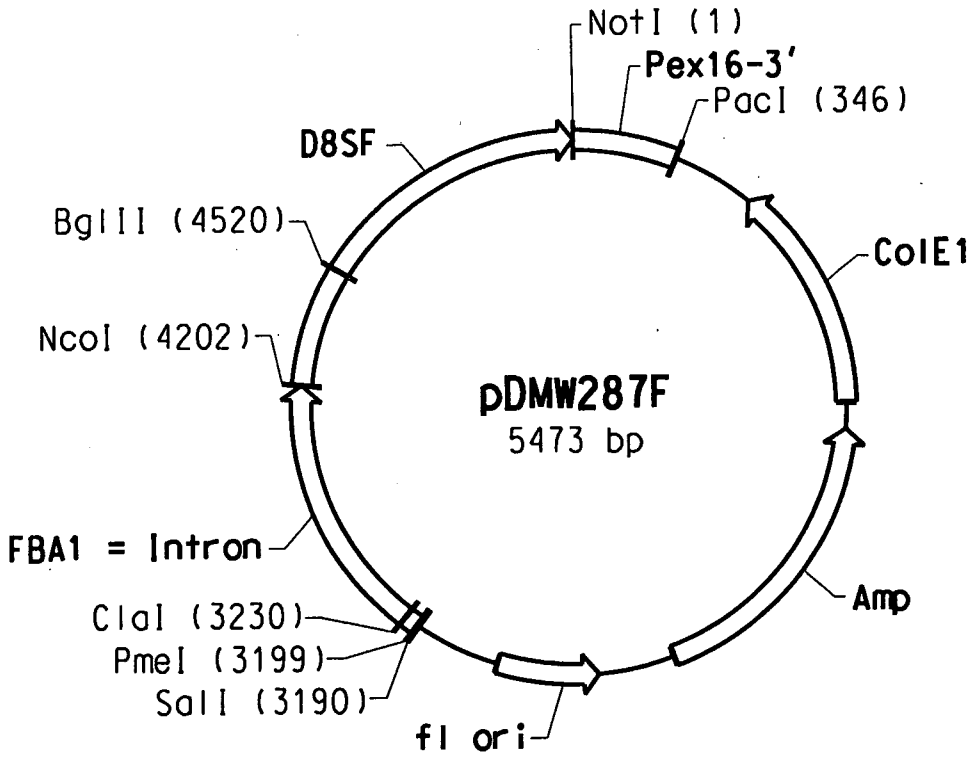
NotI (1)
Pex16-3'
PacI (346)
D8SF
BglII (4520)
ColE1
NcoI (4202)
pDMW287F
5473 bp
FBA1 = Intron
Amp
ClaI (3230)
PmeI (3199)
SalI (3190)
f1 ori

FIG. 6B

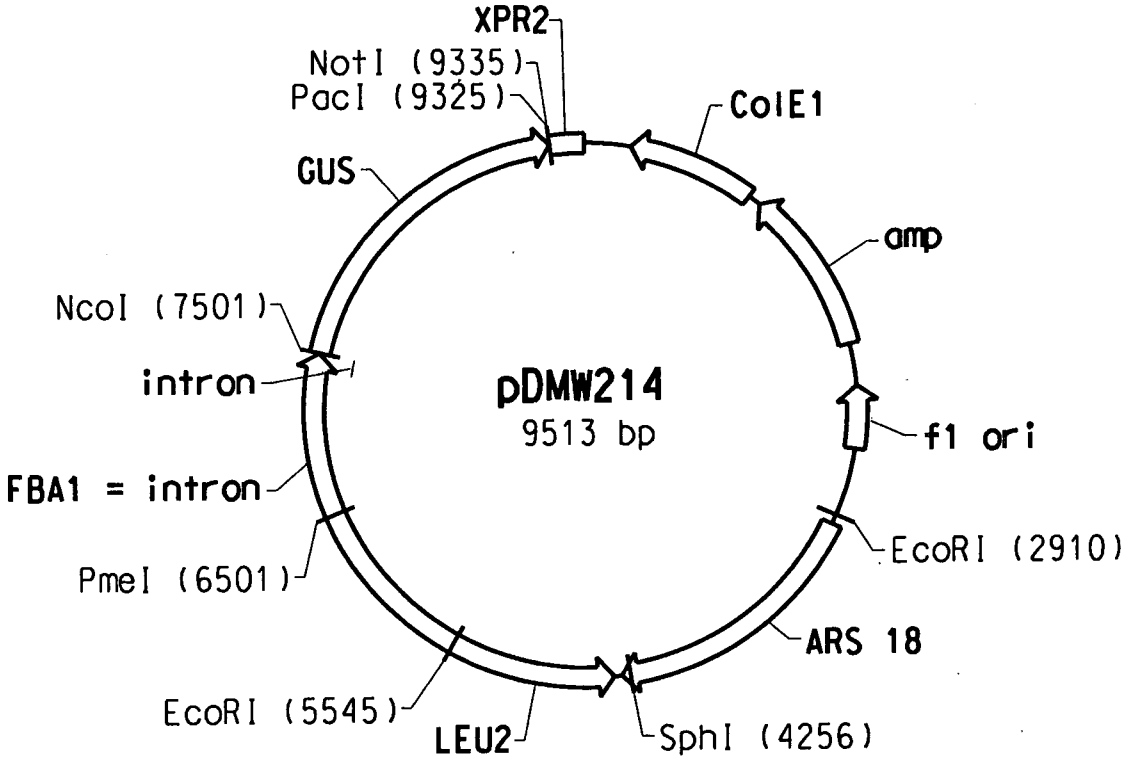
XPR2
NotI (9335)
PacI (9325)
ColE1
GUS
amp
NcoI (7501)
intron
pDMW214
9513 bp
f1 ori
FBA1 = intron
EcoRI (2910)
PmeI (6501)
ARS 18
EcoRI (5545)
LEU2
SphI (4256)

FIG. 6C

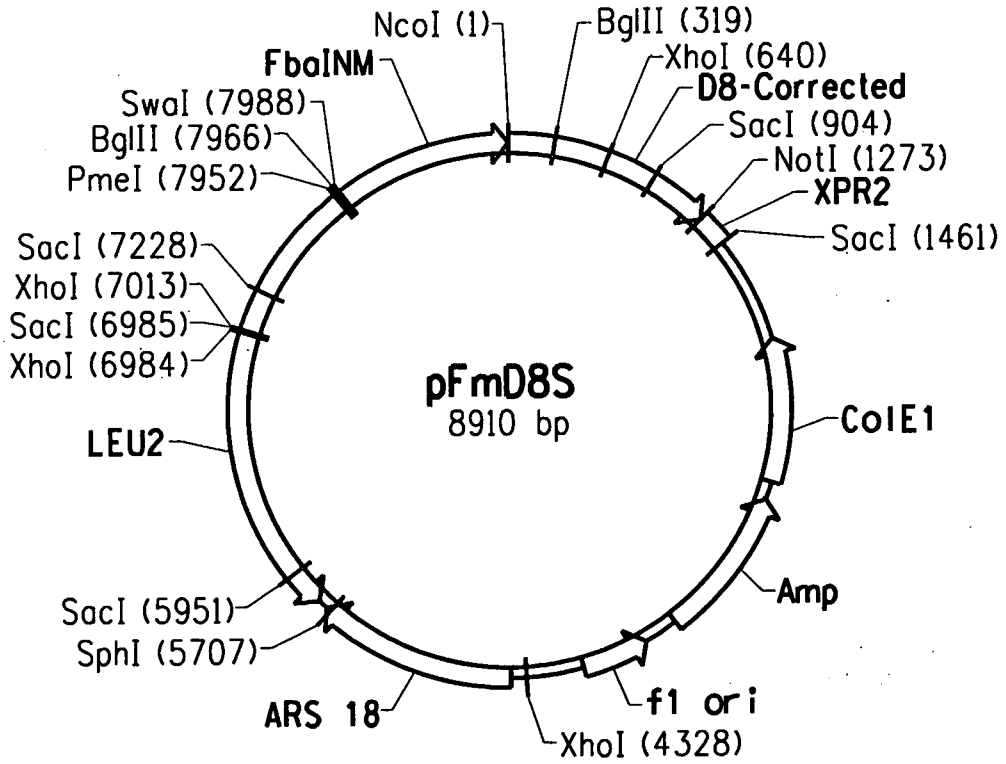
NcoI (1)
BglII (319)
FbaINM
XhoI (640)
D8-Corrected
SwaI (7988)
SacI (904)
BglII (7966)
NotI (1273)
PmeI (7952)
XPR2
SacI (1461)
SacI (7228)
XhoI (7013)
SacI (6985)
XhoI (6984)
pFmD8S
8910 bp
ColE1
LEU2
Amp
SacI (5951)
SphI (5707)
f1 ori
ARS 18
XhoI (4328)

FIG. 6D

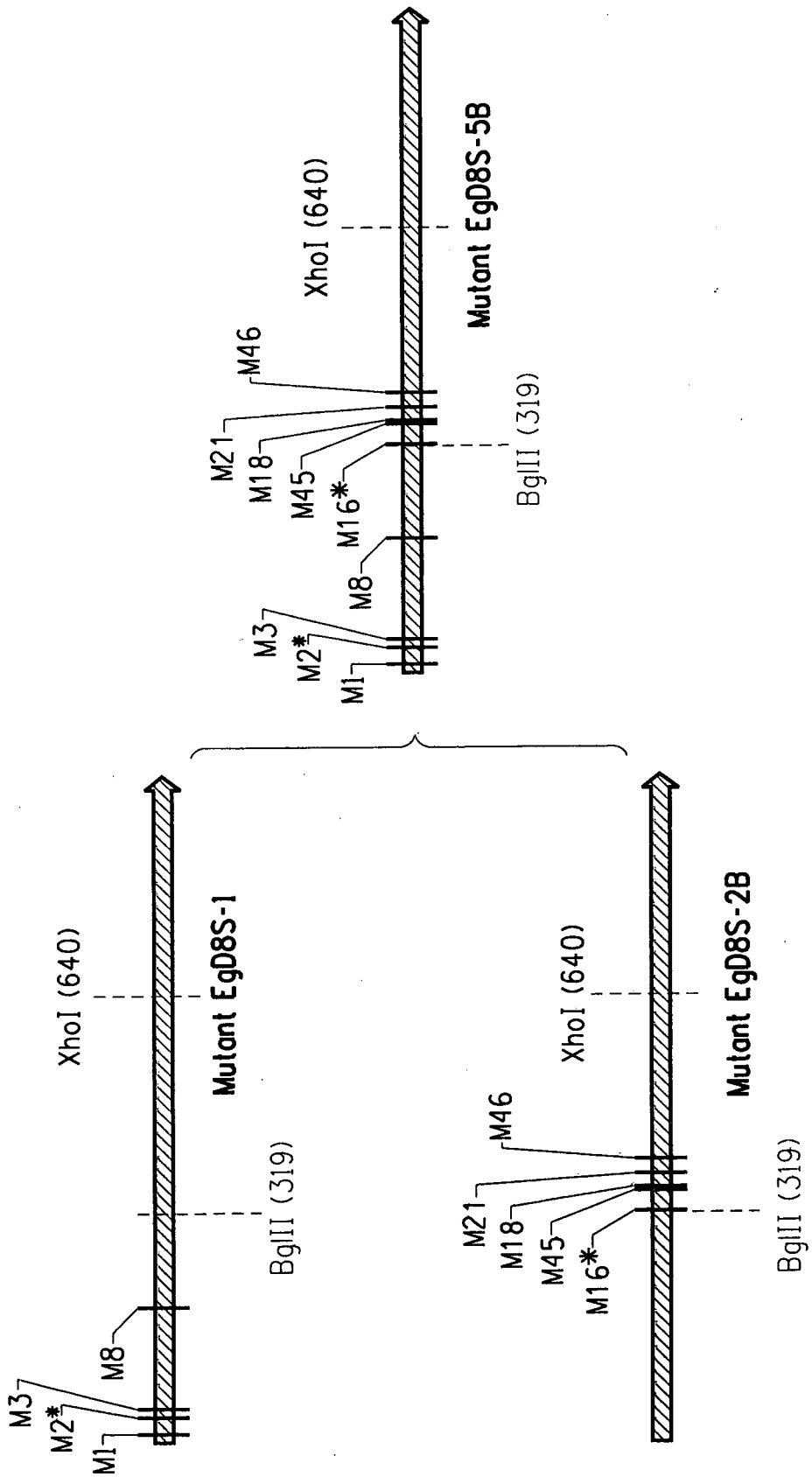
Mutant EgD8S-5B
XhoI (640)
M46
M21
M18
M45
M16*
M8
M3
M2*
M1
BglII (319)
Mutant EgD8S-1
XhoI (640)
M8
M3
M2*
M1
BglII (319)
Mutant EgD8S-2B
XhoI (640)
M46
M21
M18
M45
M16*
BglII (319)

FIG. 7

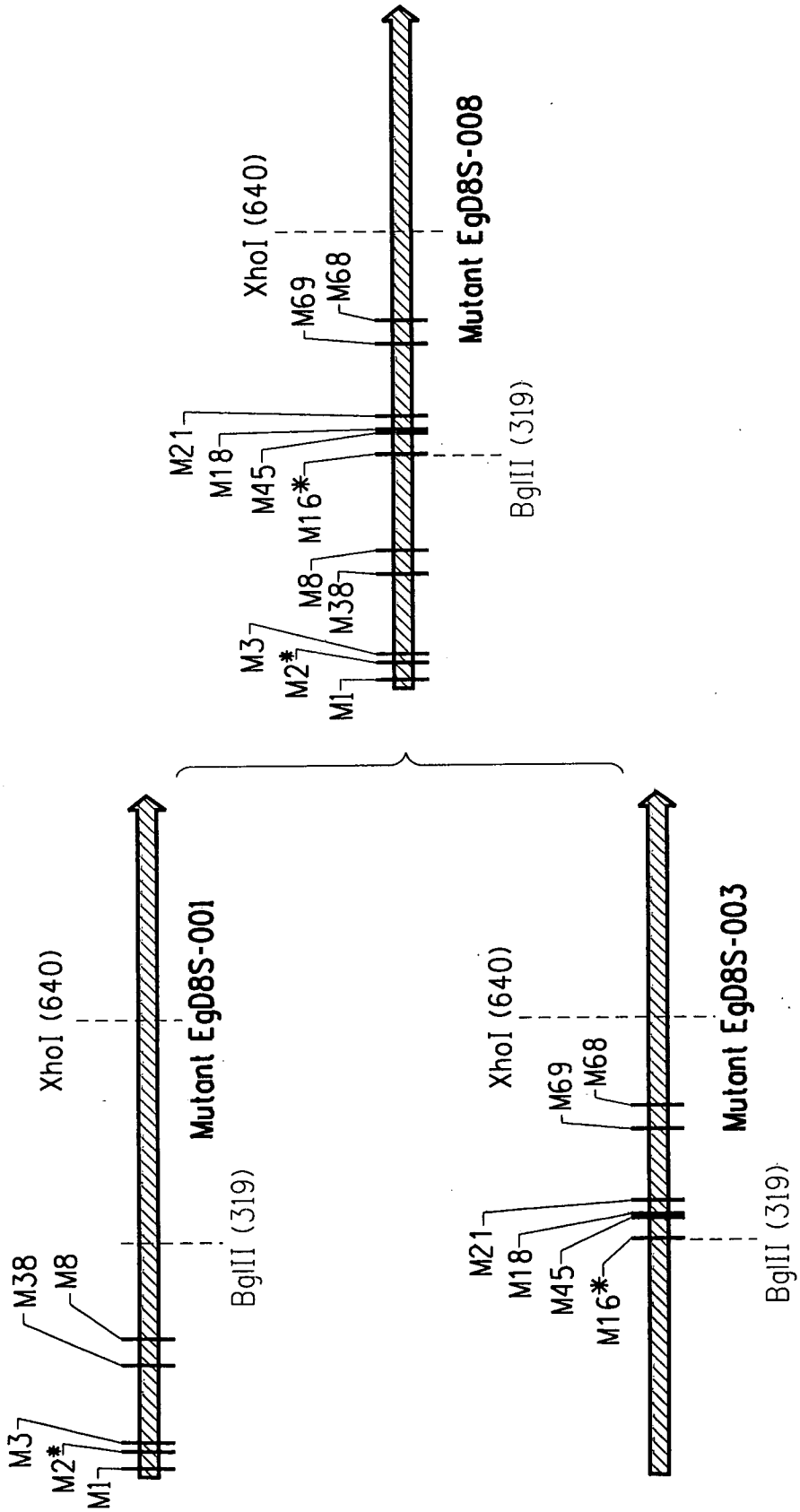
Mutant EgD8S-008
XhoI (640)
M69
M68
M21
M18
M45
M16*
BglII (319)
M8
M38
M3
M2*
M1
Mutant EgD8S-001
XhoI (640)
BglII (319)
M38
M8
M3
M2*
M1
Mutant EgD8S-003
XhoI (640)
M69
M68
M21
M18
M45
M16*
BglII (319)

FIG. 8A

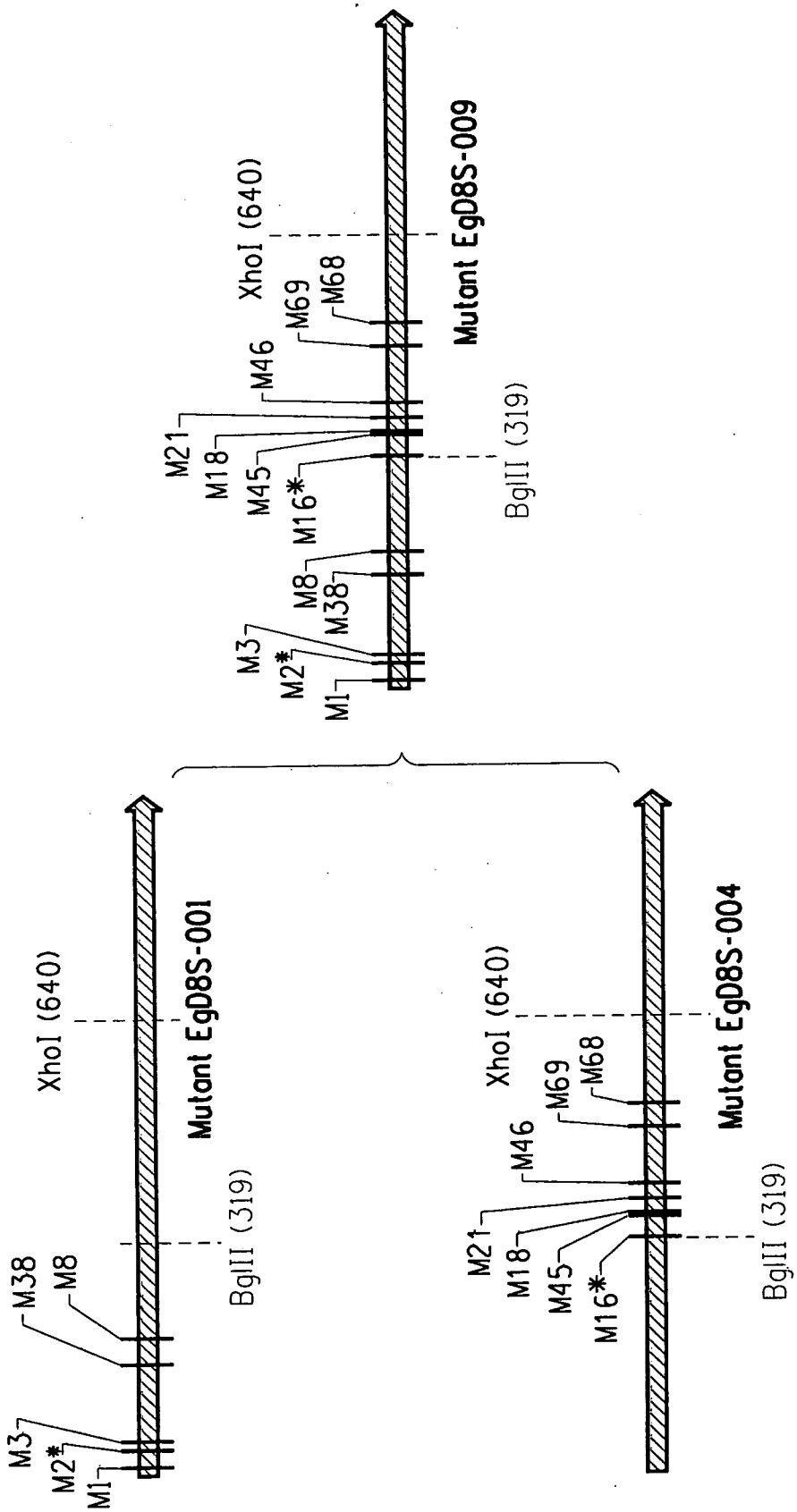
Mutant EgD8S-009
XhoI (640)
M69
M68
M46
M21
M18
M45
M16*
BglII (319)
M8
M38
M3
M2*
M1
Mutant EgD8S-001
XhoI (640)
BglII (319)
M38
M8
M3
M2*
M1
Mutant EgD8S-004
XhoI (640)
M69
M68
M46
M21
M18
M45
M16*
BglII (319)

FIG. 8B

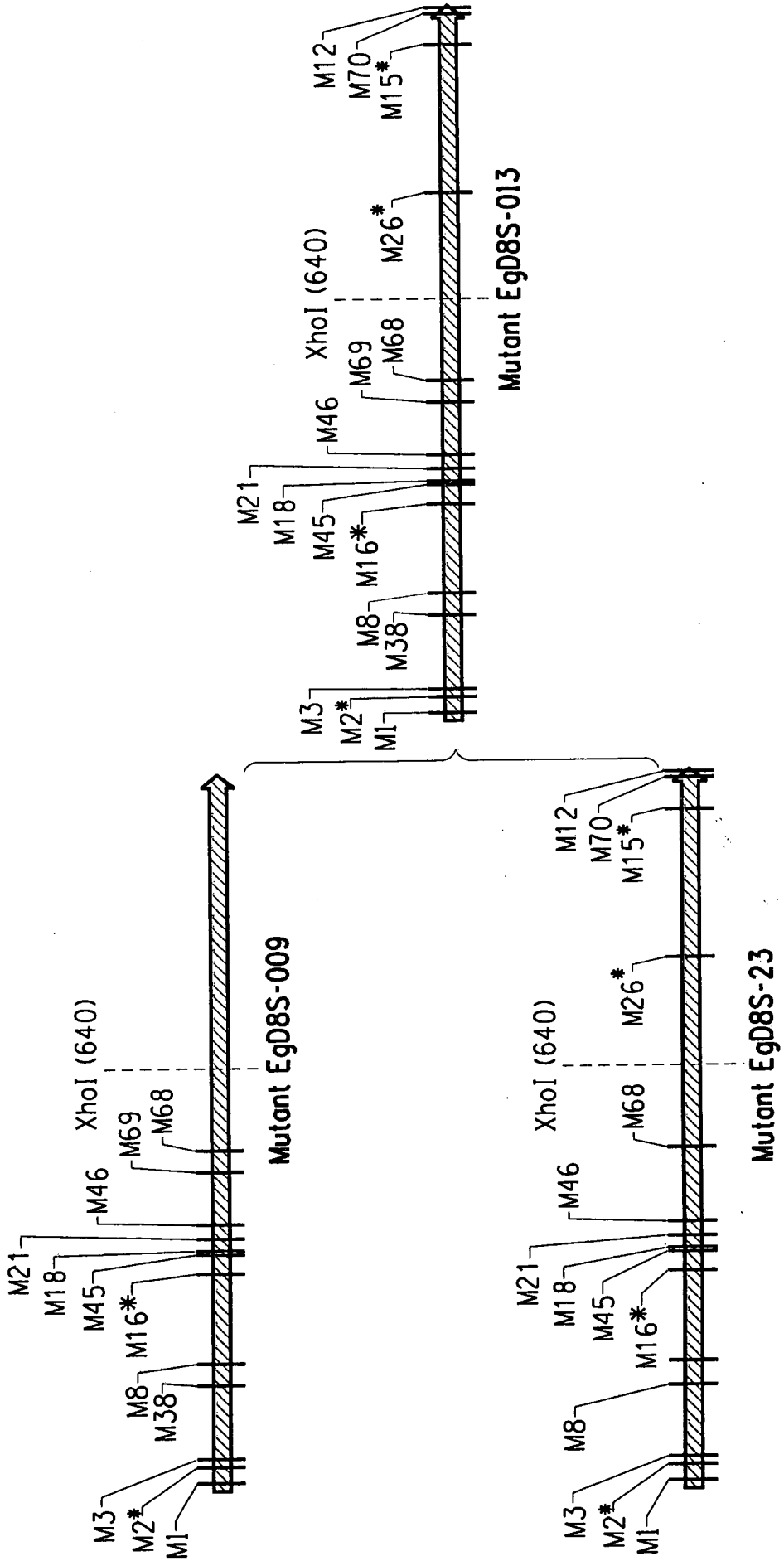
M12
M70
M15*
M26*
XhoI (640)
M69
M68
M46
M21
M18
M45
M16*
M8
M38
M3
M2*
M1
Mutant EgD8S-013
Mutant EgD8S-009
Mutant EgD8S-23

FIG. 9A

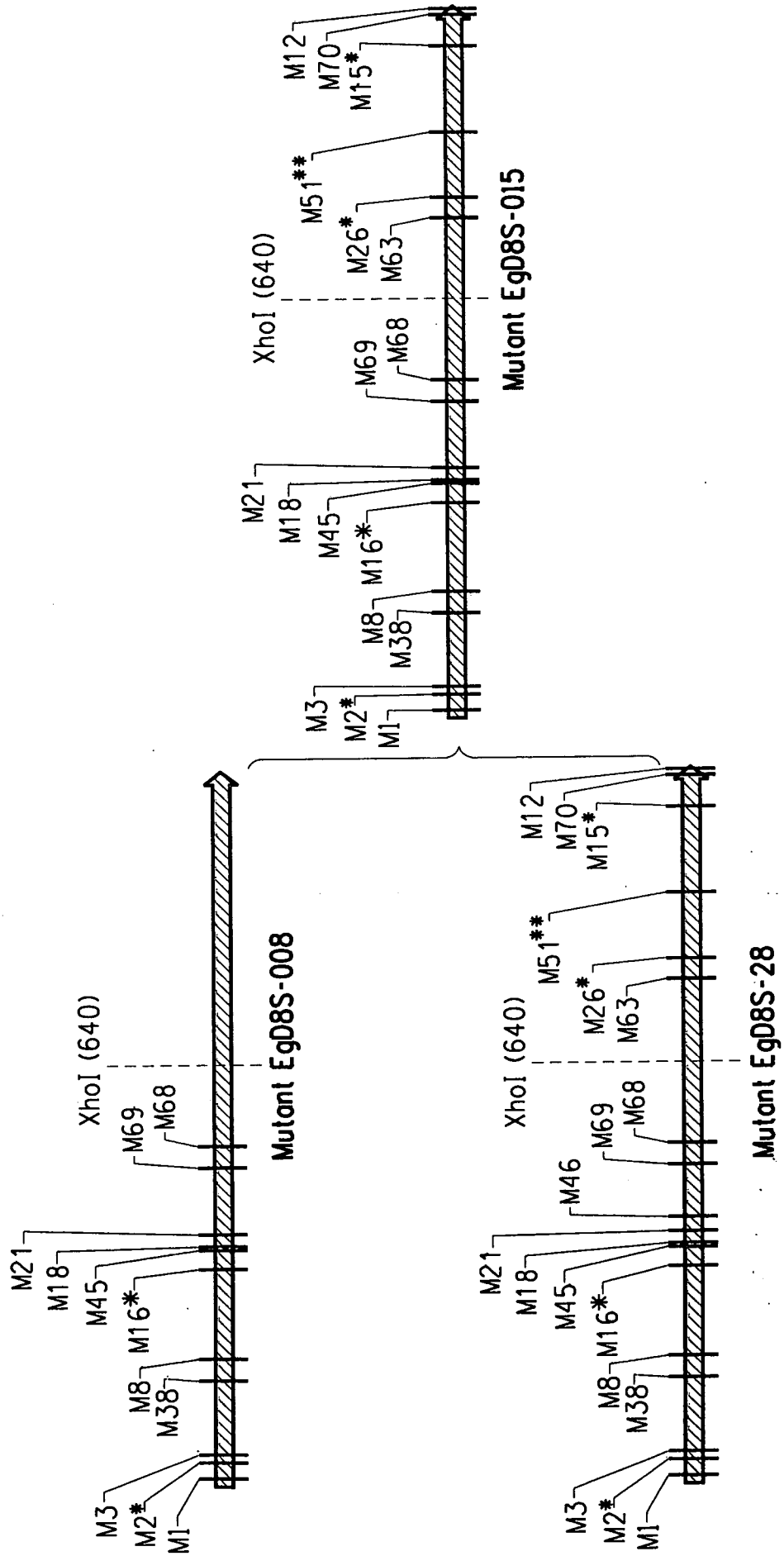
Mutant EgD8S-008
XhoI (640)
M21
M18
M45
M16*
M8
M38
M69
M68
M3
M2*
M1
Mutant EgD8S-015
XhoI (640)
M12
M70
M15*
M51**
M26*
M63
M69
M68
M21
M18
M45
M16*
M8
M38
M3
M2*
M1
Mutant EgD8S-28
XhoI (640)
M12
M70
M15*
M51**
M26*
M63
M69
M68
M46
M21
M18
M45
M16*
M8
M38
M3
M2*
M1

FIG. 9B


1 MVKSKRQALPLTIDGTTYDVSAWVNFHPGGAEIIENYQGRDATDAFMVMHSQEAFDKLKRMPKIN
1 MVKASRQALPLVIDGKVYDVSAWVNFHPGGAEIIENYQGRDATDAFMVMHSQEAFDKLKRMPKIN
1 MVKASRQALPLVIDGKVYDVSAWVNFHPGGAEIIENYQGRDATDAFMVMHSQEGYDKLKRMPKIN
1 MVKASRQALPLVIDGKVYDVSAWVNFHPGGAEIIENYQGRDATDAFMVMHSQEGYDKLKRMPKIN
EgD8S (SEQ ID NO:10)
EgD8S-23 (SEQ ID NO:4)
EgD8S-013 (SEQ ID NO:6)
EgD8S-015 (SEQ ID NO:8)
66 PSSELPPQAAVNEAQEDFRKLREELIATGMFDASPLWYSYKISTTLGLGVIGYFLMVQYQMYFIG
66 QASELPPQAAVNEAQEDFRKLREELIATGMFDASPLWYSYKILTTLGLGVLAFFMLVQYHLYFIG
66 QASELPPQAAVNEAQEDFRKLREELIATGMFDASPLWYSYKILTTLGLGVLAFFMLVQYHLYFIG
66 QASELPPQAAVNEAQEDFRKLREELIATGMFDASPLWYSYKILTTLGLGVLAFFMLVQYHLYFIG
EgD8S (SEQ ID NO:10)
EgD8S-23 (SEQ ID NO:4)
EgD8S-013 (SEQ ID NO:6)
EgD8S-015 (SEQ ID NO:8)
131 AVLLGMHYQQMGWLSHDICHHQTFKNRNWNNLVGLVFGNGLQGFSVTWWKDRHNAHHSATNVQGH
131 ALVLGMHYQQMGWLSHDICHHQTFKNRNWNNVLGLVFGNGLQGFSVTWWKDRHNAHHSATNVQGH
131 ALVLGMHYQQMGWLSHDICHHQTFKNRNWNNVLGLVFGNAVQGFSVTWWKDRHNAHHSATNVQGH
131 AVLLGMHYQQMGWLSHDICHHQTFKNRNWNNVLGLVFGNAVQGFSVTWWKDRHNAHHSATNVQGH
EgD8S (SEQ ID NO:10)
EgD8S-23 (SEQ ID NO:4)
EgD8S-013 (SEQ ID NO:6)
EgD8S-015 (SEQ ID NO:8)
196 DPDIDNLPLLAWSEDDVTRASPISRKLIQFQQYYFLVICILLRFIWCFQSVLTVRSLKDRDNQFY
196 DPDIDNLPLLAWSEDDVTRASPISRKLIQFQQYYFLVICILLRFIWCFQSVLTVRSLKDRDNQFY
196 DPDIDNLPLLAWSEDDVTRASPISRKLIQFQQYYFLVICILLRFIWCFQSVLTVRSLKDRDNQFY
196 DPDIDNLPLLAWSEDDVTRASPISRKLIQFQQYYFLVICILLRFIWCFQSVLTVRSLKDRDNQFY
EgD8S (SEQ ID NO:10)
EgD8S-23 (SEQ ID NO:4)
EgD8S-013 (SEQ ID NO:6)
EgD8S-015 (SEQ ID NO:8)

FIG. 10A

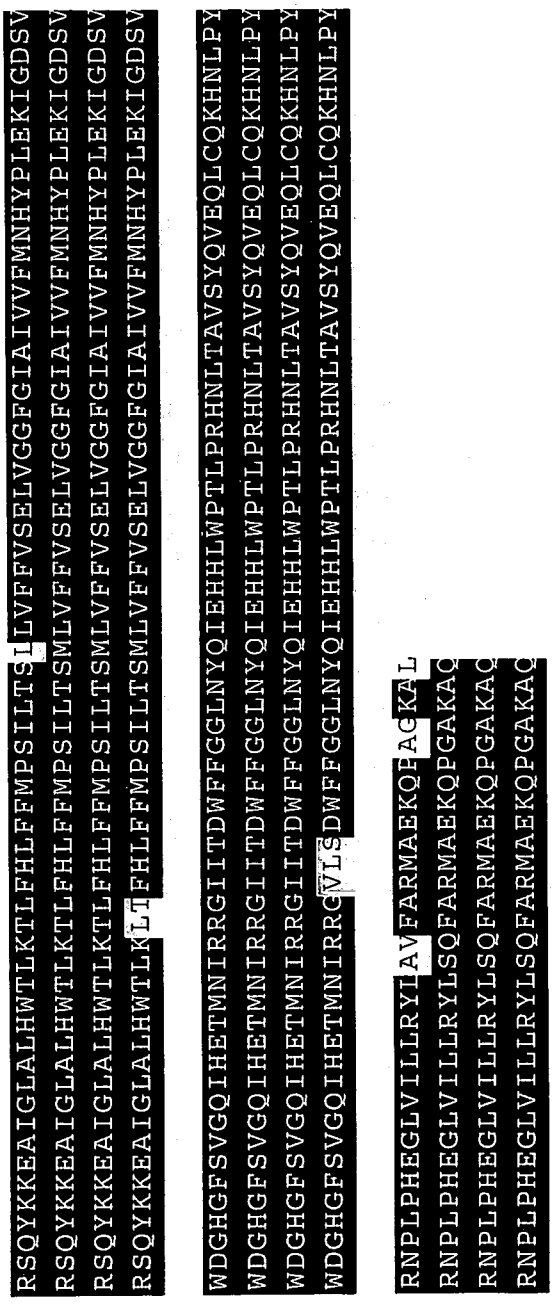
261 RSQYKKEAIGLALHWTLKTLFHLFFMPSILTSLVFFVSELVGGFGIAIVVFMNHYPLEKIGDSV EgD8S (SEQ ID NO:10)
261 RSQYKKEAIGLALHWTLKTLFHLFFMPSILTSMLVFFVSELVGGFGIAIVVFMNHYPLEKIGDSV EgD8S-23 (SEQ ID NO:4)
261 RSQYKKEAIGLALHWTLKTLFHLFFMPSILTSMLVFFVSELVGGFGIAIVVFMNHYPLEKIGDSV EgD8S-013 (SEQ ID NO:6)
261 RSQYKKEAIGLALHWTLKTLFHLFFMPSILTSMLVFFVSELVGGFGIAIVVFMNHYPLEKIGDSV EgD8S-015 (SEQ ID NO:8)
326 WDGHGFSVGQIHETMNIRRGIITDWFFGGLNYQIEHHLWPTLPRHNLTAVSYQVEQLCQKHNLPY EgD8S (SEQ ID NO:10)
326 WDGHGFSVGQIHETMNIRRGIITDWFFGGLNYQIEHHLWPTLPRHNLTAVSYQVEQLCQKHNLPY EgD8S-23 (SEQ ID NO:4)
326 WDGHGFSVGQIHETMNIRRGIITDWFFGGLNYQIEHHLWPTLPRHNLTAVSYQVEQLCQKHNLPY EgD8S-013 (SEQ ID NO:6)
326 WDGHGFSVGQIHETMNIRRGVLSDWFFGGLNYQIEHHLWPTLPRHNLTAVSYQVEQLCQKHNLPY EgD8S-015 (SEQ ID NO:8)
391 RNPLPHEGLVILLRYLAVFARMAEKQPAGKAL EgD8S (SEQ ID NO:10)
391 RNPLPHEGLVILLRYLSQFARMAEKQPGAKAQ EgD8S-23 (SEQ ID NO:4)
391 RNPLPHEGLVILLRYLSQFARMAEKQPGAKAQ EgD8S-013 (SEQ ID NO:6)
391 RNPLPHEGLVILLRYLSQFARMAEKQPGAKAQ EgD8S-015 (SEQ ID NO:8)

FIG. 10B

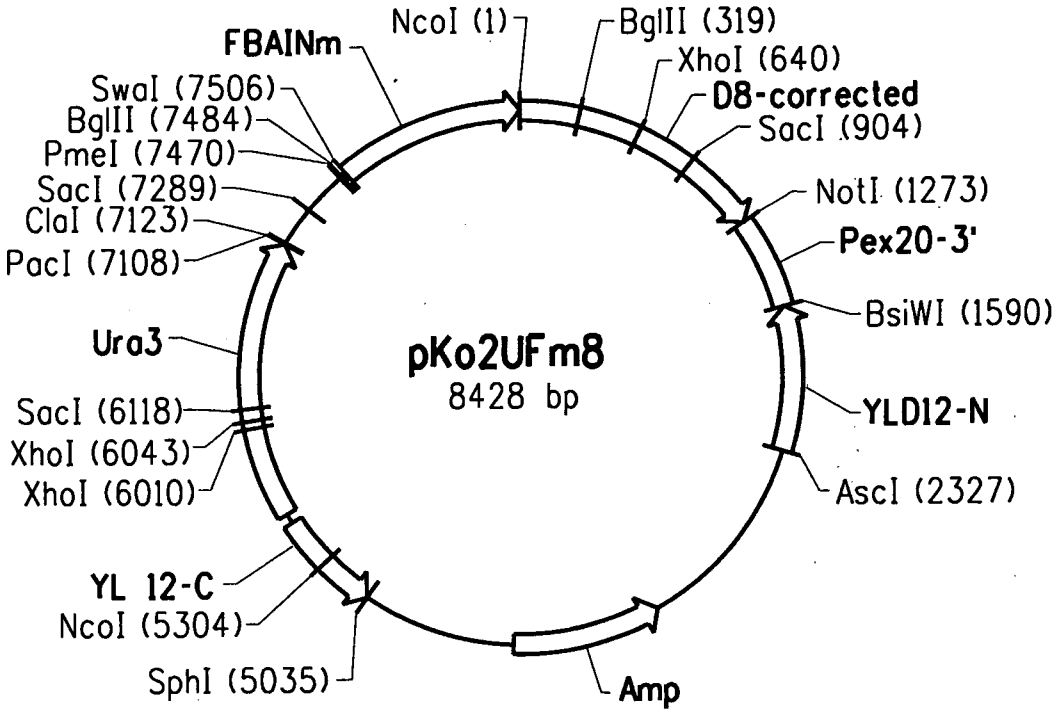

FIG. 11A

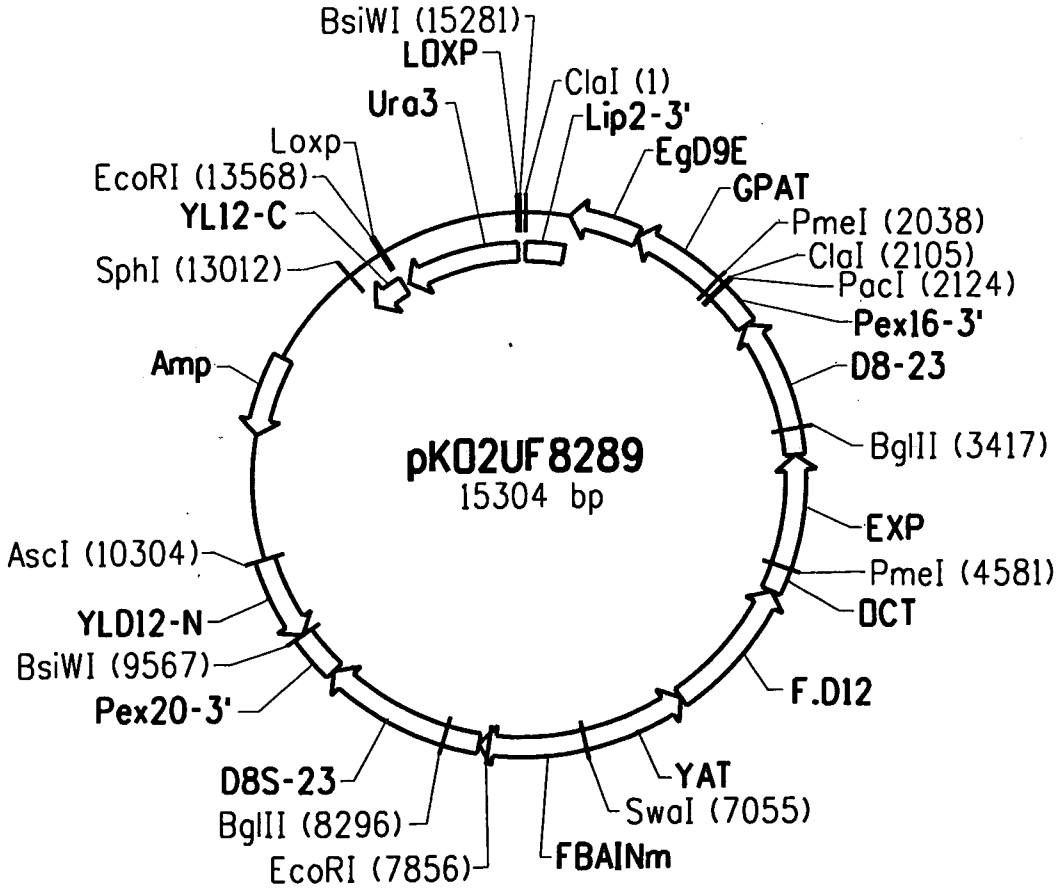

FIG. 11B

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2007/025001

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N9/12

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)
C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, PAJ, Sequence Search, WPI Data, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 2007/127381 A (DU PONT [US]; DAMUDE HOWARD GLENN [US]; ZHU QUINN QUN [US]) 8 November 2007 (2007-11-08) the whole document	1-23
X	OTTEN L.G. ET AL.: "Directed evolution: selecting todays biocatalysts" BIOMOLECULAR ENGINEERING, vol. 22, 2005, pages 1-9, XP002476123 the whole document	1-24
X	WO 2006/012326 A (DU PONT [US]; DAMUDE HOWARD G [US]; ZHU QUINN QUN [US]) 2 February 2006 (2006-02-02) page 43 - page 53; claims 1-22; example 14 examples 1-5	1-24

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* 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 document 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
11 April 2008	06/05/2008

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Keller, Yves

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2007/025001

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2006/012325 A (DU PONT [US]; DAMUDE HOWARD G [US]; ZHU QUINN QUN [US]) 2 February 2006 (2006-02-02)	
A	WALLIS J G ET AL: "The DELTA8-desaturase of Euglena gracilis: An alternate pathway for synthesis of 20-carbon polyunsaturated fatty acids" ARCHIVES OF BIOCHEMISTRY AND BIOPHYSICS, NEW YORK, US, US, vol. 365, no. 2, 15 May 1999 (1999-05-15), pages 307-316, XP002291433 ISSN: 0003-9861	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2007/025001

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2007127381 A	08-11-2007	NONE	
WO 2006012326 A	02-02-2006	AU 2005267202 A1	02-02-2006
		BR PI0512481 A	11-03-2008
		CA 2568624 A1	02-02-2006
		EP 1758991 A1	07-03-2007
		WO 2006012325 A1	02-02-2006
WO 2006012325 A	02-02-2006	AU 2005267202 A1	02-02-2006
		BR PI0512481 A	11-03-2008
		CA 2568624 A1	02-02-2006
		EP 1758991 A1	07-03-2007
		WO 2006012326 A1	02-02-2006