



US 20050273885A1

(19) **United States**(12) **Patent Application Publication** (10) **Pub. No.: US 2005/0273885 A1**
(43) **Pub. Date: Dec. 8, 2005**

(54) **SYNTHESIS OF LONG-CHAIN
POLYUNSATURATED FATTY ACIDS BY
RECOMBINANT CELLS**(76) **Inventors:** **Surinder Pal Singh**, Downer (AU);
Stanley Suresh Robert, Oyster Cove
(AU); **Peter David Nichols**, West
Hobart (AU); **Susan Irene Ellis**
Blackburn, Battery Point (AU);
Xue-Rong Zhou, Evatt (AU); **James**
Robertson Petrie, Bywong (AU);
Allan Graham Green, Braddon (AU)**Correspondence Address:**
COOPER & DUNHAM, LLP
1185 AVENUE OF THE AMERICAS
NEW YORK, NY 10036(21) **Appl. No.: 11/112,882**(22) **Filed: Apr. 22, 2005****Related U.S. Application Data**(60) **Provisional application No. 60/564,627**, filed on Apr.
22, 2004. **Provisional application No. 60/613,861**,
filed on Sep. 27, 2004. **Provisional application No.**
60/668,705, filed on Apr. 5, 2005.(30) **Foreign Application Priority Data**

Apr. 5, 2005 (AU)..... 2005901673

Publication Classification(51) **Int. Cl.⁷** **A01H 1/00**; C12N 15/82;
C12P 7/64; C12N 5/04(52) **U.S. Cl.** **800/281**; 435/134; 435/419;
435/468(57) **ABSTRACT**

The present invention relates to methods of synthesizing long-chain polyunsaturated fatty acids, especially eicosapentaenoic acid, docosapentaenoic acid and docosahexaenoic acid, in recombinant cells such as yeast or plant cells. Also provided are recombinant cells or plants which produce long-chain polyunsaturated fatty acids. Furthermore, the present invention relates to a group of new enzymes which possess desaturase or elongase activity that can be used in methods of synthesizing long-chain poly unsaturated fatty acids.

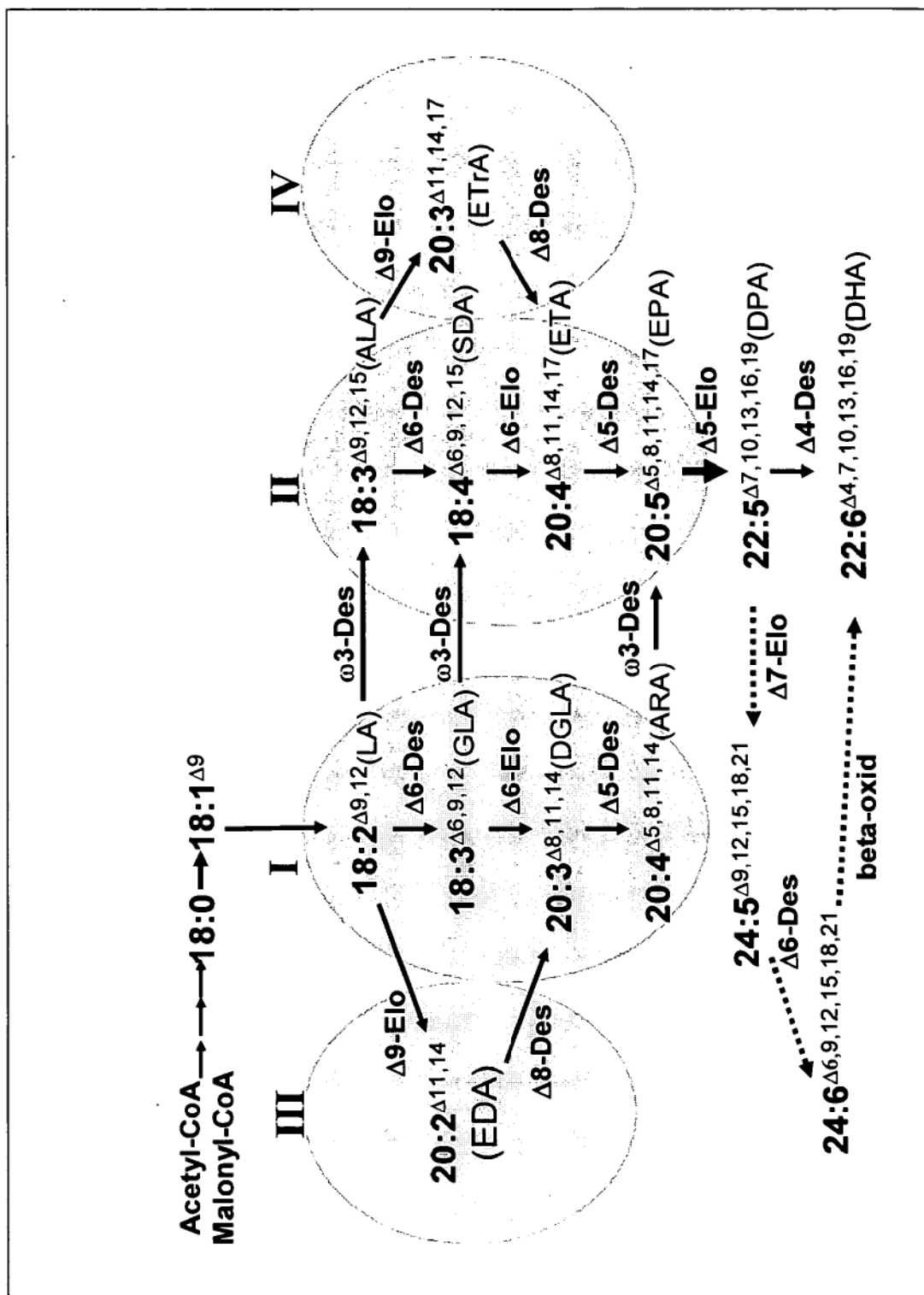


Figure 1

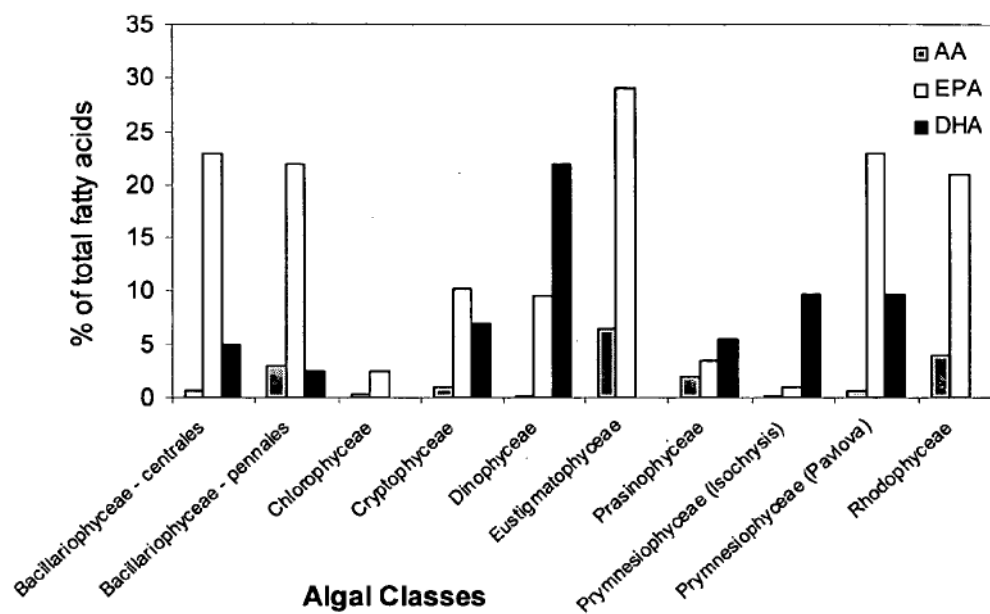


Figure 2

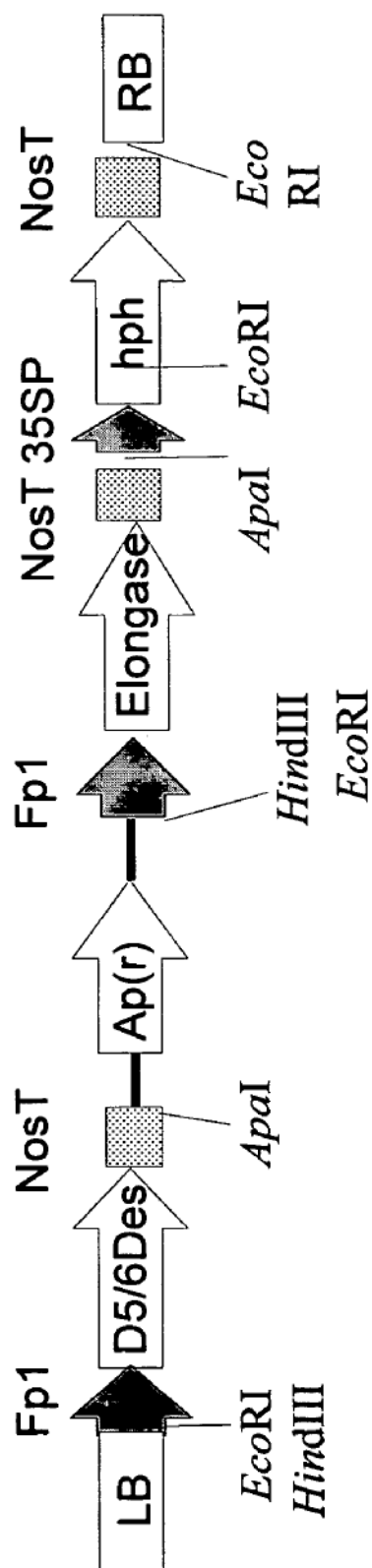


Figure 3

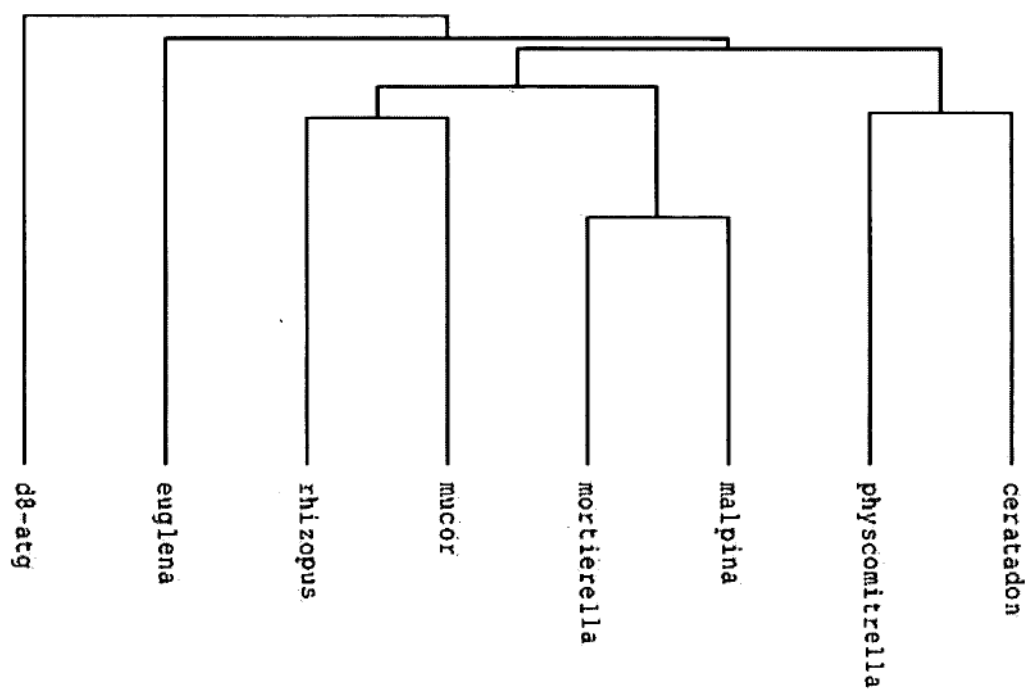


Figure 4

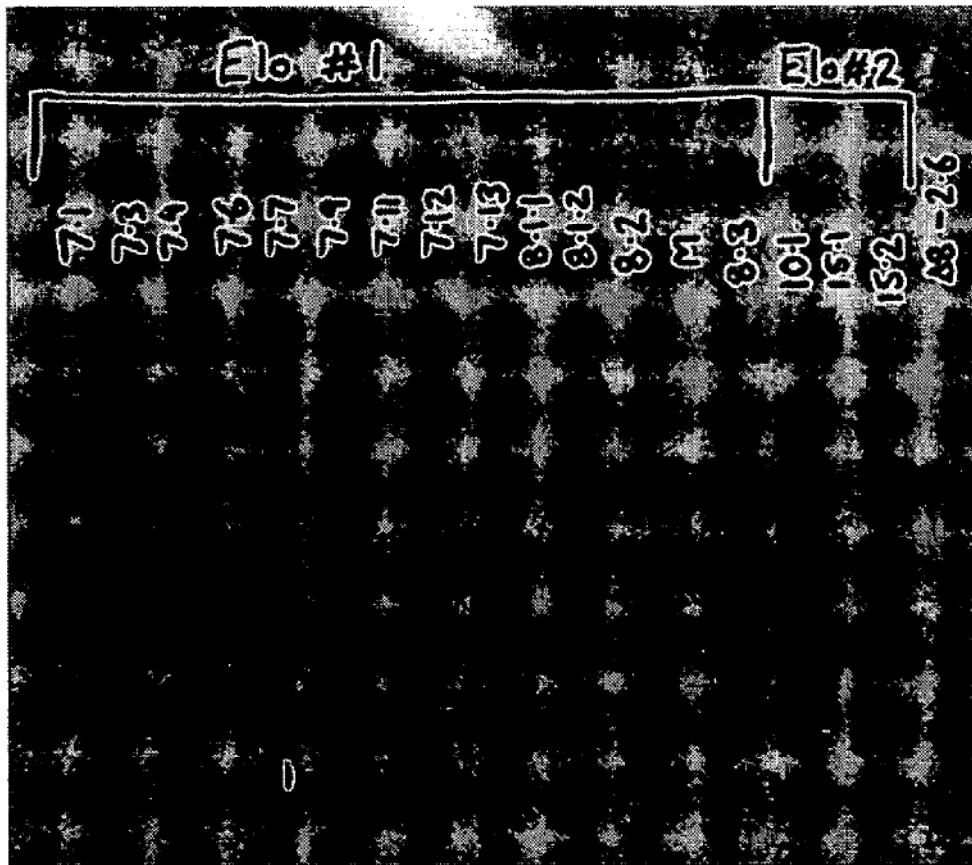


Figure 5

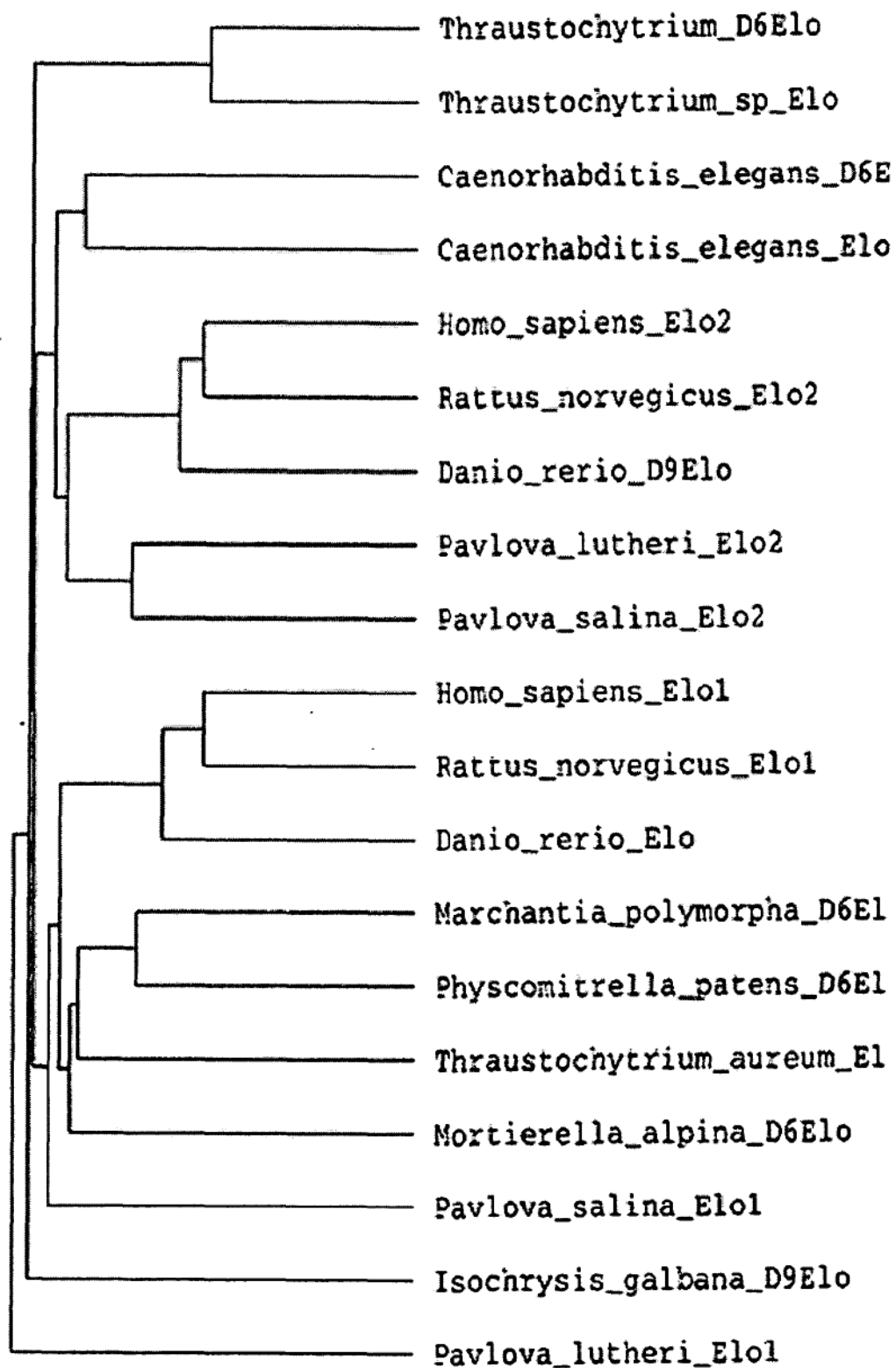


Figure 6

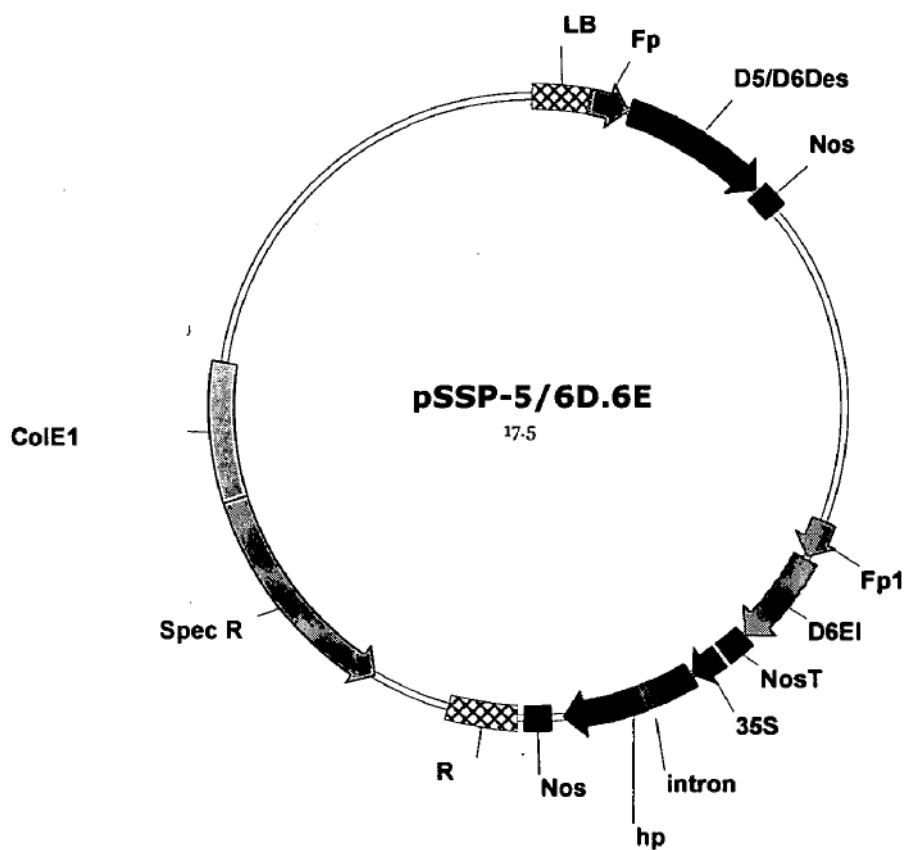


Figure 7A

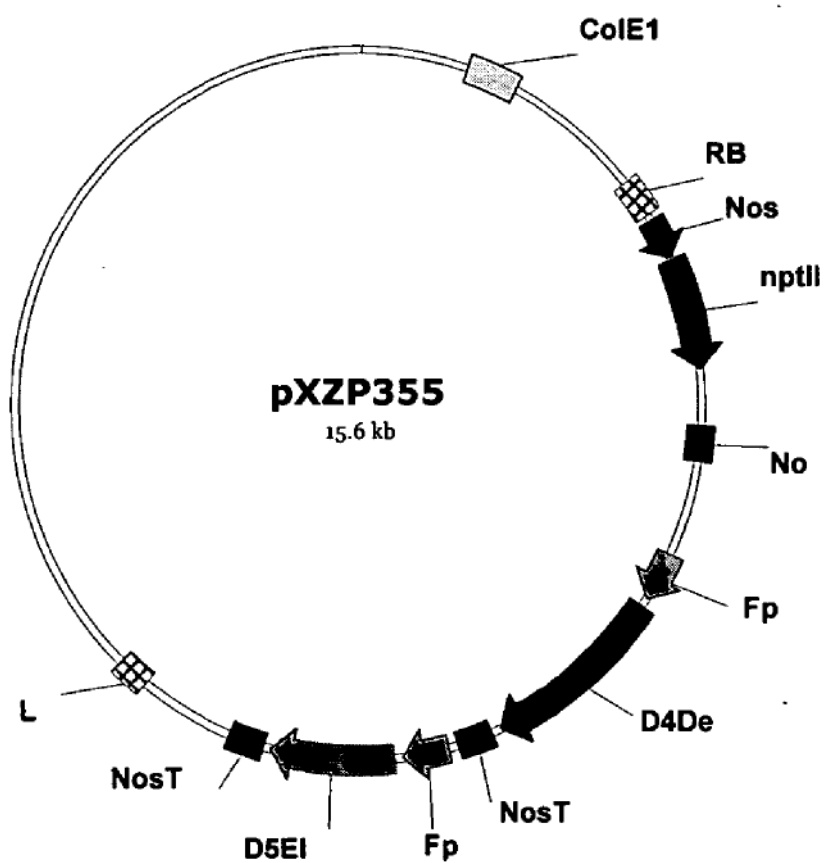


Figure 7B

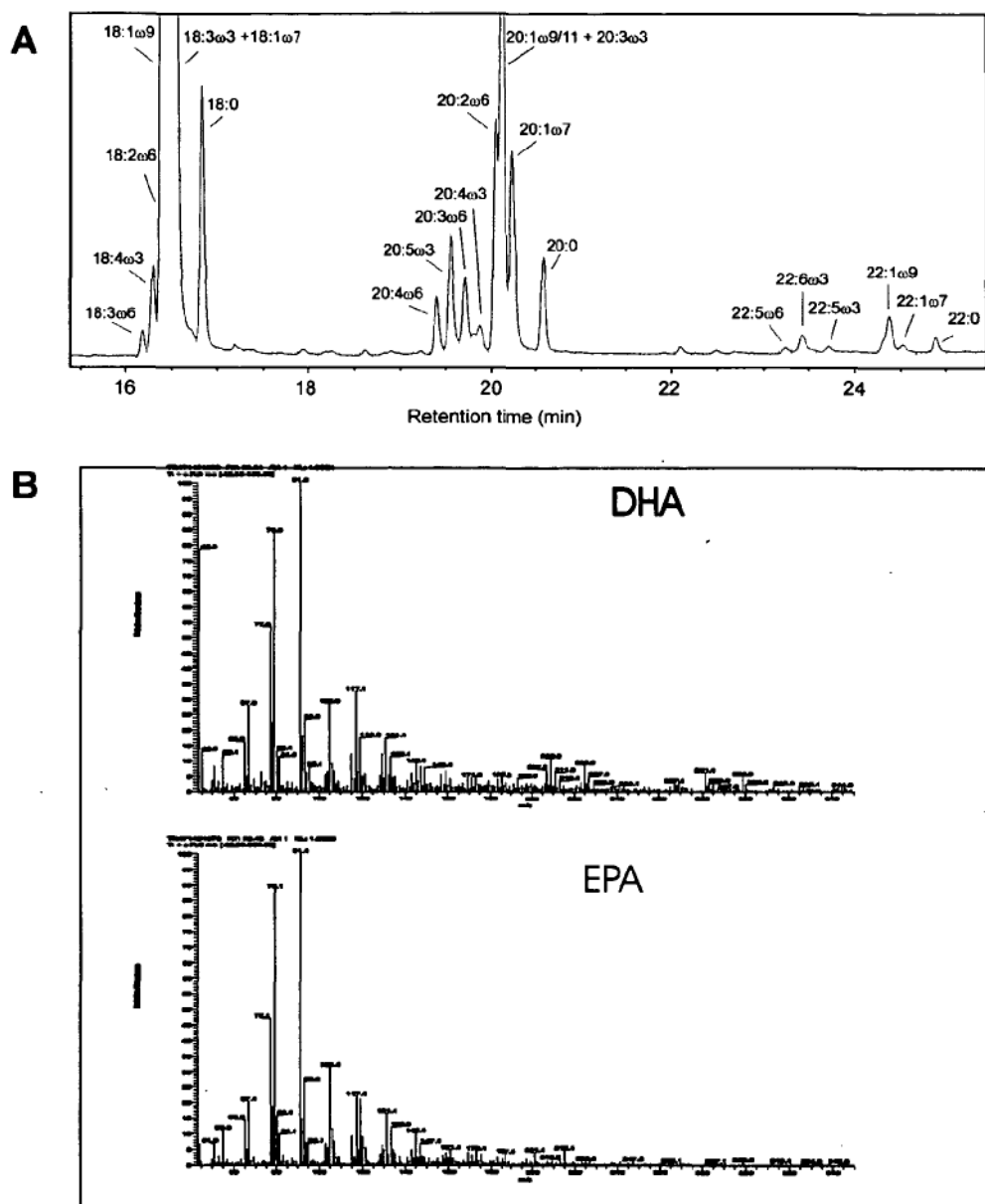


Figure 8

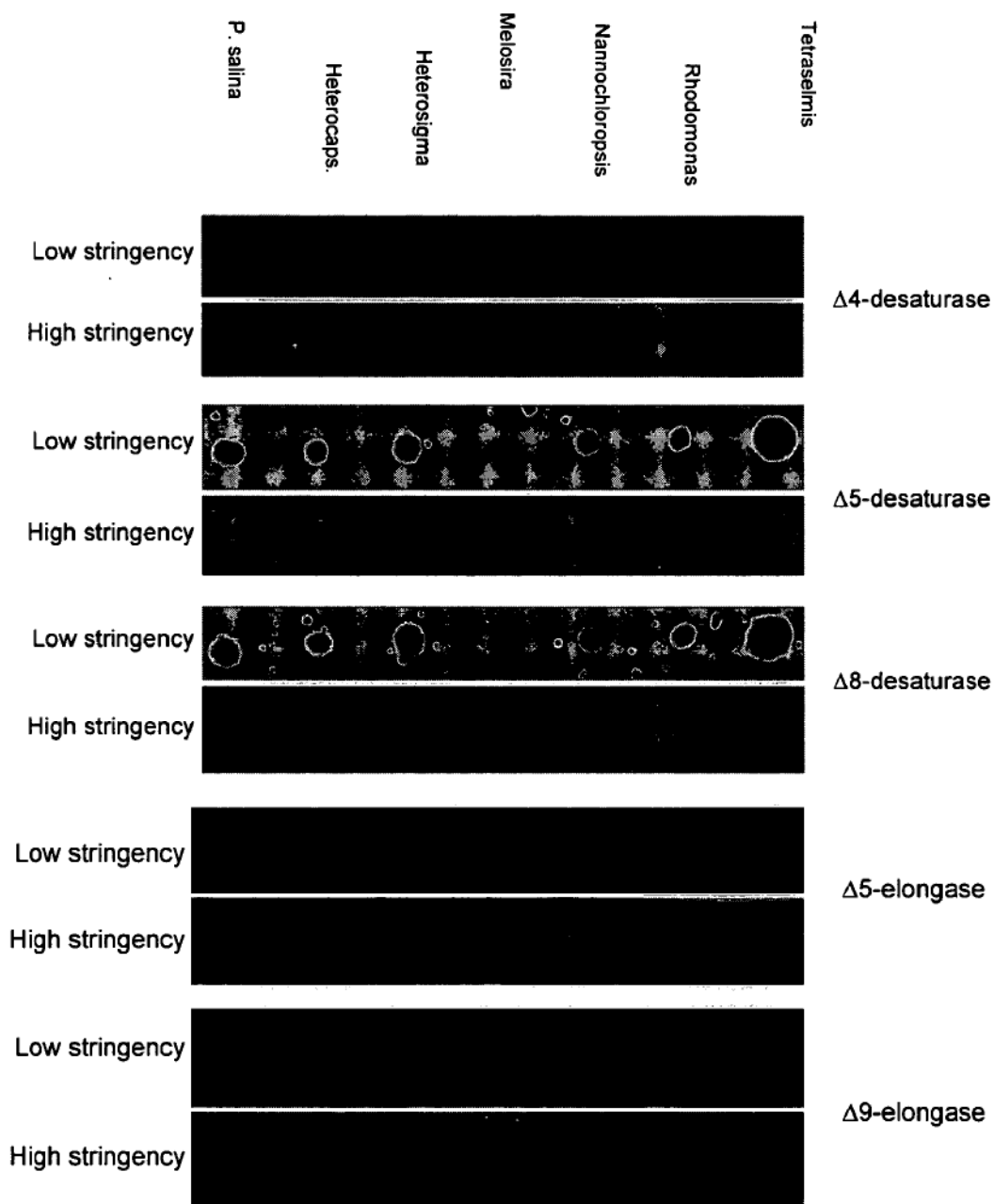


Figure 9

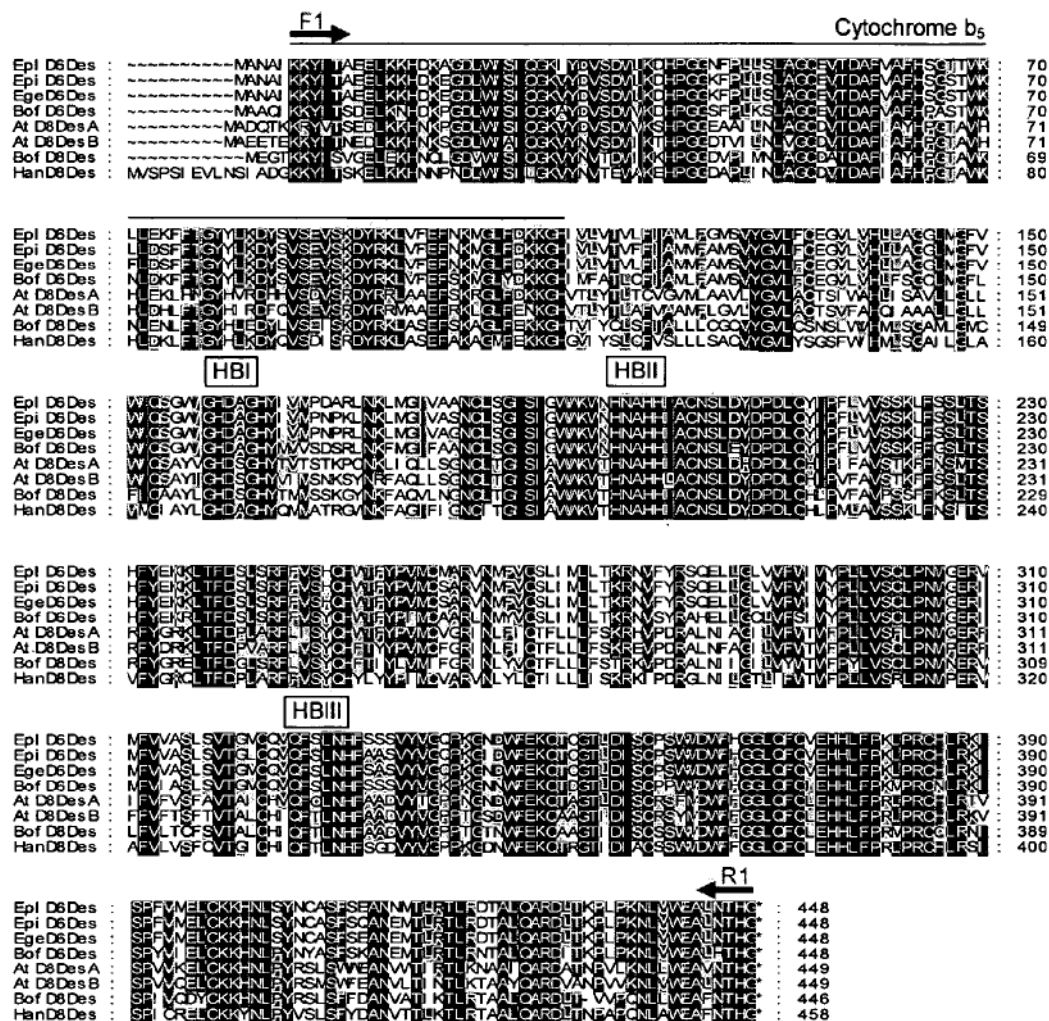
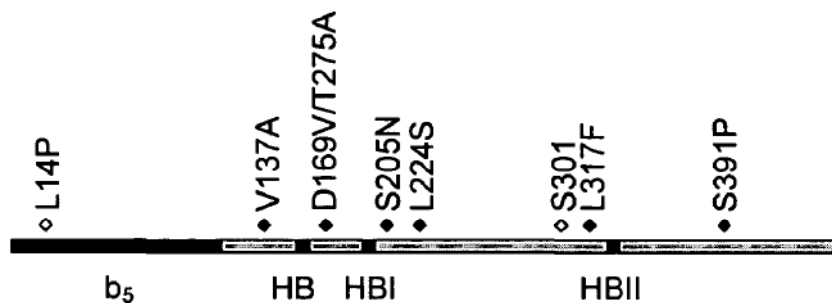


Figure 10

A



B

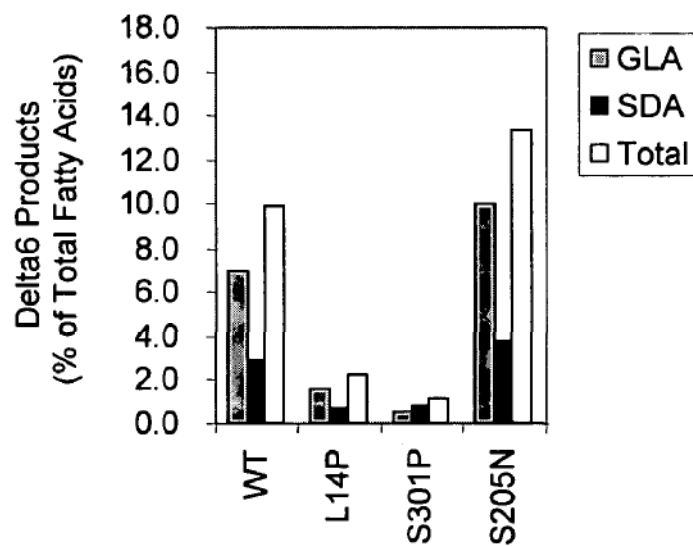


Figure 11

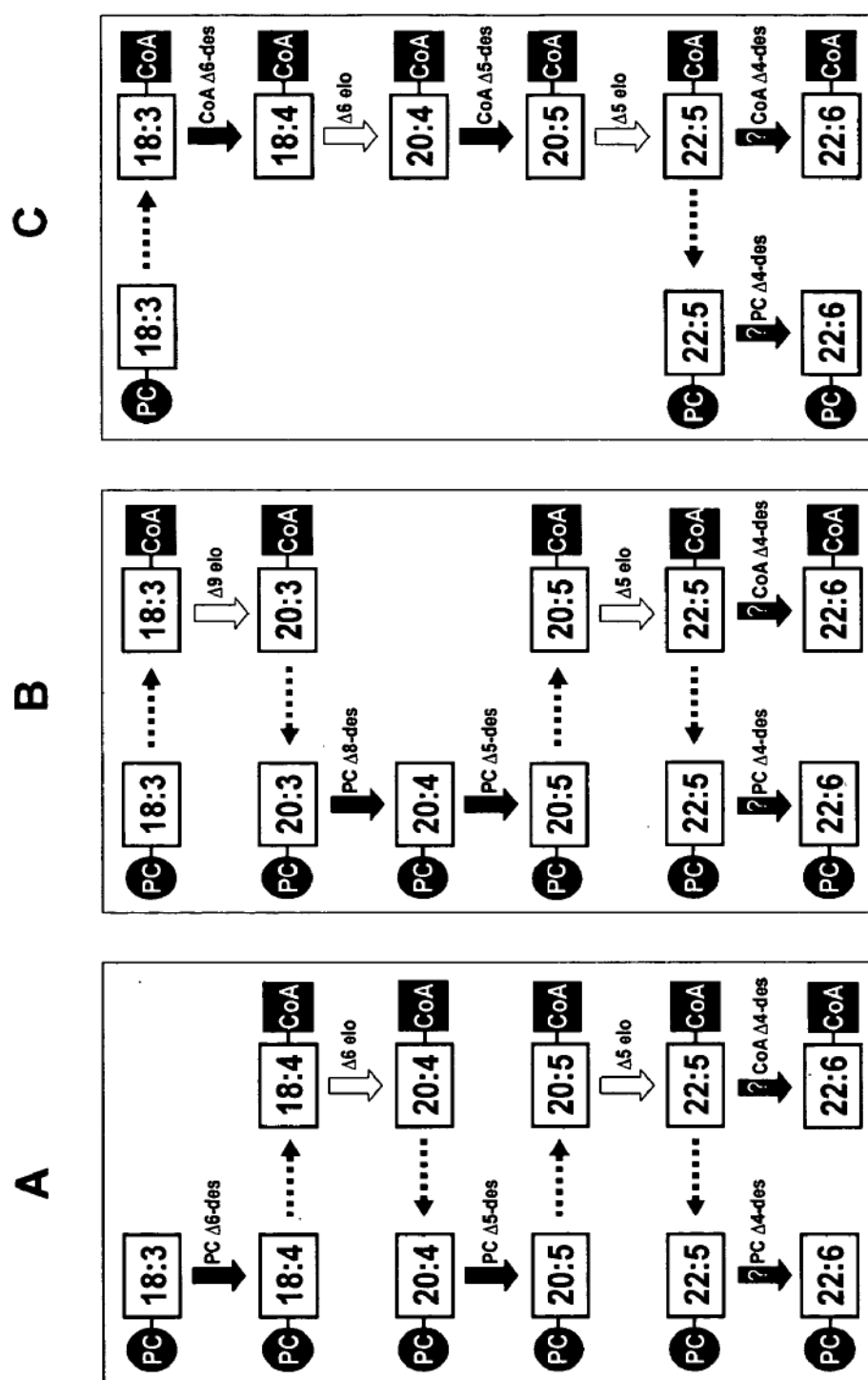


Figure 12

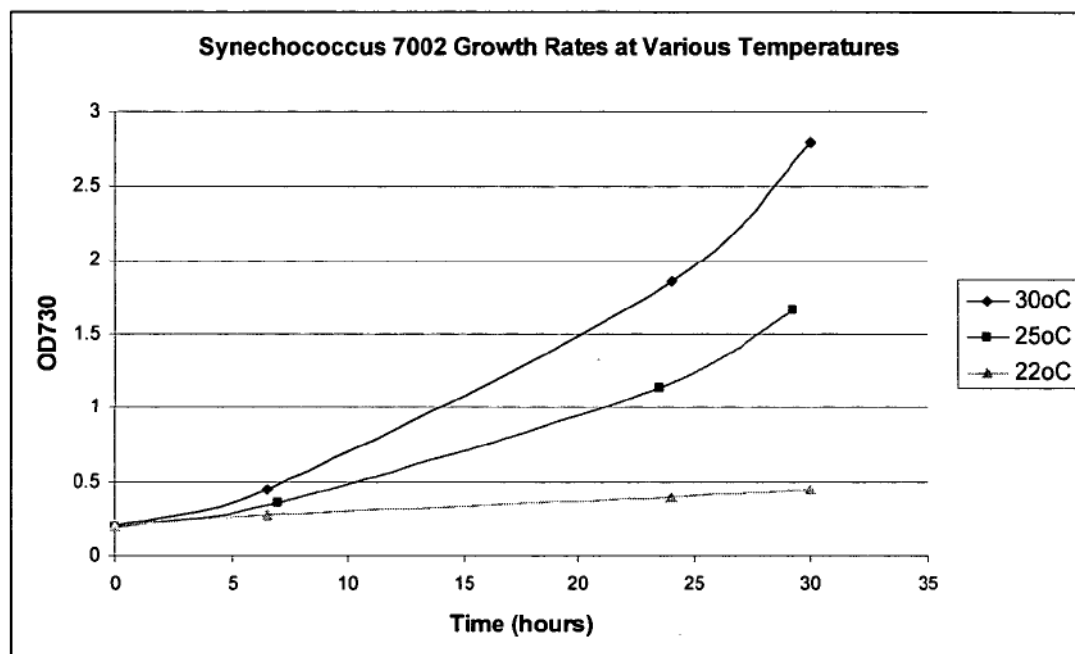


Figure 13

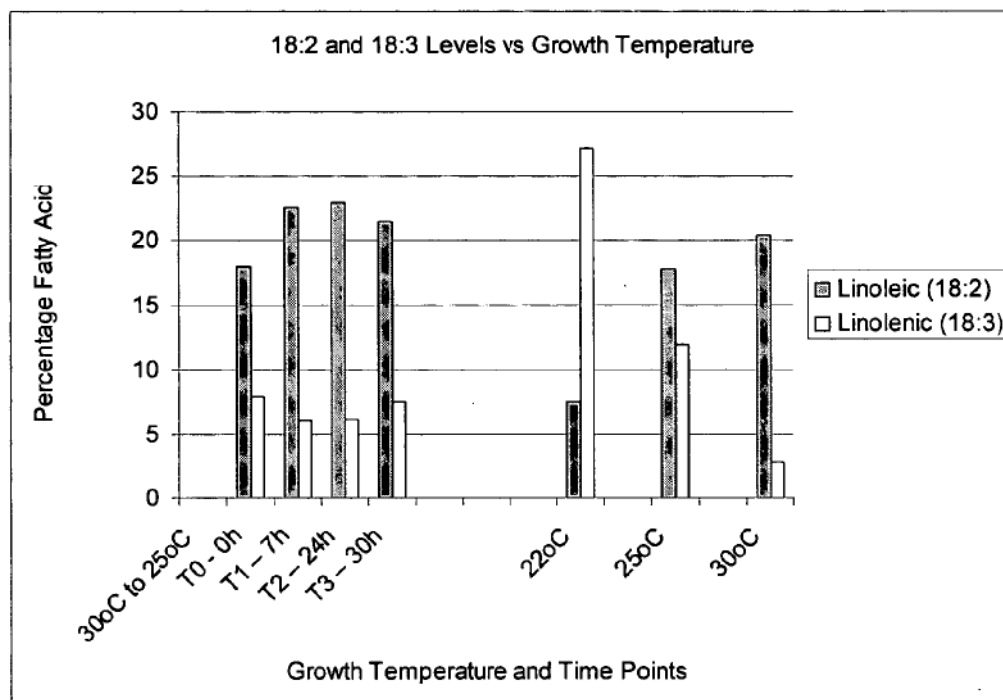


Figure 14

SYNTHESIS OF LONG-CHAIN POLYUNSATURATED FATTY ACIDS BY RECOMBINANT CELLS

FIELD OF THE INVENTION

[0001] The present invention relates to methods of synthesizing long-chain polyunsaturated fatty acids, especially eicosapentaenoic acid, docosapentaenoic acid and docosahexaenoic acid, in recombinant cells such as yeast or plant cells. Also provided are recombinant cells or plants which produce long-chain polyunsaturated fatty acids. Furthermore, the present invention relates to a group of new enzymes which possess desaturase or elongase activity that can be used in methods of synthesizing long-chain polyunsaturated fatty acids.

BACKGROUND OF THE INVENTION

[0002] Omega-3 long-chain polyunsaturated fatty acid(s) (LC-PUFA) are now widely recognized as important compounds for human and animal health. These fatty acids may be obtained from dietary sources or by conversion of linoleic (LA, omega-6) or α -linolenic (ALA, omega-3) fatty acids, both of which are regarded as essential fatty acids in the human diet. While humans and many other vertebrate animals are able to convert LA or ALA, obtained from plant sources, to LC-PUFA, they carry out this conversion at a very low rate. Moreover, most modern societies have imbalanced diets in which at least 90% of polyunsaturated fatty acid(s) (PUFA) consist of omega-6 fatty acids, instead of the 4:1 ratio or less for omega-6:omega-3 fatty acids that is regarded as ideal (Trautwein, 2001). The immediate dietary source of LC-PUFA such as eicosapentaenoic acid (EPA, 20:5) and docosahexaenoic acid (DHA, 22:6) for humans is mostly from fish or fish oil. Health professionals have therefore recommended the regular inclusion of fish containing significant levels of LC-PUFA into the human diet. Increasingly, fish-derived LC-PUFA oils are being incorporated into food products and in infant formula. However, due to a decline in global and national fisheries, alternative sources of these beneficial health-enhancing oils are needed.

[0003] Inclusion of omega-3 LC-PUFA such as EPA and DHA in the human diet has been linked with numerous health-related benefits. These include prevention or reduction of coronary heart disease, hypertension, type-2 diabetes, renal disease, rheumatoid arthritis, ulcerative colitis and chronic obstructive pulmonary disease, and aiding brain development and growth (Simopoulos, 2000). More recently, a number of studies have also indicated that omega-3 PUFA may be beneficial in infant nutrition and development and against various mental disorders such as schizophrenia, attention deficit hyperactive disorder and Alzheimer's disease.

[0004] Higher plants, in contrast to animals, lack the capacity to synthesise polyunsaturated fatty acids with chain lengths longer than 18 carbons. In particular, crop and horticultural plants along with other angiosperms do not have the enzymes needed to synthesize the longer chain omega-3 fatty acids such as EPA, DPA and DHA that are derived from ALA. An important goal in plant biotechnology is therefore the engineering of crop plants, particularly oilseed crops, that produce substantial quantities of LC-PUFA, thus providing an alternative source of these compounds.

[0005] Pathways of LC-PUFA Synthesis

[0006] Biosynthesis of LC-PUFA from linoleic and α -linolenic fatty acids in organisms such as microalgae, mosses and fungi may occur by a series of alternating oxygen-dependent desaturations and elongation reactions as shown schematically in FIG. 1. In one pathway (FIG. 1, II), the desaturation reactions are catalysed by $\Delta 6$, $\Delta 5$, and $\Delta 4$ desaturases, each of which adds an additional double bond into the fatty acid carbon chain, while each of a $\Delta 6$ and a $\Delta 5$ elongase reaction adds a two-carbon unit to lengthen the chain. The conversion of ALA to DHA in these organisms therefore requires three desaturations and two elongations. Genes encoding the enzymes required for the production of DHA in this aerobic pathway have been cloned from various microorganisms and lower plants including microalgae, mosses, fungi. Genes encoding some of the enzymes including one that catalyses the fifth step, the $\Delta 5$ elongase, have been isolated from vertebrate animals including mammals (reviewed in Sayanova and Napier, 2004). However, the $\Delta 5$ elongase isolated from human cells is not specific for the EPA to DPA reaction, having a wide specificity for fatty acid substrates (Leonard et al., 2002).

[0007] Alternative routes have been shown to exist for two sections of the ALA to DHA pathway in some groups of organisms. The conversion of ALA to EPA may be carried out by a combination of a $\Delta 9$ elongase and a $\Delta 8$ desaturase (the so-called $\Delta 8$ desaturation route, see FIG. 1, IV) in certain protists and thraustochytrids, as evidenced by the isolated of genes encoding such enzymes (Wallis and Browse, 1999; Qi et al., 2002). In mammals, the so-called "Sprecher" pathway converts DPA to DHA by three reactions, independent of a $\Delta 4$ desaturase (Sprecher et al., 1995).

[0008] Besides these desaturase/elongase systems, EPA and DHA can also be synthesized through an anaerobic pathway in a number of organisms such as *Shewanella*, *Mortierella* and *Schizochytrium* (Abbadi et al., 2001). The operons encoding these polyketide synthase (PKS) enzyme complexes have been cloned from some bacteria (Morita et al., 2000; Metz et al., 2001; Tanaka et al., 1999; Yazawa, 1996; Yu et al., 2000; WO 00/42195). The EPA PKS operon isolated from *Shewanella* spp has been expressed in *Synechococcus* allowing it to synthesize EPA (Takeyama et al., 1997). The genes encoding these enzymes are arranged in relatively large operons, and their expression in transgenic plants has not been reported. Therefore it remains to be seen if the anaerobic PKS-like system is a possible alternative to the more classic aerobic desaturase/elongase for the transgenic synthesis of LC-PUFA.

[0009] Desaturases

[0010] The desaturase enzymes that have been shown to participate in LC-PUFA biosynthesis all belong to the group of so-called "front-end" desaturases which are characterised by the presence of a cytochrome b_5 domain at the N-terminus of each protein. The cyt b_5 domain presumably acts as a receptor of electrons required for desaturation (Napier et al., 1999; Sperling and Heinz, 2001).

[0011] The enzyme $\Delta 5$ desaturase catalyses the further desaturation of C20 LC-PUFA leading to arachidonic acid (ARA, 20:4 ω 6) and EPA (20:5 ω 3). Genes encoding this enzyme have been isolated from a number of organisms, including algae (*Thraustochytrium* sp. Qiu et al., 2001),

fungi (*M. alpine*, *Pythium irregulare*, Michaelson et al., 1998; Hong et al., 2002), *Caenorhabditis elegans* and mammals. A gene encoding a bifunctional $\Delta 5$ -/ $\Delta 6$ -desaturase has also been identified from zebrafish (Hasting et al., 2001). The gene encoding this enzyme might represent an ancestral form of the "front-end desaturase" which later duplicated and evolved distinct functions. The last desaturation step to produce DHA is catalysed by a $\Delta 4$ desaturase and a gene encoding this enzyme has been isolated from the freshwater protist species *Euglena gracilis* and the marine species *Thraustochytrium* sp. (Qiu et al., 2001; Meyer et al., 2003).

[0012] Elongases

[0013] Several genes encoding PUFA-elongation enzymes have also been isolated (Sayanova and Napier, 2004). The members of this gene family were unrelated to the elongase genes present in higher plants, such as FAE1 of *Arabidopsis*, that are involved in the extension of saturated and monounsaturated fatty acids. An example of the latter is erucic acid (22:1) in *Brassicas*. In some protist species, LC-PUFA are synthesized by elongation of linoleic or α -linolenic acid with a C2 unit, before desaturation with $\Delta 6$ desaturase (FIG. 1 part IV; " $\Delta 8$ -desaturation" pathway). $\Delta 6$ desaturase and $\Delta 6$ elongase activities were not detected in these species. Instead, a $\Delta 9$ -elongase activity would be expected in such organisms, and in support of this, a C18 $\Delta 9$ -elongase gene has recently been isolated from *Isochrysis galbana* (Qi et al., 2002).

[0014] Engineered Production of LC-PUFA

[0015] Transgenic oilseed crops that are engineered to produce major LC-PUFA by the insertion of these genes have been suggested as a sustainable source of nutritionally important fatty acids. However, the requirement for coordinate expression and activity of five new enzymes encoded by genes from possibly diverse sources has made this goal difficult to achieve and the proposal remained speculative until now.

[0016] The LC-PUFA oxygen-dependent biosynthetic pathway to form EPA (FIG. 1) has been successfully constituted in yeast by the co-expression of a $\Delta 6$ -elongase with $\Delta 6$ - and $\Delta 5$ fatty acid desaturases, resulting in small but significant accumulation of ARA and EPA from exogenously supplied linoleic and α -linolenic acids (Beaudoin et al., 2000; Zank et al., 2000). This demonstrated the ability of the genes belonging to the LC-PUFA synthesis pathway to function in heterologous organisms. However, the efficiency of producing EPA was very low. For example, three genes obtained from *C. elegans*, *Borago officinalis* and *Mortierella alpina* were expressed in yeast (Beaudoin et al., 2000). When the transformed yeast were supplied with 18:2 ω -3 (LA) or 18:3 ω -3 (ALA), there was slight production of 20:4 ω -6 or 20:5 ω -3, at conversion efficiencies of 0.65% and 0.3%, respectively. Other workers similarly obtained very low efficiency production of EPA by using genes expressing two desaturases and one elongase in yeast (Domergue et al., 2003a; Zank et al., 2002). There remains, therefore, a need to improve the efficiency of production of EPA in organisms such as yeast, let alone the production of the C22 PUFA which requires the provision of additional enzymatic steps.

[0017] Some progress has been made in the quest for introducing the aerobic LC-PUFA biosynthetic pathway into higher plants including oilseed crops (reviewed by Sayanova

and Napier, 2004; Drexler et al., 2003; Abbadi et al., 2001). A gene encoding a $\Delta 6$ -fatty acid desaturase isolated from borage (*Borago officinalis*) was expressed in transgenic tobacco and *Arabidopsis*, resulting in the production of GLA (18:3 ω 6) and SDA (18:4 ω 3), the direct precursors for LC-PUFA, in the transgenic plants (Sayanova et al., 1997; 1999). However, this provides only a single, first step.

[0018] Domergue et al. (2003a) used a combination of three genes, encoding $\Delta 6$ - and $\Delta 5$ fatty acid desaturases and a $\Delta 6$ -elongase in both yeast and transgenic linseed. The desaturase genes were obtained from the diatom *Phaeodactylum tricornutum* and the elongase gene from the moss *Physcomitrella patens*. Low elongation yields were obtained for endogenously produced $\Delta 6$ -fatty acids in yeast cells (i.e. combining the first and second enzymatic steps), and the main C20 PUFA product formed was 20:2 $\Delta^{11,14}$ representing an unwanted side reaction. Domergue et al. (2003a) also state, without presenting data, that the combination of the three genes were expressed in transgenic linseed which consequently produced ARA and EPA, but that production was inefficient. They commented that the same problem as had been observed in yeast existed in the seeds of higher plants and that the "bottleneck" needed to be circumvented for production of LC-PUFA in oil seed crops.

[0019] WO 2004/071467 (DuPont) reported the expression of various desaturases and elongases in soybean cells but did not show the synthesis of DHA in regenerated plants or in seeds.

[0020] Abbadi et al. (2004) described attempts to express combinations of desaturases and elongases in transgenic linseed, but achieved only low levels of synthesis of EPA. Abbadi et al. (2004) indicated that their low levels of EPA production were also due to an unknown "bottleneck".

[0021] Qi et al. (2004) achieved synthesis in leaves but did not report results in seeds. This is an important issue as the nature of LC-PUFA synthesis can vary between leaves and seeds. In particular, oilseeds store lipid in seeds mostly as TAG while leaves synthesize the lipid mostly as phosphatidyl lipids. Furthermore, Qi et al. (2004) only produced AA and EPA.

[0022] As a result, there is a need for further methods of producing long-chain polyunsaturated, particularly EPA, DPA and DHA, in recombinant cells.

SUMMARY OF THE INVENTION

[0023] In a first aspect, the present invention provides a recombinant cell which is capable of synthesising a long chain polyunsaturated fatty acid(s) (LC-PUFA), comprising one or more polynucleotides which encode at least two enzymes each of which is a $\Delta 5$ / $\Delta 6$ bifunctional desaturase, $\Delta 5$ desaturase, $\Delta 6$ desaturase, $\Delta 5$ / $\Delta 6$ bifunctional elongase, $\Delta 5$ elongase, $\Delta 6$ elongase, $\Delta 4$ desaturase, $\Delta 9$ elongase, or $\Delta 8$ desaturase, wherein the one or more polynucleotides are operably linked to one or more promoters that are capable of directing expression of said polynucleotides in the cell, wherein said recombinant cell is derived from a cell that is not capable of synthesising said LC-PUFA.

[0024] In a second aspect, the present invention provides a recombinant cell with an enhanced capacity to synthesize a LC-PUFA relative to an isogenic non-recombinant cell, comprising one or more polynucleotides which encode at

least two enzymes each of which is a $\Delta 5/\Delta 6$ bifunctional desaturase, $\Delta 5$ desaturase, $\Delta 6$ desaturase, $\Delta 5/\Delta 6$ bifunctional elongase, $\Delta 5$ elongase, $\Delta 6$ elongase, $\Delta 4$ desaturase, $\Delta 9$ elongase, or $\Delta 8$ desaturase, wherein the one or more polynucleotides are operably linked to one or more promoters that are capable of expressing said polynucleotides in said recombinant cell.

[0025] In one embodiment, at least one of the enzymes is a $\Delta 5$ elongase.

[0026] The present inventors are the first to identify an enzyme which has greater $\Delta 5$ elongase activity than $\Delta 6$ elongase activity. As a result, this enzyme provides an efficient means of producing DPA in a recombinant cell as the $\Delta 5$ elongation of EPA is favoured over the $\Delta 6$ elongation of SDA. Thus, in an embodiment, the $\Delta 5$ elongase is relatively specific, that is, where the $\Delta 5$ elongase also has $\Delta 6$ elongase activity the elongase is more efficient at synthesizing DPA from EPA than it is at synthesizing ETA from SDA.

[0027] In another embodiment, the $\Delta 5$ elongase comprises

[0028] i) an amino acid sequence as provided in SEQ ID NO:2,

[0029] ii) an amino acid sequence which is at least 50%, more preferably at least 80%, even more preferably at least 90%, identical to SEQ ID NO:2, or

[0030] iii) a biologically active fragment of i) or ii).

[0031] In another embodiment, the $\Delta 5$ elongase can be purified from algae.

[0032] In another embodiment, at least one of the enzymes is a $\Delta 9$ elongase.

[0033] The present inventors are the first to identify an enzyme which has both $\Delta 9$ elongase activity and $\Delta 6$ elongase activity. When expressed in a cell with a $\Delta 6$ desaturase and a $\Delta 8$ desaturase this enzyme can use the two available pathways to produce ETA from ALA, DGLA from LA, or both (see FIG. 1), thus increasing the efficiency of ETA and/or DGLA production. Thus, in an embodiment, the $\Delta 9$ elongase also has $\Delta 6$ elongase activity. Preferably, the $\Delta 9$ elongase is more efficient at synthesizing ETrA from ALA than it is at synthesizing ETA from SDA. Furthermore, in another embodiment the $\Delta 9$ elongase is able to elongate SDA to ETA, GLA to DGLA, or both, in a yeast cell.

[0034] In a further embodiment, the $\Delta 9$ elongase comprises

[0035] i) an amino acid sequence as provided in SEQ ID NO:3, SEQ ID NO:85 or SEQ ID NO:86,

[0036] ii) an amino acid sequence which is at least 50%, more preferably at least 80%, even more preferably at least 90%, identical to SEQ ID NO:3, SEQ ID NO:85 or SEQ ID NO:86, or

[0037] iii) a biologically active fragment of i) or ii).

[0038] Preferably, the $\Delta 9$ elongase can be purified from algae or fungi.

[0039] It is well known in the art that the greater the number of transgenes in an organism, the greater the likelihood that at least one fitness parameter of the organism,

such as expression level of at least one of the transgenes, growth rate, oil production, reproductive capacity etc, will be compromised. Accordingly, it is desirable to minimize the number of transgenes in a recombinant cell. To this end, the present inventors have devised numerous strategies for producing LC-PUFA's in a cell which avoid the need for a gene to each step in the relevant pathway.

[0040] Thus, in another embodiment, at least one of the enzymes is a $\Delta 5/\Delta 6$ bifunctional desaturase or a $\Delta 5/\Delta 6$ bifunctional elongase. The $\Delta 5/\Delta 6$ bifunctional desaturase may be naturally produced by a freshwater species of fish.

[0041] In a particular embodiment, the $\Delta 5/\Delta 6$ bifunctional desaturase comprises

[0042] i) an amino acid sequence as provided in SEQ ID NO:15,

[0043] ii) an amino acid sequence which is at least 50%, more preferably at least 80%, even more preferably at least 90%, identical to SEQ ID NO:15, or

[0044] iii) a biologically active fragment of i) or ii).

[0045] Preferably, the $\Delta 5/\Delta 6$ bifunctional desaturase is naturally produced by a freshwater species of fish.

[0046] Preferably, the $\Delta 5/\Delta 6$ bifunctional elongase comprises

[0047] i) an amino acid sequence as provided in SEQ ID NO:2 or SEQ ID NO:14,

[0048] ii) an amino acid sequence which is at least 50%, more preferably at least 80%, even more preferably at least 90%, identical to SEQ ID NO:2 or SEQ ID NO:14, or

[0049] iii) a biologically active fragment of i) or ii).

[0050] In another embodiment, at least one of the enzymes is a $\Delta 5$ desaturase.

[0051] In a further embodiment, at least one of the enzymes is a $\Delta 8$ desaturase.

[0052] In another embodiment, the LC-PUFA is docosahexaenoic acid (DHA).

[0053] Preferably, the introduced polynucleotide(s) encode three or four enzymes each of which is a $\Delta 5/\Delta 6$ bifunctional desaturase, $\Delta 5$ desaturase, $\Delta 6$ desaturase, $\Delta 5/\Delta 6$ bifunctional elongase, $\Delta 5$ elongase, $\Delta 6$ elongase, or $\Delta 4$ desaturase. More preferably, the enzymes are any one of the following combinations;

[0054] i) a $\Delta 5/\Delta 6$ bifunctional desaturase, a $\Delta 5/\Delta 6$ bifunctional elongase, and a $\Delta 4$ desaturase,

[0055] ii) a $\Delta 5/\Delta 6$ bifunctional desaturase, a $\Delta 5$ elongase, a $\Delta 6$ elongase, and a $\Delta 4$ desaturase, or

[0056] iii) a $\Delta 5$ desaturase, a $\Delta 6$ desaturase, a $\Delta 5/\Delta 6$ bifunctional elongase, and a $\Delta 4$ desaturase.

[0057] In another embodiment, the LC-PUFA is DHA and the introduced polynucleotide(s) encode five enzymes wherein the enzymes are any one of the following combinations;

[0058] i) a $\Delta 4$ desaturase, a $\Delta 5$ desaturase, a $\Delta 6$ desaturase, a $\Delta 5$ elongase and a $\Delta 6$ elongase, or

- [0059] ii) a $\Delta 4$ desaturase, a $\Delta 5$ desaturase, a $\Delta 8$ desaturase, a $\Delta 5$ elongase and a $\Delta 9$ elongase.
- [0060] In a further embodiment, the cell is of an organism suitable for fermentation, and the enzymes are at least a $\Delta 5/\Delta 6$ bifunctional desaturase, a $\Delta 5$ elongase, a $\Delta 6$ elongase, and a $\Delta 4$ desaturase.
- [0061] In another embodiment, the LC-PUFA is docosapentaenoic acid (DPA).
- [0062] Preferably, the introduced polynucleotide(s) encode two or three enzymes each of which is a $\Delta 5/\Delta 6$ bifunctional desaturase, $\Delta 5$ desaturase, $\Delta 6$ desaturase, $\Delta 5/\Delta 6$ bifunctional elongase, $\Delta 5$ elongase, or $\Delta 6$ elongase. More preferably, the enzymes are any one of the following combinations;
- [0063] i) a $\Delta 5/\Delta 6$ bifunctional desaturase and a $\Delta 5/\Delta 6$ bifunctional elongase,
- [0064] ii) a $\Delta 5/\Delta 6$ bifunctional desaturase, a $\Delta 5$ elongase, and a $\Delta 6$ elongase, or
- [0065] iii) a $\Delta 5$ desaturase, a $\Delta 6$ desaturase, and a $\Delta 5/\Delta 6$ bifunctional elongase.
- [0066] In a further embodiment, the LC-PUFA is DPA and the introduced polynucleotide(s) encode four enzymes wherein the enzymes are any one of the following combinations;
- [0067] i) a $\Delta 5$ desaturase, a $\Delta 6$ desaturase, a $\Delta 5$ elongase and a $\Delta 6$ elongase, or
- [0068] ii) a $\Delta 5$ desaturase, a $\Delta 8$ desaturase, a $\Delta 5$ elongase and a $\Delta 9$ elongase.
- [0069] In another embodiment, the cell is of an organism suitable for fermentation, and the enzymes are at least a $\Delta 5/\Delta 6$ bifunctional desaturase, a $\Delta 5$ elongase, and a $\Delta 6$ elongase.
- [0070] In a further embodiment, the LC-PUFA is eicosapentaenoic acid (EPA).
- [0071] Preferably, the introduced polynucleotide(s) encode a $\Delta 5/\Delta 6$ bifunctional desaturase and a $\Delta 5/\Delta 6$ bifunctional elongase.
- [0072] In another embodiment, the introduced polynucleotide(s) encode three enzymes wherein the enzymes are any one of the following combinations;
- [0073] i) a $\Delta 5$ desaturase, a $\Delta 6$ desaturase, and a $\Delta 6$ elongase, or
- [0074] ii) a $\Delta 5$ desaturase, a $\Delta 8$ desaturase, and a $\Delta 9$ elongase.
- [0075] Evidence to date suggests that desaturases expressed in at least some recombinant cells, particularly yeast, have relatively low activity. However, the present inventors have identified that this may be a function of the capacity of the desaturase to use acyl-CoA as a substrate in LC-PUFA synthesis. In this regard, it has also been determined that desaturase of vertebrate origin are particularly useful for the production of LC-PUFA in recombinant cells, for example, plant cells, seeds, or yeast. Thus, in another preferred embodiment, the recombinant cell comprises either
- [0076] i) at least one $\Delta 5$ elongase catalyses the conversion of EPA to DPA in the cell,
- [0077] ii) at least one desaturase which is able to act on an acyl-CoA substrate,
- [0078] iii) at least one desaturase from a vertebrate or a variant desaturase thereof, or
- [0079] iv) any combination of i), ii) or iii).
- [0080] In a particular embodiment, the $\Delta 5$ elongase comprises
- [0081] i) an amino acid sequence as provided in SEQ ID NO:2,
- [0082] ii) an amino acid sequence which is at least 50% identical to SEQ ID NO:2, or
- [0083] iii) a biologically active fragment of i) or ii).
- [0084] The desaturase able to act on an acyl-CoA substrate or from a vertebrate may be a $\Delta 5$ desaturase, a $\Delta 6$ desaturase, or both. In a particular embodiment, the desaturase comprises
- [0085] i) an amino acid sequence as provided in SEQ ID NO:16, SEQ ID NO:21 or SEQ ID NO:22,
- [0086] ii) an amino acid sequence which is at least 50% identical to SEQ ID NO:16, SEQ ID NO:21 or SEQ ID NO:22, or
- [0087] iii) a biologically active fragment of i) or ii).
- [0088] Preferably, the at least one desaturase is naturally produced by a vertebrate.
- [0089] Alternatively, when the cell is a yeast cell, the LC-PUFA is DHA, and the enzymes are at least a $\Delta 5/\Delta 6$ bifunctional desaturase, a $\Delta 5$ elongase, a $\Delta 6$ elongase, and a $\Delta 4$ desaturase.
- [0090] In a further alternative, when the cell is a yeast cell, the LC-PUFA is DPA, and the enzymes are at least a $\Delta 5/\Delta 6$ bifunctional desaturase, a $\Delta 5$ elongase, and a $\Delta 6$ elongase.
- [0091] Although the cell may be any cell type, preferably, said cell is capable of producing said LC-PUFA from endogenously produced linoleic acid (LA), α -linolenic acid (ALA), or both. More preferably, the ratio of the endogenously produced ALA to LA is at least 1:1 or at least 2:1.
- [0092] In one embodiment, the cell is a plant cell, a plant cell from an angiosperm, an oilseed plant cell, or a cell in a seed. Preferably, at least one promoter is a seed specific promoter.
- [0093] In another embodiment, the cell is of a unicellular microorganism. Preferably, the unicellular microorganism is suitable for fermentation. Preferably, the microorganism is a yeast.
- [0094] In a further embodiment, the cell is a non-human animal cell or a human cell in vitro.
- [0095] In a further embodiment, the recombinant cell produces a LC-PUFA which is incorporated into triacylglycerols in said cell. More preferably, at least 50% of the LC-PUFA that is produced in said cell is incorporated into triacylglycerols.

[0096] In another embodiment, at least the protein coding region of one, two or more of the polynucleotides is obtained from an algal gene. Preferably, the algal gene is from the genus *Pavlova* such as from the species *Pavlova salina*.

[0097] In another aspect, the present invention provides a recombinant cell that is capable of producing DHA from a fatty acid which is ALA, LA, GLA, ARA, SDA, ETA, EPA, or any combination or mixture of these, wherein said recombinant cell is derived from a cell that is not capable of synthesising DHA.

[0098] In a further aspect, the present invention provides a recombinant cell that is capable of producing DPA from a fatty acid which is ALA, LA, GLA, ARA, SDA, ETA, EPA, or any combination or mixture of these, wherein said recombinant cell is derived from a cell that is not capable of synthesising DPA.

[0099] In yet a further aspect, the present invention provides a recombinant cell that is capable of producing EPA from a fatty acid which is ALA, LA, GLA, SDA, ETA or any combination or mixture of these, wherein said recombinant cell is derived from a cell that is not capable of synthesising EPA.

[0100] In another aspect, the present invention provides a recombinant cell that is capable of producing both ETrA from ALA and ETA from SDA, and which produces EPA from a fatty acid which is ALA, LA, GLA, SDA, ETA, or any combination or mixture of these, wherein said recombinant cell is derived from a cell that is not capable of synthesising ETrA, ETA or both.

[0101] In a further aspect, the present invention provides a recombinant cell of an organism useful in fermentation processes, wherein the cell is capable of producing DPA from LA, ALA, arachidonic acid (ARA), eicosatetraenoic acid (ETA), or any combination or mixture of these, wherein said recombinant cell is derived from a cell that is not capable of synthesising DPA.

[0102] In another aspect, the present invention provides a recombinant plant cell capable of producing DPA from LA, ALA, EPA, or any combination or mixture of these, wherein the plant cell is from an angiosperm.

[0103] In an embodiment, the plant cell is also capable of producing DHA.

[0104] In yet another aspect, the present invention provides a recombinant cell which is capable of synthesising DGLA, comprising a polynucleotide(s) encoding one or both of:

[0105] a) a polypeptide which is an $\Delta 9$ elongase, wherein the $\Delta 9$ elongase is selected from the group consisting of:

[0106] i) a polypeptide comprising an amino acid sequence as provided in SEQ ID NO:3, SEQ ID NO:85 or SEQ ID NO:86,

[0107] ii) a polypeptide comprising an amino acid sequence which is at least 40% identical to SEQ ID NO:3, SEQ ID NO:85 or SEQ ID NO:86, and

[0108] iii) a biologically active fragment of i) or ii), and/or

[0109] b) a polypeptide which is an $\Delta 8$ desaturase, wherein the $\Delta 8$ desaturase is selected from the group consisting of:

[0110] i) a polypeptide comprising an amino acid sequence as provided in SEQ ID NO:1,

[0111] ii) a polypeptide comprising an amino acid sequence which is at least 40% identical to SEQ ID NO:1, and

[0112] iii) a biologically active fragment of i) or ii),

[0113] wherein the polynucleotide(s) is operably linked to one or more promoters that are capable of directing expression of said polynucleotide(s) in the cell, and wherein said recombinant cell is derived from a cell that is not capable of synthesising DGLA.

[0114] In an embodiment, the cell is capable of converting DGLA to ARA.

[0115] In another embodiment, the cell further comprises a polynucleotide which encodes a $\Delta 5$ desaturase, wherein the polynucleotide encoding the $\Delta 5$ desaturase is operably linked to one or more promoters that are capable of directing expression of said polynucleotide in the cell, and wherein the cell is capable of producing ARA.

[0116] In a particular embodiment, the cell lacks $\omega 3$ desaturase activity and is not capable of producing ALA. Such cells may be naturally occurring, or produced by reducing the $\omega 3$ desaturase activity of the cell using techniques well known in the art.

[0117] Preferably, the cell is a plant cell or a cell of an organism suitable for fermentation.

[0118] In a further embodiment, a recombinant cell of the invention also possesses the enzyme required to perform the "Sprecher" pathway of converting EPA to DHA. These enzymes may be native to the cell or produced recombinantly. Such enzymes at least include a $\Delta 7$ elongase, $\Delta 6$ desaturase and enzymes required for the peroxisomal β -oxidation of tetracosahexaenoic acid to produce DHA.

[0119] The present inventors have also identified a group of new desaturases and elongases. As a result, further aspects of the invention relate to these enzymes, as well as homologs/variants/derivatives thereof.

[0120] The polypeptide may be a fusion protein further comprising at least one other polypeptide sequence.

[0121] The at least one other polypeptide may be a polypeptide that enhances the stability of a polypeptide of the present invention, or a polypeptide that assists in the purification of the fusion protein.

[0122] Also provided are isolated polynucleotides which, inter alia, encode polypeptides of the invention.

[0123] In a further aspect, the present invention provides a vector comprising or encoding a polynucleotide according to the invention. Preferably, the polynucleotide is operably linked to a seed specific promoter.

[0124] In another aspect, the present invention provides a recombinant cell comprising an isolated polynucleotide according to the invention.

[0125] In a further aspect, the present invention provides a method of producing a cell capable of synthesising one or more LC-PUFA, the method comprising introducing into the cell one or more polynucleotides which encode at least two enzymes each of which is a $\Delta 5/\Delta 6$ bifunctional desaturase, $\Delta 5$ desaturase, $\Delta 6$ desaturase, $\Delta 5/\Delta 6$ bifunctional elongase, $\Delta 5$ elongase, $\Delta 6$ elongase, $\Delta 4$ desaturase, $\Delta 9$ elongase, or $\Delta 8$ desaturase, wherein the one or more polynucleotides are operably linked to one or more promoters that are capable of directing expression of said polynucleotides in the cell.

[0126] In another aspect, the present invention provides a method of producing a recombinant cell with an enhanced capacity to synthesize one or more LC-PUFA, the method comprising introducing into a first cell one or more polynucleotides which encode at least two enzymes each of which is a $\Delta 5/\Delta 6$ bifunctional desaturase, $\Delta 5$ desaturase, $\Delta 6$ desaturase, $\Delta 5/\Delta 6$ bifunctional elongase, $\Delta 5$ elongase, $\Delta 6$ elongase, $\Delta 4$ desaturase, $\Delta 9$ elongase, or $\Delta 8$ desaturase, wherein the one or more polynucleotides are operably linked to one or more promoters that are capable of directing expression of said polynucleotides in the recombinant cell, and wherein said recombinant cell has an enhanced capacity to synthesize said one or more LC-PUFA relative to said first cell.

[0127] Naturally, it will be appreciated that each of the embodiments described herein in relation to the recombinant cells of the invention will equally apply to methods for the production of said cells.

[0128] In a further aspect, the present invention provides a cell produced by a method of the invention.

[0129] In another aspect, the present invention provides a transgenic plant comprising at least one recombinant cell according to the invention.

[0130] Preferably, the plant is an angiosperm. More preferably, the plant is an oilseed plant.

[0131] In a further embodiment, the transgenic plant, or part thereof including a transgenic seed, does not comprise a transgene which encodes an enzyme which preferentially converts an $\omega 6$ LC-PUFA into an $\omega 3$ LC-PUFA.

[0132] In yet a further embodiment, the transgenic plant, or part thereof including a transgenic seed, comprises a transgene encoding a $\Delta 8$ desaturase and/or a $\Delta 9$ elongase.

[0133] In a further aspect, the present invention provides a method of producing an oilseed, the method comprising

[0134] i) growing a transgenic oilseed plant according to the invention under suitable conditions, and

[0135] ii) harvesting the seed of the plant.

[0136] In a further aspect, the invention provides a part of the transgenic plant of the invention, wherein said part comprises an increased level of LC-PUFA in its fatty acid relative to the corresponding part from an isogenic non-transformed plant.

[0137] Preferably, said plant part is selected from, but not limited to, the group consisting of: a seed, leaf, stem, flower, pollen, roots or specialised storage organ (such as a tuber).

[0138] Previously, it has not been shown that LC-PUFA can be produced in plant seeds, nor that these LC-PUFA can be incorporated into plant oils such as triacylglycerol.

[0139] Thus, in another aspect the present invention provides a transgenic seed comprising a LC-PUFA.

[0140] Preferably, the LC-PUFA is selected from the group consisting of:

[0141] i) EPA,

[0142] ii) DPA,

[0143] iii) DHA,

[0144] iv) EPA and DPA, and

[0145] v) EPA, DHA, and DPA.

[0146] More preferably, the LC-PUFA is selected from the group consisting of:

[0147] i) DPA,

[0148] ii) DHA, or

[0149] iii) DHA and DPA.

[0150] Even more preferably, the LC-PUFA is EPA, DHA, and DPA.

[0151] Preferably, the seed is derived from an isogenic non-transgenic seed which produces LA and/or ALA. More preferably, the isogenic non-transgenic seed comprises a higher concentration of ALA than LA in its fatty acids. Even more preferably, the isogenic non-transgenic seed comprises at least about 13% ALA or at least about 27% ALA or at least about 50% ALA in its fatty acid.

[0152] Preferably, the total fatty acid in the oil of the seed comprises at least 9% C20 fatty acids.

[0153] Preferably, the seed is derived from an oilseed plant. More preferably, the oilseed plant is oilseed rape (*Brassica napus*), maize (*Zea mays*), sunflower (*Helianthus annuus*), soybean (*Glycine max*), sorghum (*Sorghum bicolor*), flax (*Linum usitatissimum*), sugar (*Saccharum officinarum*), beet (*Beta vulgaris*), cotton (*Gossypium hirsutum*), peanut (*Arachis hypogaea*), poppy (*Papaver somniferum*), mustard (*Sinapis alba*), castor bean (*Ricinus communis*), sesame (*Sesamum indicum*), or safflower (*Carthamus tinctorius*).

[0154] It is preferred that the seed has a germination rate which is substantially the same as that of the isogenic non-transgenic seed.

[0155] It is further preferred that the timing of germination of the seed is substantially the same as that of the isogenic non-transgenic seed.

[0156] Preferably, at least 25%, or at least 50%, or at least 75% of the LC-PUFA in the seed form part of triacylglycerols.

[0157] Surprisingly, the present inventors have found that transgenic seeds produced using the methods of the invention have levels of ALA and LA which are substantially the same as those of an isogenic non-transgenic seed. As a result, it is preferred that the transgenic seed has levels of ALA and LA which are substantially the same as those of an isogenic non-transgenic seed. Furthermore, it was surprising to note that the levels of monounsaturated fatty acids were decreased in transgenic seeds produced using the methods of the invention. Accordingly, in a further preferred embodiment

ment, the transgenic seed has decreased levels of monounsaturated fatty acids when compared to an isogenic non-transgenic seed.

[0158] In another aspect, the present invention provides a method of producing a transgenic seed according to the invention, the method comprising

[0159] i) introducing into a progenitor cell of a seed one or more polynucleotides which encode at least two enzymes each of which is a $\Delta 5/\Delta 6$ bifunctional desaturase, $\Delta 5$ desaturase, $\Delta 6$ desaturase, $\Delta 5/\Delta 6$ bifunctional elongase, $\Delta 5$ elongase, $\Delta 6$ elongase, $\Delta 4$ desaturase, $\Delta 9$ elongase, or $\Delta 8$ desaturase, wherein the one or more polynucleotides are operably linked to one or more promoters that are capable of directing expression of said polynucleotides in the cell, thereby producing a recombinant progenitor cell,

[0160] ii) culturing said recombinant progenitor cell to produce a plant which comprises said transgenic seed, and

[0161] iii) recovering the seed from the plant so produced.

[0162] In yet a further aspect, the present invention provides a method of producing a transgenic seed comprising cultivating a transgenic plant which produces the transgenic seed of the invention, and harvesting said transgenic seed from the plant.

[0163] In a further aspect, the invention provides an extract from the transgenic plant of the invention, or a plant part of the invention, or a seed of the invention, wherein said extract comprises an increased level of LC-PUFA in its fatty acid relative to a corresponding extract from an isogenic non-transformed plant.

[0164] Preferably, the extract is substantially purified oil comprising at least 50% triacylglycerols.

[0165] In a further aspect, the present invention provides a non-human transgenic animal comprising at least one recombinant cell according to the invention.

[0166] Also provided is a method of producing a LC-PUFA, the method comprising culturing, under suitable conditions, a recombinant cell according to the invention.

[0167] In one embodiment, the cell is of an organism suitable for fermentation and the method further comprises exposing the cell to at least one LC-PUFA precursor. Preferably, the LC-PUFA precursor is at least one of linoleic acid or α -linolenic acid. In a particular embodiment, the LC-PUFA precursor is provided in a vegetable oil.

[0168] In another embodiment, the cell is an algal cell and the method further comprises growing the algal cell under suitable conditions for production of said LC-PUFA.

[0169] In a further aspect, the present invention provides a method of producing one or more LC-PUFA, the method comprising cultivating, under suitable conditions, a transgenic plant of the invention.

[0170] In another aspect, the present invention provides a method of producing oil comprising at least one LC-PUFA, comprising obtaining the transgenic plant of the invention, or the plant part of the invention, or the seed of the invention, and extracting oil from said plant, plant part or seed.

[0171] Preferably, said oil is extracted from the seed by crushing said seed.

[0172] In another aspect, the present invention provides a method of producing DPA from EPA, the method comprising exposing EPA to a polypeptide of the invention and a fatty acid precursor, under suitable conditions.

[0173] In an embodiment, the method occurs in a cell which uses the polyketide-like system to produce EPA.

[0174] In yet another aspect, the present invention provides a fermentation process comprising the steps of:

[0175] i) providing a vessel containing a liquid composition comprising a cell of the invention and constituents required for fermentation and fatty acid biosynthesis; and

[0176] ii) providing conditions conducive to the fermentation of the liquid composition contained in said vessel.

[0177] Preferably, a constituent required for fermentation and fatty acid biosynthesis is LA.

[0178] Preferably, the cell is a yeast cell.

[0179] In another aspect, the present invention provides a composition comprising a cell of the invention, or an extract or portion thereof comprising LC-PUFA, and a suitable carrier.

[0180] In another aspect, the present invention provides a composition comprising the transgenic plant of the invention, or the plant part of the invention, or the seed of the invention, or an extract or portion thereof comprising LC-PUFA, and a suitable carrier.

[0181] In yet another aspect, the present invention provides a feedstuff comprising a cell of the invention, a plant of the invention, the plant part of the invention, the seed of the invention, an extract of the invention, the product of the method of the invention, the product of the fermentation process of the invention, or a composition of the invention.

[0182] Preferably, the feedstuff at least comprises DPA, wherein at least one enzymatic reaction in the production of DPA was performed by a recombinant enzyme in a cell.

[0183] Furthermore, it is preferred that the feedstuff comprises at least comprises DHA, wherein at least one enzymatic reaction in the production of DHA was performed by a recombinant enzyme in a cell.

[0184] In a further aspect, the present invention provides a method of preparing a feedstuff, the method comprising admixing a cell of the invention, a plant of the invention, the plant part of the invention, the seed of the invention, an extract of the invention, the product of the method of the invention, the product of the fermentation process of the invention, or a composition of the invention, with a suitable carrier. Preferably, the feedstuff is for consumption by a mammal or a fish.

[0185] In a further aspect, the present invention provides a method of increasing the levels of a LC-PUFA in an organism, the method comprising administering to the organism a cell of the invention, a plant of the invention, the plant part of the invention, the seed of the invention, an extract of the invention, the product of the method of the

invention, the product of the fermentation process of the invention, or a composition of the invention, or a feedstuff of the invention.

[0186] Preferably, the administration route is oral.

[0187] Preferably, the organism is a vertebrate. More preferably, the vertebrate is a human, fish, companion animal or livestock animal.

[0188] In a further aspect, the present invention provides a method of treating or preventing a condition which would benefit from a LC-PUFA, the method comprising administering to a subject a cell of the invention, a plant of the invention, the plant part of the invention, the seed of the invention, an extract of the invention, the product of the method of the invention, the product of the fermentation process of the invention, or a composition of the invention, or a feedstuff of the invention.

[0189] Preferably, the condition is arrhythmia's, angioplasty, inflammation, asthma, psoriasis, osteoporosis, kidney stones, AIDS, multiple sclerosis, rheumatoid arthritis, Crohn's disease, schizophrenia, cancer, foetal alcohol syndrome, attention deficient hyperactivity disorder, cystic fibrosis, phenylketonuria, unipolar depression, aggressive hostility, adrenoleukodystrophy, coronary heart disease, hypertension, diabetes, obesity, Alzheimer's disease, chronic obstructive pulmonary disease, ulcerative colitis, restenosis after angioplasty, eczema, high blood pressure, platelet aggregation, gastrointestinal bleeding, endometriosis, premenstrual syndrome, myalgic encephalomyelitis, chronic fatigue after viral infections or ocular disease.

[0190] Whilst providing the subject with any amount of LC-PUFA will be beneficial to the subject, it is preferred that an effective amount to treat the condition is administered.

[0191] In another aspect, the present invention provides for the use of a cell of the invention, a plant of the invention, the plant part of the invention, the seed of the invention, an extract of the invention, the product of the method of the invention, the product of the fermentation process of the invention, or a composition of the invention, or a feedstuff of the invention, for the manufacture of a medicament for treating or preventing a condition which would benefit from a LC-PUFA.

[0192] The *Caenorhabditis elegans* $\Delta 6$ elongase has previously been expressed in yeast and been shown to convert octadecatetraenoic acid to eicosatetraenoic acid. However, the present inventors have surprisingly found that this enzyme also possesses $\Delta 5$ elongase activity, being able to convert eicosapentaenoic acid to docosapentaenoic acid.

[0193] In a further aspect, the present invention provides a method of producing an unbranched LC-PUFA comprising 22 carbon atoms, the method comprising incubating an unbranched 20 carbon atom LC-PUFA with a polypeptide selected from the group consisting of:

[0194] i) a polypeptide comprising an amino acid sequence as provided in SEQ ID NO:2 or SEQ ID NO:14,

[0195] ii) a polypeptide comprising an amino acid sequence which is at least 50% identical to SEQ ID NO:2 or SEQ ID NO:14, and

[0196] iii) a biologically active fragment of i) or ii),

[0197] wherein the polypeptide also has $\Delta 6$ elongase activity.

[0198] Preferably, the unbranched LC-PUFA comprising 22 carbon atoms is DPA, and the unbranched 20 carbon atom LC-PUFA is EPA.

[0199] Preferably, the method is performed within a recombinant cell which produces the polypeptide and EPA.

[0200] In yet a further aspect, the present invention provides a substantially purified antibody, or fragment thereof, that specifically binds a polypeptide of the invention.

[0201] In another aspect, the present invention provides a method of identifying a recombinant cell, tissue or organism capable of synthesising one or more LC-PUFA, the method comprising detecting the presence in said cell, tissue or organism of one or more polynucleotides which encode at least two enzymes each of which is a $\Delta 5/\Delta 6$ bifunctional desaturase, $\Delta 5$ desaturase, $\Delta 6$ desaturase, $\Delta 5/\Delta 6$ bifunctional elongase, $\Delta 5$ elongase, $\Delta 6$ elongase, $\Delta 4$ desaturase, $\Delta 9$ elongase, or $\Delta 8$ desaturase, wherein the one or more polynucleotides are operably linked to one or more promoters that are capable of directing expression of said polynucleotides in the cell, tissue or organism.

[0202] Preferably, the method comprises a nucleic acid amplification step, a nucleic acid hybridisation step, a step of detecting the presence of a transgene in the cell, tissue or organism, or a step of determining the fatty acid content or composition of the cell, tissue or organism.

[0203] Preferably, the organism is an animal, plant, angiosperm plant or microorganism:

[0204] In another aspect, the present invention provides a method of producing DPA from EPA, the method comprising exposing EPA to a $\Delta 5$ elongase of the invention and a fatty acid precursor, under suitable conditions.

[0205] Preferably, the method occurs in a cell which uses the polyketide-like system to produce EPA.

[0206] Naturally, recombinant (transgenic) cells, plants, non-human animals comprising a new polynucleotide provided herein may also produce other elongase and/or desaturases such as those defined herein.

[0207] As will be apparent, preferred features and characteristics of one aspect of the invention are applicable to many other aspects of the invention.

[0208] Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

[0209] The invention is hereinafter described by way of the following non-limiting Examples and with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE ACCOMPANYING DRAWINGS

[0210] FIG. 1. Possible pathways of $\omega 3$ and $\omega 6$ LC-PUFA synthesis. The sectors labelled I, II, III, and IV correspond to the $\omega 6$ ($\Delta 6$), $\omega 3$ ($\Delta 6$), $\omega 6$ ($\Delta 8$), and $\omega 3$ (Δ) pathways, respectively. Compounds in sectors I and III are $\omega 6$ com-

pounds, while those in sectors II and IV are ω 3 compounds. "Des" refers to desaturase steps in the pathway catalysed by desaturases as indicated, while "Elo" refers to elongase steps catalysed by elongases as indicated. The thickened arrow indicates the Δ 5 elongase step. The dashed arrows indicate the steps in the "Sprecher" pathway that operates in mammalian cells for the production of DHA from DPA.

[0211] **FIG. 2.** Distribution of LC-PUFA in microalgal classes. Chlorophyceae and Prasinophyceae are described as "green algae", Eustigmatophyceae as "yellow-green algae", Rhodophyceae as "red algae", and Bacillariophyceae and Prymnesiophyceae as diatoms and golden brown algae.

[0212] **FIG. 3.** Genetic construct for expression of LC-PUFA biosynthesis genes in plant cells.

[0213] **FIG. 4.** PILEUP of desaturase enzymes. d8-atg—*Pavlova salina* Δ 8 desaturase; euglena—AAD45877 (Δ 8 desaturase, *Euglena gracilis*); rhizopus—AAP83964 (Δ 6 desaturase, *Rhizopus* sp. NK030037); mucor—BAB69055 (Δ 6 desaturase, *Mucor circinelloides*); mortierella—AAL73948 (Δ 6 desaturase, *Mortierella isabellina*); malpina—BAA85588 (Δ 6 desaturase, *Mortierella alpina*); physcomitrella—CAA11032 (Δ 6 acyl-lipid desaturase, *Physcomitrella patens*); ceratodon—CAB94992 (Δ 6 fatty acid acetylase, *Ceratodon purpureus*).

[0214] **FIG. 5.** Southern blot of PCR products, hybridized to Elo1 or Elo2 probes.

[0215] **FIG. 6.** PILEUP of elongase enzymes.

[0216] **FIG. 7.** Transgene constructs used to express genes encoding LC-PUFA biosynthetic enzymes in *Arabidopsis*. The "EPA construct" pSSP-5/6D.6E (also called pZebdesat-CeloPWvec8 in Example 5) (**FIG. 7A**) contained the zebra-fish dual function Δ 5/ Δ 6-desaturase (D5/D6Des) and the nematode Δ 6-elongase (D6Elo) both driven by the truncated napin promoter (Fp1), and the hygromycin resistance selectable marker gene (hph) driven by the CaMV-35S (35SP) promoter. The "DHA construct" pXZP355 (**FIG. 7B**) comprised the *Pavlova salina* Δ 4-desaturase (D4Des) and Δ 5-elongase (D5Elo) genes both driven by the truncated napin promoter (Fp1), and the kanamycin resistance selectable marker gene (nptII) driven by the nopaline synthase promoter (NosP). All genes were flanked at the 3' end by the nopaline synthase terminator (NosT).

[0217] **FIG. 8.** A. Gas chromatogram (GLC) showing fatty acid profile for *Arabidopsis thaliana* line DO11 carrying EPA and DHA gene constructs. B. Mass spectra for EPA and DHA obtained from *Arabidopsis thaliana* line DO11.

[0218] **FIG. 9.** Autoradiograms of dot-blot hybridisations carried out under low stringency or high stringency conditions, as described in Example 12, to DNA from various microalgal species indicated at the top, using radiolabelled probes consisting of *P. salina* LC-PUFA gene coding regions as indicated on the right.

[0219] **FIG. 10.** Amino acid sequence alignment of Δ 6- and Δ 8-desaturases from higher plants. The amino acid sequences of Δ 6-desaturases from *E. plantagineum* (EplD6Des) (SEQ ID NO:64), *E. gentianoides* (EgeD6Des, accession number AY055117) (SEQ ID NO:65), *E. pitardii* (EpiD6Des, AY055118) (SEQ ID NO:66), *Borago officinalis* (BofD6Des, U79010) (SEQ ID NO:67) and Δ 8-desaturases from *B. officinalis* (BofD8Des, AF133728) (SEQ ID

NO:68), *Helianthus annuus* (HanD8Des, S68358) (SEQ ID NO:69), and *Arabidopsis thaliana* (AtD8DesA, AAC62885.1; and AtD8DesB, CAB71088.1) (SEQ ID NO:70 and SEQ ID NO:71 respectively) were aligned by PILEUP (GCG, Wisconsin, USA). HBI, HBII, HBIII are three conserved histidine boxes. F1 and R1 are the corresponding regions for the degenerate primers EpD6Des-F1 and EpD6Des-R1 used to amplify the cDNA. The N-terminal cytochrome b₅ domain with conserved HPGG motif is also indicated.

[0220] **FIG. 11.** Variant EplD6Des enzymes isolated and representative enzyme activities. EplD6Des with cytochrome b₅, histidine boxes I, II, and III are shown as b₅, HBI, HBII, HBIII respectively. Variants isolated are shown in panel A in the format: wild-type amino acid—position number—variant amino acid. Empty diamonds indicate mutants with significant reduction of enzyme activity, while solid diamonds indicate the variants with no significant effect on enzyme activity. Panel B shows the comparison of GLA and SDA production in transgenic tobacco leaves from two variants with that of wild-type enzyme.

[0221] **FIG. 12.** Alternative pathways for synthesis of the ω 3 LC-PUFA SDA (18:4), EPA (20:5) and DHA (22:6) from ALA (18:3). Desaturases, elongases and acyltransferases are shown as solid, open and dashed arrows respectively. Chain elongation occurs only on acyl-CoA substrates, whereas desaturation can occur on either acyl-PC [A&B] or acyl-CoA substrates [C]. The acyl-PC or acyl-CoA substrate preference of the final Δ 4-desaturase step has not yet been determined. Pathways involving acyl-PC desaturases require acyltransferase-mediated shuttling of acyl groups between the PC and CoA substrates. Panels A and B show the " Δ 6 pathway" and " Δ 8 pathway" variants of the acyl-PC desaturase pathway respectively. Panel C shows the pathway expressed in the current study in which the acyl-CoA Δ 6 and Δ 5 desaturase activities were encoded by the zebra-fish Δ 6/ Δ 5 dual-function desaturase. Synthesis of ω 6 LC-PUFA such as ARA (20:4) occurs by the same set of reactions but commencing with LA (18:2) as the initial substrate.

[0222] **FIG. 13.** Growth rates of *Synechococcus* 7002 at 22° C., 25° C., 30° C.

[0223] **FIG. 14.** *Synechococcus* 7002 linoleic and linolenic acid levels at various growth temperatures.

KEY TO THE SEQUENCE LISTING

[0224] SEQ ID NO:1— Δ 8 desaturase from *Pavlova salina*.

[0225] SEQ ID NO:2— Δ 5 elongase from *Pavlova salina*.

[0226] SEQ ID NO:3— Δ 9 elongase from *Pavlova salina*.

[0227] SEQ ID NO:4— Δ 4 desaturase from *Pavlova salina*.

[0228] SEQ ID NO:5—cDNA encoding open reading frame of Δ 8 desaturase from *Pavlova salina*.

[0229] SEQ ID NO:6—Full length cDNA encoding of Δ 8 desaturase from *Pavlova salina*.

[0230] SEQ ID NO:7—cDNA encoding open reading frame of Δ 5 elongase from *Pavlova salina*.

- [0231] SEQ ID NO:8—Full length cDNA encoding of $\Delta 5$ elongase from *Pavlova salina*.
- [0232] SEQ ID NO:9—cDNA encoding open reading frame of $\Delta 9$ elongase from *Pavlova salina*.
- [0233] SEQ ID NO:10—Full length cDNA encoding of $\Delta 9$ elongase from *Pavlova salina*.
- [0234] SEQ ID NO:11—Partial cDNA encoding N-terminal portion of $\Delta 4$ desaturase from *Pavlova salina*.
- [0235] SEQ ID NO:12—cDNA encoding open reading frame of $\Delta 4$ desaturase from *Pavlova salina*.
- [0236] SEQ ID NO:13—Full length cDNA encoding $\Delta 4$ desaturase from *Pavlova salina*.
- [0237] SEQ ID NO:14— $\Delta 5/\Delta 6$ bifunctional elongase from *Caenorhabditis elegans*.
- [0238] SEQ ID NO:15— $\Delta 5/\Delta 6$ bifunctional desaturase from *Danio rerio* (zebrafish).
- [0239] SEQ ID NO:16— $\Delta 5$ desaturase from humans (Genbank Accession No: AAF29378).
- [0240] SEQ ID NO:17— $\Delta 5$ desaturase from *Pythium irregulare* (Genbank Accession No: AAL13311).
- [0241] SEQ ID NO:18— $\Delta 5$ desaturase from *Thraustochytrium* sp. (Genbank Accession No: AAM09687).
- [0242] SEQ ID NO:19— $\Delta 5$ desaturase from *Mortierella alpina* (Genbank Accession No: 074212).
- [0243] SEQ ID NO:20— $\Delta 5$ desaturase from *Caenorhabditis elegans* (Genbank Accession No: T43319).
- [0244] SEQ ID NO:21— $\Delta 6$ desaturase from humans (Genbank Accession No: AAD20018).
- [0245] SEQ ID NO:22— $\Delta 6$ desaturase from mouse (Genbank Accession No: NP_062673).
- [0246] SEQ ID NO:23— $\Delta 6$ desaturase from *Pythium irregulare* (Genbank Accession No: AAL13310).
- [0247] SEQ ID NO:24— $\Delta 6$ desaturase from *Borago officinalis* (Genbank Accession No: AAD01410).
- [0248] SEQ ID NO:25— $\Delta 6$ desaturase from *Anemone leveillei* (Genbank Accession No: AAQ10731).
- [0249] SEQ ID NO:26— $\Delta 6$ desaturase from *Ceratodon purpureus* (Genbank Accession No: CAB94993).
- [0250] SEQ ID NO:27— $\Delta 6$ desaturase from *Physcomitrella patens* (Genbank Accession No: CAA 1033).
- [0251] SEQ ID NO:28— $\Delta 6$ desaturase from *Mortierella alpina* (Genbank Accession No: BAC82361).
- [0252] SEQ ID NO:29— $\Delta 6$ desaturase from *Caenorhabditis elegans* (Genbank Accession No: AAC15586).
- [0253] SEQ ID NO:30— $\Delta 5$ elongase from humans (Genbank Accession No: NP_068586).
- [0254] SEQ ID NO:31— $\Delta 6$ elongase from *Physcomitrella patens* (Genbank Accession No: AAL84174).
- [0255] SEQ ID NO:32— $\Delta 6$ elongase from *Mortierella alpina* (Genbank Accession No: AAF70417).
- [0256] SEQ ID NO:33— $\Delta 4$ desaturase from *Thraustochytrium* sp. (Genbank Accession No: AAM09688).
- [0257] SEQ ID NO:34— $\Delta 4$ desaturase from *Euglena gracilis* (Genbank Accession No: AAQ 19605).
- [0258] SEQ ID NO:35— $\Delta 9$ elongase from *Isochrysis galbana* (Genbank Accession No: AAL37626).
- [0259] SEQ ID NO:36— $\Delta 8$ desaturase from *Euglena gracilis* (Genbank Accession No: AAD45877).
- [0260] SEQ ID NO:37—cDNA encoding $\Delta 5/\Delta 6$ bifunctional elongase from *Caenorhabditis elegans*.
- [0261] SEQ ID NO:38—cDNA encoding $\Delta 5/\Delta 6$ bifunctional desaturase from *Danio rerio* (zebrafish).
- [0262] SEQ ID NO's:39 to 42, 46, 47, 50, 51, 53, 54, 56, 57, 81, 82, 83, 84 and 87—Oligonucleotide primers.
- [0263] SEQ ID NO's:43 to 45, 48, 49 and 52—Conserved motifs of various desaturases/elongases.
- [0264] SEQ ID NO:55—Partial cDNA encoding *Pavlova salina* FAE-like elongase.
- [0265] SEQ ID NO:58—Full length cDNA encoding $\Delta 5$ desaturase from *Pavlova salina*.
- [0266] SEQ ID NO:59—cDNA encoding open reading frame of $\Delta 5$ desaturase from *Pavlova salina*.
- [0267] SEQ ID NO:60— $\Delta 5$ desaturase from *Pavlova salina*.
- [0268] SEQ ID NO's 61 and 62—Fragments of *Echium pitardii* $\Delta 6$ desaturase.
- [0269] SEQ ID NO:63—cDNA encoding open reading frame of $\Delta 6$ desaturase from *Echium plantagineum*.
- [0270] SEQ ID NO:64— $\Delta 6$ desaturase from *Echium plantagineum*.
- [0271] SEQ ID NO:65— $\Delta 6$ desaturase from *Echium gentianoides* (Genbank Accession No: AY055117).
- [0272] SEQ ID NO:66— $\Delta 6$ desaturase from *Echium pitardii* (Genbank Accession No: AY055118).
- [0273] SEQ ID NO:67— $\Delta 6$ desaturase from *Borago officinalis* (Genbank Accession No: U79010).
- [0274] SEQ ID NO:68— $\Delta 8$ desaturase from *Borago officinalis* (Genbank Accession No: AF133728).
- [0275] SEQ ID NO:69— $\Delta 8$ desaturase from *Helianthus annuus* (Genbank Accession No: S68358).
- [0276] SEQ ID NO:70— $\Delta 8$ desaturaseA from *Arabidopsis thaliana* (Genbank Accession No: AAC62885.1).
- [0277] SEQ ID NO:71— $\Delta 8$ desaturaseB from *Arabidopsis thaliana* (Genbank Accession No: CAB71088.1).
- [0278] SEQ ID NO:72 and 73—Conserved motifs of $\Delta 6$ —and $\Delta 8$ —desaturases.
- [0279] SEQ ID NO:74— $\Delta 6$ elongase from *Thraustochytrium* sp. (Genbank Accession No: AX951565).
- [0280] SEQ ID NO:75— $\Delta 9$ elongase from *Danio rerio* (Genbank Accession No: NM_199532).
- [0281] SEQ ID NO:76— $\Delta 9$ elongase from *Pavlova lutheri*.

- [0282] SEQ ID NO:77— $\Delta 5$ elongase from *Danio rerio* (Genbank Accession No: AF532782).
- [0283] SEQ ID NO:78— $\Delta 5$ elongase from *Pavlova lutheri*.
- [0284] SEQ ID NO:79—Partial gene sequence from *Heterocapsa niei* encoding an elongase.
- [0285] SEQ ID NO:80—Protein encoded by SEQ ID NO:79, presence of stop codon suggests an intron in SEQ ID NO:79.
- [0286] SEQ ID NO:85— $\Delta 9$ elongase from *Pavlova salina*, encoded by alternate start codon at position 31 of SEQ ID NO:9.
- [0287] SEQ ID NO:86— $\Delta 9$ elongase from *Pavlova salina*, encoded by alternate start codon at position 85 of SEQ ID NO:9.
- [0288] SEQ ID NO:88—Partial elongase amino acid sequence from *Melosira* sp.
- [0289] SEQ ID NO:89—cDNA sequence encoding partial elongase from *Melosira* sp.

DETAILED DESCRIPTION OF THE INVENTION

[0290] General Techniques and Definitions

[0291] Unless specifically defined otherwise, all technical and scientific terms used herein shall be taken to have the same meaning as commonly understood by one of ordinary skill in the art (e.g., in cell culture, plant biology, molecular genetics, immunology, immunohistochemistry, protein chemistry, fatty acid synthesis, and biochemistry).

[0292] Unless otherwise indicated, the recombinant nucleic acid, recombinant protein, cell culture, and immunological techniques utilized in the present invention are standard procedures, well known to those skilled in the art. Such techniques are described and explained throughout the literature in sources such as, J. Perbal, *A Practical Guide to Molecular Cloning*, John Wiley and Sons (1984), J. Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory Press (1989), T. A. Brown (editor), *Essential Molecular Biology: A Practical Approach*, Volumes 1 and 2, IRL Press (1991), D. M. Glover and B. D. Hames (editors), *DNA Cloning: A Practical Approach*, Volumes 1-4, IRL Press (1995 and 1996), and F. M. Ausubel et al. (editors), *Current Protocols in Molecular Biology*, Greene Pub. Associates and Wiley-Interscience (1988, including all updates until present), Ed Harlow and David Lane (editors) *Antibodies: A Laboratory Manual*, Cold Spring Harbour Laboratory, (1988), and J. E. Coligan et al. (editors) *Current Protocols in Immunology*, John Wiley & Sons (including all updates until present), and are incorporated herein by reference.

[0293] As used herein, the terms “long-chain polyunsaturated fatty acid”, “LC-PUFA” or “C20+polyunsaturated fatty acid” refer to a fatty acid which comprises at least 20 carbon atoms in its carbon chain and at least three carbon-carbon double bonds. As used herein, the term “very long-chain polyunsaturated fatty acid”, “VLC-PUFA” or “C22+ polyunsaturated fatty acid” refers to a fatty acid which comprises at least 22 carbon atoms in its carbon chain and at least three carbon-carbon double bonds. Ordinarily, the

number of carbon atoms in the carbon chain of the fatty acids refers to an unbranched carbon chain. If the carbon chain is branched, the number of carbon atoms excludes those in sidegroups. In one embodiment, the long-chain polyunsaturated fatty acid is an $\omega 3$ fatty acid, that is, having a desaturation (carbon-carbon double bond) in the third carbon-carbon bond from the methyl end of the fatty acid. In another embodiment, the long-chain polyunsaturated fatty acid is an $\omega 6$ fatty acid, that is, having a desaturation (carbon-carbon double bond) in the sixth carbon-carbon bond from the methyl end of the fatty acid. In a further embodiment, the long-chain polyunsaturated fatty acid is selected from the group consisting of; arachidonic acid (ARA, 20:4 $\Delta 5,8,11,14$; $\omega 6$), eicosatetraenoic acid (ETA, 20:4 $\Delta 8,11,14,17$, $\omega 3$) eicosapentaenoic acid (EPA, 20:5 $\Delta 5,8,11,14,17$; $\omega 3$), docosapentaenoic acid (DPA, 22:5 $\Delta 7,10,13,16,19$, $\omega 3$), or docosahexaenoic acid (DHA, 22:6 $\Delta 4,7,10,13,16,19$, $\omega 3$). The LC-PUFA may also be dihomo- γ -linoleic acid (DGLA) or eicosatrienoic acid (ETra, 20:3 $\Delta 11,14,17$, $\omega 3$). It would readily be apparent that the LC-PUFA that is produced according to the invention may be a mixture of any or all of the above and may include other LC-PUFA or derivatives of any of these LC-PUFA. In a preferred embodiment, the $\omega 3$ fatty acid is EPA, DPA, or DHA, or even more preferably DPA or DHA.

[0294] Furthermore, as used herein the terms “long-chain polyunsaturated fatty acid” or “very long-chain polyunsaturated fatty acid” refer to the fatty acid being in a free state (non-esterified) or in an esterified form such as part of a triglyceride, diacylglyceride, monoacylglyceride, acyl-CoA bound or other bound form. The fatty acid may be esterified as a phospholipid such as a phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, phosphatidylglycerol, phosphatidylinositol or diphosphatidylglycerol forms. Thus, the LC-PUFA may be present as a mixture of forms in the lipid of a cell or a purified oil or lipid extracted from cells, tissues or organisms. In preferred embodiments, the invention provides oil comprising at least 75% or 85% triacylglycerols, with the remainder present as other forms of lipid such as those mentioned, with at least said triacylglycerols comprising the LC-PUFA. The oil may be further purified or treated, for example by hydrolysis with a strong base to release the free fatty acid, or by fractionation, distillation or the like.

[0295] As used herein, the abbreviations “LC-PUFA” and “VLC-PUFA” can refer to a single type of fatty acid, or to multiple types of fatty acids. For example, a transgenic plant of the invention which produces LC-PUFA may produce EPA, DPA and DHA.

[0296] The desaturase and elongase proteins and genes encoding them that may be used in the invention are any of those known in the art or homologues or derivatives thereof. Examples of such genes and the encoded protein sizes are listed in Table 1. The desaturase enzymes that have been shown to participate in LC-PUFA biosynthesis all belong to the group of so-called “front-end” desaturases which are characterised by the presence of a cytochrome b_5 -like domain at the N-terminus of each protein. The

TABLE 1

Cloned genes involved in LC-PUFA biosynthesis.					
Enzyme	Type of organism	Species	Accession Nos.	Protein size (aa's)	References
$\Delta 4$ -desaturase	Algae	<i>Euglena gracilis</i>	AY278558	541	Meyer et al., 2003
		<i>Pavlova lutherii</i>	AY332747	445	Tonon et al., 2003
		<i>Thraustochytrium</i> sp.	AF489589	519	Qiu et al., 2001
		<i>Thraustochytrium aureum</i>	AF391543-5	515	(NCBI)
$\Delta 5$ -desaturase	Mammals	<i>Homo sapiens</i>	AF199596	444	Cho et al., 1999b Leonard et al., 2000b
		<i>Caenorhabditis elegans</i>	AF11440, NM_069350	447	Michaelson et al., 1998b; Watts and Browne, 1999b
	Fungi	<i>Mortierella alpina</i>	AF067654	446	Michaelson et al., 1998a; Knutzon et al., 1998
		<i>Pythium irregulare</i>	AF419297	456	Hong et al., 2002a
		<i>Dictyostelium discoideum</i>	AB022097	467	Saito et al., 2000
		<i>Saprolegnia diclina</i>		470	WO02081668
	Diatom	<i>Phaeodactylum tricornutum</i>	AY082392	469	Domergue et al., 2002
	Algae	<i>Thraustochytrium</i> sp.	AF489588	439	Qiu et al., 2001
		<i>Thraustochytrium aureum</i>		439	WO02081668
		<i>Isochrysis galbana</i>		442	WO02081668
	Moss	<i>Marchantia polymorpha</i>	AY583465	484	Kajikawa et al., 2004
	Mammals	<i>Homo sapiens</i>	NM_013402	444	Cho et al., 1999a; Leonard et al., 2000
		<i>Mus musculus</i>	NM_019699	444	Cho et al., 1999a
	Nematode	<i>Caenorhabditis elegans</i>	Z70271	443	Napier et al., 1998
$\Delta 6$ -desaturase	Plants	<i>Borago officinales</i>	U79010	448	Sayanova et al., 1997
		<i>Echium</i>	AY055117		Garcia-Maroto et al., 2002
		<i>Primula vialii</i>	AY055118 AY234127	453	Sayanova et al., 2003
		<i>Anemone leveillei</i>	AF536525	446	Whitney et al., 2003
	Mosses	<i>Ceratodon purpureus</i>	AJ250735	520	Sperling et al., 2000
		<i>Marchantia polymorpha</i>	AY583463	481	Kajikawa et al., 2004
		<i>Physcomitrella patens</i>			Girke et al., 1998
	Fungi	<i>Mortierella alpina</i>	AF110510 AB020032	457	Huang et al., 1999; Sakuradani et al., 1999
		<i>Pythium irregulare</i>	AF419296	459	Hong et al., 2002a
		<i>Mucor circinelloides</i>	AB052086	467	NCBI*
		<i>Rhizopus</i> sp.	AY320288	458	Zhang et al., 2004
	Diatom	<i>Saprolegnia diclina</i>		453	WO02081668
		<i>Phaeodactylum tricornutum</i>	AY082393	477	Domergue et al., 2002
	Bacteria	<i>Synechocystis</i>	L11421	359	Reddy et al., 1993
	Algae	<i>Thraustochytrium aureum</i>	—	456	WO02081668

TABLE 1-continued

Cloned genes involved in LC-PUFA biosynthesis.					
Enzyme	Type of organism	Species	Accession Nos.	Protein size (aa's)	References
Bifunctional $\Delta 5/\Delta 6$ desaturase	Fish	<i>Danio rerio</i>	AF309556	444	Hastings et al., 2001
	Algae	<i>Euglena gracilis</i>	AF139720	419	Wallis and Browse, 1999
$\Delta 6$ -elongase	Plants	<i>Borago officinales</i>	AF133728		
	Nematode	<i>Caenorhabditis elegans</i>	NM_069288	288	Beaudoin et al., 2000
	Mosses	<i>Physcomitrella patens</i>	AF428243	290	Zank et al., 2002
		<i>Marchantia polymorpha</i>	AY583464	290	Kajikawa et al., 2004
	Fungi	<i>Mortierella alpina</i>	AF206662	318	Parker-Barnes et al., 2000
	Algae	<i>Pavlova lutheri</i> **		501	WO 03078639
		<i>Thraustochytrium</i>	AX951565	271	WO 03093482
		<i>Thraustochytrium</i> sp**	AX214454	271	WO 0159128
PUFA-elongase	Mammals	<i>Homo sapiens</i>	AF231981	299	Leonard et al., 2000b; Leonard et al., 2002
		<i>Rattus norvegicus</i>	AB071985	299	Inagaki et al., 2002
		<i>Rattus norvegicus</i>	AB071986	267	Inagaki et al., 2002
		<i>Mus musculus</i>	AF170907	279	Tvrđik et al., 2000
		<i>Mus musculus</i>	AF170908	292	Tvrđik et al., 2000
	Fish	<i>Danio rerio</i>	AF532782	291 (282)	Agaba et al., 2004
	Worm	<i>Danio rerio</i> **	NM_199532	266	Lo et al., 2003
		<i>Caenorhabditis elegans</i>	Z68749	309	Abbott et al 1998
					Beaudoin et al 2000
	Algae	<i>Thraustochytrium aureum</i> **	AX464802	272	WO 0208401-A2
	$\Delta 9$ -elongase	Algae	<i>Pavlova lutheri</i> **		?
<i>Isochrysis galbana</i>			AF390174	263	Qi et al., 2002

*http://www.ncbi.nlm.nih.gov/

**Function not proven/not demonstrated

[0297] cytochrome b_5 -like domain presumably acts as a receptor of electrons required for desaturation (Napier et al., 1999; Sperling and Heinz, 2001).

[0298] Activity of any of the elongases or desaturases for use in the invention may be tested by expressing a gene encoding the enzyme in a cell such as, for example, a yeast cell or a plant cell, and determining whether the cell has an increased capacity to produce LC-PUFA compared to a comparable cell in which the enzyme is not expressed.

[0299] Unless stated to the contrary, embodiments of the present invention which relate to cells, plants, seeds, etc., and methods for the production thereof, and that refer to at least "two enzymes" (or at least "three enzymes" etc) of the list that is provided means that the polynucleotides encode at least two "different" enzymes from the list provided and not two identical (or very similar with only a few differences as to not substantially alter the activity of the encoded enzyme) open reading frames encoding essentially the same enzyme.

[0300] As used herein, unless stated to the contrary, the term "substantially the same", or variations thereof, means that two samples being analysed, for example two seeds from different sources, are substantially the same if they only vary about $\pm 10\%$ in the trait being investigated.

[0301] As used herein, the term "an enzyme which preferentially converts an $\omega 6$ LC-PUFA into an $\omega 3$ LC-PUFA" means that the enzyme is more efficient at performing said conversion than it is at performing a desaturation reaction outlined in pathways II or III of FIG. 1.

[0302] Whilst certain enzymes are specifically described herein as "bifunctional", the absence of such a term does not necessarily imply that a particular enzyme does not possess an activity other than that specifically defined.

[0303] Desaturases

[0304] As used herein, a " $\Delta 5/\Delta 6$ bifunctional desaturase" or " $\Delta 5/\Delta 6$ desaturase" is at least capable of i) converting α -linolenic acid to octadecatetraenoic acid, and ii) convert-

ing eicosatetraenoic acid to eicosapentaenoic acid. That is, a $\Delta 5/\Delta 6$ bifunctional desaturase is both a $\Delta 5$ desaturase and a $\Delta 6$ desaturase, and $\Delta 5/\Delta 6$ bifunctional desaturases may be considered a sub-class of each of these. A gene encoding a bifunctional $\Delta 5/\Delta 6$ -desaturase has been identified from zebrafish (Hasting et al., 2001). The gene encoding this enzyme might represent an ancestral form of the "front-end desaturase" which later duplicated and the copies evolved distinct $\Delta 5$ - and $\Delta 6$ -desaturase functions. In one embodiment, the $\Delta 5/\Delta 6$ bifunctional desaturase is naturally produced by a freshwater species of fish. In a particular embodiment, the $\Delta 5/\Delta 6$ bifunctional desaturase comprises

[0305] i) an amino acid sequence as provided in SEQ ID NO:15,

[0306] ii) an amino acid sequence which is at least 50% identical to SEQ ID NO:15, or

[0307] iii) a biologically active fragment of i) or ii).

[0308] As used herein, a " $\Delta 5$ desaturase" is at least capable of converting eicosatetraenoic acid to eicosapentaenoic acid. In one embodiment, the enzyme $\Delta 5$ desaturase catalyses the desaturation of C20 LC-PUFA, converting DGLA to arachidonic acid (ARA, 20:4 ω 6) and ETA to EPA (20:5 ω 3). Genes encoding this enzyme have been isolated from a number of organisms, including algae (*Thraustochytrium* sp. Qiu et al., 2001), fungi (*M. alpine*, *Pythium irregulare*, *P. tricornutum*, *Dictyostelium*), *Caenorhabditis elegans* and mammals (Table 1). In another embodiment, the $\Delta 5$ desaturase comprises (i) an amino acid sequence as provided in SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:60, (ii) an amino acid sequence which is at least 50% identical to any one of SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20 or SEQ ID NO:60, or (iii) a biologically active fragment of i) or ii). In a further embodiment, the $\Delta 5$ desaturase comprises (i) an amino acid sequence as provided in SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:20 or SEQ ID NO:60, (ii) an amino acid sequence which is at least 90% identical to any one of SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:20 or SEQ ID NO:60, or (iii) a biologically active fragment of i) or ii). In a further embodiment, the $\Delta 5$ desaturase is encoded by the protein coding region of one of the $\Delta 5$ desaturase genes listed in Table 1 or gene substantially identical thereto.

[0309] As used herein, a " $\Delta 6$ desaturase" is at least capable of converting α -linolenic acid to octadecatetraenoic acid. In one embodiment, the enzyme $\Delta 6$ desaturase catalyses the desaturation of C18 LC-PUFA, converting LA to GLA and ALA to SDA. In another embodiment, the $\Delta 6$ desaturase comprises (i) an amino acid sequence as provided in SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66 or SEQ ID NO:67, (ii) an amino acid sequence which is at least 50% identical to any one of SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66 or SEQ ID NO:67, or (iii) a biologically active fragment of i) or ii). In a further embodiment, the $\Delta 6$ desaturase comprises an amino acid sequence which is at least 90% identical to any one of SEQ ID NO:21, SEQ ID

NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66 or SEQ ID NO:67. In a further embodiment, the $\Delta 6$ desaturase is encoded by the protein coding region of one of the $\Delta 6$ desaturase genes listed in Table 1 or gene substantially identical thereto.

[0310] As used herein, a " $\Delta 4$ desaturase" is at least capable of converting docosapentaenoic acid to docosahexaenoic acid. The desaturation step to produce DHA from DPA is catalysed by a $\Delta 4$ desaturase in organisms other than mammals, and a gene encoding this enzyme has been isolated from the freshwater protist species *Euglena gracilis* and the marine species *Thraustochytrium* sp. (Qiu et al., 2001; Meyer et al., 2003). In one embodiment, the $\Delta 4$ desaturase comprises (i) an amino acid sequence as provided in SEQ ID NO:4, SEQ ID NO:33 or SEQ ID NO:34, (ii) an amino acid sequence which is at least 50% identical to SEQ ID NO:4, SEQ ID NO:33 or SEQ ID NO:34, or (iii) a biologically active fragment of i) or ii). In a further embodiment, the $\Delta 4$ desaturase is encoded by the protein coding region of one of the $\Delta 4$ desaturase genes listed in Table 1 or gene substantially identical thereto.

[0311] As used herein, a " $\Delta 8$ desaturase" is at least capable of converting 20:3 $\Delta^{11,14,17}\omega 3$ to eicosatetraenoic acid. In one embodiment, the $\Delta 8$ desaturase is relatively specific for $\Delta 8$ substrates. That is, it has greater activity in desaturating $\Delta 8$ substrates than other substrates, in particular $\Delta 6$ desaturated substrates. In a preferred embodiment, the $\Delta 8$ desaturase has little or no $\Delta 6$ desaturase activity when expressed in yeast cells. In another embodiment, the $\Delta 8$ desaturase comprises (i) an amino acid sequence as provided in SEQ ID NO:1, SEQ ID NO:36, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70 or SEQ ID NO:71, (ii) an amino acid sequence which is at least 50% identical to SEQ ID NO:1, SEQ ID NO:36, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70 or SEQ ID NO:71, or (iii) a biologically active fragment of i) or ii). In further embodiment, the $\Delta 8$ desaturase comprises (i) an amino acid sequence as provided in SEQ ID NO:1, (ii) an amino acid sequence which is at least 90% identical to SEQ ID NO:1, or (iii) a biologically active fragment of i) or ii).

[0312] As used herein, an " $\omega 3$ desaturase" is at least capable of converting LA to ALA and/or GLA to SDA and/or ARA to EPA. Examples of $\omega 3$ desaturase include those described by Pereira et al. (2004), Horiguchi et al. (1998), Berberich et al. (1998) and Spychalla et al. (1997). In one embodiment, a cell of the invention is a plant cell which lacks $\omega 3$ desaturase activity. Such cells can be produced using gene knockout technology well known in the art. These cells can be used to specifically produce large quantities of $\omega 6$ LC-PUFA such as DGLA.

[0313] Elongases

[0314] Biochemical evidence suggests that the fatty acid elongation consists of 4 steps: condensation, reduction, dehydration and a second reduction. In the context of this invention, an "elongase" refers to the polypeptide that catalyses the condensing step in the presence of the other members of the elongation complex, under suitable physiological conditions. It has been shown that heterologous or homologous expression in a cell of only the condensing component ("elongase") of the elongation protein complex

is required for the elongation of the respective acyl chain. Thus the introduced elongase is able to successfully recruit the reduction and dehydration activities from the transgenic host to carry out successful acyl elongations. The specificity of the elongation reaction with respect to chain length and the degree of desaturation of fatty acid substrates is thought to reside in the condensing component. This component is also thought to be rate limiting in the elongation reaction.

[0315] Two groups of condensing enzymes have been identified so far. The first are involved in the extension of saturated and monounsaturated fatty acids (C18-22) such as, for example, the FAE1 gene of *Arabidopsis*. An example of a product formed is erucic acid (22:1) in *Brassicas*. This group are designated the FAE-like enzymes and do not appear to have a role in LC-PUFA biosynthesis. The other identified class of fatty acid elongases, designated the ELO family of elongases, are named after the ELO genes whose activities are required for the synthesis of the very long-chain fatty acids of sphingolipids in yeast. Apparent paralogs of the ELO-type elongases isolated from LC-PUFA synthesizing organisms like algae, mosses, fungi and nematodes have been shown to be involved in the elongation and synthesis of LC-PUFA. Several genes encoding such PUFA-elongation enzymes have also been isolated (Table 1). Such genes are unrelated in nucleotide or amino acid sequence to the FAE-like elongase genes present in higher plants.

[0316] As used herein, a “ $\Delta 5/\Delta 6$ bifunctional elongase” or “ $\Delta 5/\Delta 6$ elongase” is at least capable of i) converting octadecatetraenoic acid to eicosatetraenoic acid, and ii) converting eicosapentaenoic acid to docosapentaenoic acid. That is, a $\Delta 5/\Delta 6$ bifunctional elongase is both a $\Delta 5$ elongase and a $\Delta 6$ elongase, and $\Delta 5/\Delta 6$ bifunctional elongases may be considered a sub-class of each of these. In one embodiment, the $\Delta 5/\Delta 6$ bifunctional elongase is able to catalyse the elongation of EPA to form DPA in a plant cell such as, for example, a higher plant cell, when that cell is provided with a source of EPA. The EPA may be provided exogenously or preferably endogenously. A gene encoding such an elongase has been isolated from an invertebrate, *C. elegans* (Beaudoin et al., 2000) although it was not previously known to catalyse the $\Delta 5$ -elongation step. In one embodiment, the $\Delta 5/\Delta 6$ bifunctional elongase comprises (i) an amino acid sequence as provided in SEQ ID NO:2 or SEQ ID NO:14, (ii) an amino acid sequence which is at least 50% identical to SEQ ID NO:2 or SEQ ID NO:14, or (iii) a biologically active fragment of i) or ii).

[0317] As used herein, a “ $\Delta 5$ elongase” is at least capable of converting eicosapentaenoic acid to docosapentaenoic acid. In one embodiment, the $\Delta 5$ elongase is from a non-vertebrate source such as, for example, an algal or fungal source. Such elongases can have advantages in terms of the specificity of the elongation reactions carried out (for example the $\Delta 5$ elongase provided as SEQ ID NO:2). In a preferred embodiment, the $\Delta 5$ elongase is relatively specific for C20 substrates over C22 substrates. For example, it may have at least 10-fold lower activity toward C22 substrates (elongated to C24 fatty acids) relative to the activity toward a corresponding C20 substrate when expressed in yeast cells. It is preferred that the activity when using C20 $\Delta 5$ desaturated substrates is high, such as for example, providing an efficiency for the conversion of 20:5 ω 3 into 22:5 ω 3 of at least 7% when expressed in yeast cells. In another embodiment, the $\Delta 5$ elongase is relatively specific for $\Delta 5$

desaturated substrates over $\Delta 6$ desaturated substrates. For example, it may have at least 10-fold lower activity toward $\Delta 6$ desaturated C18 substrates relative to $\Delta 5$ desaturated C20 substrates when expressed in yeast cells. In a further embodiment, the $\Delta 5$ elongase comprises (i) an amino acid sequence as provided in SEQ ID NO:2, SEQ ID NO:30, SEQ ID NO:77 or SEQ ID NO:78, (ii) an amino acid sequence which is at least 50% identical to SEQ ID NO:2, SEQ ID NO:30, SEQ ID NO:77 or SEQ ID NO:78, or (iii) a biologically active fragment of i) or ii). In another embodiment, the $\Delta 5$ elongase comprises (i) an amino acid sequence as provided in SEQ ID NO:2, (ii) an amino acid sequence which is at least 90% identical to SEQ ID NO:2, or (iii) a biologically active fragment of i) or ii). In a further embodiment, the $\Delta 5$ elongase is encoded by the protein coding region of one of the $\Delta 5$ elongase genes listed in Table 1 or gene substantially identical thereto.

[0318] As used herein, a “ $\Delta 6$ elongase” is at least capable of converting octadecatetraenoic acid to eicosatetraenoic acid. In one embodiment, the $\Delta 6$ elongase comprises (i) an amino acid sequence as provided in SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:74, SEQ ID NO:85, SEQ ID NO:86 or SEQ ID NO:88, (ii) an amino acid sequence which is at least 50% identical to SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:74, SEQ ID NO:85, SEQ ID NO:86 or SEQ ID NO:88, or (iii) a biologically active fragment of i) or ii). In another embodiment, the $\Delta 6$ elongase comprises (i) an amino acid sequence as provided in SEQ ID NO:2, SEQ ID NO:3 or SEQ ID NO:32, SEQ ID NO:85, SEQ ID NO:86 or SEQ ID NO:88, (ii) an amino acid sequence which is at least 90% identical to SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:32, SEQ ID NO:85, SEQ ID NO:86 or SEQ ID NO:88, or (iii) a biologically active fragment of i) or ii). In a further embodiment, the $\Delta 6$ elongase is encoded by the protein coding region of one of the $\Delta 6$ elongase genes listed in Table 1 or gene substantially identical thereto.

[0319] In some protist species, LC-PUFA are synthesized by elongation of linoleic or α -linolenic acid with a C2 unit, before desaturation with $\Delta 8$ desaturase (FIG. 1 part IV; “ $\Delta 8$ -desaturation” pathway). $\Delta 6$ desaturase and $\Delta 6$ elongase activities were not detected in these species. Instead, a $\Delta 9$ -elongase activity would be expected in such organisms, and in support of this, a C18 $\Delta 9$ -elongase gene has recently been isolated from *Isochrysis galbana* (Qi et al., 2002). As used herein, a “ $\Delta 9$ elongase” is at least capable of converting α -linolenic acid to 20:3 ω 3. In one embodiment, the $\Delta 9$ elongase comprises (i) an amino acid sequence as provided in SEQ ID NO:3, SEQ ID NO:35, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:85 or SEQ ID NO:86, (ii) an amino acid sequence which is at least 50% identical to SEQ ID NO:3, SEQ ID NO:35, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:85 or SEQ ID NO:86, or (iii) a biologically active fragment of i) or ii). In another embodiment, the $\Delta 9$ elongase comprises (i) an amino acid sequence as provided in SEQ ID NO:3, SEQ ID NO:85 or SEQ ID NO:86, (ii) an amino acid sequence which is at least 90% identical to SEQ ID NO:3, SEQ ID NO:85 or SEQ ID NO:86, or (iii) a biologically active fragment of i) or ii). In a further embodiment, the $\Delta 9$ elongase is encoded by the protein coding region of the $\Delta 9$ elongase gene listed in Table 1 or gene substantially identical thereto. In another embodiment, the $\Delta 9$ elongase also has $\Delta 6$ elongase activity. The elongase in this embodiment is able to convert SDA to ETA and/or GLA

to DGLA ($\Delta 6$ elongase activity) in addition to converting ALA to ETrA ($\Delta 9$ elongase). In a preferred embodiment, such an elongase is from an algal or fungal source such as, for example, the genus *Pavlova*.

[0320] As used herein, a " $\Delta 4$ elongase" is at least capable of converting docosahexaenoic acid to 24:6 ^{$\Delta 6,9,12,15,18,21$} _{$\omega 3$} .

[0321] Cells

[0322] Suitable cells of the invention include any cell that can be transformed with a polynucleotide encoding a polypeptide/enzyme described herein, and which is thereby capable of being used for producing LC-PUFA. Host cells into which the polynucleotide(s) are introduced can be either untransformed cells or cells that are already transformed with at least one nucleic acid molecule. Such nucleic acid molecule may related to LC-PUFA synthesis, or unrelated. Host cells of the present invention either can be endogenously (i.e., naturally) capable of producing proteins of the present invention or can be capable of producing such proteins only after being transformed with at least one nucleic acid molecule.

[0323] As used herein, the term "cell with an enhanced capacity to synthesize a long chain polyunsaturated fatty acid" is a relative term where the recombinant cell of the invention is compared to the native cell, with the recombinant cell producing more long chain polyunsaturated fatty acids, or a greater concentration of LC-PUFA such as EPA, DPA or DHA (relative to other fatty acids), than the native cell.

[0324] The cells may be prokaryotic or eukaryotic. Host cells of the present invention can be any cell capable of producing at least one protein described herein, and include bacterial, fungal (including yeast), parasite, arthropod, animal and plant cells. Preferred host cells are yeast and plant cells. In a preferred embodiment, the plant cells are seed cells.

[0325] In one embodiment, the cell is an animal cell or an algal cell. The animal cell may be of any type of animal such as, for example, a non-human animal cell, a non-human vertebrate cell, a non-human mammalian cell, or cells of aquatic animals such as fish or crustacea, invertebrates, insects, etc.

[0326] An example of a bacterial cell useful as a host cell of the present invention is *Synechococcus* spp. (also known as *Synechocystis* spp.), for example *Synechococcus elongatus*.

[0327] The cells may be of an organism suitable for fermentation. As used herein, the term the "fermentation process" refers to any fermentation process or any process comprising a fermentation step. A fermentation process includes, without limitation, fermentation processes used to produce alcohols (e.g., ethanol, methanol, butanol); organic acids (e.g., citric acid, acetic acid, itaconic acid, lactic acid, gluconic acid); ketones (e.g., acetone); amino acids (e.g., glutamic acid); gases (e.g., H₂ and CO₂); antibiotics (e.g., penicillin and tetracycline); enzymes; vitamins (e.g., riboflavin, beta-carotene); and hormones. Fermentation processes also include fermentation processes used in the consumable alcohol industry (e.g., beer and wine), dairy industry (e.g., fermented dairy products), leather industry and tobacco industry. Preferred fermentation processes

include alcohol fermentation processes, as are well known in the art. Preferred fermentation processes are anaerobic fermentation processes, as are well known in the art.

[0328] Suitable fermenting cells, typically microorganisms are able to ferment, i.e., convert, sugars, such as glucose or maltose, directly or indirectly into the desired fermentation product. Examples of fermenting microorganisms include fungal organisms, such as yeast. As used herein, "yeast" includes *Saccharomyces* spp., *Saccharomyces cerevisiae*, *Saccharomyces carlsbergensis*, *Candida* spp., *Kluveromyces* spp., *Pichia* spp., *Hansenula* spp., *Trichoderma* spp., *Lipomyces starkey*, and *Yarrowia lipolytica*. Preferred yeast include strains of the *Saccharomyces* spp., and in particular, *Saccharomyces cerevisiae*. Commercially available yeast include, e.g., Red Star/Lesaffre Ethanol Red (available from Red Star/Lesaffre, USA) FALI (available from Fleischmann's Yeast, a division of Burns Philp Food Inc., USA), SUPERSTART (available from Ailtech), GERT STRAND (available from Gert Strand AB, Sweden) and FERMIO (available from DSM Specialties).

[0329] Evidence to date suggests that some desaturases expressed heterologously in yeast have relatively low activity in combination with some elongases. However, the present inventors have identified that this may be alleviated by providing a desaturase with the capacity of to use an acyl-CoA form of the fatty acid as a substrate in LC-PUFA synthesis, and this is thought to be advantageous in recombinant cells other than yeast as well. In this regard, it has also been determined that desaturases of vertebrate origin are particularly useful for the production of LC-PUFA. Thus in embodiments of the invention, either (i) at least one of the enzymes is a $\Delta 5$ elongase that catalyses the conversion of EPA to DPA in the cell, (ii) at least one of the desaturases is able to act on an acyl-CoA substrate, (iii) at least one desaturase is from vertebrate or is a variant thereof, or (iv) a combination of ii) and iii).

[0330] In a particularly preferred embodiment, the host cell is a plant cell, such as those described in further detail herein.

[0331] As used herein, a "progenitor cell of a seed" is a cell that divides and/or differentiates into a cell of a transgenic seed of the invention, and/or a cell that divides and/or differentiates into a transgenic plant that produces a transgenic seed of the invention.

[0332] Levels of LC-PUFA Produced

[0333] The levels of the LC-PUFA that are produced in the recombinant cell are of importance. The levels may be expressed as a composition (in percent) of the total fatty acid that is a particular LC-PUFA or group of related LC-PUFA, for example the $\omega 3$ LC-PUFA or the $\omega 6$ LC-PUFA, or the C22+ PUFA, or other which may be determined by methods known in the art. The level may also be expressed as a LC-PUFA content, such as for example the percentage of LC-PUFA in the dry weight of material comprising the recombinant cells, for example the percentage of the dry weight of seed that is LC-PUFA. It will be appreciated that the LC-PUFA that is produced in an oilseed may be considerably higher in terms of LC-PUFA content than in a vegetable or a grain that is not grown for oil production, yet both may have similar LC-PUFA compositions, and both may be used as sources of LC-PUFA for human or animal consumption.

[0334] The levels of LC-PUFA may be determined by any of the methods known in the art. For example, total lipid may be extracted from the cells, tissues or organisms and the fatty acid converted to methyl esters before analysis by gas chromatography (GC). Such techniques are described in Example 1. The peak position in the chromatogram may be used to identify each particular fatty acid, and the area under each peak integrated to determine the amount. As used herein, unless stated to the contrary, the percentage of particular fatty acid in a sample is determined as the area under the peak for that fatty acid as a percentage of the total area for fatty acids in the chromatogram. This corresponds essentially to a weight percentage (w/w). The identity of fatty acids may be confirmed by GC-MS, as described in Example 1.

[0335] In certain embodiments, where the recombinant cell is useful in a fermentation process such as, for example, a yeast cell, the level of EPA that is produced may be at least 0.21% of the total fatty acid in the cell, preferably at least 0.82% or at least 2% and even more preferably at least 5%.

[0336] In other embodiments, the total fatty acid of the recombinant cell may comprise at least 1.5% EPA, preferably at least 2.1% EPA, and more preferably at least 2.5%, at least 3.1%, at least 4% or at least 5.1% EPA.

[0337] In further embodiments, where the recombinant cell is useful in a fermentation process or is a plant cell and DPA is produced, the total fatty acid in the cell may comprise at least 0.1% DPA, preferably at least 0.13% or at least 0.15% and more preferably at least 0.5% or at least 1% DPA.

[0338] In further embodiments, the total fatty acid of the cell may comprise at least 2% C20 LC-PUFA, preferably at least 3% or at least 4% C20 LC-PUFA, more preferably at least 4.7% or at least 7.9% C20 LC-PUFA and most preferably at least 10.2% C20 LC-PUFA.

[0339] In further embodiments, the total fatty acid of the cell may comprise at least 2.5% C20 ω 3 LC-PUFA, preferably at least 4.1% or more preferably at least 5% C20 ω 3 LC-PUFA.

[0340] In other embodiments, where both EPA and DPA are synthesized in a cell, the level of EPA reached is at least 1.5%, at least 2.1% or at least 2.5% and the level of DPA at least 0.13%, at least 0.5% or at least 1.0%.

[0341] In each of these embodiments, the recombinant cell may be a cell of an organism that is suitable for fermentation such as, for example, a unicellular microorganism which may be a prokaryote or a eukaryote such as yeast, or a plant cell. In a preferred embodiment, the cell is a cell of an angiosperm (higher plant). In a further preferred embodiment, the cell is a cell in a seed such as, for example, an oilseed or a grain or cereal.

[0342] The level of production of LC-PUFA in the recombinant cell may also be expressed as a conversion ratio, i.e. the amount of the LC-PUFA formed as a percentage of one or more substrate PUFA or LC-PUFA. With regard to EPA, for example, this may be expressed as the ratio of the level of EPA (as a percentage in the total fatty acid) to the level of a substrate fatty acid (ALA, SDA, ETA or ETrA). In a preferred embodiment, the conversion efficiency is for ALA to EPA. In particular embodiments, the conversion ratio for

production of EPA in a recombinant cell may be at least 0.5%, at least 1%, or at least 2%. In another embodiment, the conversion efficiency for ALA to EPA is at least 14.6%. In further embodiments, the conversion ratio for production of DPA from EPA in a recombinant cell is at least 5%, at least 7%, or at least 10%. In other embodiments, the total ω 3 fatty acids produced that are products of Δ 6 desaturation (i.e. downstream of 18:3 ω 3 (ALA), calculated as the sum of the percentages for 18:4 ω 3 (SDA), 20:4 ω 3 (ETA), 20:5 ω 3 (EPA) and 22:5 ω 3 (DPA)) is at least 4.2%. In a particular embodiment, the conversion efficiency of ALA to ω 3 products through a Δ 6 desaturation step and/or an Δ 9 elongation step in a recombinant cell, preferably a plant cell, more preferably a seed cell, is at least 22% or at least 24%. Stated otherwise, in this embodiment the ratio of products derived from ALA to ALA (products:ALA) in the cell is at least 1:3.6.

[0343] The content of the LC-PUFA in the recombinant cell may be maximized if the parental cell used for introduction of the genes is chosen such that the level of fatty acid substrate that is produced or provided exogenously is optimal. In particular embodiments, the cell produces ALA endogenously at levels of at least 30%, at least 50%, or at least 66% of the total fatty acid. The level of LC-PUFA may also be maximized by growing or incubating the cells under optimal conditions, for example at a slightly lower temperature than the standard temperature for that cell, which is thought to favour accumulation of polyunsaturated fatty acid.

[0344] There are advantages to maximizing production of a desired LC-PUFA while minimizing the extent of side-reactions. In a particular embodiment, there is little or no ETrA detected (less than 0.1%) while the level of EPA is at least 2.1%.

[0345] Turning to transgenic plants of the invention, in one embodiment, at least one plant part synthesizes EPA, wherein the total fatty acid of the plant part comprises at least 1.5%, at least 2.1%, or at least 2.5% EPA.

[0346] In another embodiment, at least one plant part synthesizes DPA, wherein the total fatty acid of the plant part comprises at least 0.1%, at least 0.13%, or at least 0.5% DPA.

[0347] In a further embodiment, at least one plant part synthesizes DHA.

[0348] In another embodiment, at least one plant part synthesizes DHA, wherein the total fatty acid of the plant part comprises at least 0.1%, at least 0.2%, or at least 0.5% DHA.

[0349] In another embodiment, at least one plant part synthesizes at least one ω 3 C20 LC-PUFA, wherein the total fatty acid of the plant part comprises at least 2.5%, or at least 4.1% ω 3 C20 LC-PUFA.

[0350] In yet another embodiment, at least one plant part synthesizes EPA, wherein the efficiency of conversion of ALA to EPA in the plant part is at least 2% or at least 14.6%.

[0351] In a further embodiment, at least one plant part synthesizes ω 3 polyunsaturated fatty acids that are the products of Δ 6-desaturation of ALA and/or the products of Δ 9 elongation of ALA, wherein the efficiency of conversion of ALA to said products in the plant part is at least 22% or at least 24%.

[0352] In yet another embodiment, at least one plant part synthesizes DPA from EPA, wherein the efficiency of conversion of EPA to DPA in the plant part is at least 5% or at least 7%.

[0353] With regard to transgenic seeds of the invention, in one embodiment EPA is synthesized in the seed and the total fatty acid of the seed comprises at least 1.5%, at least 2.1%, or at least 2.5% EPA.

[0354] In another embodiment, DPA is synthesized in the seed and the total fatty acid of the seed comprises at least 0.1%, at least 0.13%, or at least 0.5% DPA.

[0355] In a further embodiment, DHA is synthesized in the seed.

[0356] In another embodiment, DHA is synthesized in the seed and the total fatty acid of the seed comprises at least 0.1%, at least 0.2%, or at least 0.5% DHA.

[0357] In yet a further embodiment, at least one ω 3 C20 LC-PUFA is synthesized in the seed and the total fatty acid of the seed comprises at least 2.5%, or at least 4.1% ω 3 C20 LC-PUFA.

[0358] In a further embodiment, EPA is synthesized in the seed and the efficiency of conversion of ALA to EPA in the seed is at least 2% or at least 14.6%.

[0359] In another embodiment, ω 3 polyunsaturated fatty acids that are the products of Δ 6-desaturation of ALA and/or the products of Δ 9 elongation of ALA, are synthesized in the seed, and the efficiency of conversion of ALA to said products in the seed is at least 22% or at least 24%.

[0360] In a further embodiment, DPA is synthesized from EPA in the seed and the efficiency of conversion of EPA to DPA in the seed is at least 5% or at least 7%.

[0361] Referring to extracts of the invention, in one embodiment, the total fatty acid content of the extract comprises at least 1.5%, at least 2.1%, or at least 2.5% EPA.

[0362] In another embodiment, the total fatty acid content of the extract comprises at least 0.1%, at least 0.13%, or at least 0.5% DPA.

[0363] In a further embodiment, the extract comprises DHA.

[0364] In another embodiment, the total fatty acid content of the extract comprises at least 0.1%, at least 0.2%, or at least 0.5% DHA.

[0365] In another embodiment, the total fatty acid content of the extract comprises at least 2.5%, or at least 4.1% ω 3 C20 LC-PUFA.

[0366] In yet a further embodiment, the extract comprises ARA, EPA, DPA, DHA, or any mixture of these in the triacylglycerols.

[0367] With regard to methods of the invention for producing a LC-PUFA, in one embodiment, the cell comprises at least one C20 LC-PUFA, and the total fatty acid of the cell comprises at least 2%, at least 4.7%, or at least 7.9% C20 LC-PUFA.

[0368] In another embodiment, the cell comprises at least one ω 3 C20 LC-PUFA, and the total fatty acid of the cell comprises at least 2.5%, or at least 4.1% ω 3 C20 LC-PUFA.

[0369] In a further embodiment, the cell comprises ω 3 polyunsaturated fatty acids that are the products of Δ 6-desaturation of ALA and/or the products of Δ 9 elongation of ALA, and the efficiency of conversion of ALA to said products in the cell is at least 22% or at least 24%.

[0370] In yet another embodiment, the cell comprises DPA, and the total fatty acid of the cell comprises at least 0.1%, at least 0.13%, or at least 0.5% DPA.

[0371] In a further embodiment, the cell comprises DPA, and the efficiency of conversion of EPA to DPA in the cell is at least 5% or at least 7%.

[0372] In another embodiment, the cell comprises EPA, and wherein the total fatty acid of the cell comprises at least 1.5%, at least 2.1%, or at least 2.5% EPA.

[0373] In a further embodiment, the cell comprises EPA, and the efficiency of conversion of ALA to EPA in the cell is at least 2% or at least 14.6%.

[0374] Polypeptides

[0375] In one aspect, the present invention provides a substantially purified polypeptide selected from the group consisting of:

[0376] i) a polypeptide comprising an amino acid sequence as provided in SEQ ID NO:1,

[0377] ii) a polypeptide comprising an amino acid sequence which is at least 40% identical to SEQ ID NO:1, and

[0378] iii) a biologically active fragment of i) or ii),

[0379] wherein the polypeptide has Δ 8 desaturase activity.

[0380] Preferably, the Δ 8 desaturase does not also have Δ 6 desaturase activity.

[0381] In another aspect, the present invention provides a substantially purified polypeptide selected from the group consisting of:

[0382] i) a polypeptide comprising an amino acid sequence as provided in SEQ ID NO:2,

[0383] ii) a polypeptide comprising an amino acid sequence which is at least 60% identical to SEQ ID NO:2, and

[0384] iii) a biologically active fragment of i) or ii),

[0385] wherein the polypeptide has Δ 5 elongase and/or Δ 6 elongase activity.

[0386] Preferably, the polypeptide has Δ 5 elongase and Δ 6 elongase activity, and wherein the polypeptide is more efficient at synthesizing DPA from EPA than it is at synthesizing ETA from SDA. More preferably, the polypeptide can be purified from algae. Furthermore, when expressed in yeast cells, is more efficient at elongating C20 LC-PUFA than C22 LC-PUFA.

[0387] In another aspect, the invention provides a substantially purified polypeptide selected from the group consisting of:

[0388] i) a polypeptide comprising an amino acid sequence as provided in SEQ ID NO:3, SEQ ID NO:85 or SEQ ID NO:86,

[0389] ii) a polypeptide comprising an amino acid sequence which is at least 40% identical to SEQ ID NO:3, SEQ ID NO:85 or SEQ ID NO:86, and

[0390] iii) a biologically active fragment of i) or ii),

[0391] wherein the polypeptide has $\Delta 9$ elongase and/or $\Delta 6$ elongase activity.

[0392] Preferably, the polypeptide has $\Delta 9$ elongase and $\Delta 6$ elongase activity. Preferably, the polypeptide is more efficient at synthesizing ETrA from ALA than it is at synthesizing ETA from SDA. Further, it is preferred that the polypeptide can be purified from algae or fungi.

[0393] In yet another aspect, the present invention provides a substantially purified polypeptide selected from the group consisting of:

[0394] i) a polypeptide comprising an amino acid sequence as provided in SEQ ID NO:4,

[0395] ii) a polypeptide comprising an amino acid sequence which is at least 70% identical to SEQ ID NO:4, and

[0396] iii) a biologically active fragment of i) or ii),

[0397] wherein the polypeptide has $\Delta 4$ desaturase activity.

[0398] In a further aspect, the present invention provides a substantially purified polypeptide selected from the group consisting of:

[0399] i) a polypeptide comprising an amino acid sequence as provided in SEQ ID NO:60,

[0400] ii) a polypeptide comprising an amino acid sequence which is at least 55% identical to SEQ ID NO:60, and

[0401] iii) a biologically active fragment of i) or ii), wherein the polypeptide has $\Delta 5$ desaturase activity.

[0402] In yet another aspect, the present invention provides a substantially purified polypeptide selected from the group consisting of:

[0403] i) a polypeptide comprising an amino acid sequence as provided in SEQ ID NO:64,

[0404] ii) a polypeptide comprising an amino acid sequence which is at least 90% identical to SEQ ID NO:64, and

[0405] iii) a biologically active fragment of i) or ii),

[0406] wherein the polypeptide has $\Delta 6$ desaturase activity.

[0407] In yet another aspect, the present invention provides a substantially purified polypeptide selected from the group consisting of:

[0408] i) a polypeptide comprising an amino acid sequence as provided in SEQ ID NO:88,

[0409] ii) a polypeptide comprising an amino acid sequence which is at least 76% identical to SEQ ID NO:88, and

[0410] iii) a biologically active fragment of i) or ii),

[0411] wherein the polypeptide has $\Delta 6$ elongase activity.

[0412] Preferably, in relation to any one of the above aspects, it is preferred that the polypeptide can be isolated from a species selected from the group consisting of *Pavlova* and *Melosira*.

[0413] By "substantially purified polypeptide" we mean a polypeptide that has been at least partially separated from the lipids, nucleic acids, other polypeptides, and other contaminating molecules with which it is associated in its native state. Preferably, the substantially purified polypeptide is at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated. Furthermore, the term "polypeptide" is used interchangeably herein with the term "protein".

[0414] The % identity of a polypeptide is determined by GAP (Needleman and Wunsch, 1970) analysis (GCG program) with a gap creation penalty=5, and a gap extension penalty=0.3. Unless stated otherwise, the query sequence is at least 15 amino acids in length, and the GAP analysis aligns the two sequences over a region of at least 15 amino acids. More preferably, the query sequence is at least 50 amino acids in length, and the GAP analysis aligns the two sequences over a region of at least 50 amino acids. Even more preferably, the query sequence is at least 100 amino acids in length and the GAP analysis aligns the two sequences over a region of at least 100 amino acids.

[0415] With regard to the defined polypeptides/enzymes, it will be appreciated that % identity figures higher than those provided above will encompass preferred embodiments. Thus, where applicable, in light of the minimum % identity figures, it is preferred that the polypeptide comprises an amino acid sequence which is at least 60%, more preferably at least 65%, more preferably at least 70%, more preferably at least 75%, more preferably at least 76%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 91%, more preferably at least 92%, more preferably at least 93%, more preferably at least 94%, more preferably at least 95%, more preferably at least 96%, more preferably at least 97%, more preferably at least 98%, more preferably at least 99%, more preferably at least 99.1%, more preferably at least 99.2%, more preferably at least 99.3%, more preferably at least 99.4%, more preferably at least 99.5%, more preferably at least 99.6%, more preferably at least 99.7%, more preferably at least 99.8%, and even more preferably at least 99.9% identical to the relevant nominated SEQ ID NO.

[0416] In a further embodiment, the present invention relates to polypeptides which are substantially identical to those specifically described herein. As used herein, with reference to a polypeptide the term "substantially identical" means the deletion, insertion and/or substitution of one or a few (for example 2, 3, or 4) amino acids whilst maintaining at least one activity of the native protein.

[0417] As used herein, the term "biologically active fragment" refers to a portion of the defined polypeptide/enzyme which still maintains desaturase or elongase activity (whichever is relevant). Such biologically active fragments can readily be determined by serial deletions of the full length protein, and testing the activity of the resulting fragment.

[0418] Amino acid sequence mutants/variants of the polypeptides/enzymes defined herein can be prepared by introducing appropriate nucleotide changes into a nucleic acid encoding the polypeptide, or by in vitro synthesis of the desired polypeptide. Such mutants include, for example, deletions, insertions or substitutions of residues within the amino acid sequence. A combination of deletion, insertion and substitution can be made to arrive at the final construct, provided that the final protein product possesses the desired characteristics.

[0419] In designing amino acid sequence mutants, the location of the mutation site and the nature of the mutation will depend on characteristic(s) to be modified. The sites for mutation can be modified individually or in series, e.g., by (1) substituting first with conservative amino acid choices and then with more radical selections depending upon the results achieved, (2) deleting the target residue, or (3) inserting other residues adjacent to the located site.

[0420] Amino acid sequence deletions generally range from about 1 to 30 residues, more preferably about 1 to 10 residues and typically about 1 to 5 contiguous residues.

[0421] Substitution mutants have at least one amino acid residue in the polypeptide molecule removed and a different residue inserted in its place. The sites of greatest interest for substitutional mutagenesis include sites identified as the active or binding site(s). Other sites of interest are those in which particular residues obtained from various strains or species are identical. These positions may be important for biological activity. These sites, especially those falling within a sequence of at least three other identically conserved sites, are preferably substituted in a relatively conservative manner. Such conservative substitutions are shown in Table 2.

[0422] Furthermore, if desired, unnatural amino acids or chemical amino acid analogues can be introduced as a substitution or addition into the polypeptides of the present invention. Such amino acids include, but are not limited to, the D-isomers of the common amino acids, 2,4-diaminobutyric acid, α -amino isobutyric acid, 4-aminobutyric acid, 2-aminobutyric acid, 6-amino hexanoic acid, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteine, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β -alanine, fluoro-amino acids, designer amino acids such as β -methyl amino acids, α -methyl amino acids, $N\alpha$ -methyl amino acids, and amino acid analogues in general.

[0423] Also included within the scope of the invention are polypeptides of the present invention which are differentially modified during or after synthesis, e.g., by biotinylation, benzylation, glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. These modifications may serve to increase the stability and/or bioactivity of the polypeptide of the invention.

TABLE 2

Exemplary substitutions.	
Original Residue	Exemplary Substitutions
Ala (A)	val; leu; ile; gly
Arg (R)	lys
Asn (N)	gln; his
Asp (D)	glu
Cys (C)	ser
Gln (Q)	asn; his
Glu (E)	asp
Gly (G)	pro, ala
His (H)	asn; gln
Ile (I)	leu; val; ala
Leu (L)	ile; val; met; ala; phe
Lys (K)	arg
Met (M)	leu; phe
Phe (F)	leu; val; ala
Pro (P)	gly
Ser (S)	thr
Thr (T)	ser
Trp (W)	tyr
Tyr (Y)	trp; phe
Val (V)	ile; leu; met; phe, ala

[0424] Polypeptides of the present invention can be produced in a variety of ways, including production and recovery of natural proteins, production and recovery of recombinant proteins, and chemical synthesis of the proteins. In one embodiment, an isolated polypeptide of the present invention is produced by culturing a cell capable of expressing the polypeptide under conditions effective to produce the polypeptide, and recovering the polypeptide. A preferred cell to culture is a recombinant cell of the present invention. Effective culture conditions include, but are not limited to, effective media, bioreactor, temperature, pH and oxygen conditions that permit protein production. An effective medium refers to any medium in which a cell is cultured to produce a polypeptide of the present invention. Such medium typically comprises an aqueous medium having assimilable carbon, nitrogen and phosphate sources, and appropriate salts, minerals, metals and other nutrients, such as vitamins. Cells of the present invention can be cultured in conventional fermentation bioreactors, shake flasks, test tubes, microtiter dishes, and petri plates. Culturing can be carried out at a temperature, pH and oxygen content appropriate for a recombinant cell. Such culturing conditions are within the expertise of one of ordinary skill in the art.

[0425] Polynucleotides

[0426] In one aspect, the present invention provides an isolated polynucleotide comprising a sequence of nucleotides selected from the group consisting of:

[0427] i) a sequence of nucleotides as provided in SEQ ID NO:5 or SEQ ID NO:6;

[0428] ii) a sequence encoding a polypeptide of the invention;

[0429] iii) a sequence of nucleotides which is at least 50% identical to SEQ ID NO:5 or SEQ ID NO:6; and

[0430] iv) a sequence which hybridizes to any one of i) to iii) under high stringency conditions.

[0431] In another aspect, the present invention provides an isolated polynucleotide comprising a sequence of nucleotides selected from the group consisting of:

[0432] i) a sequence of nucleotides as provided in SEQ ID NO:7 or SEQ ID NO:8;

[0433] ii) a sequence encoding a polypeptide of the invention;

[0434] iii) a sequence of nucleotides which is at least 51% identical to SEQ ID NO:7 or SEQ ID NO:8; and

[0435] iv) a sequence which hybridizes to any one of i) to iii) under high stringency conditions.

[0436] In yet another aspect, the present invention provides an isolated polynucleotide comprising a sequence of nucleotides selected from the group consisting of:

[0437] i) a sequence of nucleotides as provided in SEQ ID NO:9 or SEQ ID NO:10;

[0438] ii) a sequence encoding a polypeptide of the invention;

[0439] iii) a sequence of nucleotides which is at least 51% identical to SEQ ID NO:9 or SEQ ID NO:10; and

[0440] iv) a sequence which hybridizes to any one of i) to iii) under high stringency conditions.

[0441] In a preferred embodiment, the sequence encoding a polypeptide of the invention is nucleotides 31 to 915 or SEQ ID NO:9 or nucleotides 85 to 915 of SEQ ID NO:9.

[0442] In a further aspect, the present invention provides an isolated polynucleotide comprising a sequence of nucleotides selected from the group consisting of:

[0443] i) a sequence of nucleotides as provided in SEQ ID NO:11, SEQ ID NO:12 or SEQ ID NO:13;

[0444] ii) a sequence encoding a polypeptide of the invention;

[0445] iii) a sequence of nucleotides which is at least 70% identical to SEQ ID NO:11, SEQ ID NO:12 or SEQ ID NO:13; and

[0446] iv) a sequence which hybridizes to any one of i) to iii) under high stringency conditions.

[0447] In another aspect, the present invention provides an isolated polynucleotide comprising a sequence of nucleotides selected from the group consisting of:

[0448] i) a sequence of nucleotides as provided in SEQ ID NO:58 or SEQ ID NO:59;

[0449] ii) a sequence encoding a polypeptide of the invention;

[0450] iii) a sequence of nucleotides which is at least 55% identical to SEQ ID NO:58 or SEQ ID NO:59; and

[0451] iv) a sequence which hybridizes to any one of i) to iii) under high stringency conditions.

[0452] In another aspect, the present invention provides an isolated polynucleotide comprising a sequence of nucleotides selected from the group consisting of:

[0453] i) a sequence of nucleotides as provided in SEQ ID NO:63;

[0454] ii) a sequence encoding a polypeptide of the invention;

[0455] iii) a sequence of nucleotides which is at least 90% identical to SEQ ID NO:63; and

[0456] iv) a sequence which hybridizes to any one of i) to iii) under high stringency conditions.

[0457] In another aspect, the present invention provides an isolated polynucleotide comprising a sequence of nucleotides selected from the group consisting of:

[0458] i) a sequence of nucleotides as provided in SEQ ID NO:89;

[0459] ii) a sequence encoding a polypeptide of the invention;

[0460] iii) a sequence of nucleotides which is at least 76% identical to SEQ ID NO:89; and

[0461] iv) a sequence which hybridizes to any one of i) to iii) under high stringency conditions.

[0462] The present inventors are also the first to isolate polynucleotide encoding a keto-acyl synthase-like fatty acid elongase from a non-higher plant.

[0463] Accordingly, in a further aspect the present invention provides an isolated polynucleotide comprising a sequence of nucleotides selected from the group consisting of:

[0464] i) a sequence of nucleotides as provided in SEQ ID NO:55;

[0465] ii) a sequence of nucleotides which is at least 40% identical to SEQ ID NO:55; and

[0466] iii) a sequence which hybridizes to i) or ii) under high stringency conditions.

[0467] By an "isolated polynucleotide", including DNA, RNA, or a combination of these, single or double stranded, in the sense or antisense orientation or a combination of both, dsRNA or otherwise, we mean a polynucleotide which is at least partially separated from the polynucleotide sequences with which it is associated or linked in its native state. Preferably, the isolated polynucleotide is at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated. Furthermore, the term "polynucleotide" is used interchangeably herein with the term "nucleic acid molecule".

[0468] The % identity of a polynucleotide is determined by GAP (Needleman and Wunsch, 1970) analysis (GCG program) with a gap creation penalty=5, and a gap extension penalty=0.3. Unless stated otherwise, the query sequence is at least 45 nucleotides in length, and the GAP analysis aligns the two sequences over a region of at least 45 nucleotides. Preferably, the query sequence is at least 150 nucleotides in length, and the GAP analysis aligns the two sequences over a region of at least 150 nucleotides. More preferably, the query sequence is at least 300 nucleotides in length and the GAP analysis aligns the two sequences over a region of at least 300 nucleotides.

[0469] With regard to the defined polynucleotides, it will be appreciated that % identity figures higher than those provided above will encompass preferred embodiments. Thus, where applicable, in light of the minimum % identity figures, it is preferred that the polynucleotide comprises a nucleotide sequence which is at least 60%, more preferably at least 65%, more preferably at least 70%, more preferably at least 75%, more preferably at least 76%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 91%, more preferably at least 92%, more preferably at least 93%, more preferably at least 94%, more preferably at least 95%, more preferably at least 96%, more preferably at least 97%, more preferably at least 98%, more preferably at least 99%, more preferably at least 99.1%, more preferably at least 99.2%, more preferably at least 99.3%, more preferably at least 99.4%, more preferably at least 99.5%, more preferably at least 99.6%, more preferably at least 99.7%, more preferably at least 99.8%, and even more preferably at least 99.9% identical to the relevant nominated SEQ ID NO.

[0470] In a further embodiment, the present invention relates to polynucleotides which are substantially identical to those specifically described herein. As used herein, with reference to a polynucleotide the term "substantially identical" means the substitution of one or a few (for example 2, 3, or 4) nucleotides whilst maintaining at least one activity of the native protein encoded by the polynucleotide. In addition, this term includes the addition or deletion of nucleotides which results in the increase or decrease in size of the encoded native protein by one or a few (for example 2, 3, or 4) amino acids whilst maintaining at least one activity of the native protein encoded by the polynucleotide.

[0471] Oligonucleotides of the present invention can be RNA, DNA, or derivatives of either. The minimum size of such oligonucleotides is the size required for the formation of a stable hybrid between an oligonucleotide and a complementary sequence on a nucleic acid molecule of the present invention. Preferably, the oligonucleotides are at least 15 nucleotides, more preferably at least 18 nucleotides, more preferably at least 19 nucleotides, more preferably at least 20 nucleotides, even more preferably at least 25 nucleotides in length. The present invention includes oligonucleotides that can be used as, for example, probes to identify nucleic acid molecules, or primers to produce nucleic acid molecules. Oligonucleotide of the present invention used as a probe are typically conjugated with a label such as a radioisotope, an enzyme, biotin, a fluorescent molecule or a chemiluminescent molecule.

[0472] Polynucleotides and oligonucleotides of the present invention include those which hybridize under stringent conditions to a sequence provided as SEQ ID NO's: 5 to 13. As used herein, stringent conditions are those that (1) employ low ionic strength and high temperature for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate/0.1% NaDodSO₄ at 50° C.; (2) employ during hybridisation a denaturing agent such as formamide, for example, 50% (vol/vol) formamide with 0.1% bovine serum albumin, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, 75 mM sodium citrate at 42° C.; or (3) employ 50% formamide, 5×SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5× Den-

hardt's solution, sonicated salmon sperm DNA (50 g/ml), 0.1% SDS and 10% dextran sulfate at 42° C. in 0.2×SSC and 0.1% SDS.

[0473] Polynucleotides of the present invention may possess, when compared to naturally occurring molecules, one or more mutations which are deletions, insertions, or substitutions of nucleotide residues. Mutants can be either naturally occurring (that is to say, isolated from a natural source) or synthetic (for example, by performing site-directed mutagenesis on the nucleic acid).

[0474] Also provided are antisense and/or catalytic nucleic acids (such as ribozymes) which hybridize to a polynucleotide of the invention, and hence inhibit the production of an encoded protein. Furthermore, provided are dsRNA molecules, particularly small dsRNA molecules with a double stranded region of about 21 nucleotides, which can be used in RNA interference to inhibit the production of a polypeptide of the invention in a cell. Such inhibitory molecules can be used to alter the types of fatty acids produced by a cell, such as an animal cell, moss, or algal cell. The production of such antisense, catalytic nucleic acids and dsRNA molecules is well with the capacity of the skilled person (see for example, G. Hartmann and S. Endres, *Manual of Antisense Methodology*, Kluwer (1999); Haseloff and Gerlach, 1988; Perriman et al., 1992; Shippy et al., 1999; Waterhouse et al. (1998); Smith et al. (2000); WO 99/32619, WO 99/53050, WO 99/49029, and WO 01/34815).

[0475] Gene Constructs and Vectors

[0476] One embodiment of the present invention includes a recombinant vector, which includes at least one isolated polynucleotide molecule encoding a polypeptide/enzyme defined herein, inserted into any vector capable of delivering the nucleic acid molecule into a host cell. Such a vector contains heterologous nucleic acid sequences, that is nucleic acid sequences that are not naturally found adjacent to nucleic acid molecules of the present invention and that preferably are derived from a species other than the species from which the nucleic acid molecule(s) are derived. The vector can be either RNA or DNA, either prokaryotic or eukaryotic, and typically is a virus or a plasmid.

[0477] One type of recombinant vector comprises a nucleic acid molecule of the present invention operatively linked to an expression vector. As indicated above, the phrase operatively linked refers to insertion of a nucleic acid molecule into an expression vector in a manner such that the molecule is able to be expressed when transformed into a host cell. As used herein, an expression vector is a DNA or RNA vector that is capable of transforming a host cell and effecting expression of a specified nucleic acid molecule. Preferably, the expression vector is also capable of replicating within the host cell. Expression vectors can be either prokaryotic or eukaryotic, and are typically viruses or plasmids. Expression vectors of the present invention include any vectors that function (i.e., direct gene expression) in recombinant cells of the present invention, including in bacterial, fungal, endoparasite, arthropod, other animal, and plant cells. Preferred expression vectors of the present invention can direct gene expression in yeast, animal or plant cells.

[0478] In particular, expression vectors of the present invention contain regulatory sequences such as transcription

control sequences, translation control sequences, origins of replication, and other regulatory sequences that are compatible with the recombinant cell and that control the expression of nucleic acid molecules of the present invention. In particular, recombinant molecules of the present invention include transcription control sequences. Transcription control sequences are sequences which control the initiation, elongation, and termination of transcription. Particularly important transcription control sequences are those which control transcription initiation, such as promoter, enhancer, operator and repressor sequences. Suitable transcription control sequences include any transcription control sequence that can function in at least one of the recombinant cells of the present invention. A variety of such transcription control sequences are known to those skilled in the art.

[0479] Another embodiment of the present invention includes a recombinant cell comprising a host cell transformed with one or more recombinant molecules of the present invention. Transformation of a nucleic acid molecule into a cell can be accomplished by any method by which a nucleic acid molecule can be inserted into the cell. Transformation techniques include, but are not limited to, transfection, electroporation, microinjection, lipofection, adsorption, and protoplast fusion. A recombinant cell may remain unicellular or may grow into a tissue, organ or a multicellular organism. Transformed nucleic acid molecules can remain extrachromosomal or can integrate into one or more sites within a chromosome of the transformed (i.e., recombinant) cell in such a manner that their ability to be expressed is retained.

[0480] Transgenic Plants and Parts Thereof

[0481] The term "plant" as used herein as a noun refers to whole plants, but as used as an adjective refers to any substance which is present in, obtained from, derived from, or related to a plant, such as for example, plant organs (e.g. leaves, stems, roots, flowers), single cells (e.g. pollen), seeds, plant cells and the like. Plants provided by or contemplated for use in the practice of the present invention include both monocotyledons and dicotyledons. In preferred embodiments, the plants of the present invention are crop plants (for example, cereals and pulses, maize, wheat, potatoes, tapioca, rice, sorghum, millet, cassava, barley, or pea), or other legumes. The plants may be grown for production of edible roots, tubers, leaves, stems, flowers or fruit. The plants may be vegetables or ornamental plants. The plants of the invention may be: corn (*Zea mays*), canola (*Brassica napus*, *Brassica rapa* ssp.), flax (*Linum usitatissimum*), alfalfa (*Medicago sativa*), rice (*Oryza sativa*), rye (*Secale cereale*), sorghum (*Sorghum bicolor*, *Sorghum vulgare*), sunflower (*Helianthus annuus*), wheat (*Triticum aestivum*), soybean (*Glycine max*), tobacco (*Nicotiana tabacum*), potato (*Solanum tuberosum*), peanuts (*Arachis hypogaea*), cotton (*Gossypium hirsutum*), sweet potato (*Lopmoea batatas*), cassava (*Manihot esculenta*), coffee (*Cofea* spp.), coconut (*Cocos nucifera*), pineapple (*Anana comosus*), citrus tree (*Citrus* spp.), cocoa (*Theobroma cacao*), tea (*Camellia senensis*), banana (*Musa* spp.), avocado (*Persea americana*), fig (*Ficus casica*), guava (*Psidium guajava*), mango (*Mangifer indica*), olive (*Olea europaea*), papaya (*Carica papaya*), cashew (*Anacardium occidentale*), macadamia (*Macadamia intergrifolia*), almond (*Prunus amygdalus*), sugar beets (*Beta vulgaris*), oats, or barley.

[0482] In one embodiment, the plant is an oilseed plant, preferably an oilseed crop plant. As used herein, an "oilseed plant" is a plant species used for the commercial production of oils from the seeds of the plant. The oilseed plant may be oil-seed rape (such as canola), maize, sunflower, soybean, sorghum, flax (linseed) or sugar beet. Furthermore, the oilseed plant may be other *Brassicas*, cotton, peanut, poppy, mustard, castor bean, sesame, safflower, or nut producing plants. The plant may produce high levels of oil in its fruit, such as olive, oil palm or coconut. Horticultural plants to which the present invention may be applied are lettuce, endive, or vegetable brassicas including cabbage, broccoli, or cauliflower. The present invention may be applied in tobacco, cucurbits, carrot, strawberry, tomato, or pepper.

[0483] When the production of ω 3 LC-PUFA is desired it is preferable that the plant species which is to be transformed has an endogenous ratio of ALA to LA which is at least 1:1, more preferably at least 2:1. Examples include most, if not all, oilseeds such as linseed. This maximizes the amount of ALA substrate available for the production of SDA, ETA, ETrA, EPA, DPA and DHA.

[0484] The plants produced using the methods of the invention may already be transgenic, and/or transformed with additional genes to those described in detail herein. In one embodiment, the transgenic plants of the invention also produce a recombinant ω 3 desaturase. The presence of a recombinant ω 3 desaturase increases the ratio of ALA to LA in the plants which, as outlined in the previous paragraph, maximizes the production of LC-PUFA such as SDA, ETA, ETrA, EPA, DPA and DHA.

[0485] Grain plants that provide seeds of interest include oil-seed plants and leguminous plants. Seeds of interest include grain seeds, such as corn, wheat, barley, rice, sorghum, rye, etc. Leguminous plants include beans and peas. Beans include guar, locust bean, fenugreek, soybean, garden beans, cowpea, mungbean, lima bean, fava bean, lentils, chickpea, etc.

[0486] The term "extract or portion thereof" refers to any part of the plant. "Portion" generally refers to a specific tissue or organ such as a seed or root, whereas an "extract" typically involves the disruption of cell walls and possibly the partial purification of the resulting material. Naturally, the "extract or portion thereof" will comprise at least one LC-PUFA. Extracts can be prepared using standard techniques of the art.

[0487] Transgenic plants, as defined in the context of the present invention include plants and their progeny which have been genetically modified using recombinant techniques. This would generally be to cause or enhance production of at least one protein/enzyme defined herein in the desired plant or plant organ. Transgenic plant parts include all parts and cells of said plants such as, for example, cultured tissues, callus, protoplasts. Transformed plants contain genetic material that they did not contain prior to the transformation. The genetic material is preferably stably integrated into the genome of the plant. The introduced genetic material may comprise sequences that naturally occur in the same species but in a rearranged order or in a different arrangement of elements, for example an antisense sequence. Such plants are included herein in "transgenic plants". A "non-transgenic plant" is one which has not been genetically modified with the introduction of genetic mate-

rial by recombinant DNA techniques. In a preferred embodiment, the transgenic plants are homozygous for each and every gene that has been introduced (transgene) so that their progeny do not segregate for the desired phenotype.

[0488] Several techniques exist for introducing foreign genetic material into a plant cell. Such techniques include acceleration of genetic material coated onto microparticles directly into cells (see, for example, U.S. Pat. No. 4,945,050 and U.S. Pat. No. 5,141,131). Plants may be transformed using *Agrobacterium* technology (see, for example, U.S. Pat. No. 5,177,010, U.S. Pat. No. 5,104,310, U.S. Pat. No. 5,004,863, U.S. Pat. No. 5,159,135). Electroporation technology has also been used to transform plants (see, for example, WO 87/06614, U.S. Pat. Nos. 5,472,869, 5,384,253, WO 92/09696 and WO 93/21335). In addition to numerous technologies for transforming plants, the type of tissue which is contacted with the foreign genes may vary as well. Such tissue would include but would not be limited to embryogenic tissue, callus tissue type I and II, hypocotyl, meristem, and the like. Almost all plant tissues may be transformed during development and/or differentiation using appropriate techniques described herein.

[0489] A number of vectors suitable for stable transfection of plant cells or for the establishment of transgenic plants have been described in, e.g., Pouwels et al., *Cloning Vectors: A Laboratory Manual*, 1985, supp. 1987; Weissbach and Weissbach, *Methods for Plant Molecular Biology*, Academic Press, 1989; and Gelvin et al., *Plant Molecular Biology Manual*, Kluwer Academic Publishers, 1990. Typically, plant expression vectors include, for example, one or more cloned plant genes under the transcriptional control of 5' and 3' regulatory sequences and a dominant selectable marker. Such plant expression vectors also can contain a promoter regulatory region (e.g., a regulatory region controlling inducible or constitutive, environmentally- or developmentally-regulated, or cell- or tissue-specific expression), a transcription initiation start site, a ribosome binding site, an RNA processing signal, a transcription termination site, and/or a polyadenylation signal.

[0490] Examples of plant promoters include, but are not limited to ribulose-1,6-bisphosphate carboxylase small subunit, beta-conglycinin promoter, phaseolin promoter, high molecular weight glutenin (HMW-GS) promoters, starch biosynthetic gene promoters, ADH promoter, heat-shock promoters and tissue specific promoters. Promoters may also contain certain enhancer sequence elements that may improve the transcription efficiency. Typical enhancers include but are not limited to Adh-intron 1 and Adh-intron 6.

[0491] Constitutive promoters direct continuous gene expression in all cells types and at all times (e.g., actin, ubiquitin, CaMV 35S). Tissue specific promoters are responsible for gene expression in specific cell or tissue types, such as the leaves or seeds (e.g., zein, oleosin, napin, ACP, globulin and the like) and these promoters may also be used. Promoters may also be active during a certain stage of the plants' development as well as active in plant tissues and organs. Examples of such promoters include but are not limited to pollen-specific, embryo specific, corn silk specific, cotton fibre specific, root specific, seed endosperm specific promoters and the like.

[0492] In a particularly preferred embodiment, the promoter directs expression in tissues and organs in which lipid

and oil biosynthesis take place, particularly in seed cells such as endosperm cells and cells of the developing embryo. Promoters which are suitable are the oilseed rape napin gene promoter (U.S. Pat. No. 5,608,152), the *Vicia faba* USP promoter (Baumlein et al., 1991), the *Arabidopsis* oleosin promoter (WO 98/45461), the *Phaseolus vulgaris* phaseolin promoter (U.S. Pat. No. 5,504,200), the *Brassica* Bce4 promoter (WO 91/13980) or the legumin B4 promoter (Baumlein et al., 1992), and promoters which lead to the seed-specific expression in monocots such as maize, barley, wheat, rye, rice and the like. Notable promoters which are suitable are the barley Ipt2 or Ipt1 gene promoter (WO 95/15389 and WO 95/23230) or the promoters described in WO 99/16890 (promoters from the barley hordein gene, the rice glutelin gene, the rice oryza gene, the rice prolamin gene, the wheat gliadin gene, the wheat glutelin gene, the maize zein gene, the oat glutelin gene, the sorghum kasirin gene, the rye secalin gene). Other promoters include those described by Broun et al. (1998) and U.S. 20030159173.

[0493] Under certain circumstances it may be desirable to use an inducible promoter. An inducible promoter is responsible for expression of genes in response to a specific signal, such as: physical stimulus (heat shock genes); light (RUBP carboxylase); hormone (Em); metabolites; and stress. Other desirable transcription and translation elements that function in plants may be used.

[0494] In addition to plant promoters, promoters from a variety of sources can be used efficiently in plant cells to express foreign genes. For example, promoters of bacterial origin, such as the octopine synthase promoter, the nopaline synthase promoter, the mannopine synthase promoter; promoters of viral origin, such as the cauliflower mosaic virus (35S and 19S) and the like may be used.

[0495] It will be apparent that transgenic plants adapted for the production of LC-PUFA as described herein, in particular DHA, can either be eaten directly or used as a source for the extraction of essential fatty acids, of which DHA would be a constituent.

[0496] As used herein, "germination" refers to the emergence of the root tip from the seed coat after imbibition. "Germination rate" refers to the percentage of seeds in a population which have germinated over a period of time, for example 7 or 10 days, after imbibition. A population of seeds can be assessed daily over several days to determine the germination percentage over time.

[0497] With regard to seeds of the present invention, as used herein the term "germination rate which is substantially the same" means that the germination rate of the transgenic seeds is at least 60%, more preferably at least 80%, and even more preferably at least 90%, that of isogenic non-transgenic seeds. Germination rates can be calculated using techniques known in the art.

[0498] With further regard to seeds of the present invention, as used herein the term "timing of germination of the seed is substantially the same" means that the timing of germination of the transgenic seeds is at least 60%, more preferably at least 80%, and even more preferably at least 90%, that of isogenic non-transgenic seeds. Timing of germination can be calculated using techniques known in the art.

[0499] The present inventors have found that at least in some circumstances the production of LC-PUFA in recom-

binant plant cells is enhanced when the cells are homozygous for the transgene. As a result, it is preferred that the recombinant plant cell, preferably the transgenic plant, is homozygous for at least one desaturase and/or elongase gene. In one embodiment, the cells/plant are homozygous for the zebrafish $\Delta 6/\Delta 5$ desaturase and/or the *C. elegans* elongase.

[0500] Transgenic Non-Human Animals

[0501] Techniques for producing transgenic animals are well known in the art. A useful general textbook on this subject is Houdebine, *Transgenic animals—Generation and Use* (Harwood Academic, 1997).

[0502] Heterologous DNA can be introduced, for example, into fertilized mammalian ova. For instance, totipotent or pluripotent stem cells can be transformed by microinjection, calcium phosphate mediated precipitation, liposome fusion, retroviral infection or other means, the transformed cells are then introduced into the embryo, and the embryo then develops into a transgenic animal. In a highly preferred method, developing embryos are infected with a retrovirus containing the desired DNA, and transgenic animals produced from the infected embryo. In a most preferred method, however, the appropriate DNAs are co-injected into the pronucleus or cytoplasm of embryos, preferably at the single cell stage, and the embryos allowed to develop into mature transgenic animals.

[0503] Another method used to produce a transgenic animal involves microinjecting a nucleic acid into pro-nuclear stage eggs by standard methods. Injected eggs are then cultured before transfer into the oviducts of pseudopregnant recipients.

[0504] Transgenic animals may also be produced by nuclear transfer technology. Using this method, fibroblasts from donor animals are stably transfected with a plasmid incorporating the coding sequences for a binding domain or binding partner of interest under the control of regulatory. Stable transfectants are then fused to enucleated oocytes, cultured and transferred into female recipients.

[0505] Feedstuffs

[0506] The present invention includes compositions which can be used as feedstuffs. For purposes of the present invention, "feedstuffs" include any food or preparation for human or animal consumption (including for enteral and/or parenteral consumption) which when taken into the body (a) serve to nourish or build up tissues or supply energy; and/or (b) maintain, restore or support adequate nutritional status or metabolic function. Feedstuffs of the invention include nutritional compositions for babies and/or young children.

[0507] Feedstuffs of the invention comprise, for example, a cell of the invention, a plant of the invention, the plant part of the invention, the seed of the invention, an extract of the invention, the product of the method of the invention, the product of the fermentation process of the invention, or a composition along with a suitable carrier(s). The term "carrier" is used in its broadest sense to encompass any component which may or may not have nutritional value. As the skilled addressee will appreciate, the carrier must be suitable for use (or used in a sufficiently low concentration) in a feedstuff such that it does not have deleterious effect on an organism which consumes the feedstuff.

[0508] The feedstuff of the present invention comprises an oil, fatty acid ester, or fatty acid produced directly or indirectly by use of the methods, cells or plants disclosed herein. The composition may either be in a solid or liquid form. Additionally, the composition may include edible macronutrients, vitamins, and/or minerals in amounts desired for a particular use. The amounts of these ingredients will vary depending on whether the composition is intended for use with normal individuals or for use with individuals having specialized needs, such as individuals suffering from metabolic disorders and the like.

[0509] Examples of suitable carriers with nutritional value include, but are not limited to, macronutrients such as edible fats, carbohydrates and proteins. Examples of such edible fats include, but are not limited to, coconut oil, borage oil, fungal oil, black current oil, soy oil, and mono- and diglycerides. Examples of such carbohydrates include (but are not limited to): glucose, edible lactose, and hydrolyzed starch. Additionally, examples of proteins which may be utilized in the nutritional composition of the invention include (but are not limited to) soy proteins, electro-dialysed whey, electro-dialysed skim milk, milk whey, or the hydrolysates of these proteins.

[0510] With respect to vitamins and minerals, the following may be added to the feedstuff compositions of the present invention: calcium, phosphorus, potassium, sodium, chloride, magnesium, manganese, iron, copper, zinc, selenium, iodine, and Vitamins A, E, D, C, and the B complex. Other such vitamins and minerals may also be added.

[0511] The components utilized in the feedstuff compositions of the present invention can be of semi-purified or purified origin. By semi-purified or purified is meant a material which has been prepared by purification of a natural material or by de novo synthesis.

[0512] A feedstuff composition of the present invention may also be added to food even when supplementation of the diet is not required. For example, the composition may be added to food of any type, including (but not limited to): margarine, modified butter, cheeses, milk, yogurt, chocolate, candy, snacks, salad oils, cooking oils, cooking fats, meats, fish and beverages.

[0513] The genus *Saccharomyces* spp is used in both brewing of beer and wine making and also as an agent in baking, particularly bread. Yeast is a major constituent of vegetable extracts. Yeast is also used as an additive in animal feed. It will be apparent that genetically engineered yeast strains can be provided which are adapted to synthesise LC-PUFA as described herein. These yeast strains can then be used in food stuffs and in wine and beer making to provide products which have enhanced fatty acid content and in particular DHA content.

[0514] Additionally, LC-PUFA produced in accordance with the present invention or host cells transformed to contain and express the subject genes may also be used as animal food supplements to alter an animal's tissue or milk fatty acid composition to one more desirable for human or animal consumption. Examples of such animals include sheep, cattle, horses and the like.

[0515] Furthermore, feedstuffs of the invention can be used in aquaculture to increase the levels of LC-PUFA in fish for human or animal consumption.

[0516] In mammals, the so-called "Sprecher" pathway converts DPA to DHA by three reactions, independent of a $\Delta 7$ elongase, $\Delta 4$ desaturase, and a beta-oxidation step (Sprecher et al., 1995) (FIG. 1). Thus, in feedstuffs for mammal consumption, for example formulations for the consumption by human infants, it may only be necessary to provide DPA produced using the methods of the invention as the mammalian subject should be able to fulfill its nutritional needs for DHA by using the "Sprecher" pathway to convert DPA to DHA. As a result, in an embodiment of the present invention, a feedstuff described herein for mammalian consumption at least comprises DPA, wherein at least one enzymatic reaction in the production of DPA was performed by a recombinant enzyme in a cell.

[0517] Compositions

[0518] The present invention also encompasses compositions, particularly pharmaceutical compositions, comprising one or more of the fatty acids and/or resulting oils produced using the methods of the invention.

[0519] A pharmaceutical composition may comprise one or more of the LC-PUFA and/or oils, in combination with a standard, well-known, non-toxic pharmaceutically-acceptable carrier, adjuvant or vehicle such as phosphate-buffered saline, water, ethanol, polyols, vegetable oils, a wetting agent or an emulsion such as a water/oil emulsion. The composition may be in either a liquid or solid form. For example, the composition may be in the form of a tablet, capsule, ingestible liquid or powder, injectible, or topical ointment or cream. Proper fluidity can be maintained, for example, by the maintenance of the required particle size in the case of dispersions and by the use of surfactants. It may also be desirable to include isotonic agents, for example, sugars, sodium chloride, and the like. Besides such inert diluents, the composition can also include adjuvants, such as wetting agents, emulsifying and suspending agents, sweetening agents, flavoring agents and perfuming agents.

[0520] Suspensions, in addition to the active compounds, may comprise suspending agents such as ethoxylated isostearyl alcohols, polyoxyethylene sorbitol and sorbitan esters, microcrystalline cellulose, aluminum metahydroxide, bentonite, agar-agar, and tragacanth or mixtures of these substances.

[0521] Solid dosage forms such as tablets and capsules can be prepared using techniques well known in the art. For example, LC-PUFA produced in accordance with the present invention can be tableted with conventional tablet bases such as lactose, sucrose, and cornstarch in combination with binders such as acacia, cornstarch or gelatin, disintegrating agents such as potato starch or alginic acid, and a lubricant such as stearic acid or magnesium stearate. Capsules can be prepared by incorporating these excipients into a gelatin capsule along with antioxidants and the relevant LC-PUFA(s).

[0522] For intravenous administration, the PUFA produced in accordance with the present invention or derivatives thereof may be incorporated into commercial formulations.

[0523] A typical dosage of a particular fatty acid is from 0.1 mg to 20 g, taken from one to five times per day (up to 100 g daily) and is preferably in the range of from about 10 mg to about 1, 2, 5, or 10 g daily (taken in one or multiple

doses). As known in the art, a minimum of about 300 mg/day of LC-PUFA is desirable. However, it will be appreciated that any amount of LC-PUFA will be beneficial to the subject.

[0524] Possible routes of administration of the pharmaceutical compositions of the present invention include, for example, enteral (e.g., oral and rectal) and parenteral. For example, a liquid preparation may be administered orally or rectally. Additionally, a homogenous mixture can be completely dispersed in water, admixed under sterile conditions with physiologically acceptable diluents, preservatives, buffers or propellants to form a spray or inhalant.

[0525] The dosage of the composition to be administered to the patient may be determined by one of ordinary skill in the art and depends upon various factors such as weight of the patient, age of the patient, overall health of the patient, past history of the patient, immune status of the patient, etc.

[0526] Additionally, the compositions of the present invention may be utilized for cosmetic purposes. It may be added to pre-existing cosmetic compositions such that a mixture is formed or a LC-PUFA produced according to the subject invention may be used as the sole "active" ingredient in a cosmetic composition.

[0527] Medical, Veterinary, Agricultural and Aquacultural Uses

[0528] The present invention also includes the treatment of various disorders by use of the pharmaceutical and/or feedstuff compositions described herein. In particular, the compositions of the present invention may be used to treat restenosis after angioplasty. Furthermore, symptoms of inflammation, rheumatoid arthritis, asthma and psoriasis may also be treated with the compositions (including feedstuffs) of the invention. Evidence also indicates that LC-PUFA may be involved in calcium metabolism; thus, the compositions of the present invention may be utilized in the treatment or prevention of osteoporosis and of kidney or urinary tract stones.

[0529] Additionally, the compositions of the present invention may also be used in the treatment of cancer. Malignant cells have been shown to have altered fatty acid compositions. Addition of fatty acids has been shown to slow their growth, cause cell death and increase their susceptibility to chemotherapeutic agents. Moreover, the compositions of the present invention may also be useful for treating cachexia associated with cancer.

[0530] The compositions of the present invention may also be used to treat diabetes as altered fatty acid metabolism and composition have been demonstrated in diabetic animals.

[0531] Furthermore, the compositions of the present invention, comprising LC-PUFA produced either directly or indirectly through the use of the cells of the invention, may also be used in the treatment of eczema and in the reduction of blood pressure. Additionally, the compositions of the present invention may be used to inhibit platelet aggregation, to induce vasodilation, to reduce cholesterol levels, to inhibit proliferation of vessel wall smooth muscle and fibrous tissue, to reduce or to prevent gastrointestinal bleeding and other side effects of non-steroidal anti-inflammatory drugs (U.S. Pat. No. 4,666,701), to prevent or to treat endometriosis and premenstrual syndrome (U.S. Pat. No.

4,758,592), and to treat myalgic encephalomyelitis and chronic fatigue after viral infections (U.S. Pat. No. 5,116,871).

[0532] Further uses of the compositions of the present invention include, but are not limited to, use in the treatment or prevention of cardiac arrhythmia's, angioplasty, AIDS, multiple sclerosis, Crohn's disease, schizophrenia, foetal alcohol syndrome, attention deficient hyperactivity disorder, cystic fibrosis, phenylketonuria, unipolar depression, aggressive hostility, adrenoleukodystrophy, coronary heart disease, hypertension, obesity, Alzheimer's disease, chronic obstructive pulmonary disease, ulcerative colitis or an ocular disease, as well as for maintenance of general health.

[0533] Furthermore, the above-described pharmaceutical and nutritional compositions may be utilized in connection with animals (i.e., domestic or non-domestic, including mammals, birds, reptiles, lizards, etc.), as well as humans, as animals experience many of the same needs and conditions as humans. For example, the oil or fatty acids of the present invention may be utilized in animal feed supplements, animal feed substitutes, animal vitamins or in animal topical ointments.

[0534] Compositions such as feedstuffs of the invention can also be used in aquaculture to increase the levels of LC-PUFA in fish for human or animal consumption.

[0535] Any amount of LC-PUFA will be beneficial to the subject. However, it is preferred that an "amount effective to treat" the condition of interest is administered to the subject. Such dosages to effectively treat a condition which would benefit from administration of a LC-PUFA are known those skilled in the art. As an example, a dose of at least 300 mg/day of LC-PUFA for at least a few weeks, more preferably longer would be suitable in many circumstances.

[0536] Antibodies

[0537] The invention also provides monoclonal and/or polyclonal antibodies which bind specifically to at least one polypeptide of the invention or a fragment thereof. Thus, the present invention further provides a process for the production of monoclonal or polyclonal antibodies to polypeptides of the invention.

[0538] The term "binds specifically" refers to the ability of the antibody to bind to at least one protein of the present invention but not other proteins present in a recombinant cell, particularly a recombinant plant cell, of the invention.

[0539] As used herein, the term "epitope" refers to a region of a protein of the invention which is bound by the antibody. An epitope can be administered to an animal to generate antibodies against the epitope, however, antibodies of the present invention preferably specifically bind the epitope region in the context of the entire protein.

[0540] If polyclonal antibodies are desired, a selected mammal (e.g., mouse, rabbit, goat, horse, etc.) is immunised with an immunogenic polypeptide. Serum from the immunised animal is collected and treated according to known procedures. If serum containing polyclonal antibodies contains antibodies to other antigens, the polyclonal antibodies can be purified by immunoaffinity chromatography. Techniques for producing and processing polyclonal antisera are known in the art. In order that such antibodies may be made, the invention also provides polypeptides of the invention or

fragments thereof haptenised to another polypeptide for use as immunogens in animals or humans.

[0541] Monoclonal antibodies directed against polypeptides of the invention can also be readily produced by one skilled in the art. The general methodology for making monoclonal antibodies by hybridomas is well known. Immortal antibody-producing cell lines can be created by cell fusion, and also by other techniques such as direct transformation of B lymphocytes with oncogenic DNA, or transfection with Epstein-Barr virus. Panels of monoclonal antibodies produced can be screened for various properties; i.e., for isotype and epitope affinity.

[0542] An alternative technique involves screening phage display libraries where, for example the phage express scFv fragments on the surface of their coat with a large variety of complementarity determining regions (CDRs). This technique is well known in the art.

[0543] For the purposes of this invention, the term "antibody", unless specified to the contrary, includes fragments of whole antibodies which retain their binding activity for a target antigen. Such fragments include Fv, F(ab') and F(ab')₂ fragments, as well as single chain antibodies (scFv). Furthermore, the antibodies and fragments thereof may be humanised antibodies, for example as described in EP-A-239400.

[0544] Antibodies of the invention may be bound to a solid support and/or packaged into kits in a suitable container along with suitable reagents, controls, instructions and the like.

[0545] Preferably, antibodies of the present invention are detectably labeled. Exemplary detectable labels that allow for direct measurement of antibody binding include radio-labels, fluorophores, dyes, magnetic beads, chemiluminescers, colloidal particles, and the like. Examples of labels which permit indirect measurement of binding include enzymes where the substrate may provide for a coloured or fluorescent product. Additional exemplary detectable labels include covalently bound enzymes capable of providing a detectable product signal after addition of suitable substrate. Examples of suitable enzymes for use in conjugates include horseradish peroxidase, alkaline phosphatase, malate dehydrogenase and the like. Where not commercially available, such antibody-enzyme conjugates are readily produced by techniques known to those skilled in the art. Further exemplary detectable labels include biotin, which binds with high affinity to avidin or streptavidin; fluorochromes (e.g., phycoerythrin, phycoerythrin and allophycocyanins; fluorescein and Texas red), which can be used with a fluorescence activated cell sorter; haptens; and the like. Preferably, the detectable label allows for direct measurement in a plate luminometer, e.g., biotin. Such labeled antibodies can be used in techniques known in the art to detect proteins of the invention.

EXAMPLES

Example 1

Materials and Methods

[0546] Culturing *Pavlova salina*

[0547] *Pavlova salina* isolates including strain CS-49 from the CSIRO Collection of Living Microalgae was

cultivated under standard culture conditions (<http://www.marine.csiro.au/microalgae>). A stock culture from the Collection was sub-cultured and scaled-up in a dilution of 1 in 10 over consecutive transfers in 1 L Erlenmeyer flasks and then into 10 L polycarbonate carboys. The culture medium was f/2, a modification of Guillard and Ryther's (1962) f medium containing half-strength nutrients, with a growth temperature of $20 \pm 1^\circ \text{C}$. Other culturing conditions included a light intensity of $100 \mu\text{mol. photons PAR.m}^{-2}.\text{s}^{-1}$, 12:12 hour light:dark photoperiod, and bubbling with 1% CO_2 in air at a rate of $200 \text{ mL.L}^{-1}.\text{min}^{-1}$.

[0548] Yeast Culturing and Feeding with Precursor Fatty Acids

[0549] Plasmids were introduced into yeast by heat shock and transformants were selected on yeast minimal medium (YMM) plates containing 2% raffinose as the sole carbon source. Clonal inoculum cultures were established in liquid YMM with 2% raffinose as the sole carbon source. Experimental cultures were inoculated from these, in YMM+1% NP-40, to an initial OD_{600} of ~ 0.3 . Cultures were grown at 30°C . with shaking ($\sim 60 \text{ rpm}$) until OD_{600} was approximately 1.0. At this point galactose was added to a final concentration of 2% and precursor fatty acids were added to a final concentration of 0.5 mM. Cultures were incubated at 20°C . with shaking for a further 48 hours prior to harvesting by centrifugation. Cell pellets were washed with 1% NP-40, 0.5% NP-40 and water to remove any unincorporated fatty acids from the surface of the cells.

[0550] Gas Chromatography (GC) Analysis of Fatty Acids

[0551] Fatty Acid Preparation

[0552] Fatty acid methyl esters (FAME) were formed by transesterification of the centrifuged yeast pellet or *Arabidopsis* seeds by heating with $\text{MeOH}-\text{CHCl}_3-\text{HCl}$ (10:1:1, v/v/v) at $90-100^\circ \text{C}$. for 2 h in a glass test tube fitted with a Teflon-lined screw-cap. FAME were extracted into hexane-dichloromethane (4:1, v/v) and analysed by GC and GC-MS.

[0553] Capillary Gas-Liquid Chromatography (GC)

[0554] FAME were analysed with Hewlett Packard (HP) 5890 GC or Agilent 6890 gas chromatograph fitted with HP 7673A or 6980 series automatic injectors respectively and a flame-ionization detector (FID). Injector and detector temperatures were 290°C . and 310°C . respectively. FAME samples were injected at 50°C . onto a non-polar cross-linked methyl-silicone fused-silica capillary column (HP-5; $50 \text{ m} \times 0.32 \text{ mm i.d.}$; $0.17 \mu\text{m}$ film thickness.). After 1 min, the oven temperature was raised to 210°C . at $30^\circ \text{C. min}^{-1}$, then to a final temperature of 280°C . at $3^\circ \text{C. min}^{-1}$ where it was kept for 5 min. Helium was the carrier gas with a column head pressure of 65 KPa and the purge opened 2 min after injection. Identification of peaks was based on comparison of relative retention time data with standard FAME with confirmation using mass-spectrometry. For quantification Empower software (Waters) or Chemstation (Agilent) was used to integrate peak areas.

[0555] Gas Chromatography-Mass Spectrometry (GC-MS)

[0556] GC-MS was carried out on a Finnigan GCQ Plus GC-MS ion-trap fitted with on-column injection set at 4°C . Samples were injected using an AS2000 auto sampler onto a retention gap attached to an HP-5 Ultra 2 bonded-phase

column ($50 \text{ m} \times 0.32 \text{ mm i.d.} \times 0.17 \mu\text{m}$ film thickness). The initial temperature of 45°C . was held for 1 min, followed by temperature programming at $30^\circ \text{C. min}^{-1}$ to 140°C . then at $3^\circ \text{C. min}^{-1}$ to 310°C . where it was held for 12 min. Helium was used as the carrier gas. Mass spectrometer operating conditions were: electron impact energy 70 eV; emission current 250 μamp , transfer line 310°C .; source temperature 240°C .; scan rate 0.8 scans.s^{-1} and mass range 40-650 Dalton. Mass spectra were acquired and processed with Xcalibur™ software.

[0557] Construction of *P. salina* cDNA Library

[0558] mRNA, for the construction of a cDNA library, was isolated from *P. salina* cells using the following method. 2 g (wet weight) of *P. salina* cells were powdered using a mortar and pestle in liquid nitrogen and sprinkled slowly into a beaker containing 22 ml of extraction buffer that was being stirred constantly. To this, 5% insoluble polyvinylpyrrolidone, 90 mM 2-mercaptoethanol, and 10 mM dithiothreitol were added and the mixture stirred for a further 10 minutes prior to being transferred to a Corex™ tube. 18.4 ml of 3M ammonium acetate was added and mixed well. The sample was then centrifuged at $6000 \times g$ for 20 minutes at 4°C . The supernatant was transferred to a new tube and nucleic acid precipitated by the addition of 0.1 volume of 3M NaAc (pH 5.2) and 0.5 volume of cold isopropanol. After a 1 hour incubation at -20°C ., the sample was centrifuged at $6000 \times g$ for 30 minutes in a swing rotor. The pellet was resuspended in 1 ml of water extracted with phenol/chloroform. The aqueous layer was transferred to a new tube and nucleic acids were precipitated once again by the addition of 0.1 volume 3M NaAc (pH 5.2) and 2.5 volume of ice cold ethanol. The pellet was resuspended in water, the concentration of nucleic acid determined and then mRNA was isolated using the Oligotex mRNA system (Qiagen).

[0559] First strand cDNA was synthesised using an oligo-dT primer supplied with the ZAP-cDNA synthesis kit (Stratagene—cat # 200400) and the reverse transcriptase SuperscriptIII (Invitrogen). Double stranded cDNA was ligated to EcoRI/XhoI adaptors and from this a library was constructed using the ZAP-cDNA synthesis kit as described in the accompanying instruction manual (Stratagene—cat # 200400). The titer of the primary library was 2.5×10^5 plaque forming units (pfu)/ml and that of the amplified library was 2.5×10^9 pfu/ml. The average insert size of cDNA inserts in the library was 1.3 kilobases and the percentage of recombinants in the library was 74%.

Example 2

Microalgae and Polyunsaturated Fatty Acid Contents Thereof

[0560] The CSIRO Collection of Living Microalgae

[0561] CSIRO established and maintained a Collection of Living Microalgae (CLM) containing over 800 strains from 140 genera representing the majority of marine and some freshwater microalgal classes (list of strains available downloadable from <http://www.marine.csiro.au>). Selected microheterotrophic strains were also maintained.

[0562] This collection is the largest and most diverse microalgal culture collection in Australia. The CLM focused on isolates from Australian waters—over 80% of the strains

were isolated from diverse localities and climatic zones, from tropical northern Australia to the Australian Antarctic Territory, from oceanic, inshore coastal, estuarine, intertidal and freshwater environments. Additionally, emphasis has been placed on representation of different populations of a single species, usually by more than one strain. All strains in the culture collection were unialgal and the majority were clonal. A subset of strains were axenic. Another collection is the NIES-Collection (National Institute for Environmental Studies, Environment Agency) maintained in Japan.

[0563] Microalgae are known for their cosmopolitanism at the morphological species level, with very low endemism being shown. However this morphological cosmopolitanism can hide a plethora of diversity at the intra-specific level. There have been a number of studies of genetic diversity on different microalgae using approaches such as interbreeding, isozymes, growth rates and a range of molecular techniques. The diversity identified by these studies ranges from large regional and global scales (Chinain et al., 1997) to between and within populations (Gallagher, 1980; Medlin et al., 1996; Bolch et al., 1999a,b). Variation at the intra-specific level, between morphologically indistinguishable microalgae, can usually only be identified using strains isolated from the environment and cultured in the laboratory.

[0564] It is essential to have identifiable and stable genotypes within culture collections. While there are recorded instances of change or loss of particular characteristics in long term culture (Coleman, 1977), in general, culturing guarantees genetic continuity and stability of a particular strain. Cryopreservation strategies could also be used to limit the potential for genetic drift.

[0565] Microalgae and Their Use in Aquaculture

[0566] Because of their chemical/nutritional composition including PUFAs, microalgae are utilized in aquaculture as live feeds for various marine organisms. Such microalgae must be of an appropriate size for ingestion and readily digested. They must have rapid growth rates, be amenable to mass culture, and also be stable in culture to fluctuations in temperature, light and nutrients as may occur in hatchery systems. Strains fulfilling these attributes and used widely in aquaculture include northern hemisphere strains such as *Isochrysis* sp. (TISO) CS-177, *Pavlova lutheri* CS-182, *Chaetoceros calcitrans* CS-178, *C. muelleri* CS-176, *Skeletonema costatum* CS-181, *Thalassiosira pseudonana* CS-173, *Tetraselmis suecica* CS-187 and *Nannochloropsis oculata* CS-189. Australian strains used include *Pavlova pinguis* CS-375, *Skeletonema* sp. CS-252, *Nannochloropsis* sp. CS-246, *Rhodomonas salina* CS-24 and *Navicula jeffreyi* CS-46. Biochemical assessment of over 50 strains of microalgae used (or of potential use) in aquaculture found that cells grown to late-logarithmic growth phase typically contained 30 to 40% protein, 10 to 20% lipid and 5 to 15% carbohydrate (Brown et al., 1997).

[0567] Lipid Composition Including PUFA Content of Microalgae

[0568] There is considerable interest in microalgae containing a high content of the nutritionally important long-chain polyunsaturated fatty acids (LC-PUFA), in particular EPA [eicosapentaenoic acid, 20:5(ω 3)] and DHA [docosahexaenoic acid, 22:6(ω 3)] as these are essential for the health of both humans and aquacultured animals. While

these PUFA are available in fish oils, microalgae are the primary producers of EPA and DHA.

[0569] The lipid composition of a range of microalgae (46 strains) and particularly the proportion and content of important PUFA in the lipid of the microalgae were profiled. C₁₈-C₂₂ PUFA composition of microalgal strains from different algal classes varied considerably across the range of classes of phototrophic algae (Table 3, FIG. 2, see also Dunstan et al., 1994, Volkman et al., 1989; Mansour et al., 1999a). Diatoms and eustigmatophytes were rich in EPA and produced small amounts of the less common PUFA, ARA [arachidonic acid, 20:4(ω 6)] with negligible amounts of DHA. In addition, diatoms made unusual C₁₆ PUFA such as 16:4(ω 1) and 16:3(ω 4). In contrast, dinoflagellates had high concentrations of DHA and moderate to high proportions of EPA and precursor C₁₈ PUFA [18:5(ω 3) and 18:4(ω 3) SDA, stearidonic acid]. Prymnesiophytes also contained EPA and DHA, with EPA the dominant PUFA. Cryptomonads were a rich source of the C₁₈ PUFA 18:3(ω 3) (ALA α -linolenic acid) and SDA, as well as EPA and DHA. Green algae (e.g. Chlorophytes such as *Dunaliella* spp. and *Chlorella* spp.) were relatively deficient in both C20 and C22 PUFA, although some species had small amounts of EPA (up to 3%) and typically contained abundant ALA and 18:2(ω 6), and were also able to make 16:4(ω 3). The biochemical or nutritional significance of uncommon C₁₆ PUFA [e.g. 16:4(ω 3), 16:4(ω 1), 16:3(ω 4)] and C₁₈ PUFA (e.g. 18:5(ω 3) and STA1) is unclear. However there is current interest in C₁₈ PUFA such as SDA that are now being increasingly recognized as precursors for the beneficial EPA and DHA, unlike ALA which has only limited conversion to EPA and DHA.

[0570] New strains of Australian thraustochytrids were isolated. When examined, these thraustochytrids showed great morphological diversity from single cells to clusters of cells, complex reticulate forms and motile stages. Thraustochytrids are a group of single cell organisms that produce both high oil and LC-PUFA content. They were initially thought to be primitive fungi, although more recently have been assigned to the subclass Thraustochytridae (Chromista, Heterokonta), which aligns them more closely with other heterokont algae (e.g. diatoms and brown algae). Under culture, thraustochytrids can achieve considerably higher biomass yield (>20 g/L) than other microalgae. In addition, thraustochytrids can be grown in fermenters with an organic carbon source and therefore represent a highly attractive, renewable and contaminant-free, source of omega-3 oils.

TABLE 3

Distribution of selected PUFA and LC-PUFA in microalgae and other groups, and areas of application.			
Group	Genus/Species	PUFA	Application
Eustigmatophytes	<i>Nannochloropsis</i>	EPA	Aquaculture
Diatoms	<i>Chaetoceros</i>		
Dinoflagellates	<i>Cryptocodinium cohnii</i>	DHA	Aquaculture, health supplements, infant formula
Thraustochytrids	<i>Schizochytrium</i>		Aquaculture, infant formula
Red algae	<i>Porphyridium</i>	ARA	Pharmaceutical industry (precursor to prostaglandins)
Thraustochytrids	undescribed species		
Fungi	<i>Mortierella</i>		

TABLE 3-continued

Distribution of selected PUFA and LC-PUFA in microalgae and other groups, and areas of application.			
Group	Genus/Species	PUFA	Application
Blue green algae	<i>Spirulina</i>	GLA	health supplements

Abbreviations: γ -linolenic acid, GLA, 18:3 ω 6; 20:5 ω 3, eicosapentaenoic acid, EPA, 20:5 ω 3; docosahexaenoic acid, DHA, 22:6 ω 3; arachidonic acid, ARA, 20:4 ω 6.

[0571] Representative fatty acid profiles for selected Australian thraustochytrids are shown in Table 4. Strain O was particularly attractive as it contained very high levels of DHA (61%). Other PUFA were present at less than 5% each. High DHA-containing thraustochytrids often also contained high proportions of 22:5 ω 6, docosapentaenoic acid (DPA), as was observed for strains A, C and H. DPA was only a minor component in strain O under the culture conditions employed, making this strain particularly interesting. Strain A contained both DHA (28%) and EPA (16%) as the main LC-PUFA. Strains C and H differed from the other strains with ARA (10-13%) also being present as a major LC-PUFA. A number of other LC-PUFA were present in the thraustochytrids including DPA(3) and 22:4 ω 6 and other components.

TABLE 4

Fatty acid composition (% of total) of thraustochytrid strains.				
Fatty acid	Percentage composition			
	Strain			
	A	C	H	O
16:0	18.0	16.4	13.5	22.1
20:4 ω 6 ARA	4.0	10.5	13.4	0.7
20:5 ω 3 EPA	15.8	7.7	5.2	4.1
22:5 ω 6 DPA(6)	16.6	9.3	12.7	3.4
22:6 ω 3 DHA	28.2	21.6	19.2	61.0

[0572] The microalgal and thraustochytrid isolates in the CLM that may be used for isolation of genes involved in the synthesis of LC-PUFA are of the genera or species as follows:

[0573] Class Bacillariophyceae (Diatoms)

[0574] *Attheya septentrionalis*, *Aulacoseira* sp., *Chaetoceros affinis*, *Chaetoceros calcitrans*, *Chaetoceros calcitrans* f. *pumilum*, *Chaetoceros* cf. *mitra*, *Chaetoceros* cf. *peruvianus*, *Chaetoceros* cf. *radians*, *Chaetoceros didymus*, *Chaetoceros difficile*, *Chaetoceros gracilis*, *Chaetoceros muelleri*, *Chaetoceros simplex*, *Chaetoceros socialis*, *Chaetoceros* sp., *Chaetoceros* cf. *minus*, *Chaetoceros* cf. *tenuissimus*, *Coscinodiscus wailesii*, other *Coscinodiscus* spp., *Dactyliosolen fragilissimus*, *Detonula pumila*, *Ditylum brightwellii*, *Eucampia zodiacus*, *Extubocellulus spinifera*, *Lauderia annulata*, *Leptocylindrus danicus*, *Melosira moniliformis*, *Melosira* sp., *Minidiscus trioculatus*, *Minutocellus polymorphus*, *Odontella aurita*, *Odontella mobilensis*, *Odontella regia*, *Odontella rhombus*, *Odontella* sp., *Papiliocellulus simplex*, *Planktosphaerium* sp., *Proboscia alata*, *Rhizosolenia imbricata*, *Rhizosolenia setigera*, *Rhizosolenia* sp., *Skeletonema costatum*, *Skeletonema pseudocostatum*, *Skeletonema* sp., *Skeletonema tropicum*, other *Skel-*

etonema spp., *Stephanopyxis turris*, *Streptotheca* sp., *Streptotheca tamesis*, *Streptotheca* spp., *Striatella* sp., *Thalassiosira delicatula*, *Thalassiosira eccentrica*, *Thalassiosira mediterranea*, *Thalassiosira oceanica*, *Thalassiosira oestrupii*, *Thalassiosira profunda*, *Thalassiosira pseudonana*, *Thalassiosira rotula*, *Thalassiosira stellaris*, other *Thalassiosira* spp., *Achnanthes* cf. *amoena*, *Amphiprora* cf. *alata*, *Amphiprora hyalina*, *Amphora* spp., *Asterionella glacialis*, *Asterionellopsis glacialis*, *Biddulphia* sp., *Cocconeis* sp., *Cylindrotheca closterium*, *Cylindrotheca fusiformis*, *Delphineis* sp., *Diploneis* sp., *Entomoneis* sp., *Fallacia carpentariae*, *Grammatophora oceanica*, *Haslea ostrearia*, *Licmophora* sp., *Manguinea* sp., *Navicula* cf. *jeffreyi*, *Navicula jeffreyi*, other *Navicula* spp., *Nitzschia* cf. *bilobata*, *Nitzschia* cf. *constricta*, *Nitzschia* cf. *cylindrus* *Nitzschia* cf. *frustulum*, *Nitzschia* cf. *paleacea*, *Nitzschia closterium*, *Nitzschia fraudulenta*, *Nitzschia frustulum*, *Nitzschia* sp., *Phaeodactylum tricornutum*, *Pleurosigma delicatulum*, other *Pleurosigma* spp., *Pseudonitzschia australis*, *Pseudonitzschia delicatissima*, *Pseudonitzschia fraudulenta*, *Pseudonitzschia pseudodelicatissima*, *Pseudonitzschia pungens*, *Pseudonitzschia* sp., *Pseudostaurosira shiloi*, *Thalassionema nitzschioides*, or *Thalassiothrix heteromorphia*.

[0575] Class Chrysophyceae

[0576] *Chrysolepidomonas* cf. *marina*, *Hibberdia* spp., *Ochromonas danica*, *Pelagococcus subviridis*, *Phaeoplaca* spp., *Synura shagnicola* or other *Chrysophyte* spp.

[0577] Class Cryptophyceae

[0578] *Chroomonas placoidea*, *Chroomonas* sp., *Gemini-gera cryophila*, *Hemiselmis simplex*, *Hemiselmis* sp., *Rhodomonas baltica*, *Rhodomonas maculata*, *Rhodomonas salina*, *Rhodomonas* sp. or other *Cryptomonad* spp.

[0579] Class Dinophyceae (Dinoflagellates)

[0580] *Alexandrium affine*, *Alexandrium catenella*, *Alexandrium margalefi*, *Alexandrium minutum*, *Alexandrium protogonyaulax*, *Alexandrium tamarense*, *Amphidinium carterae*, *Amphidinium* cf. *britannicum*, *Amphidinium klebsii*, *Amphidinium* sp., *Amphidinium steinii*, *Amylax tricantha*, *Cryptothecodinium cohnii*, *Ensiculifera* sp., *Fragilidium* spp., *Gambierdiscus toxicus*, *Gymnodinium catenatum*, *Gymnodinium galathaneum*, *Gymnodinium galatheanum*, *Gymnodinium nolleri*, *Gymnodinium sanguineum*, or other *Gymnodinium* spp., *Gyrodinium pulchellum*, or other *Gyrodinium* spp., *Heterocapsa niei*, *Heterocapsa rotundata*, *Katodinium* cf. *rotundatum*, *Kryptoperidinium foliaceum*, *Peridinium balticum*, *Prorocentrum gracile*, *Prorocentrum mexicanum*, *Prorocentrum micans*, *Protoceratium reticulatum*, *Pyrodinium bahamense*, *Scrippsiella* cf. *precaria*, or other *Scrippsiella* spp. *Symbiodinium microadriaticum*, or *Woloszynskia* sp.

[0581] Class Euglenophyceae

[0582] *Euglena gracilis*.

[0583] Class Prasinophyceae

[0584] *Pycnococcus* sp., *Mantoniella squamata*, *Micromonas pusilla*, *Nephroselmis minuta*, *Nephroselmis pyriformis*, *Nephroselmis rotunda*, *Nephroselmis* spp., or other *Prasinophyte* spp., *Pseudoscurfieldia marina*, *Pycnococcus provasolii*, *Pyramimonas cordata*, *Pyramimonas gelidicola*, *Pyramimonas grossii*, *Pyramimonas oltmansii*, *Pyramimonas propulsa*, other *Pyramimonas* spp., *Tetraselmis antarctica*, *Tetraselmis chuii*, *Tetraselmis* sp., *Tetraselmis suecica*, or other *Tetraselmis* spp.

[0585] Class Prymnesiophyceae

[0586] *Chrysochromulina acantha*, *Chrysochromulina apheles*, *Chrysochromulina brevifilum*, *Chrysochromulina camella*, *Chrysochromulina hirta*, *Chrysochromulina kappa*, *Chrysochromulina minor*, *Chrysochromulina pienaar*, *Chrysochromulina simplex*, *Chrysochromulina sp.*, *Chrysochromulina spinifera*, *Chrysochromulina strobilus*, and other *Chrysophyte* spp., *Chrysotila lamellosa*, *Cricosphaera carterae*, *Crystallolithus hyalinus*, *Diacronema vlikianum*, *Dicrateria inornata*, *Dicrateria sp.*, *Emiliania huxleyi*, *Gephyrocapsa oceanica*, *Imantonia rotunda*, and other *Isochrysis* spp., *Ochrosphaera neapolitana*, *Pavlova* cf. *pinguis*, *Pavlova gyrans*, *Pavlova lutheri*, *Pavlova pinguis*, *Pavlova salina*, *Pavlova sp.*, *Phaeocystis* cf. *pouchetii*, *Phaeocystis globosa*, *Phaeocystis pouchetii*, other *Phaeocystis* spp., *Pleurochrysis* aff. *carterae*, *Prymnesium parvum*, *Prymnesium patelliferum*, other *Prymnesium* spp., or *Pseudoisochrysis paradoxa*.

[0587] Class Raphidophyceae

[0588] *Chattonella antiqua*, other *Chattonella* spp., *Fibrocapsa japonica*, other *Fibrocapsa* spp., *Heterosigma akashiwo*, *Heterosigma carterae*, or other *Heterosigma* spp.

[0589] Class Thraustochytridae

[0590] *Schizochytrium* spp., *Thraustochytrium aureum*, *Thraustochytrium roseum*, or other *Thraustochytrium* spp.

[0591] Class Eustigmatophytae as a source of genes for EPA production:

[0592] *Eustigmatois vischeri*, *Monodus subterraneus*, *Nannochloropsis oculata*, *Nannochloropsis salina*, *Vischeria helvetica*, *Vischeria punctata*, *Chloridella neglecta*, *Chloridella simplex*, *Chlorobotrys regularis*, *Ellipsoidon parvum*, *Ellipsoidon solitare*, *Eustigmatois magnus*, *Eustigmatois polyphem*, *Goniocloris sculpta*, *Monodus subterraneus*, *Monodus unipapilla*, *Nannochloropsis gaditana*, *Nannochloropsis granulata*, *Nannochloropsis limnetica*, *Pseudocharaciopsis ovalis*, *Pseudocharaciopsis texensis*, *Pseudostaurastrum limneticum*, or *Vischeria stellata*

Example 3

Isolation of Zebrafish $\Delta 5/6$ Desaturase and Functional Characterization in Yeast

[0593] As well as microalgae, some other organisms have the capacity to synthesise LC-PUFA from precursors such as α -linolenic acid (18:3, ALA) (see FIG. 1) and some of the genes responsible for such synthesis have been isolated (see Sayanova and Napier, 2004). The genes involved in omega-3 C_{20} +PUFA biosynthesis have been cloned from various organisms including algae, fungi, mosses, plants, nematodes and mammals. Based on the current understanding of genes involved in the synthesis of omega-3 C_{20} +PUFA, synthesis of EPA in plants would require the transfer of genes encoding at least two desaturases and one PUFA elongase. The synthesis of DHA from EPA in plants would require the additional transfer of a further desaturase and a further elongase (Sayanova and Napier, 2004). These enzymes are: for the synthesis of EPA, the sequential activities of a $\Delta 6$ desaturase, $\Delta 6$ elongase and a $\Delta 5$ desaturase is required. Based on an alternative pathway operative in some algae, EPA may also be synthesised by the sequential activities of a $\Delta 9$ elongase, a $\Delta 8$ desaturase and a $\Delta 5$ desaturase (Wallis and Browse, 1999; Qi et al., 2002). For the further conversion of EPA to DHA in plants, a further

transfer of a $\Delta 5$ elongase and $\Delta 4$ desaturase will be required (Sayanova and Napier, 2004).

[0594] Hastings et al. (2001) isolated a gene encoding a $\Delta 5/\Delta 6$ bifunctional desaturase from zebra fish (*Danio rerio*) and showed that, when expressed in yeast, the desaturase was able to catalyse the synthesis of both $\Delta 6$ (GLA and SDA) and $\Delta 5$ (20:4 and EPA) fatty acids. The desaturase was therefore able to act on both $\omega 6$ and $\omega 3$ substrates.

[0595] Isolation of the Zebrafish $\Delta 5/\Delta 6$ Desaturase

[0596] RNA was extracted using the RNeasy system according to the manufacturers instructions (Qiagen) from freshly dissected zebrafish livers. Based on the published sequence (Hastings et al. 2001), primers, sense, 5'-CCCAAGCTTACTATGGGTGGCGGAGGACAGC-3' (SEQ ID NO:39) and antisense 5'-CCGCTGGAGT-TATTTGTTGAGATACGC-3' (SEQ ID NO:40) at the 5' and 3' extremities of the zebrafish $\Delta 5/6$ ORF were designed and used in a one-step reverse transcription-PCR (RT-PCR, Promega) with the extracted RNA and using buffer conditions as recommended by the manufacturer. A single amplicon of size 1335 bp was obtained, ligated into pGEM-T easy (Promega) and the sequence confirmed as identical to that published.

[0597] A fragment containing the entire coding region (SEQ ID NO:38) was excised and ligated into the yeast shuttle vector pYES2 (Invitrogen). The vector pYES2 carried the URA3 gene, which allowed selection for yeast transformants based on uracil prototrophy. The inserted coding region was under the control of the inducible GAL1 promoter and polyadenylation signal of pYES2. The resultant plasmid was designated pYES2-zf $\Delta 5/6$, for introduction and expression in yeast (*Saccharomyces cerevisiae*).

[0598] Expression of Zebrafish $\Delta 5/\Delta 6$ Desaturase in Yeast

[0599] The gene construct pYES2-zf $\Delta 5/6$ was introduced into yeast strain S288. Yeast was a good host for analysing heterologous potential LC-PUFA biosynthesis genes including desaturases and elongases for several reasons. It was easily transformed. It synthesised no LC-PUFA of its own and therefore any new PUFA made was easily detectable without any background problems. Furthermore, yeast cells readily incorporated fatty acids from growth media into cellular lipids, thereby allowing the presentation of appropriate precursors to transformed cells containing genes encoding new enzymes, allowing for confirmation of their enzymatic activities.

[0600] Biochemical Analyses

[0601] Yeast cells transformed with pYES2-zf $\Delta 5/6$ were grown in YMM medium and induced by the addition of galactose. The fatty acids 18:3 $\omega 3$ (ALA, 0.5 mM) or 20:4 $\omega 3$ (ETA, 0.5 mM) were added to the medium as described above. After 48 hours incubation, the cells were harvested and fatty acid analysis carried out by capillary gas-liquid chromatography (GC) as described in Example 1. The analysis showed that 18:4 $\omega 3$ (1.9% of total fatty acid) was formed from 18:3 $\omega 3$ and 20:5 $\omega 3$ [0003] (0.24% of fatty acids) from 20:4 $\omega 3$, demonstrating $\Delta 6$ desaturase activity and $\Delta 5$ desaturase activity, respectively. These data are summarized in Table 5 and confirm the results of Hastings et al (2001).

Example 4

Isolation of *C. elegans* Elongase and Functional Characterization in Yeast[0602] Cloning of *C. elegans* Elongase Gene

[0603] Beaudoin and colleagues isolated a gene encoding an ELO-type fatty acid elongase from the nematode *Caenorhabditis elegans* (Beaudoin et al., 2000) and this gene was isolated as follows. Oligonucleotide primers having the sequences 5'-GCGGGTACCATGGCTCAGCATCCGCTC-3' (SEQ ID NO:41) (sense orientation) and 5'-GCGGGATCCTTAGTTGTTCTTCTTCTT-3' (SEQ ID NO:42) (anti-sense orientation) were designed and synthesized, based on the 5' and 3' ends of the elongase coding region. These primers were used in a PCR reaction to amplify the 867 basepair coding region from a *C. elegans* N2 mixed-stage gene library, using an annealing temperature of 58° C. and an extension time of 1 minute. The PCR amplification was carried out for 30 cycles. The amplification product was inserted into the vector PGEM™ T-easy (Promega) and the nucleotide sequence confirmed (SEQ ID NO:37). An EcoRI/BamHI fragment including the entire coding region was excised and inserted into the EcoRI/BglII sites of pSEC-TRP (Stratagene), generating pSEC-Ceelo, for introduction and expression in yeast. pSEC-TRP contains the TRP1 gene, which allowed for the selection of transformants in yeast by tryptophan prototrophy, and the GAL1 promoter for expression of the chimeric gene in an inducible fashion in the presence of galactose in the growth medium.

[0604] Functional Characterization of *C. elegans* Elongase Gene in Yeast

[0605] Yeast strain S288 was transformed, using the method described in Example 1, with both vectors pYES2-zfΔ5/6 and pSEC-Ceelo simultaneously and double transformants were selected on YMM medium that lacked tryptophan and uracil. The transformants grew well on both minimal and enriched media, in contrast to transformants of strain S288 carrying pSEC-Ceelo alone, in the absence of pYES2-zfΔ5/6, which grew quite poorly. Double transformants were grown in YMM medium and induced by the addition of galactose. The fatty acid 18:3ω3 (ALA, 0.5 mM) was added to the medium and, after 48 hours incubation, cells were harvested and fatty acid analysis carried out by capillary gas-liquid chromatography (GC) as described in Example 1. The analysis showed that 18:4ω3 (0.82% of total fatty acid) and 20:3ω3 (0.20%) were formed from 18:3ω3, and 20:4ω3 (0.02% of fatty acids) from either of those, demonstrating the concerted action of an elongase activity in addition to the Δ6 desaturase activity and Δ5 desaturase activity of the zebrafish desaturase (Table 5). The concerted action of a bifunctional Δ5/6 desaturase gene and an elongase gene has not been reported previously. In particular, the use of a bifunctional enzyme, if showing the same activities in plant cells, would reduce the number of genes that would need to be introduced and expressed. This also has not been reported previously.

TABLE 5

Enzymatic activities in yeast and <i>Arabidopsis</i>				
Clone	Precursor PUFA	Synthesised PUFA	% (of total FA)	Observed activity
pYES2-zfΔ5/6	18:3ω3	18:4ω3	1.9	Δ6 desaturase
pYES2-zfΔ5/6	20:4ω3	20:5ω3	0.24	Δ5 desaturase
pYES2zfΔ5/6, pSEC-Ceelo	18:3ω3	18:4ω3	0.82	Δ6 desaturase
		20:3ω3	0.20	Δ9 elongase
		20:4ω3	0.02	Δ6 elongase
				NOT
pYES2-psΔ8	18:3ω3	18:4ω3	—	Δ6 desaturase
pYES2-psΔ8	20:3ω3	20:4ω3	0.12	Δ8 desaturase
pYES2-psELO1	18:2ω6	20:2ω6	—	
pYES2-psELO1	18:3ω3	—	—	
pYES2-psELO1	20:3ω3	22:3ω3	—	
pYES2-psELO1	20:4ω3	22:4ω3	—	
pYES2-psELO1	20:5ω3	22:5ω3	0.82	Δ5 elongase
pYES-psELO2	18:2ω6	20:2ω6	0.12	Δ9 elongase
pYES-psELO2	18:3ω3	20:3ω3	0.20	Δ9 elongase
pYES-psELO2	20:3ω3	22:3ω3	—	
pYES-psELO2	20:4ω3	22:4ω3	—	
pYES-psELO2	20:5ω3	22:5ω3	—	
<i>Arabidopsis</i> + zfΔ5/6 & Ceelo (plant #1)	—	18:3ω6	0.32	Δ5/6 desaturase,
	—	18:4ω3	1.1	Δ5/6/9 elongase
	—	20:4ω6	1.1	
	—	20:5ω3	2.1	
	—	20:3ω6	1.1	
	—	20:4ω3	0.40	
	—	20:2ω6	3.2	
	—	20:3ω3	TR	
	—	22:4ω6	0.06	
	—	22:5ω3	0.13	
	—	22:3ω6	0.03	

TR, trace, not accurately determined.

Example 5

Coordinate Expression of Fatty Acid Desaturase and Elongase in Plants

[0606] Genetic Construct for Co-Expression of the Zebrafish $\Delta 6/\Delta 5$ Desaturase and *C. elegans* Elongase in Plant Cells

[0607] Beaudoin and colleagues (2000) showed that the *C. elegans* $\Delta 6$ elongase protein, when expressed in yeast, could elongate the C18 $\Delta 6$ desaturated fatty acids GLA and SDA, i.e. that it had $\Delta 6$ elongase activity on C18 substrates. They also showed that the protein did not have $\Delta 5$ elongase activity on a C20 substrate in yeast. We tested, therefore, whether this elongase would be able to elongate the $\Delta 6$ desaturated fatty acids GLA and SDA in *Arabidopsis* seed. *Arabidopsis thaliana* seed have been shown to contain both omega-6 (18:2, LA) and omega-3 (18:3, ALA) fatty acids (Singh et al, 2001). The presence of 18:3 in particular makes *Arabidopsis* seed an excellent system to study the expression of genes that could lead to the synthesis of omega-3 C₂₀+ PUFA like EPA and DHA.

[0608] The test for elongase activity in *Arabidopsis* required the coordinate expression of a $\Delta 6$ desaturase in the seed to first form GLA or SDA. We chose to express the elongase gene in conjunction with the zebrafish desaturase gene described above. There were no previous reports of the expression of the zebra fish $\Delta 6/\Delta 5$ desaturase and *C. elegans* elongase genes in plant cells, either individually or together.

[0609] Seed-specific co-expression of the zebra fish $\Delta 6/\Delta 5$ desaturase and *C. elegans* elongase genes was achieved by placing the genes independently under the control of a -309 napin promoter fragment, designated Fp1 (Stalberg et al., 1993). For plant transformation, the genes were inserted into the binary vector pWvec8 that comprised an enhanced hygromycin resistance gene as selectable marker (Wang et al., 1997). To achieve this, the *C. elegans* elongase coding region from Example 4 was inserted as a blunt-end fragment between the Fp1 and Nos 3' polyadenylation/terminator

fragment in the binary vector pWvec8, forming pCeloP-Wvec8. The zebrafish $\Delta 5/\Delta 6$ desaturase coding region from Example 3 was initially inserted as a blunt end fragment between the Fp1 and Nos 3' terminator sequences and this expression cassette assembled between the HindIII and ApaI cloning sites of the pBluescript cloning vector (Stratagene). Subsequently, the entire vector containing the desaturase expression cassette was inserted into the HindIII site of pCeloP-Wvec8, forming pZebdesatCeloP-Wvec8. The construct, shown schematically in FIG. 3, was introduced into *Agrobacterium* strain AGL1 (Valvekens et al., 1988) by electroporation prior to transformation into *Arabidopsis thaliana*, ecotype Columbia. The construct was also designated the "DO" construct, and plants obtained by transformation with this construct were indicated by the prefix "DO".

[0610] Plant Transformation and Analysis

[0611] Plant transformation was carried out using the floral dipping method (Clough and Bent, 1998). Seeds (T1 seeds) from the treated plants (T0 plants) were plated out on hygromycin (20 mg/l) selective media and transformed plants selected and transferred to soil to establish T1 plants. One hygromycin resistant plant was recovered from a first screen and established in soil. The transformation experiment was repeated and 24 further confirmed T1 transgenic plants were recovered and established in soil. Most of these T1 plants were expected to be heterozygous for the introduced transgenes.

[0612] T2 seed from the 25 transgenic plants were collected at maturity and analysed for fatty acid composition. As summarised in Table 6, seed of untransformed *Arabidopsis* (Columbia ecotype) contained significant amounts of both the $\omega 6$ and $\omega 3$, C18 fatty acid precursors LA and ALA but did not contain any $\Delta 6$ -desaturated C18 (18:3 $\omega 6$ or 18:4 $\omega 3$), $\omega 6$ -desaturated C20 PUFA or $\omega 3$ -desaturated C20 PUFA. In contrast, fatty acids of the seed oil of the transformed plants comprising the zebra fish $\Delta 5/\Delta 6$ desaturase and *C. elegans* elongase gene constructs contained 18:3 $\omega 6$, 18:4 $\omega 3$ and a whole series of $\omega 6$.

TABLE 6

Fatty acid composition in transgenic seed (% of the total fatty acid in seed oil).

Plant number	Fatty acid										
	GLA 18:3 $\omega 6$	SDA 18:4 $\omega 3$	ARA 20:4 $\omega 6$	EPA 20:5 $\omega 3$	DGLA 20:3 $\omega 6$	ETA 20:4 $\omega 3$	EDA 20:2 $\omega 6$	ETrA 20:3 $\omega 3$	22:4 $\omega 6$	DPA 22:5 $\omega 3$	22:3 $\omega 6$
Wt	—	—	—	—	—	—	—	—	—	—	—
DO1	0.32	1.10	1.10	2.10	1.10	0.40	3.20	TR	0.06	0.13	0.03
DO2	0.20	0.70	0.60	1.20	0.80	0.40	1.60	—	0.10	TR	—
DO3	0.20	0.50	0.40	0.80	0.60	0.30	1.90	—	TR	TR	—
DO4	0.30	0.90	0.80	1.30	1.10	0.50	1.90	—	—	0.10	—
DO5	0.10	0.50	0.20	0.40	0.40	—	0.30	—	TR	TR	—
DO6	0.30	1.00	1.00	1.70	1.20	0.50	2.50	—	0.10	0.10	—
DO7	0.10	0.40	0.40	0.70	0.70	0.30	1.60	—	TR	TR	—
DO8	0.30	1.20	1.10	2.10	1.40	0.60	2.80	—	0.10	0.10	—
DO9	0.30	1.30	0.90	2.20	1.30	0.60	3.10	—	0.10	0.10	—
DO10	0.10	0.40	0.30	0.70	0.50	0.30	0.10	—	TR	TR	—
DO11	0.30	1.00	1.40	2.30	1.50	0.60	3.20	—	0.10	0.20	—
DO12	0.40	1.40	1.10	1.90	1.20	0.60	2.30	—	0.10	0.10	—
DO13	0.20	0.60	0.60	0.90	0.80	0.40	0.40	—	TR	0.10	—
DO14	0.30	1.00	0.70	1.70	1.10	0.60	2.50	—	TR	TR	—
DO15	0.30	1.30	1.00	2.30	1.50	0.60	2.60	—	0.10	0.10	—
DO17	0.20	0.40	0.40	0.70	0.70	0.30	1.80	—	TR	TR	—
DO18	0.20	0.60	0.50	0.90	0.80	0.40	1.70	—	TR	TR	—

TABLE 6-continued

Fatty acid composition in transgenic seed (% of the total fatty acid in seed oil).											
Plant number	Fatty acid										
	GLA 18:3 ω 6	SDA 18:4 ω 3	ARA 20:4 ω 6	EPA 20:5 ω 3	DGLA 20:3 ω 6	ETA 20:4 ω 3	EDA 20:2 ω 6	ETRA 20:3 ω 3	22:4 ω 6	DPA 22:5 ω 3	22:3 ω 6
DO19	0.20	0.40	0.40	0.80	0.70	0.30	2.00	—	TR	0.10	—
DO20	0.30	1.00	0.50	0.90	0.70	0.30	1.60	—	TR	TR	—
DO21	0.30	1.20	0.90	2.00	1.30	0.60	2.50	—	—	0.10	—
DO22	0.30	0.90	0.70	1.20	1.00	0.40	0.30	—	TR	TR	—
DO23	—	—	—	—	0.10	0.10	1.80	—	—	—	—
DO24	0.30	1.10	0.70	1.50	1.10	0.50	2.90	—	TR	0.10	—
DO25	0.10	0.50	0.30	0.70	0.50	0.20	1.60	—	TR	0.10	—

Wt = untransformed *Arabidopsis* (Columbia).

TR indicates less than 0.05%.

Dash (—) indicates not detected.

[0613] and ω 3-C20 PUFA. These resulted from the sequential action of the desaturase and elongase enzymes on the respective C18 precursors. Most importantly and unexpectedly, the transgenic seed contained both 20:5 ω 3 (EPA), reaching at least 2.3% of the total fatty acid in the seedoil, and 22:5 ω 3 (DPA), reaching at least 0.2% of this omega-3 LC-PUFA in the fatty acid of the seedoil. The total C20 fatty acids produced in the transgenic seed oil reached at least 9.0%. The total ω 3 fatty acids produced that were a product of Δ 6 desaturation (i.e. downstream of 18:3 ω 3 (ALA), calculated as the sum of the percentages for 18:4 ω 3 (SDA), 20:4 ω 3 (ETA), 20:5 ω 3 (EPA) and 22:5 ω 3 (DPA)) reached at least 4.2%. These levels represent a conversion efficiency of ALA, which is present in seed oil of the wild-type *Arabidopsis* plants used for the transformation at a level of about 13-15%, to ω 3 products through a Δ 6 desaturation step of at least 28%. Stated otherwise, the ratio of ALA products to ALA (products:ALA) in the seed oil was at least 1:3.6. Of significance here, *Arabidopsis* has a relatively low amount of ALA in its seed oil compared to some commercial oilseed crops.

[0614] The T2 lines described above included lines that were homozygous for the transgenes as well as heterozygotes. To distinguish homozygotes and heterozygotes for lines expressing the transgenes at the highest levels, T2 plants were established from the T2 seed for the 5 lines

containing the highest EPA levels, using selection on MS medium containing hygromycin (15 mg/L) to determine the presence of the transgenes. For example, the T2 seed was used from the T1 plant designated DO11, containing 2.3% EPA and showing a 3:1 segregation ratio of resistant to susceptible progeny on the hygromycin medium, indicating that DO11 contained the transgenes at a single genetic locus. Homozygous lines were identified. For example, T2 progeny plant DO11-5 was homozygous as shown by the uniformly hygromycin resistance in its T3 progeny. Other T2 plants were heterozygous for the hygromycin marker.

[0615] The fatty acid profiles of T3 seed lots from DO11-5 and other T2 progeny of DO11 were analysed and the data are presented in Table 7. As expected, the EPA contents reflected segregation of the DO construct. The levels of EPA in the fatty acid of the seedoil obtained from the T3 lines were in three groups: negligible (nulls for the DO construct), in the range 1.6-2.3% (heterozygotes for the DO construct) and reaching at least 3.1% (homozygotes for the DO construct). The levels obtained were higher in homozygotes than heterozygotes, indicating a gene dosage effect. T3 seed from the DO11-5 plant synthesized a total of 9.6% new ω 3 and ω 6 PUFAs, including 3.2% EPA, 1.6% ARA, 0.1% DPA, 0.6% SDA and 1.8% GLA (Table 7). This level of EPA synthesis in seed was four fold higher than the 0.8% level previously achieved in linseed

TABLE 7

Fatty acid composition in transgenic seed (% of the total fatty acid in seed oil).														
Fatty acid	Wild-type	DO 11-5	DO 11-6	DO 11-7	DO 11-8	DO 11-10	DO 11-11	DO 11-12	DO 11-13	DO 11-16	DO 11-18	DO 11-19	DO 11-20	DO 11-21
14:0	0.3	0.0	0.1	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1
15:0	0.0	0.0	0.2	0.2	0.2	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.2	0.1
16:1 ω 7	0.5	0.4	0.6	0.7	0.6	0.5	0.4	0.6	0.5	0.6	0.4	0.4	0.7	0.5
16:0	8.1	7.1	7.9	7.8	7.6	7.0	7.1	7.8	7.7	7.6	6.8	6.7	7.6	7.3
17:1 ω 8	0.0	0.0	0.0	0.1	0.1	0.1	0.0	0.1	0.1	0.0	0.1	0.1	0.0	0.1
17:0	0.3	0.1	0.0	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.0	0.1
18:3 ω 6 GLA	0.0	0.6	0.0	0.0	0.0	0.3	0.3	0.0	0.4	0.0	0.2	0.3	0.0	0.4
18:4 ω 3 SDA	0.0	1.8	0.0	0.0	0.0	1.0	1.1	0.0	1.3	0.0	0.7	1.1	0.0	1.2
18:2 ω 6 LA	26.6	25.8	29.8	28.6	28.8	25.6	25.4	28.6	25.6	29.0	25.7	25.2	29.4	27.3
18:1 ω 9	17.9	18.7	15.6	19.6	18.2	22.0	18.6	18.6	20.4	15.5	20.1	19.8	16.6	14.8
18:1 ω 7/18:3 ω 3 ALA	16.0	11.5	15.3	14.7	15.9	10.6	11.6	14.5	11.1	16.0	13.7	13.6	14.8	13.1
18:0	3.4	4.2	2.9	2.7	2.8	3.5	3.9	2.8	3.9	2.9	3.3	3.4	2.9	3.7

TABLE 7-continued

Fatty acid composition in transgenic seed (% of the total fatty acid in seed oil).															
Fatty acid		Wild-type	DO 11-5	DO 11-6	DO11-7	DO11-8	DO11-10	DO11-11	DO11-12	DO11-13	DO11-16	DO11-18	DO11-19	DO11-20	DO11-21
19:0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.1	0.0	0.1
20:4ω6	ARA	0.0	1.6	0.0	0.0	0.0	0.9	0.9	0.0	1.3	0.0	0.4	0.8	0.0	1.3
20:5ω3	EPA	0.0	3.2	0.0	0.1	0.0	1.6	2.1	0.0	2.1	0.0	1.1	1.8	0.0	2.3
20:3ω6	DGLA	0.0	1.9	0.0	0.0	0.0	1.2	1.5	0.0	1.4	0.0	0.7	1.0	0.0	1.5
20:4ω3	ETA	0.0	0.4	0.0	0.0	0.0	0.4	0.6	0.0	0.2	0.0	0.3	0.4	0.0	0.5
20:2ω6		0.0	3.4	0.2	0.1	0.2	2.2	3.1	0.1	2.4	0.2	1.7	2.1	0.1	2.8
20:1ω9/ω11		17.4	10.9	17.8	18.1	17.3	14.8	12.5	18.2	13.2	18.0	15.4	14.0	18.6	12.4
20:1ω7		1.9	2.7	2.2	1.9	2.2	2.2	2.3	2.0	2.0	2.3	2.2	2.2	2.3	2.7
20:0		1.8	1.8	2.1	1.8	2.0	2.0	2.0	2.0	1.9	2.2	2.0	2.0	2.3	2.1
22:4ω6		0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.1
22:5ω3	DPA	0.0	0.1	0.0	0.0	0.0	0.1	0.1	0.0	0.1	0.0	0.1	0.1	0.0	0.2
22:1ω11/ω13		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
22:1ω9		1.3	0.8	1.9	1.7	1.7	1.5	1.1	1.7	1.1	2.0	1.6	1.4	2.1	1.5
22:1ω7		0.0	0.0	0.2	0.1	0.2	0.1	0.0	0.1	0.0	0.2	0.1	0.1	0.2	0.2
22:0		0.2	0.3	0.3	0.3	0.3	0.3	0.4	0.3	0.3	0.4	0.3	0.3	0.4	0.4
24:1ω9		0.6	0.4	0.2	0.2	0.3	0.2	0.2	0.2	0.2	0.3	0.2	0.2	0.2	0.3
24:1ω7		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
24:0		0.0	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.3

Wild-type here refers to untransformed *Arabidopsis thaliana*, ecotype Columbia

[0616] (Abbadi et al., 2004). Considering also that the level of ALA precursor for EPA synthesis in *Arabidopsis* seed was less than a third of that present in linseed, it appeared that the LC-PUFA pathway as described above which included a desaturase that was capable of using an acyl-CoA substrate, was operating with significantly greater efficiency than the acyl-PC dependent desaturase pathway expressed in linseed.

[0617] The relative efficiencies of the individual enzymatic steps encoded by the EPA construct can be assessed by examining the percentage conversion of substrate fatty acid to product fatty acids (including subsequent derivatives) in DO11-5. The zebra-fish $\Delta 5/\Delta 6$ desaturase exhibited strong $\Delta 5$ desaturation, with 89% of 20:4 ω 3 being converted to EPA and DPA, and 45% of 20:3 ω 6 being converted to ARA, consistent with the previously reported preference of this enzyme for $\omega 3$ PUFA over $\omega 6$ PUFA substrates (Hastings et al., 2001). In comparison, $\Delta 6$ -desaturation occurred at significantly lower levels, with 32% of ALA and 14% LA being converted to $\Delta 6$ -desaturated PUFA. Given that previous studies in yeast showed this enzyme to actually have higher $\Delta 6$ -desaturase activity than $\Delta 5$ -desaturase activity, the lower $\Delta 6$ -desaturation levels achieved in *Arabidopsis* seeds could be reflect a limited availability of ALA and LA substrates in the acyl-CoA pool (Singh et al., in press). The $\Delta 6$ -elongase operated highly efficiently, with 86% of GLA and 67% of SDA being elongated, suggesting that this enzyme may have a slight preference for elongation of $\omega 6$ —PUFA substrate.

[0618] The germination ability of the T2 (segregating) and T3 seed (homozygous population) was assessed on MS medium and on soil. Seed from the EPA and DPA containing lines DO11 and DO11-5 showed the same timing and frequency of germination as wild-type seed, and the T2 and T3 plants did not have any apparent abnormal morphological features. Plant growth rates in vitro or in soil and the quantities of seed obtained from the plants were also unaffected. Including the germination of the T1 seed from which plant DO11 was obtained, the normal germination of seed of

the DO11 line was thus observed over three generations. In addition, normal germination rates and timing were also observed for the other EPA and DPA containing seed. This feature was both important and not predictable, as higher plants do not naturally produce EPA or DPA and their seed therefore has never previously contained these LC-PUFA. Germination requires the catabolism of stored seed oils and use for growth and as an energy supply. The observed normal germination rates showed that plant seed were able to carry out these processes using EPA and DPA, and that these compounds were not toxic.

[0619] It has been reported that a $\Delta 4$ desaturase encoded by a gene isolated from *Thraustochytrium* spp and expressed in *Brassica juncea* leaves was able to convert exogenously supplied DPA to DHA (Qiu et al., 2001). DPA produced in the plant seed described herein can serve as a precursor for DHA production. This conversion of DPA to DHA may be achieved in plant cells by the introduction of a $\Delta 4$ desaturase gene into the DPA producing plant cells (Example 11).

[0620] Discussion

[0621] The presence of 22:5 ω 3 in the *Arabidopsis* seed oil implied that the *C. elegans* elongase gene not only had $\Delta 6$ elongase activity, but also $\Delta 5$ elongase activity in plant cells. This result was most surprising given that the gene had been demonstrated to lack $\Delta 5$ elongase activity in yeast. Furthermore, this demonstrated that only two genes could be used for the synthesis of DPA from ALA in plant cells. The synthesis of DPA in a higher plant has not previously been reported. Furthermore, the conversion efficiency of ALA to its $\omega 3$ products in seed, including EPA, DPA or both, of at least 28% was striking.

[0622] Synthesis of LC-PUFA such as EPA and DHA in cells such as plant cells by the $\Delta 6$ desaturation pathway required the sequential action of PUFA desaturases and elongases. The required desaturases in one pathway had $\Delta 6$, $\Delta 5$ and $\Delta 4$ desaturating activity, in that order, and the

required PUFA elongases had elongating activity on $\Delta 6$ and $\Delta 5$ substrates. This conventional pathway operates in algae, mosses, fungi, diatoms, nematodes and some freshwater fish (Sayanova and Napier, 2004). The PUFA desaturases from algae, fungi, mosses and worms are selective for desaturation of fatty acids esterified to the sn-2 position of phosphatidylcholine (PC) while the PUFA elongases act on fatty acids in the form of acyl-CoA substrates represented in the acyl-CoA pool of tissues. In contrast, vertebrate $\Delta 6$ desaturases have been shown to be able to desaturate acyl-CoA substrates (Domergue et al., 2003a).

[0623] Attempts to reconstitute LC-PUFA pathways in plant cells and other cells have to take into account the different sites of action and substrate requirements of the desaturases and elongase enzymes. For example, PUFA elongases are membrane bound, and perhaps even integral membrane proteins, which use acyl-CoAs which are present as a distinct pool in the endoplasmic reticulum (ER). This acyl-CoA pool is physiologically separated from the PC component of the ER, hence for a PUFA fatty acid to be sequentially desaturated and elongated it has to be transferred between PC and acyl-CoA pools in the ER. Therefore, earlier reported attempts to constitute LC-PUFA biosynthesis in yeast using desaturases and elongase from lower and higher plants, fungi and worms, have been inefficient, at best. In addition, the constituted pathways have led to the synthesis of only C20 PUFA such as ARA and EPA. There is no previous report of the synthesis of C22 PUFA such as DPA and DHA in yeast (Beaudoin et al., 2000, Domergue et al., 2003a).

[0624] The strategy described above of using a vertebrate desaturase, in this example a $\Delta 5/\Delta 6$ desaturase from zebra fish, with a $\Delta 6$ PUFA elongase from *C. elegans* had the advantage that both the desaturase and the elongase have activity on acyl-CoA substrates in the acyl-CoA pool. This may explain why this strategy was more efficient in the synthesis of LC-PUFA. Furthermore, using a bifunctional desaturase displaying dual $\Delta 5/\Delta 6$ desaturase activities allowed the synthesis of EPA by the action of only 2 genes instead of the 3 genes used by other researchers (Beaudoin et al., 2000, Domergue et al., 2003a). The use of a bifunctional $\Delta 5/\Delta 6$ elongase in plant cells also allowed the formation of DPA from ALA by the insertion of only three genes (one elongase and two desaturases) or, as exemplified, of only two genes (bifunctional elongase and bifunctional desaturase). Both of these aspects were surprising and unexpected.

[0625] Biochemical evidence suggests that the fatty acid elongation consists of 4 steps: condensation, reduction, dehydration and a second reduction. Two groups of condensing enzymes have been identified so far. The first are involved in the synthesis of saturated and monosaturated fatty acids (C18-22). These are the FAE-like enzymes and do not appear to have a role in LC-PUFA biosynthesis. The other class of elongases identified belong to the ELO family of elongases named after the ELO gene family whose activities are required for the synthesis of the very long-chain fatty acids of sphingolipids in yeast. Apparent paralogs of the ELO-type elongases isolated from LC-PUFA synthesizing organisms like algae, mosses, fungi and nematodes have been shown to be involved in the elongation and synthesis of LC-PUFA. It has been shown that only the expression of the condensing component of the elongase is

required for the elongation of the respective acyl chain. Thus the introduced condensing component of the elongase is able to successfully recruit the reduction and dehydration activities from the transgenic host to carry out successful acyl elongations. Thus far, successful elongations of C16 and C18 PUFA have been demonstrated in yeast by the heterologous expression of ELO type elongases. In this regard, the *C. elegans* elongase used as described above was unable to elongate C20 PUFA when expressed in yeast (Beaudoin et al., 2000). Our demonstration that the *C. elegans* elongase, when expressed in plants, was able to elongate the C20:5 fatty acid EPA as evidenced by the production of DPA in *Arabidopsis* seed was a novel and unexpected result. One explanation as to why the *C. elegans* elongase was able to elongate C20 PUFA in plants, but not in yeast, might reside in its ability to interact successfully with the other components of the elongation machinery of plants to bind and act on C20 substrates.

[0626] This example showed that an ELO-type elongase from a non-vertebrate organism was able to elongate C20 PUFA in plant cells. Leonard et al. (2002) reported that an ELO-type elongase gene isolated from humans, when expressed in yeast, was able to elongate EPA to DPA but in a non-selective fashion.

Example 6

Isolation of a $\Delta 8$ Desaturase Gene from *P. salina* and Functional Characterization in Yeast

[0627] Microalgae are the only organisms which have been reported to contain $\Delta 8$ desaturases, aside from the $\Delta 8$ sphingolipid desaturases in higher plants that are not involved in LC-PUFA biosynthesis. A gene encoding a $\Delta 8$ desaturase has been isolated from *Euglena gracilis* (Wallis and Browse, 1999). The existence of a $\Delta 8$ desaturase in *Isochrysis galbana* may be presumed because it contains a $\Delta 9$ elongase (Qi et al., 2002), the product of which, 20:3n-3, is the precursor for a $\Delta 8$ desaturase (see FIG. 1). The fatty acid profiles of microalgae alone, however, do not provide sufficient basis for identifying which microalgae will contain $\Delta 8$ desaturase genes since multiple pathways may operate to produce the LC-PUFA.

[0628] Isolation of a $\Delta 8$ Desaturase Gene Fragment

[0629] An alignment of $\Delta 6$ desaturase amino acid sequences with those from the following Genbank accession numbers, AF465283, AF007561, AAC15586 identified the consensus amino acid sequence blocks DHPGGS (SEQ ID NO:43), WWKDKHN (SEQ ID NO:44) and QIEHHLF (SEQ ID NO:45) corresponding to amino acid positions 204-210 and 394-400, respectively, of AF465283. DHPGGS corresponded to the "cytochrome b5 domain" block that had been identified previously (Mitchell and Martin, 1995). WWKDKHN was a consensus block that had not previously been identified or used to design degenerate primers for the isolation of desaturase genes. The QIEHHLF block, or variants thereof, corresponded to a required histidine-containing motif that was conserved in desaturases. It had been identified and used before as the "third His box" to design degenerate oligonucleotides for desaturase gene isolation (Michaelson et al., 1998). This combination of blocks had not been used previously to isolate desaturase genes.

[0630] Based on the second and third conserved amino acid blocks, the degenerate primers 5'-TGGTGAAR-

CAYAARCAAY-3' (SEQ ID NO:46) and 5'-GCGAGG-GATCCAAGGRAANARRTGRTGYTC-3' (SEQ ID NO:47) were synthesised. Genomic DNA from *P. salina* was isolated using the DNAeasy system (Qiagen). PCR amplifications were carried out in reaction volumes of 20 μ L using 20 pmol of each primer, 200 ng of *P. salina* genomic DNA and Hotstar Taq DNA polymerase (Qiagen) with buffer and nucleotide components as specified. The cycling conditions were: 1 cycle of 95° C. for 15 minutes; 5 cycles of 95° C. 1 min; 38° C., 1 min; 72° C., 1 min; followed by 35 cycles of 95° C., 35 sec; 52° C., 30 sec; 72° C., 1 min; and finishing with 1 cycle of 72° C., 10 min. A 515 basepair amplicon was generated, ligated into pGEM-T easy (Promega), sequenced and used as a probe to screen a *P. salina* cDNA library.

[0631] Isolation of a cDNA Encoding a $\Delta 8$ Desaturase from *P. salina*

[0632] A *P. salina* cDNA library in λ -bacteriophage was constructed using the Zap-cDNA Synthesis Kit (Stratagene) (see Example 1). The library was plated out at a concentration of ~50,000 plaques per plate and lifts taken with Hybond N+ membrane and treated using standard methods (Ausubel et al., 1988, supra). The 515 bp desaturase fragment, generated by PCR, was radio-labelled with ^{32} P-dCTP and used to probe the lifts under the following high stringency conditions: Overnight hybridisation at 65° C. in 6 $\times$ SSC with shaking, a 5 minute wash with 2 $\times$ SSC/0.1% SDS followed by two 10 minute washes with 0.2 $\times$ SSC/0.1% SDS.

[0633] Fifteen primary library plates (150 mm) were screened for hybridization to the labeled 515 bp fragment. Forty strongly hybridizing plaques were identified and ten of these were carried through to a secondary screen. Plasmids from five secondary plaques hybridizing to the 515 bp probe were excised with ExAssist Helper phage according to the suppliers protocol (Stratagene). The nucleotide sequences of the inserts were obtained using the ABI Prism Big Dye Terminator kit (PE Applied Biosystems). The nucleotide sequences were identical where they overlapped, indicating that all five inserts were from the same gene. One of the five inserts was shown to contain the entire coding region, shown below to be from a $\Delta 8$ desaturase gene. This sequence is provided as SEQ ID NO:6.

[0634] The full-length amino acid sequence (SEQ ID NO:1) revealed that the isolated cDNA encoded a putative $\Delta 6$ or $\Delta 8$ desaturase, based on BLAST analysis. These two types of desaturases are very similar at the amino acid level and it was therefore not possible to predict on sequence alone which activity was encoded. The maximum degree of identity between the *P. salina* desaturase and other desaturases (BLASTX) was 27-30%, while analysis using the GAP program which allows the insertions of "gaps" in the alignment showed that the maximum overall amino acid identity over the entire coding regions of the *P. salina* desaturase and AAD45877 from *Euglena gracilis* was 45%. A Pileup diagram of other sequences similar to the *Pavlova salina* desaturase is provided in FIG. 4.

[0635] The entire coding region of this clone, contained within an EcoRI/XhoI fragment, was inserted into pYES2 (Invitrogen), generating pYES2-ps $\Delta 8$, for introduction and functional characterisation in yeast. Cells of yeast strain S288 were transformed with pYES2-ps $\Delta 8$ as described in Example 1, and transformants were selected on medium

without uracil. The yeast cells containing pYES2-ps $\Delta 8$ were grown in culture and then induced by galactose. After the addition of 18:3 $\omega 3$ or 20:3 $\omega 3$ (0.5 mM) to the culture medium and 48 hours of further culturing at 30° C., the fatty acids in cellular lipids were analysed as described in Example 1. When 18:3 $\omega 3$ ($\Delta 9$, 12, 15) was added to the medium, no 18:4 $\omega 3$ ($\Delta 6$, 9, 12, 15) was detected. However, when 20:3 $\omega 3$ ($\Delta 11, 14, 17$) was added to the medium, the presence of 20:4 $\omega 3$ ($\Delta 8, 11, 14, 17$) in the cellular lipid of the yeast transformants was detected (0.12%). It was concluded the transgene encoded a polypeptide having $\Delta 8$ but not $\Delta 6$ desaturase activity in yeast cells.

[0636] Isolation of a gene encoding a $\Delta 8$ fatty acid desaturase that does not also have $\Delta 6$ desaturase activity has not been reported previously. The only previously reported gene encoding a $\Delta 8$ desaturase that was isolated (from *Euglena gracilis*) was able to catalyse the desaturation of both 18:3 $\omega 3$ and 20:3 $\omega 3$ (Wallis and Browse, 1999). Moreover, expression of a gene encoding a $\Delta 8$ desaturase has not previously been reported in higher plants.

[0637] As shown in FIG. 1, expression of a $\Delta 8$ desaturase in concert with a $\Delta 9$ elongase (for example the gene encoding ELO2—see below) and a $\Delta 5$ desaturase (for example, the zebrafish $\Delta 5/\Delta 6$ gene or an equivalent gene from *P. salina* or other microalgae) would cause the synthesis of EPA in plants.

[0638] Aside from providing an alternative route for the production of EPA in cells, the strategy of using a $\Delta 9$ elongase in combination with the $\Delta 8$ desaturase may provide an advantage in that the elongation, which occurs on fatty acids coupled to CoA, precedes the desaturation, which occurs on fatty acids coupled to PC, thereby ensuring the availability of the newly elongated C20 PUFA on PC for subsequent desaturations by $\Delta 8$ and $\Delta 5$ desaturases, leading possibly to a more efficient synthesis of EPA. That is, the order of reactions—an elongation followed by two desaturations—will reduce the number of substrate linking switches that need to occur. The increased specificity provided by the *P. salina* $\Delta 8$ desaturase is a further advantage.

Example 7

Isolation of *P. salina* ELO1 and ELO2 Fatty Acid Elongases

[0639] ELO-type PUFA elongases from organisms such as nematodes, fungi and mosses have been identified on the basis of EST or genome sequencing strategies. A gene encoding a $\Delta 9$ elongase with activity on 18:3 $\omega 3$ (ALA) was isolated from *Isochrysis galbana* using a PCR approach with degenerate primers, and shown to have activity in yeast cells that were supplied with exogenous 18:2 $\omega 6$ (LA) or 18:3 $\omega 3$ (ALA), forming C20 fatty acids 20:2 $\omega 6$ and 20:3 $\omega 3$ respectively. The coding region of the gene IgASE1 encoded a protein of 263 amino acids with a predicted molecular weight of about 30 kDa and with limited homology (up to 27% identity) to other elongating proteins.

[0640] Isolation of Elongase Gene Fragments from *P. salina*

[0641] Based on multiple amino acid sequence alignments for fatty acid elongases the consensus amino acid blocks FLHXYH (SEQ ID NO:48) and MYXYFY (SEQ ID

NO:49) were identified and the corresponding degenerate primers 5'-CAGGATCCTTYYTNCATNNNTAYCA-3' (SEQ ID NO:50) (sense) and 5'-GATCTAGARAARTARTANNNTACAT-3' (SEQ ID NO:51) (antisense) were synthesised. Primers designed to the motif FLHXYH or their use in combination with the MYXYF primer have not previously been described. These primers were used in PCR amplification reactions in reaction volumes of 20 μ L with 20 pmol of each primer, 200 ng of *P. salina* genomic DNA and Hotstar Taq DNA polymerase (Qiagen) with buffer and nucleotide components as specified by the supplier. The reactions were cycled as follows: 1 cycle of 95° C. for 15 minutes, 5 cycles of 95° C., 1 min, 38° C., 1 min, 72° C., 1 min, 35 cycles of 95° C., 35 sec, 52° C., 30 sec, 72° C., 1 min, 1 cycle of 72° C., 10 min. Fragments of approximately 150 bp were generated and ligated into pGEM-Teasy for sequence analysis.

[0642] Of the 35 clones isolated, two clones had nucleotide or amino acid sequence with similarity to known elongases. These were designated Elo1 and Elo2. Both gene fragments were radio-labelled with ³²P-dCTP and used to probe the *P. salina* cDNA library under the following high stringency conditions: overnight hybridisation at 65° C. in 6×SSC with shaking, 5 minute wash with 2×SSC/0.1% SDS followed by two 10 minute washes with 0.2×SSC/0.1% SDS. Ten primary library plates (150 mm) were screened using the Elo1 or Elo2 probes. Elo1 hybridized strongly to several plaques on each plate, whilst Elo2 hybridised to only three plaques in the ten plates screened. All Elo1-hybridising plaques were picked from a single plate and carried through to a secondary screen, whilst all three Elo2-hybridising plaques were carried through to a secondary screen. Each secondary plaque was then used as a PCR template using the forward and reverse primers flanking the multiple cloning site in the pBluescript phagemid and the PCR products electrophoresed on a 1% TAE gel. Following electrophoresis, the gel was blotted onto a Hybond N+ membrane and the membrane hybridised overnight with ³²P-labelled Elo1 and Elo2 probes. Six of the amplified Elo1 secondary plaques and one of the amplified Elo2 secondary plaques hybridised to the Elo1/2 probe (FIG. 5).

[0643] Two classes of elongase-like sequences were identified in the *P. salina* cDNA library on the basis of their hybridisation to the Elo1 and Elo2 probes. Phagemids that hybridised strongly to either labelled fragment were excised with ExAssist Helper phage (Stratagene), and sequenced using the ABI Prism Big Dye Terminator kit (PE Applied Biosystems). All of the 5 inserts hybridizing to the Elo1 probe were shown to be from the same gene. Similarly DNA sequencing of the 2 inserts hybridising to the Elo2 probe showed them to be from the same gene. The cDNA sequence of the Elo1 clone is provided as SEQ ID NO:8, and the encoded protein as SEQ ID NO:2, whereas the cDNA sequence of the Elo2 clone is provided as SEQ ID NO:10, and the encoded proteins as SEQ ID NO:3, SEQ ID NO:85 and SEQ ID NO:86 using three possible start methionines).

[0644] A comparison was performed of the Elo1 and Elo2 and other known PUFA elongases from the database using the PILEUP software (NCBI), and is shown in FIG. 6.

[0645] The Elo1 cDNA was 1234 nucleotides long and had an open reading frame encoding a protein of 302 amino acid residues. According to the PILEUP analysis, Elo1

clustered with other Elo-type sequences associated with the elongation of PUFA including $\Delta 6$ desaturated fatty acids (FIG. 6). The Elo1 protein showed the greatest degree of identity (33%) to an elongase from the moss, *P. patens* (Accession No. AF428243) across the entire coding regions. The Elo1 protein also displayed a conserved amino acids motifs found in all other Elo-type elongases.

[0646] The Elo2 cDNA was 1246 nucleotides long and had an open reading frame encoding a protein of 304 amino acid residues. According to PILEUP analysis, Elo2 clustered with other Elo-type sequences associated with the elongation of PUFA, including those with activity on $\Delta 6$ or $\Delta 9$ PUFA (FIG. 6). Elo2 was on the same sub-branch as the $\Delta 9$ elongase isolated from *Isochrysis galbana* (AX571775). Elo2 displayed 31% identity to the *Isochrysis* gene across its entire coding region. The Elo2 ORF also displayed a conserved amino acid motif found in all other Elo-type elongases.

Example 8

Functional Characterization of $\Delta 5$ Fatty Acid Elongase in Yeast and Plant Cells

[0647] Yeast

[0648] The entire coding region of the *P. salina* Elo1 gene was ligated into pYES2, generating pYES2-psELO1, for characterisation in yeast. This genetic construct was introduced into yeast strains and tested for activity by growth in media containing exogenous fatty acids as listed in the Table 8. Yeast cells containing pYES2-psELO1 were able to convert 20:5 ω 3 into 22:5 ω 3, confirming $\Delta 5$ elongase activity on C20 substrate. The conversion ratio of 7% indicated high activity for this substrate. The same yeast cells converted 18:4 ω 3 ($\Delta 6,9,12,15$) to 20:4 ω 3 and 18:3 ω 6 ($\Delta 6,9,12$) to 20:3 ω 6, demonstrating that the elongase also had $\Delta 6$ elongase activity in yeast cells, but at approximately 10-fold lower conversion rates (Table 8). This indicated that the Elo1 gene encodes a specific or selective $\Delta 5$ elongase in yeast cells. This represents the first report of a specific $\Delta 5$ elongase, namely an enzyme that has a greater $\Delta 5$ elongase activity when compared to $\Delta 6$ elongase activity. This molecule is also the first $\Delta 5$ elongase isolated from an algal source. This enzyme is critical in the conversion of EPA to DPA (FIG. 1).

[0649] Plants

[0650] The $\Delta 5$ elongase, Elo1 isolated from *Pavlova* is expressed in plants to confirm its ability to function in plants. Firstly, a plant expression construct is made for constitutive expression of Elo1. For this purpose, the Elo1 sequence is placed under the control of the 35S promoter in the plant binary vector pBI121 (Clontech). This construct is introduced into *Arabidopsis* using the floral dip method described above. Analysis of leaf lipids is used to determine the specificity of fatty acids elongated by the Elo1 sequence. In another approach, co expression of the Elo1 construct with the zebra fish $\Delta 5/\Delta 6$ desaturase/C. *elegans* elongase construct and the $\Delta 4$ desaturase isolated from *Pavlova*, results in DHA synthesis from ALA in *Arabidopsis* seed, demonstrating the use of the $\Delta 5$ elongase in producing DHA in cells. In a further approach, the Elo1 gene may be co-expressed with $\Delta 6$ -desaturase and $\Delta 5$ desaturase genes, or a $\Delta 6/\Delta 5$ bifunctional desaturase gene, to produce DPA

from ALA in cells, particularly plant cells. In an alternative approach, the $\Delta 5$ elongase and $\Delta 4$ elongase genes are used in combination with the PKS genes of *Shewanella* which produce EPA (Takeyama et al., 1997), in plants, for the synthesis of DHA.

TABLE 8

Conversion of fatty acids in yeast cells transformed with genetic constructs expressing Elo1 or Elo2.			
Clone	Fatty acid precursor/(% of total FA)	Fatty acid formed/(% of total FA)	Conversion ratio (%)
pYES2-psELO1	20:5n - 3/3%	22:5n - 3/0.21%	7%
pYES2-psELO1	18:4n - 3/16.9%	20:4n - 3/0.15%	0.89%
pYES2-psELO1	18:3n - 6/19.8%	20:3n - 6/0.14%	0.71%
pYES2-psELO1	20:5n - 3/2.3%	22:5n - 3/tr	—
pYES2-psELO2	18:4n - 3/32.5%	20:4n - 3/0.38%	1.2%
pYES2-psELO2	18:3n - 6/12.9%	20:3n - 6/0.08%	0.62%
pYES2-psELO2	18:2n - 6/30.3%	20:2n - 6/0.12%	0.40%
pYES2-psELO2	18:3n - 3/42.9%	18:3n - 3/0.20%	0.47%

tr: trace amounts (<0.02%) detected.

Example 9

Functional Characterization of $\Delta 9$ Fatty Acid Elongase in Yeast and Plant Cells

[0651] Expression in Yeast Cells

[0652] The entire coding region of the *P. salina* Elo2 gene encoding a protein of 304 amino acids (SEQ ID NO:3) was ligated into pYES2, generating pYES2-psELO2, for characterisation in yeast. This genetic construct was introduced into yeast strains and tested for activity by growth in media containing exogenous fatty acids. Yeast cells containing pYES2-psELO2 were able to convert 18:2 ω 6 into 20:2 ω 6 (0.12% of total fatty acids) and 18:3 ω 3 into 20:3 ω 3 (0.20%), confirming $\Delta 9$ elongase activity on C18 substrates (Table 8). These cells were also able to convert 18:3 ω 6 into 20:3 ω 6 and 18:4 ω 3 into 20:4 ω 3, confirming $\Delta 6$ elongase activity on C18 substrates in yeast. However, since the 18:3 ω 6 and 18:4 ω 3 substrates also have a desaturation in the $\Delta 9$ position, it could be that the Elo2 enzyme is specific for $\Delta 9$ -desaturated fatty acids, irrespective of whether they have a $\Delta 6$ desaturation as well. The cells were able to convert 20:5 ω 3 into the 22:5 product DPA. This is the first report of a $\Delta 9$ elongase that also has $\Delta 6$ elongase activity from a non-vertebrate source, in particular from a fungal or algal source.

[0653] As the coding region contained three possible ATG start codons corresponding to methionine (Met) amino acids at positions 1, 11 (SEQ ID NO:85) and 29 (SEQ ID NO:86) of SEQ ID NO:3, the possibility that polypeptides beginning at amino acid positions 11 or 29 would also be active was tested. Using 5' oligonucleotide (sense) primers corresponding to the nucleotide sequences of these regions, PCR amplification of the coding regions was performed, and the

resultant products digested with EcoRI. The fragments are cloned into pYES2 to form pYES2-psELO2-11 and pYES2-psELO2-29. Both plasmids are shown to encode active $\Delta 9$ -elongase enzymes in yeast. The three polypeptides may also be expressed in *Synechococcus* or other cells such as plant cells to demonstrate activity.

[0654] Expression in Plant Cells

[0655] The $\Delta 9$ elongase gene, Elo2, isolated from *Pavlova* was expressed in plants to confirm its ability to function in plants. Firstly, a plant expression construct is made for constitutive expression of Elo2. For this purpose, the Elo2 coding sequence from amino acid position 1 of SEQ ID NO:3 was placed under the control of the 35S promoter in the plant binary vector pBI121 (Clontech). This construct is introduced into *Arabidopsis* using the floral dip method described above. Analysis of leaf lipids indicates the specificity of fatty acids that are elongated by the Elo2 sequence.

[0656] Co-Expression of $\Delta 9$ Elongase and $\Delta 8$ -Desaturase Genes in Transformed Cells

[0657] The *P. salina* $\Delta 8$ -desaturase and $\Delta 9$ -elongase were cloned into a single binary vector, each under the control of the constitutive 35S promoter and nos terminator. In this gene construct, pBI121 containing the $\Delta 8$ -desaturase sequence was cut with HindIII and ClaI (blunt-ended) to release a fragment containing the 35S promoter and the $\Delta 8$ -desaturase gene, which was then ligated to the HindIII+ SacI (blunt ended) cut pXZP143/ $\Delta 9$ -elongase vector to result in the intermediate pJRP013. This intermediate was then opened with HindIII and ligated to a pWvec8/ $\Delta 9$ -elongase binary vector (also HindIII-opened) to result in the construct pJRP014, which contains both genes between the left and right T-DNA borders, together with a hygromycin selectable marker gene suitable for plant transformation.

[0658] This double-gene construct was then used to transform tobacco using a standard *Agrobacterium*-mediated transformation technique. Following introduction of the construct into *Agrobacterium* strain AGL1, a single transformed colony was used to inoculate 20 mL of LB media and incubated with shaking for 48 hours at 28° C. The cells were pelleted (1000 g for 10 minutes), the supernatant discarded, and the pellet resuspended in 20 mL of sterile MS media. This was step was then repeated before 10 ml of this *Agrobacterium* solution was added to freshly cut (1 cm squares) tobacco leaves from cultivar W38. After gentle mixing, the tobacco leaf pieces and *Agrobacterium* solution were allowed to stand at room temperature for 10 min. The leaf pieces were transferred to MS plates, sealed, and incubated (co-cultivation) for 2 days at 24° C. Transformed cells were selected on medium containing hygromycin, and shoots regenerated. These shoots were then cut off and transferred to MS-rooting media pots for root growth, and eventually transferred to soil. Both leaf and seed lipids from these plants are analysed for the presence of 20:2 ω 6, 20:3 ω 6, 20:3 ω 3 and 20:4 ω 3 fatty acids, demonstrating the co-expression of the two genes.

[0659] Discussion

[0660] Biochemical evidence suggests that the fatty acid elongation consists of 4 steps: condensation, reduction, dehydration and a second reduction, and the reaction is catalysed by a complex of four proteins, the first of which catalyses the condensation step and is commonly called the

elongase. There are 2 groups of condensing enzymes identified so far. The first are involved in the synthesis of saturated and monounsaturated fatty acids (C18-22). These are the FAE-like enzymes and do not play any role in LC-PUFA biosynthesis. The other class of elongases identified belong to the ELO family of elongases named after the ELO gene family whose activities are required for the synthesis of very LC fatty acids of sphingolipids in yeast. Apparent paralogs of the ELO-type elongases isolated from LC-PUFA synthesizing organisms like algae, mosses, fungi and nematodes have been shown to be involved in the elongation and synthesis of LC-PUFA. It has been shown that only the expression of the condensing component of the elongase is required for the elongation of the respective acyl chain. Thus the introduced condensing component of the elongase is able to successfully recruit the reduction and dehydration activities from the transgenic host to carry out successful acyl elongations. This was also true for the *P. salina* $\Delta 9$ -elongase.

Example 10

Isolation of a Gene Encoding a $\Delta 4$ -Desaturase from *P. salina*

[0661] The final step in the aerobic pathway of DHA synthesis in organisms other than vertebrates, such as micro-organisms, lower plants including algae, mosses, fungi, and possibly lower animals, is catalysed by a $\Delta 4$ -desaturase that introduces a double bond into the carbon chain of the fatty acid at the $\Delta 4$ position. Genes encoding such an enzyme have been isolated from the algae *Euglena* and *Pavlova* and from *Thraustochytrium*, using different approaches. For example, $\Delta 4$ -desaturase genes from *Pavlova lutheri* and *Euglena gracilis* were isolated by random sequencing of cloned ESTs (EST approach, Meyer et al., 2003; Tonon et al., 2003), and a $\Delta 4$ -desaturase gene from *Thraustochytrium* sp. ATCC21685 was isolated by RT-PCR using primers corresponding to a cytochrome b_5 HPGG domain and histidine box III region (Qiu et al., 2001). The cloned $\Delta 4$ -desaturase genes encoded front-end desaturases whose members are characterised by the presence of an N-terminal cytochrome b_5 -like domain (Napier et al., 1999; Sayanova and Napier, 2004).

[0662] Isolation of a Gene Fragment From a $\Delta 4$ -Desaturase Gene from *P. salina*

[0663] Comparison of known moss and microalgae $\Delta 4$ -desaturases revealed several conserved motifs including a HPGG (SEQ ID NO:52) motif within a cytochrome b_5 -like domain and three histidine box motifs that are presumed to be required for activity. Novel degenerate PCR primers PavD4Des-F3 (5'-AGCACGACGSSARCCACGGCG-3') (SEQ ID NO:53) and PavD4Des-R3 (5'-GTGGTGCA-CABCACGTGCT-3') (SEQ ID NO:54) corresponding to the conserved amino acid sequence of histidine box I and complementary to a nucleotide sequence encoding the amino acid sequence of histidine box II, respectively, were designed as to amplify the corresponding region of *P. salina* desaturase genes, particularly a $\Delta 4$ -desaturase gene. The use of degenerate PCR primers corresponding to histidine box I and histidine box II regions of $\Delta 4$ -desaturase has not been reported previously.

[0664] PCR amplification reactions using these primers were carried out using *P. salina* first strand cDNA as

template with cycling of 95° C., 5 min for 1 cycle, 94° C. 30 sec, 57° C. 30 sec, 72° C. 30 sec for 35 cycles, and 72° C. 5 min for 1 cycle. The PCR products were cloned into pGEM-T-easy (Promega) vectors, and nucleotide sequences were determined with an ABI3730 automatic sequencer using a reverse primer from the pGEM-Teasy vector. Among 14 clones sequenced, three clones showed homology to $\Delta 4$ -desaturase genes. Two of these three clones are truncated at one primer end. The nucleotide sequence of the cDNA insert of the third, clone 1803, is provided as SEQ ID NO:11.

[0665] The amino acid sequence encoded by SEQ ID NO:11 was used to search the NCBI protein sequence database using the BLASTX software. The results indicated that this sequence was homologous to known $\Delta 4$ -desaturases. The amino acid sequence of the *P. salina* gene fragment showed 65%, 49%, 46% and 46% identity to that of $\Delta 4$ -desaturases of *P. lutheri*, *Thraustochytrium* sp. ATCC21685, *Thraustochytrium aureum* and *Euglena gracilis* respectively.

[0666] Isolation of a Full-Length $\Delta 4$ -Desaturase Gene

[0667] The insert from clone 1803 was excised, and used as probe to isolate full-length cDNAs corresponding to the putative $\Delta 4$ -desaturase gene fragment. About 750,000 pfu of the *P. salina* cDNA library were screened at high stringency. The hybridization was performed at 60° C. overnight and washing was done with 2×SSC/0.1% SDS 30 min at 65° C. then with 0.2×SSC/0.1% SDS 30 min at 65° C. Eighteen hybridising clones were isolated and secondary screening with six clones was performed under the same hybridization conditions. Single plaques from secondary screening of these six clones were isolated. Plasmids from five single plaques were excised and the nucleotide sequences of the inserts determined with an ABI 3730 automatic sequencer with reverse and forward primers from the vector. Sequencing results showed that four clones each contained $\Delta 4$ -desaturase cDNA of approximately 1.7 kb in length, each with the same coding sequence and each apparently full-length. They differed slightly in the length of the 5' and 3' UTRs even though they contained identical protein coding regions. The cDNA sequence of the longest *P. salina* $\Delta 4$ -desaturase cDNA is provided as SEQ ID NO:13, and the encoded protein as SEQ ID NO:4.

[0668] The full-length cDNA was 1687 nucleotides long and had a coding region encoding 447 amino acids. The *Pavlova salina* $\Delta 4$ -desaturase showed all the conserved motifs typical of 'front-end desaturases' including the N-terminal cytochrome b_5 -like domain and three conserved histidine-rich motifs. Comparison of the nucleotide and amino acid sequences with other $\Delta 4$ -desaturase genes showed that the greatest extent of homology was for the *P. lutheri* $\Delta 4$ -desaturase (Accession No. AY332747), which was 69.4% identical in nucleotide sequence over the protein coding region, and 67.2% identical in amino acid sequence.

[0669] Demonstration of Enzyme Activity of *Pavlova salina* $\Delta 4$ -Desaturase Gene

[0670] A DNA fragment including the *Pavlova salina* $\Delta 4$ -desaturase cDNA coding region was excised as an EcoRI-SalI cDNA fragment and inserted into the pYES2 yeast expression vector using the EcoRI and XhoI sites. The resulted plasmid was transformed into yeast cells. The transformants were grown in YMM medium and the gene

induced by the addition of galactose, in the presence of added (exogenous) $\omega 6$ and $\omega 3$ fatty acids in order to demonstrate enzyme activity and the range of substrates that could be acted upon by the expressed gene. The fatty acids 22:5 $\omega 3$ (DPA, 11.0 mM), 20:4n-3 (ETA, 11.0 mM), 22:4 $\omega 6$ (DTAG, 1.0 mM) and 20:4 $\omega 6$ (ARA, 11.0 mM) were each added separately to the medium. After 72 hours incubation, the cells were harvested and fatty acid analysis carried out by capillary gas-liquid chromatography (GC) as described in Example 1. The data obtained are shown in Table 9.

TABLE 9

Yeast PUFA feeding showing activity of delta-4 desaturase gene.		
Fatty acid composition (% of total fatty acid)	Exogenous fatty acid added to growth medium	
	22:4 $\omega 6$	22:5 $\omega 3$
14:0	0.63	0.35
15:0	0.06	0.06
16:1 $\omega 7c$	43.45	40.52
16:1 $\omega 5$	0.20	0.13
16:0	18.06	15.42
17:1 $\omega 8$	0.08	0.09
17:0	0.08	—
18:1 $\omega 9$	26.73	30.07
18:1 $\omega 7$ (major) & 18:3 $\omega 3$	1.43	1.61
18:1 $\omega 5c$	0.02	tr
18:0	7.25	8.87
20:5 $\omega 3$	0.40	0.62
20:1 $\omega 9/\omega 11$	0.03	tr
20:0	0.08	0.09
22:5 $\omega 6$	0.03	0.00
22:6 $\omega 3$	—	0.04
22:4 $\omega 6$	0.97	—
22:5 $\omega 3$	0.00	1.66
22:0	0.06	0.06
24:1 $\omega 7$	0.31	0.37
24:0	0.12	0.04
Sum	100.00%	100.00%

[0671] This showed that the cloned gene encoded a $\Delta 4$ -desaturase which was able to desaturate both C22:4 $\omega 6$ (3.0% conversion to 22:5 $\omega 6$) and C22:5 $\omega 3$ (2.4% conversion to 22:6 $\omega 3$) at the $\Delta 4$ position. The enzyme did not show any $\Delta 5$ desaturation activity when the yeast transformants were fed C20:3 $\omega 6$ or C20:4 $\omega 3$.

Example 11

Expression of *P. salina* $\Delta 4$ -Desaturase Gene in Plant Cells and Production of DHA

[0672] To demonstrate activity of the $\Delta 4$ -desaturase gene in plant cells, the coding region may be expressed either separately to allow the conversion of DPA to DHA, or in the context of other LC-PUFA synthesis genes such as, for example, a $\Delta 5$ -elongase gene for the conversion of EPA to DHA. For expression as a separate gene, the $\Delta 4$ -desaturase coding region may be excised as a BamHI-SalI fragment and inserted between a seed-specific promoter and a polyadenylation/transcription termination sequence, such as, for example, in vector pGNAP (Lee et al., 1998), so that it is expressed under the control of the seed specific promoter. The expression cassette may then be inserted into a binary vector and introduced into plant cells. The plant material

used for the transformation may be either untransformed plants or transformed plants containing a construct which expressed the zebrafish $\Delta 5/\Delta 6$ -dual desaturase gene and *C. elegans* elongase gene each under the control of a seed specific promoter (Example 5). Transgenic *Arabidopsis* containing the latter, dual-gene construct had successfully produced EPA and DPA in seeds, and the combination with the $\Delta 4$ -desaturase gene would allow the conversion of the DPA to DHA in the plant cells, as demonstrated below.

[0673] To demonstrate co-expression of a $\Delta 5$ -elongase gene with the $\Delta 4$ -desaturase gene in recombinant cells, particularly plant cells, and allow the production of DHA, the $\Delta 4$ -desaturase and the $\Delta 5$ -elongase genes from *P. salina* (Example 8) were combined in a binary vector as follows. Both coding regions were placed under the control of seed-specific (napin) promoters and nos3' terminators, and the binary vector construct had a kanamycin resistance gene as a selectable marker for selection in plant cells. The coding region of the $\Delta 5$ -elongase gene was excised from its cDNA clone as a PstI-SacII fragment and inserted into an intermediate plasmid (pXZP143) between the promoter and terminator, resulting in plasmid pXZP144. The coding region of the $\Delta 4$ -desaturase gene was excised from its cDNA clone as a BamHI-SalI fragment and inserted into plasmid pXZP143 between the promoter and nos 3' transcription terminator, resulting in plasmid pXZP150. These two expression cassettes were combined in one vector by inserting the HindIII-ApaI fragment from pXZP144 (containing promoter-Elo1-nos3') between the StuI and ApaI sites of pXZP150, resulting in plasmid pXZP191. The HindIII-StuI fragment from pXZP191 containing both expression cassettes was then cloned into the binary vector pXZP330, a derivative of pBI121, resulting in plant expression vector pXZP355. This vector is shown schematically in FIG. 7.

[0674] Plant Transformation

[0675] The $\Delta 5$ -elongase and the $\Delta 4$ -desaturase genes on pXZP355 were introduced by the *Agrobacterium*-mediated floral dip transformation method into the *Arabidopsis* plants designated DO11 (Example 5) which were already transgenic for the zebrafish $\Delta 5/\Delta 6$ bifunctional desaturase and the *C. elegans* $\Delta 5/\Delta 6$ bifunctional elongase genes. Since those transgenes were linked to a hygromycin resistance gene as a selectable marker gene, the secondary transformation with pXZP355 used a kanamycin resistance selection, thus distinguishing the two sets of transgenes. Five transgenic plants are obtained, designated "DW" plants. Since the DO11 plants were segregating for the zebrafish $\Delta 5/\Delta 6$ bifunctional desaturase and the *C. elegans* $\Delta 5/\Delta 6$ bifunctional elongase genes, some of the transformed plants were expected to be heterozygous for these genes, while others were expected to be homozygous. Seed (T2 seed) of the five transformed plants were analysed and shown to contain up to at least 0.1% DPA and up to at least 0.5% DHA in the seed oils. Data are presented for two lines in Table 10. Analysis, by mass spectrometry (GC-MS), of the fatty acids in the peaks identified as EPA and DHA from the GC analysis proved that they were indeed EPA and DHA (FIG. 8).

[0676] The fatty acid analysis of the T2 seedoil demonstrated that significant conversion of EPA to DHA had occurred in the DW2 and DW5 lines, having 0.2% and 0.5% DHA, respectively. Examination of the enzyme efficiencies in plant DW5 containing the higher level of DHA showed

that 17% of the EPA produced in its seed was elongated to DPA by the *P. salina* $\Delta 5$ -elongase, and greater than 80% of this DPA was converted to DHA by the *P. salina* $\Delta 4$ -desaturase. Since the $\Delta 5$ -elongase and $\Delta 4$ -desaturase genes were segregating in the T2 seed, the fatty acid composition data represented an average of pooled null, heterozygous and homozygous genotypes for these genes. It is expected that levels of DHA in progeny lines of DW5 will be greater in seed that is uniformly homozygous for these genes.

TABLE 10

Fatty acid composition (% of total fatty acids) of seed oils from <i>Arabidopsis thaliana</i> (ecotype Columbia) and derivatives carrying EPA and DHA gene constructs - EPA, DPA and DHA synthesis in transgenic seed.					
Fatty acid	Columbia	DO11 + DHA construct			
		Wild type	DW2	DW5	
		Total	Total	TAG	PL
Usual fatty acids					
16:0	7.2	6.7	6.1	5.5	12.5
18:0	2.9	3.8	4.4	4.3	4.5
18:1 Δ^9	20.0	20.6	16.6	18.9	13.7
18:2 $\Delta^{9,12}$ (LA)	27.5	26.0	25.9	25.5	33.1
18:3 $\Delta^{9,12,15}$ (ALA)	15.1	13.2	15.0	13.6	15.1
20:0	2.2	2.1	1.8	1.9	0.6
20:1 Δ^{11}	19.8	14.8	10.5	10.5	3.2
20:1 Δ^{13}	2.2	3.0	4.2	4.8	1.4
20:2 $\Delta^{11,14}$	0.1	1.7	3.5	3.8	3.7
22:1 Δ^{13}	1.5	1.4	1.0	0.3	0.4
Other minor	1.5	2.9	2.7	2.4	3.8
Total	100.0	96.0	91.7	91.5	92.0
New $\omega 6$-PUFA					
18:3 $\Delta^{6,9,12}$ (GLA)	0	0.2	0.4	0.4	0.2
20:3 $\Delta^{8,11,14}$	0	0.8	1.5	1.5	1.7
20:4 $\Delta^{5,8,11,14}$ (ARA)	0	0.4	1.0	1.1	1.2
22:4 $\Delta^{7,10,13,16}$	0	0	0	0	0.2
22:5 $\Delta^{4,7,10,13,16}$	0	0	0.1	0.1	0.1
Total	0	1.4	3.0	3.1	3.4
New $\omega 3$-PUFA					
18:4 $\Delta^{6,9,12,15}$ (SDA)	0	0.7	1.5	1.6	0.5
20:4 $\Delta^{8,11,14,17}$	0	0.5	0.8	0.7	0.9
20:5 $\Delta^{5,8,11,14,17}$ (EPA)	0	1.1	2.4	2.5	2.3
22:5 $\Delta^{7,10,13,16,19}$ (DPA)	0	0.1	0.1	0.2	0.7
22:6 $\Delta^{4,7,10,13,16,19}$ (DHA)	0	0.2	0.5	0.4	0.2
Total	0	2.6	5.3	5.4	4.6
Total fatty acids	100.0	100.0	100.0	100.0	100.0
Total MUFA ^a	41.3	36.8	28.1	29.7	17.3
Total C ₁₈ -PUFA ^b	42.6	39.2	40.9	39.1	48.2
Total new PUFA ^c	0	4.0	8.3	8.5	8.0

^aTotal of 18:1 Δ^9 and derived LC-MUFA (=18:1 Δ^9 + 20:1 Δ^{11} + 22:1 Δ^{13})

^b18:2 + 18:3

^cTotal of all new $\omega 6$ and $\omega 3$ -PUFA

[0677] Germination of 50 T2 seed from each of DW2 and DW5 on hygromycin-containing medium showed that the DW5 T1 plant was homozygous (50/50) for the $\Delta 5/\Delta 6$ bifunctional desaturase and $\Delta 5/\Delta 6$ bifunctional elongase genes, while the DW2 seed segregating in a 3:1 ration (resistant: susceptible) for these genes and DW2 was therefore heterozygous. This was consistent with the higher levels of EPA observed in DW5 seed compared to DW2 seed, and explained the increased level of DHA produced in the seed

homozygous for these transgenes. This further demonstrated the desirability of seed that are homozygous for the trait.

[0678] We also noted the consequences of LC-PUFA synthesis on the overall fatty acid profile in these seed. Although we observed accumulation of new $\omega 6$ and $\omega 3$ PUFA (i.e. products of $\Delta 6$ -desaturation) at levels of greater than 8% in DW5 seed, these seed had levels of the precursor fatty acids LA and ALA that were almost the same as in the wild-type seed. Rather than depleting LA and ALA, the levels of monounsaturated fatty acid C18: 1 Δ^9 and its elongated derivatives (20:1 Δ^{11} and 22:1 Δ^{13}) were significantly reduced. Thus it appeared that conversion of C₁₈-PUFA to LC-PUFA resulted in increased conversion of 18:1 to LA and ALA, and a corresponding reduction in 18:1 available for elongation.

[0679] The plant expression vector pXZP355 containing the $\Delta 4$ -desaturase and the $\Delta 5$ -elongase genes was also used to introduce the genes into plants of the homozygous line DO11-5, and 20 transgenic T1 plants were obtained. The levels of DHA and DPA in T2 seed from these plants were similar to those observed in seed from DW5. Reductions in the levels of the monounsaturated fatty acids were also observed in these seed.

[0680] Fractionation of the total seed lipids of DW5 seed revealed them to be comprised of 89% TAG and 11% polar lipids (largely made up phospholipids). Furthermore, fatty acid analysis of the TAG fraction from DW5 seed showed that the newly synthesised EPA and DHA were being incorporated into the seed oil and that the proportion of EPA and DHA in the fatty acid composition of the total seed lipid essentially reflected that of the TAG fraction (Table 10).

Example 12

Isolation of Homologous Genes from Other Sources

[0681] Homologs of the desaturase and elongase genes such as the *P. salina* genes described herein may be readily detected in other microalgae or other sources by hybridization to labelled probes derived from the genes, particularly to parts or all of the coding regions, for example by Southern blot hybridization or dot-blot hybridisation methods. The homologous genes may be isolated from genomic or cDNA libraries of such organisms, or by PCR amplification using primers corresponding to conserved regions. Similarly, homologs of vertebrate desaturases with high affinity for Acyl-CoA and/or freshwater fish bifunctional desaturases can be isolated by similar means using probes to the zebrafish $\Delta 5/\Delta 6$ desaturase.

[0682] Dot Blot Hybridisations

[0683] Genomic DNA from six microalgae species was isolated using a DNAeasy kit (Qiagen) using the suppliers instructions, and used in dot blot hybridization analyses for identification of homologous genes involved in LC-PUFA synthesis in these species. This also allowed evaluation of the sequence divergence of such genes compared to those isolated from *Pavlova salina*. The species of microalga examined in this analysis were from the genera *Melosira*, *Rhodomonas*, *Heterosigma*, *Nannochloropsis*, *Heterocapsa* and *Tetraselmis*. They were identified according to Hasle, G. R. & Syvertsen, E. E. 1996 Dinoflagellates. In: Tomas, C. R.

(ed.) Identifying Marine Phytoplankton. Academic Press, San Diego, Calif. pp 531-532. These microalga were included in the analysis on the basis of the presence of EPA, DHA, or both when cultured in vitro (Example 2).

[0684] Genomic DNA (approximately 100 μ g) isolated from each of the microalga was spotted onto strips of Hybond N+ membrane (Amersham). After air drying, each membrane strip was placed on a layer of 3MM filter paper saturated with 0.4 M NaOH for 20 min, for denaturation of the DNA, and then rinsed briefly in 2xSSC solution. The membrane strips were air dried and the DNA cross linked to the membranes under UV light. Probes labeled with ^{32} P nucleotides and consisting of the coding regions without the untranslated regions of a number of *Pavlova*-derived genes, including the $\Delta 8$, $\Delta 5$ and $\Delta 4$ desaturases and $\Delta 9$ and $\Delta 5$ elongases, were prepared and hybridized to each membrane strip/DNA dot blot. The membranes were hybridized with each probe overnight in a buffer containing 50 mM Tris-HCl, pH7.5, 1M NaCl, 50% formamide, 10x Denhardt's solution, 10% dextran sulfate, 1% SDS, 0.1% sodium pyrophosphate, and 0.1 mg/ml herring sperm DNA, at 42° C., then washed three times in a solution containing 2xSSC, 0.5% SDS at 50° C. for 15 min each (low stringency wash in this experiment) or for a high stringency wash in 0.2x SSC, 0.5% SDS at 65° C. for 20 minutes each.

[0685] It is well understood that the stringency of the washing conditions employed in DNA blot/hybridizations can reveal useful information regarding the sequence relatedness of genes. Thus hybridizations maintained when subjected to a high stringency wash indicate a high level of sequence relatedness (e.g. 80% or greater nucleotide identity over at least 100-200 nucleotides), while hybridizations maintained only during low stringency washes indicate a relatively lower degree of DNA conservation between genes (e.g. 60% or greater nucleotide identity over at least 200 nucleotides).

[0686] The hybridized dot blots were exposed to BioMax X-ray film (Kodak), and the autoradiograms are shown in **FIG. 9**. The autoradiograms reveal the presence of homologs to the *P. salina* LC-PUFA genes in these species, and moreover reveal a range of homologies based on the different levels of hybridization seen under the high and low stringency conditions. It appeared that some of the microalgal species examined have LC-PUFA genes that may differ substantially from the genes in *P. salina*, while others are more related in sequence. For example, genes from *Tetraselmis* sp appeared to be highly similar to the $\Delta 4$ - and $\Delta 5$ -desaturases and the $\Delta 5$ elongase from *Pavlova salina* on the basis of the strength of hybridizations. In contrast, all of the LC-PUFA genes identified in *Melosira* sp appeared to have lower degrees of similarity to the *P. salina* genes.

[0687] Isolation of an LC-PUFA Elongase Gene from *Heterocapsa* sp.

[0688] *Heterocapsa* spp. such as *Heterocapsa niei* in the CSIRO collection (Example 2) are dinoflagellates that were identified as producers of LC-PUFA including EPA and DHA. To exemplify the isolation of LC-PUFA synthesis genes from these dinoflagellates, DNA was purified from cells of a *Heterocapsa niei* strain originally isolated in Port Hacking, NSW, Australia in 1977. DNA was isolated using a DNAeasy kit (Qiagen) using the suppliers instructions. Based on published multiple amino acid sequence align-

ments for fatty acid elongases (Qi et al., 2002; Parker-Barnes et al., 2000), the consensus amino acid blocks FLHXYH (SEQ ID NO:48) and MYXYFY (SEQ ID NO:49) were identified and corresponding degenerate primers encoding these sequences 5'-CAGGATCCTTYYTNCAT-NNNTAYCA-3' (SEQ ID NO:50) (sense) or complementary to these sequences 5'-GATCTAGARAARTARTANNRTACAT-3' (SEQ ID NO:51) (antisense) were synthesised. PCR amplification reactions were carried out in reaction volumes of 20 μ L with 20 pmol of each primer, 200 ng of *Heterocapsa* sp. genomic DNA and Hotstar Taq DNA polymerase (Qiagen) with buffer and nucleotide components as specified by the supplier. The reactions were cycled as follows: 1 cycle of 95° C. for 15 minutes, 5 cycles of 95° C., min, 38° C., 1 min, 72° C., 1 min, 35 cycles of 95° C., 35 sec, 52° C., 30 sec, 72° C., 1 min, 1 cycle of 72° C., 10 min. Fragments of approximately 350 bp were generated and ligated into pGEM-Teasy for sequence analysis.

[0689] Of eight clones isolated, two identical clones had nucleotide and encoded amino acid sequences with similarity to regions of known elongases. These were designated Het350Elo, and the nucleotide and amino acid sequences are provided as SEQ ID NO:79 and SEQ ID NO:80 respectively. BLAST analysis and the presence of an in-frame stop codon suggested the presence of an intron between approximate positions 33 and 211.

[0690] The best matches to the amino acid sequence were animal elongase sequences, see for example Meyer et al. (2004), indicating that the isolated *Heterocapsa* gene sequence was probably involved in elongation of C18 and C20 fatty acid substrates.

[0691] Full-length clones of the elongase can readily be isolated by screening a *Heterocapsa* cDNA library or by 5'- and 3' RACE techniques, well known in the art.

[0692] Construction of *Melosira* sp. cDNA Library and EST Sequencing

[0693] mRNA, for the construction of a cDNA library, was isolated from *Melosira* sp. cells using the following method. 2 g (wet weight) of *Melosira* sp. cells were powdered using a mortar and pestle in liquid nitrogen and sprinkled slowly into a beaker containing 22 ml of extraction buffer that was being stirred constantly. To this, 5% insoluble polyvinylpyrrolidone, 90 mM 2-mercaptoethanol, and 10 mM dithiothreitol were added and the mixture stirred for a further 10 minutes prior to being transferred to a Corex™ tube. 18.4 ml of 3M ammonium acetate was added and mixed well. The sample was then centrifuged at 6000xg for 20 minutes at 4° C. The supernatant was transferred to a new tube and nucleic acid precipitated by the addition of 0.1 volume of 3M NaAc (pH 5.2) and 0.5 volume of cold isopropanol. After 1 hour incubation at -20° C., the sample was centrifuged at 6000xg for 30 minutes in a swing-out rotor. The pellet was resuspended in 1 ml of water and extracted with phenol/chloroform. The aqueous layer was transferred to a new tube and nucleic acids were precipitated once again by the addition of 0.1 volume 3M NaAc (pH 5.2) and 2.5 volume of ice cold ethanol. The pellet was resuspended in water, the concentration of nucleic acid determined and then mRNA was isolated using the Oligotex mRNA system (Qiagen).

[0694] First strand cDNA was synthesised using an oligo(dT) linker-primer supplied with the ZAP-cDNA synthe-

sis kit (Stratagene—cat # 200400) and the reverse transcriptase SuperscriptIII (Invitrogen). Double stranded cDNA was ligated to EcoRI adaptors and from this a library was constructed using the ZAP-cDNA synthesis kit as described in the accompanying instruction manual (Stratagene—cat # 200400). A primary library of 1.4×10^6 plaque forming units (pfu) was obtained. The average insert size of cDNA inserts in the library was 0.9 kilobases based on 47 random plaques and the percentage of recombinants in the library was 99%.

[0695] Single pass nucleotide sequencing of 8684 expressed sequence tags (ESTs) was performed with SK primer (5'-CGCTCTAGAACTAGTGGATC-3') (SEQ ID NO:87) using the ABI BigDye system. Sequences of 6750 ESTs were longer than 400 nucleotides, showing the inserts were at least this size. ESTs showing homology to several fatty acid desaturases and one PUFA elongase were identified by BlastX analysis.

[0696] The amino acid sequence (partial) (SEQ ID NO:88) encoded by the cDNA clone Mm301461 showed 75% identity to *Thalassiosira pseudonana* fatty acid elongase 1 (Accession No. AY591337). The nucleotide sequence of EST clone Mm301461 is provided as SEQ ID NO:89. The high degree of identity to a known elongase makes it highly likely that Mm301461 encodes a *Melosira* fatty acid elongase. RACE techniques can readily be utilized to isolate the full-length clone encoding the elongase.

Example 13

Isolation of FAE-Like Elongase Gene Fragment from *P. salina*

[0697] Random cDNA clones from the *P. salina* cDNA library were sequenced by an EST approach. In an initial round of sequencing, 73 clones were sequenced. One clone, designated 11.B1, was identified as encoding a protein (partial sequence) having sequence similarity with known beta keto-acyl synthase-like fatty acid elongases, based on BLASTX analysis. The nucleotide sequence of 11.B 1 from the 3' end is provided as (SEQ ID NO:55).

[0698] These plant elongases are different to the ELO class elongase in that they are known to be involved in the elongation of C16 to C18 fatty acids and also the elongation of very-long-chain saturated and monounsaturated fatty acids. Clone 11.B 1, represents the first non-higher plant gene in this class isolated.

Example 14

Isolation of a Gene Encoding a $\Delta 5$ -Desaturase from *P. salina*

[0699] Isolation of a Gene Fragment from a $\Delta 5$ -Desaturase Gene from *P. salina*

[0700] In order to isolate a $\Delta 5$ -desaturase gene from *P. salina*, oligonucleotides were designed for a conserved region of desaturases. The oligonucleotides designated d5A and d5B shown below were made corresponding to a short DNA sequence from a $\Delta 5$ -desaturase gene from *Pavlova lutheri*. Oligo d5A: 5'-TGGGTTGAGTACTCGGCCAAC-CACACGACCAACTGCGCGCCCTCGTGGTGGT GCGACTGGTGGATGTCTTACCTCAAC-TACCAGATCGAGCATCATCTGT-3' (nucleotides 115-214

of International patent application published as WO03078639-A2, FIG. 4a) (SEQ ID NO:56) and oligo d5B:

5'-ATAGTGCAGCCCGTGCTTCTCGAAGAGCGCCTTGACGCGCGGCGCGATCGTC GGGTG-GCGGAATTGCGGCATGGACGCGGAACA-GATGATGCTCGATCTGG-3'. (corresponding to the complement of nucleotides 195-294 of WO03078639-A2, FIG. 4a) (SEQ ID NO:57). These oligonucleotides were annealed and extended in a PCR reaction. The PCR product from was inserted into pGEM-T Easy vector and the nucleotide sequence confirmed.

[0701] The cloned fragment was labelled and used as a hybridization probe for screening of a *Pavlova salina* cDNA library under moderately high stringency conditions, hybridizing at 55° C. overnight with an SSC hybridization solution and washing the blots at 60° C. with 2×SSC/0.1% SDS three times each for 10 minutes. From screening of about 500,000 plaques, 60 plaques were isolated which gave at least a weak hybridization signal. Among 13 clones that were sequenced, one clone designated p1918 contained a partial-length cDNA encoding an amino acid sequence with homology to known $\Delta 5$ -desaturase genes. For example, the amino acid sequence was 53% identical to amino acid residues 210-430 from the C-terminal region of a *Thraustochytrium* $\Delta 5$ -desaturase gene (Accession No. AF489588).

[0702] Isolation of a Full-Length $\Delta 5$ -Desaturase Gene

[0703] The partial-length sequence in p1918 was used to design a pair of sequence specific primers, which were then used in PCR screening of the 60 isolated plaques mentioned above. Nineteen of the 60 were positive, having the same or similar cDNA sequence. One of the clones that showed a strong hybridization signal using the partial-length sequence as a probe was used to determine the full-length sequence provided as SEQ ID NO:58, and the amino acid sequence (425 amino acids in length) encoded thereby is provided as SEQ ID NO:60.

[0704] The amino acid sequence was used to search the NCBI protein sequence database using the BLASTX software. The results indicated that this sequence was homologous to known $\Delta 5$ -desaturases. The amino acid sequence of the *P. salina* protein showed 81% identity to a *P. lutheri* sequence of undefined activity in WO03/078639-A2, and 50% identity to a $\Delta 5$ -desaturase from *Thraustochytrium* (Accession No. AF489588). The *Pavlova salina* $\Delta 5$ -desaturase showed all the conserved motifs typical of 'front-end desaturases' including the N-terminal cytochrome b_5 -like domain and three conserved histidine-rich motifs.

[0705] Co-expression of $\Delta 9$ Elongase, $\Delta 8$ -Desaturase and $\Delta 5$ -Desaturase Genes in Transformed Cells

[0706] Co-expression of the $\Delta 5$ -desaturase gene together with the $\Delta 9$ elongase gene (Elo2, Example 7) and the $\Delta 8$ -desaturase gene (Example 6) was achieved in cells as follows. The plant expression vector pXZP354 containing the three genes, each from *P. salina*, and each expressed from the seed specific napin promoter was constructed. The *P. salina* $\Delta 8$ -desaturase coding region from the cDNA clone (above) was first inserted as a BamHI-NcoI fragment into pXZP143 between the seed specific napin promoter and Nos terminator, resulting in plasmid pXZP146. The *P. salina* $\Delta 9$ -elongase gene was likewise inserted, as a PstI-XhoI fragment from its cDNA clone, into pXZP143 resulting in

plasmid pXZP143-Elo2. The *P. salina* $\Delta 5$ -desaturase gene was also inserted, as a PstI-BssHII fragment from its cDNA clone, into pXZP143, resulting in plasmid pXZP147. Then, the HindIII-ApaI fragment containing the $\Delta 9$ -elongase expression cassette from pXZP143-Elo2 was inserted into pXZP146 downstream of the $\Delta 8$ -desaturase expression cassette, resulting in plasmid pXZP148. The HindIII-ApaI fragment containing the $\Delta 5$ -desaturase expression cassette from pXZP147 was inserted into pXZP148 downstream of $\Delta 8$ -desaturase and $\Delta 9$ -elongase expression cassettes, resulting in plasmid pXZP149. Then, as a final step, the HindIII-ApaI fragment containing the three genes from pXZP149 was inserted into a derivative of the binary vector pART27, containing a hygromycin resistance gene selection marker, resulting in plant expression plasmid pXZP354.

[0707] Plasmid pXZP354 was introduced into *Arabidopsis* by the *Agrobacterium*-mediated floral dip method, either in the simultaneous presence or the absence of expression plasmid pXZP355 (Example 11) containing the *P. salina* $\Delta 5$ -elongase and $\Delta 4$ -desaturase genes. Co-transformation of the vectors could be achieved since they contained different selectable marker genes. In the latter case, the transgenic plants (designated "DR" plants) were selected using hygromycin as selective agent, while in the former case, the plants ("DU" plants) were selected with both hygromycin and kanamycin.

[0708] Twenty-one DR plants (T1 plants) were obtained. Fatty acid analysis of seedoil from T2 seed from ten of these plants showed the presence of low levels of 20:2 ω (EDA), 20:3 ω 6 (DGLA) and 20:4 ω 6 (ARA), including up to 0.4% ARA. Fatty acid analysis of seedoil from T2 seed from seven DU plants showed similar levels of these fatty acids. From the relative ratios of these fatty acids, it was concluded that the $\Delta 5$ -desaturase and $\Delta 8$ -desaturase genes were functioning efficiently in seed transformed with pXZP354 but that the activity of the $\Delta 9$ elongase gene was suboptimal. It is likely that shortening of the coding region at the N-terminal end, to initiate translation at amino acid position 11 or 29 of SEQ ID NO:3 (Example 9) (see SEQ ID NO's 85 and 86) will improve the level of activity of the $\Delta 9$ elongase gene. Expression of one or two of the genes from seed-specific promoters other than the napin promoter, so they are not all expressed from the napin promoter, is also expected to improve the expression level of the $\Delta 9$ elongase gene.

Example 15

Isolation of a Gene Encoding a $\Delta 6$ -Desaturase from *Echium plantagineum*

[0709] Some plant species such as evening primrose (*Oenothera biennis*), common borage (*Borago officinalis*), blackcurrant (*Ribes nigrum*), and some *Echium* species belonging to the Boraginaceae family contain the $\omega 6$ - and $\omega 3$ -desaturated C18 fatty acids, γ -linolenic acid (18:3 ω 6, GLA) and stearidonic acid (18:4 ω 3, SDA) in their leaf lipids and seed TAGs (Guil-Guerrero et al., 2000). GLA and SDA are recognized as beneficial fatty acids in human nutrition. The first step in the synthesis of LC-PUFA is a $\Delta 6$ -desaturation. GLA is synthesized by a $\Delta 6$ -desaturase that introduces a double bond into the $\Delta 6$ -position of LA. The same enzyme is also able to introduce a double bond into $\Delta 6$ -position of ALA, producing SDA. $\Delta 6$ -Desaturase genes have

been cloned from members of the Boraginaceae, like borage (Sayanova et al., 1997) and two *Echium* species (Garcia-Maroto et al., 2002).

[0710] *Echium plantagineum* is a winter annual native to Mediterranean Europe and North Africa. Its seed oil is unusual in that it has a unique ratio of $\omega 3$ and $\omega 6$ fatty acids and contains high amounts of GLA (9.2%) and SDA (12.9%) (Guil-Guerrero et al., 2000), suggesting the presence of $\Delta 6$ -desaturase activity involved in desaturation of both $\omega 3$ and $\omega 6$ fatty acids in seeds of this plant.

[0711] Cloning of *E. platangineum* EplD6Des Gene

[0712] Degenerate primers with built-in XbaI or SacI restriction sites corresponding to N- and C-termini amino acid sequences MANAIKKY (SEQ ID NO: 61) and EALNTHG (SEQ ID NO: 62) of known *Echium pitardii* and *Echium gentianoides* (Garcia-Maroto et al., 2002) $\Delta 6$ -desaturases were used for RT-PCR amplification of $\Delta 6$ -desaturase sequences from *E. platangineum* using a proofreading DNA polymerase Pfu Turbo® (Stratagene). The 1.35 kb PCR amplification product was inserted into pBluescript SK(+) at the XbaI and SacI sites to generate plasmid pXZP106. The nucleotide sequence of the insert was determined (SEQ ID NO:63). It comprised an open reading frame encoding a polypeptide of 438 amino acid residues (SEQ ID NO:64) which had a high degree of homology with other reported $\Delta 6$ - and $\Delta 8$ -desaturases from *E. gentianoides* (SEQ ID NO:65), *E. pitardii* (SEQ ID NO:66), *Borago officinalis* (SEQ ID NO:67 and 68), *Helianthus annuus* (SEQ ID NO:69) and *Arabidopsis thaliana* (SEQ ID NO:70 and SEQ ID NO:71) (FIG. 10). It has a cytochrome b₅ domain at the N-terminus, including the HPGG (SEQ ID NO:72) motif in the heme-binding region, as reported for other $\Delta 6$ - and $\Delta 8$ -desaturases (Sayanova et al. 1997; Napier et al. 1999). In addition, the *E. plantagineum* $\Delta 6$ desaturase contains three conserved histidine boxes, including the third histidine box containing the signature QXXHH (SEQ ID NO:73) motif present in majority of the 'front-end' desaturases (FIG. 10) (Napier et al., 1999). Cluster analysis including representative members of $\Delta 6$ and $\Delta 8$ desaturases showed a clear grouping of the cloned gene with other $\Delta 6$ desaturases especially those from *Echium* species.

[0713] Heterologous Expression of *E. plantagineum* $\Delta 6$ -Desaturase Gene in Yeast

[0714] Expression experiments in yeast were carried out to confirm that the cloned *E. platangineum* gene encoded a $\Delta 6$ -desaturase enzyme. The gene fragment was inserted as an XbaI-SacI fragment into the SmaI-SacI sites of the yeast expression vector pSOS (Stratagene) containing the constitutive ADH1 promoter, resulting in plasmid pXZP271. This was transformed into yeast strain S288C α by a heat shock method and transformant colonies selected by plating on minimal media plates. For the analysis of enzyme activity, 2 mL yeast clonal cultures were grown to an O.D.₆₀₀ of 1.0 in yeast minimal medium in the presence of 0.1% NP-40 at 30° C. with shaking. Precursor free-fatty acids, either linoleic or linolenic acid as 25 mM stocks in ethanol, were added so that the final concentration of fatty acid was 0.5 mM. The cultures were transferred to 20° C. and grown for 2-3 days with shaking. Yeast cells were harvested by repeated centrifugation and washing first with 0.1% NP-40, then 0.05% NP-40 and finally with water. Fatty acids were extracted and analyzed. The peak identities of fatty acids were confirmed by GC-MS.

[0715] The transgenic yeast cells expressing the *Echium* EplD6Des were able to convert LA and ALA to GLA and SDA, respectively. Around 2.9% of LA was converted to GLA and 2.3% of ALA was converted to SDA, confirming the $\Delta 6$ -desaturase activity encoded by the cloned gene.

[0716] Functional Expression of *E. platangineum* $\Delta 6$ -desaturase gene in transgenic tobacco In order to demonstrate that the EplD6Des gene could confer the synthesis of $\Delta 6$ desaturated fatty acids in transgenic plants, the gene was expressed in tobacco plants. To do this, the gene fragment was excised from pXZP106 as an XbaI-SacI fragment and cloned into the plant expression vector pBI121 (Clonotech) at the XbaI and SacI sites under the control of a constitutive 35S CaMV promoter, to generate plant expression plasmid pXZP341. This was introduced into *Agrobacterium tumefaciens* AGL1, and used for transformation of tobacco W38 plant tissue, by selection with kanamycin.

[0717] Northern blot hybridization analysis of transformed plants was carried out to detect expression of the introduced gene, and total fatty acids present in leaf lipids of wild-type tobacco W38 and transformed tobacco plants were analysed as described above. Untransformed plants contained appreciable amounts of LA (21% of total fatty acids) and ALA (37% of total fatty acids) in leaf lipids. As expected, neither GLA nor SDA, products of $\Delta 6$ -desaturation, were detected in the untransformed leaf. Furthermore, transgenic tobacco plants transformed with the pBI121 vector had similar leaf fatty acid composition to the untransformed W38 plants. In contrast, leaves of transgenic tobacco plants expressing the EplD6Des gene showed the presence of additional peaks with retention times corresponding to GLA and SDA. The identity of the GLA and SDA peaks were confirmed by GC-MS. Notably, leaf fatty acids of plants expressing the EplD6Des gene consistently contained approximately a two-fold higher concentration of GLA than SDA even when the total $\Delta 6$ -desaturated fatty acids amounted up to 30% of total fatty acids in their leaf lipids (Table 11).

TABLE 11

Fatty acid composition in lipid from transgenic tobacco leaves (%).								
Plant	16:0	18:0	18:1	18:2	GLA	18:3	SDA	Total $\Delta 6$ -desaturate products
W38	21.78	5.50	2.44	21.21	—	37.62	—	—
ET27-1	20.33	1.98	1.25	10.23	10.22	41.10	6.35	16.57
ET27-2	18.03	1.79	1.58	14.42	1.47	53.85	0.48	1.95
ET27-4	19.87	1.90	1.35	7.60	20.68	29.38	9.38	30.07
ET27-5	15.43	2.38	3.24	11.00	0.84	49.60	0.51	1.35
ET27-6	19.85	2.05	1.35	11.12	4.54	50.45	2.19	6.73
ET27-8	19.87	2.86	2.55	11.71	17.02	27.76	7.76	24.78

TABLE 11-continued

Fatty acid composition in lipid from transgenic tobacco leaves (%).								
Plant	16:0	18:0	18:1	18:2	GLA	18:3	SDA	Total $\Delta 6$ -desaturate products
ET27-11	17.78	3.40	2.24	12.62	1.11	51.56	0.21	1.32
ET27-12	16.84	2.16	1.75	13.49	2.71	50.80	1.15	3.86

[0718] Northern analysis of multiple independent transgenic tobacco lines showed variable levels of the EplD6Des transcript which generally correlated with the levels of $\Delta 6$ -desaturated products synthesized in the plants. For example, transgenic plant ET27-2 which contained low levels of the EplD6Des transcript synthesised only 1.95% of its total leaf lipids as $\Delta 6$ -desaturated fatty acids. On the other hand, transgenic plant ET27-4 contained significantly higher levels of EplD6Des transcript and also had a much higher proportion (30%) of $\Delta 6$ -desaturated fatty acids in its leaf lipids.

[0719] Analysis of the individual tobacco plants showed that, without exception, GLA was present at a higher concentration than SDA even though a higher concentration of ALA than LA was present in untransformed plants. In contrast, expression of EplD6Des in yeast had resulted in approximately equivalent levels of conversion of LA into GLA and ALA into SDA. *Echium plantagineum* seeds, on the other hand, contain higher levels of SDA than GLA. EplD6Des probably carries out its desaturation in vivo in *Echium plantagineum* seeds on LA and ALA esterified to phosphatidyl choline (PC) (Jones and Harwood 1980). In the tobacco leaf assay, the enzyme is most likely desaturating LA and ALA esterified to the chloroplast lipid monogalactosyldiacylglycerol (MGDG) (Browse and Slack, 1981). In the yeast assay, free fatty acid precursors LA and ALA added to the medium most likely enter the acyl-CoA pool and are available to be acted upon by EplD6Des in this form.

[0720] Functional Expression of *E. platangineum* $\Delta 6$ -desaturase Gene in Transgenic Seed

[0721] To show seed-specific expression of the *Echium* $\Delta 6$ -desaturase gene, the coding region was inserted into the seed-specific expression cassette as follows. An NcoI-SacI fragment including the $\Delta 6$ -desaturase coding region was inserted into pXZP6, a pBluescriptSK derivative containing a Nos terminator, resulting in plasmid pXZP157. The SmaI-ApaI fragment containing the coding region and terminator EplD6Des-NosT was cloned into pWVec8-Fp1 downstream of the Fp1 promoter, resulting in plasmid pXZP345. The plasmid pXZP345 was used for transforming wild type *Arabidopsis* plants, ecotype Columbia, and transgenic plants selected by hygromycin B selection. The transgenic plants transformed with this gene were designated "DP" plants.

[0722] Fatty acid composition analysis of the seed oil from T2 seed from eleven T1 plants transformed with the construct showed the presence of GLA and SDA in all of the lines, with levels of $\Delta 6$ -desaturation products reaching to at least 11% (Table 12). This demonstrated the efficient $\Delta 6$ -desaturation of LA and ALA in the seed.

TABLE 12

Fatty acid composition in transgenic <i>Arabidopsis</i> seeds expressing $\Delta 6$ -desaturase from <i>Echium</i> .										
Plant	Fatty acid (%)								Total $\Delta 6$ -desaturation	
	16:0	18:0	18:1 ^{Δ9}	18:2 ^{Δ9,12} (LA)	18:3 ^{Δ6,9,12} (GLA)	18:3 ^{Δ9,12,15} (ALA)	18:4 ^{Δ6,9,12,15} (SDA)	20:0	20:1	products (%)
<u>Columbia</u>										
DP-2	8.0	2.8	22.9	27.3	2.5	11.3	0.7	1.6	15.8	3.2
DP-3	7.8	2.7	20.6	25.9	3.0	12.1	0.8	1.7	17.8	3.8
DP-4	7.8	2.8	20.4	28.5	1.2	13.7	0.4	1.7	16.1	1.5
DP-5	8.2	3.2	17.4	29.3	1.2	14.2	0.3	2.1	15.6	1.6
DP-7	8.2	2.9	18.4	26.7	5.0	12.7	1.4	1.7	15.2	6.4
DP-11	9.0	3.5	17.8	28.4	3.0	13.4	0.9	2.1	13.9	3.8
DP-12	8.6	3.0	18.9	27.8	3.3	12.6	1.0	1.8	15.4	4.3
DP-13	8.7	2.9	14.4	27.3	8.5	13.7	2.6	1.7	12.4	11.1
DP-14	9.3	2.9	14.2	32.3	2.1	15.4	0.7	1.8	12.8	2.8
DP-15	8.2	2.9	17.8	30.1	0.3	15.3	0.2	1.9	15.5	0.5
DP-16	8.0	2.8	19.5	29.2	2.7	13.1	0.8	1.7	14.2	3.5

Example 16

Mutagenesis of *E. platangineum* EplD6Des Gene

[0723] To determine whether variability could be introduced into the $\Delta 6$ -desaturase gene and yet retain desaturase activity, the *E. platangineum* $\Delta 6$ -desaturase cDNA was randomly mutated by PCR using Taq polymerase and EPD6DesF1 and EPD6DesR1 primers in the presence of dITP as described by Zhou and Christie (1997). The PCR products were cloned as XbaI-SacI fragments in pBluescript SK(+) at XbaI and SacI sites, and sequences of randomly selected clones determined. Random variants with amino acid residue changes were chosen to clone as XbaI-SacI fragments into pBI121 and the enzyme activities of proteins expressed from these variants characterized in transgenic tobacco leaves as described above for the wild-type gene.

[0724] FIG. 11A represents the activity of the EplD6Des sequence variants when expressed in tobacco plants. The variants could be divided into two broad classes in terms of their ability to carry out $\Delta 6$ -desaturation. Mutations represented as empty diamonds showed substantial reductions in the $\Delta 6$ -desaturation activity while mutations denoted as solid diamonds had little or no effect on the activity of the encoded $\Delta 6$ -desaturase enzyme. FIG. 1B represents the quantitative effect that a selection of mutations in the EplD6Des gene had on the $\Delta 6$ -desaturase activity. An L14P mutation in the cytochrome b_5 domain and an S301P mutation between histidine box II and histidine box III of EplD6Des caused substantial reductions in their $\Delta 6$ -desaturase activities, resulting in a 3- to 5-fold reduction in total $\Delta 6$ -desaturated fatty acids when compared to the wild-type enzyme in W38 plants. Surprisingly, significant activity was retained for each. In contrast, most of the variants examined, as exemplified by the S205N mutation, had no effect on the $\Delta 6$ -desaturation activity of EplD6Des gene.

Example 17

Comparison of acyl-CoA and acyl-PC Substrate Dependent Desaturases for Production of LC-PUFA in Cells

[0725] As described above, the synthesis of LC-PUFA such as EPA and DHA in cells by the conventional $\Delta 6$

desaturation pathway requires the sequential action of PUFA desaturases and elongases, shown schematically in FIG. 12 part A. This conventional pathway operates in algae, mosses, fungi, diatoms, nematodes and some freshwater fish (Sayanova and Napier, 2004). The PUFA desaturases from algae, fungi, mosses and worms are selective for desaturation of fatty acids esterified to the sn-2 position of phosphatidylcholine (PC) while the PUFA elongases act on fatty acids in the form of acyl-CoA substrates represented in the acyl-CoA pool in the endoplasmic reticulum (ER), which is physiologically separated from the PC component of the ER. Therefore, sequentially desaturation and elongation reactions on a fatty acid substrate requires that the fatty acid is transferred between the acyl-PC and acyl-CoA pools in the ER. This requires acyltransferases that are able to accommodate LC-PUFA substrates. This "substrate switching" requirement may account for the low efficiency observed in earlier reported attempts to re-constitute LC-PUFA biosynthesis (Beaudoin et al., 2000, Domergue et al., 2003a). The alternative $\Delta 8$ desaturation pathway (FIG. 12 part B) suffers from the same disadvantage of requiring "substrate switching".

[0726] As described in Example 5, the strategy of using a vertebrate desaturase which was able to desaturate acyl-CoA substrates, provided relatively efficient production of LC-PUFA in plant cells including in seed. In Example 5, the combination of a $\Delta 5/\Delta 6$ desaturase from zebra fish with a $\Delta 6$ elongase from *C. elegans* had the advantage that both the desaturase and the elongase enzymes had activity on acyl-CoA substrates in the acyl-CoA pool. This may explain why this strategy was more efficient in the synthesis of LC-PUFA. To provide a direct comparison of the relative efficiencies of using an acyl-CoA substrate-dependent desaturase compared to an acyl-PC substrate-dependent desaturase, we conducted the following experiment. This compared the use of the *Echium* $\Delta 6$ desaturase (Example 15) and the *P. salina* $\Delta 5$ desaturase (Example 14), both of which are thought to use acyl-PC substrates, with the zebrafish $\Delta 6/\Delta 5$ desaturase which uses an acyl-CoA substrate (Example 5).

[0727] A construct was prepared containing two acyl-PC dependent desaturases, namely the *Echium* $\Delta 6$ desaturase and *P. salina* $\Delta 5$ desaturase, in combination with the *C.*

elegans $\Delta 6$ elongase. The *Echium* $\Delta 6$ desaturase gene on an NcoI-SacI fragment was inserted into pXZP143 (Example 15) resulting in pXZP192. The *C. elegans* $\Delta 6$ elongase gene (Fp1-CeElo-NosT expression cassette) on the HindIII-ApaI fragment of pCeloPWVec8 (Example 5) was inserted into the StuI-ApaI sites of pXZP147 (Example 14) to make pXZP193. The HindIII-ApaI fragment of pXZP193 containing both genes (Fp1-PsD5Des-NosT and Fp1-CeElo-NosT) was inserted into the ApaI-StuI sites of pXZP192, resulting in plasmid pXZP194 containing the three expression cassettes. The XbaI-ApaI fragment from pXZP194 was inserted in a pWvec8 derivative, resulting in pXZP357.

[0728] The plasmid pXZP357 was used to transform plants of wild-type *Arabidopsis* ecotype Columbia by *Agrobacterium*-mediated floral dip method, and six transgenic plants were obtained after hygromycin B (20 mg/L) selection. The transgenic T1 plants were designated "DT" plants. The hygromycin resistant transformed plants were transferred into soil and self-fertilised. The T2 seed were harvested and the seed fatty acid composition of two lines, DT1 and DT2, was analysed. The seed fatty acids of DT1 and DT2 contained low levels of 18:3 ω 6 and 18:4 ω 4 (0.9 and 0.8% of GLA, 0.3% and 0.1% of SDA, respectively, Table 13). In addition, both DT1 and DT2 seed also contained 0.3% and 0.1% of the 20:4 ω 6 (ARA). However, there was no apparent synthesis of the ω 3 fatty acid EPA in either of the T2 seed lines, which probably reflected the greater desaturation ability of the *Echium* $\Delta 6$ desaturase on the ω 6 substrate LA compared to the ω 3 substrate ALA (Example 15).

TABLE 13

Fatty acid composition of seed-oil from T2 seed of DT1 and DT2. Fatty acid values are % of total fatty acids.			
	Control	DT1	DT2
Fatty acid			
16:0	7.2	6.5	6.5
18:0	2.9	3.6	3.3
18:1 ω 9	20.0	23.2	22.3
18:2 ω 6	27.5	23.6	24.4
18:3 ω 3	15.1	15.4	16.1
20:0	2.2	2.0	1.9
20:1 ω 9/ ω 11	19.9	19.4	19.5
20:1 ω 7	2.2	3.4	3.0
20:2 ω 6	0.1	0.0	0.0
22:1 ω 7	0.0	0.0	0.0
Other minor	2.8	1.5	1.9
Total	100.0	98.6	98.9
New ω 6-PUFA			
18:3 ω 6	0.0	0.9	0.8
20:3 ω 6	0.0	0.0	0.0
20:4 ω 6	0.0	0.3	0.1
Total	0.0	1.2	0.9
New ω 3-PUFA			
18:4 ω 3	0.0	0.3	0.2
20:4 ω 3	0.0	0.0	0.0
20:5 ω 3	0.0	0.0	0.0
Total	0.0	0.3	0.2
Total fatty acids	100.0	100.0	100.0

[0729] This data was in clear contrast to Example 5, above, where expression of the acyl-CoA dependent desaturase from zebrafish in combination with a $\Delta 6$ elongase resulted in the production of at least 1.1% ARA and 2.3% EPA in T2 seed fatty acids. Thus it would appear that acyl-PC dependent desaturases were less effective than acyl-15 CoA dependent desaturases in driving the synthesis of LC-PUFA in plants cells.

Example 18

Expression of LC-PUFA Genes in *Synechococcus*

[0730] *Synechococcus* spp. (Bacteria; Cyanobacteria; Chroococcales; *Synechococcus* species for example *Synechococcus elongatus*, also known as *Synechocystis* spp.) are unicellular, photosynthetic, marine or freshwater bacteria in the order cyanobacteria that utilize chlorophyll a in the light-harvesting apparatus. The species include important primary producers in the marine environment. One distinct biochemical feature of *Synechococcus* is the presence of phycoerythrin, an orange fluorescent compound that can be detected at an excitation wavelength of 540 nm, and which can be used to identify *Synechococcus*. Members of the marine *synechococcus* group are closely related at the level of 16s rRNA. They are obligately marine and have elevated growth requirements for Na⁺, Cl⁻, Mg²⁺, and Ca²⁺, but can be grown readily in both natural and artificial seawater liquid media as well as on plates (Waterbury et al. 1988). Since they have a rapid heterotrophic or autotrophic growth rate, contain fatty acid precursors such as LA and ALA, and are relatively simple to transform, they are suitable for functional studies involving LC-PUFA synthesis genes, or for production of LC-PUFA in fermenter type production systems. Strains such as *Synechococcus* sp. strain WH8102, PCC7002 (7002, marine), or PCC7942 (freshwater) can be grown easily and are amenable to biochemical and genetic manipulation (Carr, N. G., and N. H. Mann. 1994. The oceanic cyanobacterial picoplankton, p. 27-48. In D. A. Bryant (ed.), the Molecular biology of cyanobacteria. Kluwer Academic publishers, Boston). For example, *Synechococcus* has been used as a heterologous expression system for desaturases (Domergue 2003b).

[0731] Wildtype *Synechococcus* 7002 Fatty Acid Profile and Growth Rates

[0732] To show that cyanobacterium *Synechococcus* 7002 was a suitable host for the transformation of fatty acid synthesis genes and that this expression system could be used to rapidly test the functions and specificities of fatty acid synthesis genes, the growth of the wildtype strain 7002 was first analysed at 22° C., 25° C. and 30° C. and the resultant fatty acid profiles analysed by gas chromatography for growth at 22° C. and 30° C. (Table 14).

TABLE 14

<i>Synechococcus</i> 7002 wildtype fatty acid profiles at 22° C. and 30° C. growth temperatures (% of total fatty acid).									
Temp	Myristic	Palmitic	Palmitoleic	Stearic	Oleic	18:1 iso	Linoleic	GLA	Linolenic
22° C.	0.79	42.5	10.6	0.92	8.4	1.5	7.5	0.54	27.1
30° C.	0.76	47.1	10.9	0.67	17.0	0.34	20.4		2.9

[0733] Growth at 30° C. was much more rapid than at 22° C., with intermediate rates at 25° C. (FIG. 13). The cells were found to contain both linoleic (LA, 18:2 ω 6) and linolenic (ALA, 18:3 ω 3) acids which could be used as precursors for LC-PUFA synthesis. Although some of the preferred precursor ALA was produced at the 30° C., higher levels were obtained at 22° C. Tests were also carried out to determine whether cells could be grown at 30° C., followed by reducing the incubation temperature to 22° C. after sufficient biomass had been achieved, to see if this would result in a shift to higher production of linolenic acid (FIG. 14). In this experiment, levels of ALA obtained were greater than 5%. In further experiments, 25° C. was used as the preferred temperature for strain 7002, providing adequate growth rates and suitable precursor fatty acid profile.

[0734] Transformation Strategy

[0735] Both replicative plasmid vectors and non-replicative homologous recombination vectors have been used previously to transform various cyanobacterial species, including *Synechococcus* 7002 (Williams and Szalay, 1983; Ikeda et al., 2002; Akiyama et al., 1998a). The recombination vectors may be preferred in certain applications, and have been used to inactivate a gene, rather than create an expression strain.

[0736] A recombination vector was constructed that was suitable for introduction of one or more fatty acid synthesis genes into the chromosome of *Synechococcus* strains such as strain 7002. This vector contained the *Synechococcus* 7002 *sul2* gene in a pBluescript plasmid backbone, which provided an ampicillin gene as a selectable marker and allowed bacterial replication in species such as *E. coli*. The vector was engineered to contain a *plac* promoter from *E. coli* fused to a downstream multiple cloning site, with the two elements inserted approximately in the centre of the *sul2* gene. The *sul2* gene in *Synechococcus* encodes a low affinity sulfate which is not essential under normal growth conditions. Any gene other than *sul2*, preferably a non-essential gene, could have been chosen for incorporation in the recombination vector.

[0737] The *sul2* gene was amplified from *Synechococcus* 7002 genomic DNA using gene-specific primers, based on the near-identical sequence in strain PCC6803 (Genbank Accession No. NC_000911, nucleotides 2902831 to 2904501) and inserted into the vector pGEM-T. The *plac* promoter from pBluescript was amplified using the primers 5'-gctacgccccgggtagctctcgaggctggcgcaacgcaataatgtga-3' (SEQ ID NO:81) (sense) and 5'-cacaggaaacagcttgacatgat-taccggcaattgtacggcgcgctacggatctctcgctcgagctcgccccgggtagct-3' (SEQ ID NO:82) (antisense), which also introduced a number of restriction sites at the ends of the promoter sequence. The amplified fragment was then digested with *Sma*I and ligated to the large *Pvu*II fragment of pBluescript

including the beta-lactamase gene. This intermediate vector was then digested with *Eco*RV and *Sac*I and ligated to the *Hpa*I to *Sac*I fragment (designated *sul2b*) of the *sul2* gene. The resultant plasmid was digested with *Bam*HI, treated with DNA polymerase I (Klenow fragment) to fill in the ends, and ligated to the *Sma*I to *Hpa*I fragment (designated *sul2a*) of the *sul2* gene. Excess restriction sites were then removed from this vector by digestion with *Sac*I and *Spe*I, blunting the ends with T4 DNA polymerase, and religation. Finally, a multiple cloning site was introduced downstream of the *plac* promoter by digesting the vector with *Cla*I and *Not*II, and ligating in a *Cla*I to *Not*II fragment from pBlue-script, generating the recombination vector which was designated pJRP3.2.

[0738] Various genes related to LC-PUFA synthesis were adapted by PCR methods to include flanking restriction sites as well as ribosome binding site (RBS) sequences that were suitable for expression in the prokaryote, *Synechococcus*. For example, the *Echium plantagineum* $\Delta 6$ -desaturase (Example 15) was amplified with the primers 5'-AGCA-CATCGATGAAGGAGATATACCCatggctaataatgcaatcaagaa-3' (SEQ ID NO:83) (sense) and 5'-ACGATGCGGCCGCT-CAACCATGAGTATTAAGAGCTT-3' (SEQ ID NO:84) (antisense).

[0739] The amplified product was digested with *Cla*I and *Not*II and cloned into the *Cla*I to *Not*II sites of pJRP3.2. A selectable marker gene comprising a chloramphenicol acetyl transferase coding region (CAT) (*catB3* gene, Accession No. AAC53634) downstream of a *pbsA* promoter (*psbA-CAT*) was inserted into the *Xho*I site of pJRP3.2, producing the vector pJRP3.3. The selectable marker gene was inserted within the *sulB* gene to enable easy selection for homologous recombination events after introduction of the recombination vector into *Synechococcus*.

[0740] Transformation of *Synechococcus* 7002 was achieved by mixing vector DNA with cells during the exponential phase of growth, during which DNA uptake occurred, as follows. Approximately 1 μ g of the recombination vector DNA resuspended in 100 μ L of 10 mM Tris-HCl was added to 900 μ L of mid-log phase cells growing in BG-11 broth. The cells were incubated for 90 min at 30° C. and light intensity of 20 μ mol photons.m⁻².s⁻¹. 250 μ L aliquots were then added to 2 mL BG-11 broth, mixed with 2 mL molten agar (1.5%) and poured onto BG-11 agar plates containing 50 μ g/mL chloramphenicol (Cm) for selection of recombinant cells. The plates were incubated for 10-14 days at the same temperature/light conditions before the Cm-resistant colonies were clearly visible. These colonies were then re-streaked several times onto fresh BG-11/Cm50 plates. After several rounds of restreaking on selective plates, liquid medium was inoculated with individual colonies and the cultures incubated at 25° C.

[0741] *Synechococcus* 7002 cells containing the *Echium* $\Delta 6$ -desaturase gene inserted into the *su1B* gene via the recombination vector and expressed from the *plac* promoter are shown to produce GLA (18:3 $\Delta 6,9,12$) and SDA (18:4, $\Delta 6,9,12,15$) from endogenous linoleic acid (LA) and linolenic acid (ALA), respectively, as substrates.

[0742] Episomal vectors can also be used in *Synechococcus* rather than the integrative/recombinational vectors described above. *Synechococcus* species have native plasmids that have been adapted for use in transformation, for example pAQ-EXI, where a fragment of the native plasmid pAQI (Accession No. NC_005025) was fused with an *E. coli* plasmid to form a shuttle vector with both *E. coli* and *Synechococcus* origins of replication (Ikeda et al., 2002; Akiyama et al., 1998b).

[0743] It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects as illustrative and not restrictive.

[0744] All publications discussed above are incorporated herein in their entirety.

[0745] Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is solely for the purpose of providing a context for the present invention. It is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present invention as it existed before the priority date of each claim of this application.

REFERENCES

- [0746] Abbadi, A. et al., (2001) Eur. J. Lipid. Sci. Technol. 103:106-113.
- [0747] Abbadi, A., et al. (2004) Plant Cell 16:2734-2748.
- [0748] Abbott et al., (1998) Science 282:2012-2018.
- [0749] Agaba, M. et al., (2004) Marine Biotechnol (NY) 6:251-261.
- [0750] Akiyama, H. et al. (1998a) DNA Res. 5:327-334.
- [0751] Akiyama, H. et al. (1998b) DNA Res. 5:127-129.
- [0752] Baumlein, H. et al., (1991) Mol. Gen. Genet. 225:459-467.
- [0753] Baumlein, H. et al., (1992) Plant J. 2:233-239.
- [0754] Beaudoin, F. et al., (2000) Proc. Natl. Acad. Sci U.S.A. 97:6421-6426.
- [0755] Berberich, T. et al., (1998) Plant Mol. Biol. 36:297-306.
- [0756] Bolch, C. J. et al., (1999a) J. Phycology 35:339-355.
- [0757] Bolch, C. J. et al., (1999b) J. Phycology 35:356-367.
- [0758] Broun, P. et al., (1998) Plant J. 13:201-210.
- [0759] Brown, M. R. et al., (1997) Aquaculture 151:315-331.
- [0760] Browse, J. A. and Slack, C. R. (1981) FEBS Letters 131:111-114.
- [0761] Chinain, M. et al., (1997) J. Phycology 33:36-43.
- [0762] Cho, H. P. et al., (1999a) J. Biol. Chem. 274:471-477.
- [0763] Cho, H. P. et al., (1999b) J. Biol. Chem. 274:37335-37339.
- [0764] Clough, S. J. and Bent, A. F. (1998) Plant J. 16:735-43.
- [0765] Coleman, A. W. (1977) Am. J. Bot. 64:361-368.
- [0766] Domergue, F. et al., (2002) Eur. J. Biochem. 269:4105-4113.
- [0767] Domergue, F. et al., (2003a) J. Biol. Chem. 278:35115-35126.
- [0768] Domergue, F. et al., (2003b) Plant Physiol. 131:1648-1660.
- [0769] Drexler, H. et al., (2003) J. Plant Physiol. 160:779-802.
- [0770] Dunstan, G. A. et al., (1994) Phytochemistry 35:155-161.
- [0771] Gallagher, J. C. (1980) J. Phycology 16:464-474.
- [0772] Garcia-Maroto, F. et al., (2002) Lipids 37:417-426.
- [0773] Girke, T. et al., (1998) Plant J. 15:39-48.
- [0774] Guil-Guerrero, J. L. et al., (2000). Phytochemistry 53:451-456.
- [0775] Haseloff, J. and Gerlach, W. L. (1988) Nature 334:585-591.
- [0776] Hastings, N. et al., (2001) Proc. Natl. Acad. Sci. U.S.A. 98:14304-14309.
- [0777] Hong, H. et al., (2002) Lipids 37:863-868.
- [0778] Hong, H. et al., (2002a) Lipids 37:863-868.
- [0779] Horiguchi, G. et al., (1998) Plant Cell Physiol. 39:540-544.
- [0780] Huang, Y. S. et al., (1999) Lipids 34:649-659.
- [0781] Ikeda, K. et al. (2002) World J. Microbiol. Biotech. 18:55-56.
- [0782] Inagaki, K. et al., (2002) Biosci Biotechnol Biochem 66:613-621.
- [0783] Jones, A. V. and Harwood, J. L. (1980) Biochem J. 190:851-854.
- [0784] Kajikawa, M. et al., (2004) Plant Mol Biol 54:335-52.
- [0785] Knutzon, D. S. et al., (1998) J. Biol. Chem. 273:29360-6.
- [0786] Lee, M. et al., (1998) Science 280:915-918.

- [0787] Leonard, A. E. et al., (2000) *Biochem. J.* 347:719-724.
- [0788] Leonard, A. E. et al., (2000b) *Biochem. J.* 350:765-770.
- [0789] Leonard, A. E. et al., (2002) *Lipids* 37:733-740.
- [0790] Lo, J. et al., (2003) *Genome Res.* 13:455-466.
- [0791] Mansour, M. P. et al., (1999a) *J. Phycol.* 35:710-720.
- [0792] Medlin, L. K. et al., (1996) *J. Marine Systems* 9:13-31.
- [0793] Metz, J. G. et al., (2001) *Science* 293:290-293.
- [0794] Meyer, A. et al., (2003) *Biochemistry* 42:9779-9788.
- [0795] Meyer, A. et al., (2004) *Lipid Res* 45:1899-1909.
- [0796] Michaelson, L. V. et al., (1998a) *J. Biol. Chem.* 273:19055-19059.
- [0797] Michaelson, L. V. et al., (1998b) *FEBS Lett.* 439:215-218.
- [0798] Mitchell, A. G. and Martin, C. E. (1995) *J. Biol. Chem.* 270:29766-29772.
- [0799] Morita, N. et al., (2000) *Biochem. Soc. Trans.* 28:872-879.
- [0800] Napier, J. A. et al., (1998) *Biochem J.* 330:611-614.
- [0801] Napier, J. A. et al., (1999) *Trends in Plant Sci* 4:2-4.
- [0802] Napier, J. A. et al., (1999) *Curr. Op. Plant Biol.* 2:123-127.
- [0803] Needleman, S. B. and Wunsch, C. D. (1970) *J. Mol. Biol.* 48:443-453.
- [0804] Parker-Barnes, J. F. et al., (2000) *Proc Natl Acad Sci USA* 97:8284-8289.
- [0805] Pereira, S. L. et al., (2004) *Biochem. J.* 378:665-671.
- [0806] Perriman, R. et al., (1992) *Gene* 113:157-163.
- [0807] Qi, B. et al., (2002) *FEBS Lett.* 510:159-165.
- [0808] Qiu, X. et al., (2001) *J. Biol. Chem.* 276:31561-31566.
- [0809] Reddy, A. S. et al., (1993) *Plant Mol. Biol.* 22:293-300.
- [0810] Saito, T. et al., (2000) *Eur. J. Biochem.* 267:1813-1818.
- [0811] Sakuradani, E. et al., (1999) *Gene* 238:445-453.
- [0812] Sayanova, O. V. et al., (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94:4211-4216.
- [0813] Sayanova, O. V. et al., (1999) *Plant Physiol.* 121:641-646.
- [0814] Sayanova, O. V. et al., (2003) *FEBS Lett.* 542:100-104.
- [0815] Sayanova, O. V. and Napier, J. A. (2004) *Phytochemistry* 65:147-158.
- [0816] Shippy, R. et al., (1999) *Mol. Biotech.* 12:117-129.
- [0817] Simopoulos, A. P. (2000) *Poultry Science* 79:961-970.
- [0818] Singh, S. et al., (2001) *Planta* 212:872-879.
- [0819] Smith, N. A. et al., (2000) *Nature* 407:319-320.
- [0820] Sperling, P. et al., (2000) *Eur. J. Biochem.* 267:3801-3811.
- [0821] Sperling, P. and Heinz, E. (2001) *Eur. J. Lipid Sci. Technol* 103:158-180.
- [0822] Sprecher, H. et al., (1995) *J. Lipid Res.* 36:2471-2477.
- [0823] Spsychalla, P. J. et al., (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94:1142-1147.
- [0824] Stalberg, K. et al., (1993) *Plant. Mol. Biol.* 23:671-683.
- [0825] Takeyama, H. et al., (1997) *Microbiology* 143:2725-2731.
- [0826] Tanaka, M. et al., (1999) *Biotechnol. Lett.* 21:939-945.
- [0827] Tonon, T. et al., (2003) *FEBS Lett.* 553:440-444.
- [0828] Trautwein, E. A. (2001) *Eur. J. Lipid Sci. Technol.* 103:45-55.
- [0829] Tvrdik, P. (2000) *J. Cell Biol.* 149:707-718.
- [0830] Valvekens, D. et al., (1988) *Proc. Natl. Acad. Sci. USA.* 85:5536-5540.
- [0831] Volkman, J. K. et al., (1989) *J. Exp. Mar. Biol. E. coli.* 128:219-240.
- [0832] Wallis, J. G. and Browse, J. (1999) *Arch. Biochem. Biophys.* 365:307-316.
- [0833] Wang, M. B. et al., (1997) *J. Gen. Breed.* 51:325-334.
- [0834] Waterbury, J. B. et al. (1988) *Methods Enzymol.* 167:100-105.
- [0835] Waterhouse, P. M. et al., (1998) *Proc. Natl. Acad. Sci. USA* 95:13959-13964.
- [0836] Watts, J. L. and Browse, J. (1999b) *Arch Biochem Biophys* 362:175-182.
- [0837] Williams, J. G. and Szalay, A. A. (1983) *Gene* 24:37-51.
- [0838] Whitney, H. M. et al., (2003). *Planta* 217:983-992.
- [0839] Yazawa, K. (1996) *Lipids* 31:S297-S300.
- [0840] Yu, R. et al., (2000) *Lipids* 35:1061-1064.
- [0841] Zank, T. K. et al., (2000) *Plant J.* 31:255-268.
- [0842] Zank, T. K. et al., (2002) *Plant J.* 31:255-268.
- [0843] Zhang, Q. et al., (2004) *FEBS Lett.* 556:81-85.
- [0844] Zhou, X. R. and Christie, P. J. (1997) *J. Bacteriol* 179:5835-5842.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 89

<210> SEQ ID NO 1

<211> LENGTH: 427

<212> TYPE: PRT

<213> ORGANISM: Pavlova salina

<400> SEQUENCE: 1

```

Met Gly Arg Gly Gly Asp Ser Ser Gly Gln Ala His Pro Ala Ala Glu
 1          5          10          15
Leu Ala Val Pro Ser Asp Arg Ala Glu Val Ser Asn Ala Asp Ser Lys
          20          25          30
Ala Leu His Ile Val Leu Tyr Gly Lys Arg Val Asp Val Thr Lys Phe
          35          40          45
Gln Arg Thr His Pro Gly Gly Ser Lys Val Phe Arg Ile Phe Gln Asp
          50          55          60
Arg Asp Ala Thr Glu Gln Phe Glu Ser Tyr His Ser Lys Arg Ala Ile
          65          70          75          80
Lys Met Met Glu Gly Met Leu Lys Lys Ser Glu Asp Ala Pro Ala Asp
          85          90          95
Thr Pro Leu Pro Ser Gln Ser Pro Met Gly Lys Asp Phe Lys Ala Met
          100          105          110
Ile Glu Arg His Val Ala Ala Gly Tyr Tyr Asp Pro Cys Pro Leu Asp
          115          120          125
Glu Leu Phe Lys Leu Ser Leu Val Leu Leu Pro Thr Phe Ala Gly Met
          130          135          140
Tyr Met Leu Lys Ala Gly Val Gly Ser Pro Leu Cys Gly Ala Leu Met
          145          150          155          160
Val Ser Phe Gly Trp Tyr Leu Asp Gly Trp Leu Ala His Asp Tyr Leu
          165          170          175
His His Ser Val Phe Lys Gly Ser Val Ala Arg Thr Val Gly Trp Asn
          180          185          190
Asn Ala Ala Gly Tyr Phe Leu Gly Phe Val Gln Gly Tyr Ala Val Glu
          195          200          205
Trp Trp Arg Ala Arg His Asn Thr His His Val Cys Thr Asn Glu Asp
          210          215          220
Gly Ser Asp Pro Asp Ile Lys Thr Ala Pro Leu Leu Ile Tyr Val Arg
          225          230          235          240
Asn Lys Pro Ser Ile Ala Lys Arg Leu Asn Ala Phe Gln Arg Tyr Gln
          245          250          255
Gln Tyr Tyr Tyr Val Pro Val Met Ala Ile Leu Asp Leu Tyr Trp Arg
          260          265          270
Leu Glu Ser Ile Ala Tyr Val Ala Met Arg Leu Pro Lys Met Leu Pro
          275          280          285
Gln Ala Leu Ala Leu Val Ala His Tyr Ala Ile Val Ala Trp Val Phe
          290          295          300
Ala Gly Asn Tyr His Leu Leu Pro Leu Val Thr Val Leu Arg Gly Phe
          305          310          315          320
Gly Thr Gly Ile Thr Val Phe Ala Thr His Tyr Gly Glu Asp Ile Leu
          325          330          335
Asp Ala Asp Gln Val Arg His Met Thr Leu Val Glu Gln Thr Ala Leu

```

-continued

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

<210> SEQ ID NO 2

<211> LENGTH: 302

<212> TYPE: PRT

<213> ORGANISM: Pavlova salina

<400> SEQUENCE: 2

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

-continued

Ala Tyr Lys Gly Ala Ala Lys Gly Ala Ala Ala Ser Ser Asn Gly Ala
275 280 285

Ala Ala Pro Ser Gly Ala Lys Pro Lys Ser Ile Lys Ala Asn
290 295 300

<210> SEQ ID NO 3
<211> LENGTH: 304
<212> TYPE: PRT
<213> ORGANISM: Pavlova salina

<400> SEQUENCE: 3

Met Gly Pro Leu Ser Thr Leu Leu Ala Trp Met Pro Thr Trp Gly Glu
1 5 10 15

Phe Val Ala Gly Leu Thr Tyr Val Glu Arg Gln Gln Met Ser Glu Glu
20 25 30

Leu Val Arg Ala Asn Lys Leu Pro Leu Ser Leu Ile Pro Glu Val Asp
35 40 45

Phe Phe Thr Ile Ala Ser Val Tyr Val Gly Asp His Trp Arg Ile Pro
50 55 60

Phe Thr Ala Ile Ser Ala Tyr Leu Val Leu Ile Thr Leu Gly Pro Gln
65 70 75 80

Leu Met Ala Arg Arg Pro Pro Leu Pro Ile Asn Thr Leu Ala Cys Leu
85 90 95

Trp Asn Phe Ala Leu Ser Leu Phe Ser Phe Val Gly Met Ile Val Thr
100 105 110

Trp Thr Thr Ile Gly Glu Arg Leu Trp Lys Asn Gly Ile Glu Asp Thr
115 120 125

Val Cys Gly His Pro Ile Phe Met Gly Tyr Gly Trp Ile Gly Tyr Val
130 135 140

Met Leu Ala Phe Ile Trp Ser Lys Leu Phe Glu Leu Ile Asp Thr Val
145 150 155 160

Phe Leu Val Ala Lys Lys Ala Asp Val Ile Phe Leu His Trp Tyr His
165 170 175

His Val Thr Val Leu Leu Tyr Cys Trp His Ser Tyr Ala Val Arg Ile
180 185 190

Pro Ser Gly Ile Trp Phe Ala Ala Met Asn Tyr Phe Val His Ala Ile
195 200 205

Met Tyr Ala Tyr Phe Gly Met Thr Gln Ile Gly Pro Arg Gln Arg Lys
210 215 220

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

Met Gly Val Gly Leu Ala Val Asn Gly Leu Ile Ile Arg Tyr Pro Ser
245 250 255

Ile Gly His His Cys His Ser Asn Lys Thr Asn Thr Ile Leu Ser Trp
260 265 270

Ile Met Tyr Ala Ser Tyr Phe Val Leu Phe Ala Ala Leu Tyr Val Lys
275 280 285

Asn Tyr Ile Phe Ser Lys Leu Lys Ser Pro Lys Arg Lys Lys Val Glu
290 295 300

<210> SEQ ID NO 4
<211> LENGTH: 447
<212> TYPE: PRT
<213> ORGANISM: Pavlova salina

-continued

<400> SEQUENCE: 4

```

Met Pro Pro Ser Ala Ala Lys Gln Met Gly Ala Ser Thr Gly Val His
1          5          10          15
Ala Gly Val Thr Asp Ser Ser Ala Phe Thr Arg Lys Asp Val Ala Asp
20          25          30
Arg Pro Asp Leu Thr Ile Val Gly Asp Ser Val Tyr Asp Ala Lys Ala
35          40          45
Phe Arg Ser Glu His Pro Gly Gly Ala His Phe Val Ser Leu Phe Gly
50          55          60
Gly Arg Asp Ala Thr Glu Ala Phe Met Glu Tyr His Arg Arg Ala Trp
65          70          75          80
Pro Lys Ser Arg Met Ser Arg Phe His Val Gly Ser Leu Ala Ser Thr
85          90          95
Glu Glu Pro Val Ala Ala Asp Glu Gly Tyr Leu Gln Leu Cys Ala Arg
100         105         110
Ile Ala Lys Met Val Pro Ser Val Ser Ser Gly Phe Ala Pro Ala Ser
115         120         125
Tyr Trp Val Lys Ala Gly Leu Ile Leu Gly Ser Ala Ile Ala Leu Glu
130         135         140
Ala Tyr Met Leu Tyr Ala Gly Lys Arg Leu Leu Pro Ser Ile Val Leu
145         150         155         160
Gly Trp Leu Phe Ala Leu Ile Gly Leu Asn Ile Gln His Asp Ala Asn
165         170         175
His Gly Ala Leu Ser Lys Ser Ala Ser Val Asn Leu Ala Leu Gly Leu
180         185         190
Cys Gln Asp Trp Ile Gly Gly Ser Met Ile Leu Trp Leu Gln Glu His
195         200         205
Val Val Met His His Leu His Thr Asn Asp Val Asp Lys Asp Pro Asp
210         215         220
Gln Lys Ala His Gly Ala Leu Arg Leu Lys Pro Thr Asp Ala Trp Ser
225         230         235         240
Pro Met His Trp Leu Gln His Leu Tyr Leu Leu Pro Gly Glu Thr Met
245         250         255
Tyr Ala Phe Lys Leu Leu Phe Leu Asp Ile Ser Glu Leu Val Met Trp
260         265         270
Arg Trp Glu Gly Glu Pro Ile Ser Lys Leu Ala Gly Tyr Leu Phe Met
275         280         285
Pro Ser Leu Leu Leu Lys Leu Thr Phe Trp Ala Arg Phe Val Ala Leu
290         295         300
Pro Leu Tyr Leu Ala Pro Ser Val His Thr Ala Val Cys Ile Ala Ala
305         310         315         320
Thr Val Met Thr Gly Ser Phe Tyr Leu Ala Phe Phe Phe Ile Ser
325         330         335
His Asn Phe Glu Gly Val Ala Ser Val Gly Pro Asp Gly Ser Ile Thr
340         345         350
Ser Met Thr Arg Gly Ala Ser Phe Leu Lys Arg Gln Ala Glu Thr Ser
355         360         365
Ser Asn Val Gly Gly Pro Leu Leu Ala Thr Leu Asn Gly Gly Leu Asn
370         375         380
Tyr Gln Ile Glu His His Leu Phe Pro Arg Val His His Gly Phe Tyr

```

-continued

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

<210> SEQ ID NO 5
 <211> LENGTH: 1235
 <212> TYPE: DNA
 <213> ORGANISM: Pavlova salina
 <400> SEQUENCE: 5

atgggacgcg gcggagacag cagtgggcag gcgcaccccg cgccggagct ggccgtccc	60
agcgaccgcg cggaggtgag caacgctgac agcaaagcgc tgcacatcgt gctgtatggc	120
aagcgcgtgg atgtgaccaa gttccaacgc acgcaccccg gtggtagcaa ggtcttcccg	180
atcttccagg acccgcatgc gacggagcag ttccagtcct accactcgaa gcgcgcgac	240
aagatgatgg agggcatgct caagaagtct gaggatgctc ccgcccacac gcccttgccc	300
tcccagtcac cgatggggaa ggacttcaag gcgatgatcg agcggcacgt tgcagcgggt	360
tactacgacg catgcccgct cgatgagctg ttcaagtcga gcctcgtgct cctcccagac	420
tttgcgggca tgtacatgct caaggcgggc gtcggctccc cgctctcgcg cgccctcatg	480
gtgagctttg gctggtacct cgatggctgg ctgcgcacg actatctgca ccaactccgtc	540
ttcaaggggt ccgtcgcacg caccgtcggg tggacaacag cgccgggcta cttcctcggc	600
ttcgtgcagg ggtatgcggt cgagtgggtg gcgcgcggcg ataacacgca ccacgtgtgc	660
accaatgagg acggctcggg ccccgacatc aaaacggcgc cgctgctcat atacgtgcgc	720
aacaagccga gcacgcgcaa gcgcctgaac gccttcacg gctaccagca gtactactat	780
gtgcgggtga tggcaatcct cgacctgtac tggcggctcg agtcgatcgc ctacgtcgcg	840
atgcgcctgc cgaagatgct gccgcaggcc ctgcgactcg tcgcgcacta cgccatcgtc	900
gcgtgggtct ttgcgggcaa ctaccacctg ctcccgtcgc tgacgggttct gcgcggggtt	960
ggcactggga tcaccgtttt cgcgacgcac tacggtgagg acattctcga cgcggaccag	1020
gtgcgtcaca tgacgtcgtg cgagcagacg gcactcacct cgcgcaacat ctgcggcggc	1080
tggtcgtgta acgtgctcac cggcttcacg tcaactgcga cggagcacca cctgttccc	1140
atgatgcaa cccgcaacct catgactatc cagcccagag tgcgcgcctt cttcaagaag	1200
cacggacttg agtaccgcga gggcaacctc attga	1235

<210> SEQ ID NO 6
 <211> LENGTH: 1700
 <212> TYPE: DNA
 <213> ORGANISM: Pavlova salina
 <400> SEQUENCE: 6

ggcacgaggc agctggccgc cgtcgagaga acgccagggg ggtcatggga cgcggcggag	60
acagcagtgg gcaggcgcac ccggcggcgg agctggcggg cccgagcgac cgcgcggagg	120
tgagcaacgc tgacagcaaa gcgctgcaca tcgtgctgta tggcaagcgc gtggatgtga	180
ccaagttcca acgcacgcac ccgggtggta gcaaggtctt ccgcatcttc caggaccgcg	240

-continued

atgcgacgga gcagttcgag tcctaccact cgaagcgcgc gatcaagatg atggagggca	300
tgctcaagaa gtctgaggat gctcccgcgc acacgccctt gccctcccag tcaccgatgg	360
ggaaggactt caaggcgatg atcgagcggc acgttgacgc gggttactac gatccatgcc	420
cgtctgatga gctgttaag ctcagcctcg tgcctctccc gacctttgcg ggcatgtaca	480
tgctcaagga gggcgctggc tccccgctct gcggcgccct catggtgagc tttggctggt	540
acctcgatgg ctggctcgcg cagcactatc tgcaccactc cgtcttcaag gggccgctcg	600
cacgcaccgt cgggtggaac aacgcggcgg gctacttcct cggcttcgtg caggggtatg	660
cggctcgatg gtggcgcgcg cggcataaca cgcaccacgt gtgcaccaat gaggacggct	720
cggaccccg catcaaaacg gcgccgctgc tcatatacgt gcgcaacaag ccgagcatcg	780
ccaagcgctt gaacgccttc cagcgctacc agcagtacta ctatgtgccg gtgatggcaa	840
tcctcgacct gtactggcgg ctcgagtcga tcgcctacgt cgcgatgcgc ctgccgaaga	900
tgctgccgca ggcctcgca ctgcgcgcgc actacgccat cgtcgcgtgg gtctttgcgg	960
gcaactacca cctgctcccg ctgcgtgacg tctcgcgcgg gtttgccact gggatcaccg	1020
ttttcgcgac gcactacggt gaggacattc tcgacgcgga ccagggtgcgt cacatgacgc	1080
tcgtcgagca gacggcactc acctcgcgca acatctcggg cggctggctc gtgaacgtgc	1140
tcaccggctt catctcactg cagacggagc accacctgtt cccgatgatg ccaaccggca	1200
acctcatgac tatccagccc gaggtgcgcg ccttcttcaa gaagcacgga cttgagtacc	1260
gcgagggcaa cctcattgag tgcgtgcgcgc agaacatccg tgcgcttgca ttcgagcacc	1320
tgctttgagc gctctccgct tccaaaggcg ggatcggcgc atccgttggt gctgcggcac	1380
caacgcttcc gctctcgagc gcagatgttg gttgcatcac gcatcacacc cacccccagc	1440
cggatcacc cgtctccgat gcttgcgcta cctagccgcg tccctccagc atgcactgca	1500
actcattcac cacctgtac agtttcggcc taatccatgg cccgactgct tgccgccttg	1560
cacaccgacg agtacgcgga ctcggtccaa tgccctggcg taatctgctg gcgtgctgcg	1620
gcactgtgat tcattcattg atcgcacagt tcacagcatg ttcgctgaca caacgtttgc	1680
tgacctcata gacagtagaa	1700

<210> SEQ ID NO 7

<211> LENGTH: 909

<212> TYPE: DNA

<213> ORGANISM: Pavlova salina

<400> SEQUENCE: 7

atgaaagctg cggcaggcaa ggtgcagcag gaggcggagc gcctcacggc gggcctctgg	60
ctgccgatga tgcttgccgc cggttatctg ctggttctct ctgcaaaccg cgcgagcttc	120
tacgagaaca tcaacaacga gaaggcgccc tactcgacgt cgtggttctc gctgccgtgc	180
gtcatgacgg ctgtgtacct gggcggtgtg tttggcttga ccaagtactt tgagggtcgc	240
aagccgatgc agggcctgaa ggattacatg ttacgtaca acctgtacca ggtgatcatc	300
aacgtgtggt gcatcgcggc tttcgtcgtg gaggtgaggc gcgcgggcac gagcgcggtg	360
ggcaacaagg tcgacctcgg ccccaactcc ttcaggctcg gctttgtgac gtgggtgcac	420
tacaacaaca agtacgtcga gctgctcgac acgctgtgga tgggtgctgcg caagaagacg	480
cagcaggtct ccttctcgca cgtgtaccac cactgctgc tcctctgggc gtggttctgc	540

-continued

gtagtcaaat tctgcaacgg cggcgacgcc tactttggcg gcatgctcaa ctcgatcatc	600
cacgtgatga tgtactcgta ctacacgatg gcgctgctcg gctggagttg tccatggaag	660
cgatacctca ctcaggcgca gctcgtgcag ttctgcattt gcctcgcgca cgcgacgtgg	720
gcggcgcgca cgggcggtga ccccttcac atttgcctcg tcgagatctg ggtgatggtg	780
tcgatgctgt acctgttcac caagttctac aactctgcgt acaagggcgc agcaaagggc	840
gcagcagcga gcagcaacgg tgcggcgggc cggagcggag ccaagcctaa gagcatcaag	900
gccaaactga	909

<210> SEQ ID NO 8
 <211> LENGTH: 1216
 <212> TYPE: DNA
 <213> ORGANISM: Pavlova salina

<400> SEQUENCE: 8

gaattcggca cgaggtcttc ttccagctgt ggtcgtcatg aaagctgcgg caggcaaggt	60
gcagcaggag gcggagcgcc tcacggcggg cctctggctg ccgatgatgc ttgcggccgg	120
ttatctgctg gttctctctg caaacgcgc gagcttctac gagaacatca acaacgagaa	180
gggcgcctac tcgacgtcgt ggttctcgt cccgtgcgtc atgacggctg tgtacctggg	240
cgggtgtgtt ggcttgacca agtactttga gggtcgcaag ccgatgcagg gcctgaagga	300
ttacatgttt acgtacaacc tgtaccaggt gatcatcaac gtgtggtgca tcgcggcttt	360
cgctcgtggg gtgagggcgc cgggcatgag cgcggtgggc aacaaggctc acctcggccc	420
caactccttc aggtcgcggt ttgtgacgtg ggtgcactac aacaacaagt acgtcagct	480
gctcgacacg ctgtggatgg tgcgtgcgaa gaagacgcag caggtctcct tcctgcacgt	540
gtaccaccac gtgctgctca tctgggctg gttctgcgta gtcaaattct gcaacggcgg	600
cgacgcctac tttggcggca tgcctcaact gatcatccac gtgatgatgt actcgtacta	660
cacgatggcg ctgctcggct ggagttgtcc atggaagcga tacctcactc aggcgcagct	720
cgtgcagttc tgcatttgcc tcgcgcacgc gacgtgggcg gccgcgacgg gcgtgtaccc	780
cttccacatt tgcctcgtcg agatctgggt gatggtgctg atgctgtacc tgttcaccaa	840
gttctacaac tctgcgtaca agggcgcgc aaagggcgca gcagcgagca gcaacggtgc	900
ggcggcgccg agcggagcca agcctaagag catcaaggcc aactgaggcc tggcacgcgg	960
gcgagggcgc ggcacgcgc gcagttccgg tcggcgcaac gtcgcggtg cgcgcgcta	1020
cgcaccacgc aggcagtggt tcaggtggcg aagtgtgcag cctgtctgtc gcctgcacac	1080
ccattgattg gtcccgtcg cgctactctg cgcactgcca agtcgccaag acctgtacgt	1140
gtatgatctg actgataccg catacggatg tcccgtatgc gacgactgcc atacgtgctg	1200
cacacgttgt ccaacc	1216

<210> SEQ ID NO 9
 <211> LENGTH: 915
 <212> TYPE: DNA
 <213> ORGANISM: Pavlova salina

<400> SEQUENCE: 9

atggggccgt tgagcacgct gctagcgtgg atgcccacct ggggcgagtt tgcgcgggg	60
ctgacctatg tcgagcgcca gcagatgtca gaggagctcg tgcgcgcaaa taagctccc	120

-continued

ctgtcgctca tcccggaggt ggactttctt acgatcgct cagtctacgt gggcgaccat	180
tggcggatcc cattcacggc catctcggt tatctggtct tgatcacgt cgggcgcag	240
ctcatggcca ggcggcgcc attgccaatc aacaccttg cgtgcctctg gaatttcgcg	300
ctgtcgctct ttagttttgt cggcatgatt gttacgtgga cgaccatcg cgagcgctg	360
tggaaaaatg gtatcgagga cacagtgtgc ggccatccga tattcatggg gtacggctgg	420
atcggatatg ttatgcttgc ctcatctgg tcgaagctct tcgagctgat cgacaccgta	480
ttcctcgctc cgaagaaggc cgacgtcatc ttctgcact ggtaccacca cgtgacggtg	540
ctgtataact gctggcattc gtacgtgtt cgtatccgt cggcatctg gtttgccgcg	600
atgaattatt tcgtacacgc catcatgtac gcctacttg gcacgacaca gattgggccc	660
aggcagcgca agctcgctgc accgtacgca cggctcatca ccacgttcca gctgtcgag	720
atgggcgtcg gtctggccgt caatggcctt atcatccgt acccgctgat aggccatcat	780
tgccactcga acaagacgaa caccattttg agctggatca tgtacgcgag ctactttgtg	840
cttttcgcgc cactatacgt gaagaactac atcttctcca agctgaagtc gcccaagagg	900
aagaagggtg aatga	915

<210> SEQ ID NO 10

<211> LENGTH: 1219

<212> TYPE: DNA

<213> ORGANISM: Pavlova salina

<400> SEQUENCE: 10

gaattcggca cgagggtcag ccttgagcct tacgcaggcc gactcgccc tggtagcac	60
gcagcgcgcc caaccccatg gggccgttga gcacgtgct agcgtggatg cccacctggg	120
gcgagtttgt cgccgggctg acctatgtcg agcggcagca gatgtcagag gagctcgtgc	180
gcgcaataa gctcccgtg tcgctcatcc cggagggtga cttcttcacg atcgctcag	240
tctacgtggg cgaccattgg cggatcccat tcacggccat ctcggcttat ctggtcttga	300
tcacgtctcg gccgcagctc atggccaggc ggccgccatt gccaatcaac accttggcgt	360
gcctctggaa tttgcgctg tcgctcttta gttttgtcgg catgattgtt acgtggacga	420
ccatcggcga gcgcctgtg aaaaatggtc tcgaggacac agtgtgcggc catccgatat	480
tcattgggta cggctggatc ggatatgtta tgcttgcctt catctggtcg aagctcttcg	540
agctgatcga caccgtattc ctgctcgca agaaggccga cgtcatcttc ctgcaactgt	600
accaccacgt gacggtgctg ctatactgct ggcattcgtc cgctgttcgt atcccgctcg	660
gcatctggtt tgccgcgatg aattatttcg tacacgcat catgtacgcc tactttggca	720
tgacacagat tgggcccagg cagcgcaagc tcgtgcgacc gtacgcagg ctcatcacca	780
cgttccagct gtcgcagatg ggcgtcggtc tggccgtcaa tggccttacc atccgctacc	840
cgtcgatagg ccatcattgc cactcgaaca agacgaacac cattttgagc tggatcatgt	900
acgcgagcta ctttgtgctt ttcgcccac tatacgtgaa gaactacatc ttctccaagc	960
tgaagtcgcc caagaggaag aaggtggaat gattgactcg agctcgtgtt tccgcatct	1020
gtcattttac ggctcctcct gatgcggacg gcagctctga tctttgccac aggaccgatg	1080
acggccgatg gtgtctggtt ccagctggct ctgtcattcg cggctcatgg gcacccgggt	1140
gggagcacca gctgtcagac cggactcgat gagatagcgg tgacagacgg accgtgacca	1200

-continued

cgagtctcat catctgttg 1219

<210> SEQ ID NO 11
 <211> LENGTH: 125
 <212> TYPE: DNA
 <213> ORGANISM: Pavlova salina

<400> SEQUENCE: 11

agcacgacgg gaaccacggc gcgctctcca agtcggcctc ggtcaacctg gcgctcgggt 60
 tgtgccagca ctggatcggc gggagcatga tcctctggct gcaggagcac gtgatgatgc 120
 accac 125

<210> SEQ ID NO 12
 <211> LENGTH: 1344
 <212> TYPE: DNA
 <213> ORGANISM: Pavlova salina

<400> SEQUENCE: 12

atgcctccga gcgcgcgcaa gcagatgggc gcgagcacgg gcgtgcatgc gggcgtcaca 60
 gattcgtcgg ccttcacggc caaggatgtc gccgacaggg cggacctcac gatcgtgggt 120
 gacagcgtgt acgatgcgaa ggcgttcggc tccgagcatc cgggtggcgc gcactttgtg 180
 tcgctgttcg gcggggcgga tgccacggag gcgttcattg agtaccaccg gcgcgcctgg 240
 cccaagtgcg gcattgtcgg ctccacggc ggctctcttg catcgaccga ggagcccgtc 300
 gccgcccgat agggctacct ccagctgtgc gctcgcatcg ccaagatggg gccgtcggtc 360
 agcagcgggt tcgcgcgggc gtcgtactgg gtgaaggcgg ggctgacctc cggtccgcg 420
 atcgcgctcg aggcgtacat gctgtacggc ggcaagcgcc tgctcccgtc gatcgtgctc 480
 ggggtggctgt ttgcgtgat tggcctgaac atccagcacg atgccaacca cggcgcgctc 540
 tccaagtcgg cctcggtcaa cctggcgctc ggggtgtgcc aggaactggat cggcgggagc 600
 atgacctctt ggctgcagga gcaagttgtc atgcaccact tgcacaccaa cgacgttgac 660
 aaggaccggg accagaaggc gcacggcgcc ctgcggctca agccgaccga cgcgtggagc 720
 ccgatgcact ggctgcagca cctctacctg ctgcctgggg agacgatgta cgcttcaag 780
 ctgctgtttc tcgacatcag cgagctgggt atgtggcggt gggagggcga gccatcagc 840
 aagctggcgg ggtacctctt catgccctcg ctgctcctca agctcacctt ctgggcgcgc 900
 ttgtgcgcgc tgccgctgta cctcgcgccc agcgtgcaca cggcggtgtg catcgcggcg 960
 acggtaataa cggggagctt ctacctcgcc ttctttctct tcattctcga caacttcgag 1020
 ggcgtggcga gcgtcggacc ggacggcagc atcaccagca tgacgcgcgg cgcattcttc 1080
 ctcaagcggc aggcggagac ctctgtccaa gtggggcgcc cgctgctcgc cagctcaac 1140
 ggcggcctca actaccaaact cgagcaccac ctcttcccca gggtgacca cggtttctac 1200
 cctcgccctg cgccgttggt caaggcggag ctcgaggcgc gcggcattga gtacaagcac 1260
 taccaccaca tatggagcaa cctggcatcc acgctgaggg acatgtacgc gctcggccgc 1320
 aggcgcgcga gcaaggcgga gtga 1344

<210> SEQ ID NO 13
 <211> LENGTH: 1687
 <212> TYPE: DNA
 <213> ORGANISM: Pavlova salina

-continued

<400> SEQUENCE: 13

```

ctcctgtgag accgctgtgc gccagcgcaa ggaccgacct gcacgcgcga tgccctcgag    60
cgcggcgaaag cagatggggc cgagcacggg cgtgcatgcg ggctcacag attcgtcggc    120
cttcacgcgc aaggatgtcg ccgacaggcc ggacctcacg atcgtgggtg acagcgtgta    180
cgatgcgaag gcgttccgct ccgagcatcc ggggtggcgc cactttgtgt cgtgttcgg    240
cgggcgcgat gccacggagg cgttcatgga gtaccaccgg cgcgcctggc ccaagtcgcg    300
catgtcgcgc ttccacgtcg gctctctggc atcgaccgag gagcccgctg ccgccgatga    360
gggtacacct cagctgtgcg ctgcgcatgc caagatgggt ccgtcggtca gcagcgggtt    420
cgcgcggcgc tcgtactggg tgaaggccgg gctgatcctc ggctccgcga tcgcgctcga    480
ggcgtacatg ctgtacgcgg gcaagcgccg gctcccgtcg atcgtgctcg ggtggctgtt    540
tgcgctgatt ggctgaaca tccagcacga tgccaaccac ggcgcgctct ccaagtcggc    600
ctcgggtcaac ctggcgctcg ggttgtgcca ggactggatc ggccggagca tgatcctctg    660
gctgcaggag cacgttgtca tgaccactt gcacaccaac gacgttgaca aggaccggga    720
ccagaaggcg cacggcgccc tgcggctcaa gccgaccgac gcgtggagcc cgatgcactg    780
gctgcagcac ctctacctgc tgccctggga gacgatgtac gccttcaagc tgctgtttct    840
cgacatcagc gagctgtgta tgtggcgggt ggagggcgag cccatcagca agctggccgg    900
gtacctcttc atgcctcgc tgctcctcaa gctcaccttc tgggcgcgct ttgtcgcgct    960
gccgctgtac ctgcgcacca gcgtgcacac ggcggtgtgc atcgcgcgca cgtaaatgac   1020
ggggagcttc tacctcgctt tcttcttctt catctcgcac aacttcgagg gcgtggcgag   1080
cgtcggaccg gacggcagca tcaccagcat gacgcggcgc gcctccttc tcaagcgga    1140
ggccgagacc tcgtccaacg tggggcgccc gctgctcgcc acgctcaacg gcggcctcaa   1200
ctaccaaate gagcaccacc tcttccccag ggtgcaccac ggcttctacc ctgcctcgc    1260
gccgttggtc aaggcggagc tcgaggcgcg cggcattgag tacaagcact accccaccat   1320
atggagcaac ctggcatcca cgctgaggca catgtacgcg ctcgcccgca ggccgcgcag    1380
caaggcggag tgacgagcct cccaatcggc tgcccaggct gctggcagct tgggtcgacc   1440
atcggcacgc actcggcgac gccgacgcca gccgtggctg cacaacggc tcggcggttg    1500
ttcttcggct accgcgcggg aaccggggca cgcacctgac tcggtgacga ttgcggtcgc    1560
accagcagaa gcgtcacc cgcgcccgcc tcggatcggg cggcatagag catacatcta    1620
gcgcgttctt actatatgcc actgtgtaga ctagtcttct agcgccggaa aatctgctat   1680
caacaat

```

<210> SEQ ID NO 14

<211> LENGTH: 288

<212> TYPE: PRT

<213> ORGANISM: *Caenorhabditis elegans*

<400> SEQUENCE: 14

```

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

```

-continued

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

<210> SEQ ID NO 15

<211> LENGTH: 444

<212> TYPE: PRT

<213> ORGANISM: Danio rerio

<400> SEQUENCE: 15

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

-continued

```

Phe Arg Ala Leu Arg Glu Arg Leu Glu Ala Glu Gly Cys Phe Lys Thr
   115                               120                       125

Gln Pro Leu Phe Phe Ala Leu His Leu Gly His Ile Leu Leu Leu Glu
   130                               135                       140

Ala Ile Ala Phe Met Met Val Trp Tyr Phe Gly Thr Gly Trp Ile Asn
  145                               150                       155                       160

Thr Leu Ile Val Ala Val Ile Leu Ala Thr Ala Gln Ser Gln Ala Gly
   165                               170                       175

Trp Leu Gln His Asp Phe Gly His Leu Ser Val Phe Lys Thr Ser Gly
   180                               185                       190

Met Asn His Leu Val His Lys Phe Val Ile Gly His Leu Lys Gly Ala
   195                               200                       205

Ser Ala Gly Trp Trp Asn His Arg His Phe Gln His His Ala Lys Pro
   210                               215                       220

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

Val Gly Asn Val Gln Pro Val Glu Tyr Gly Val Lys Lys Ile Lys His
   245                               250                       255

Leu Pro Tyr Asn His Gln His Lys Tyr Phe Phe Phe Ile Gly Pro Pro
   260                               265                       270

Leu Leu Ile Pro Val Tyr Phe Gln Phe Gln Ile Phe His Asn Met Ile
   275                               280                       285

Ser His Gly Met Trp Val Asp Leu Leu Trp Cys Ile Ser Tyr Tyr Val
   290                               295                       300

Arg Tyr Phe Leu Cys Tyr Thr Gln Phe Tyr Gly Val Phe Trp Ala Ile
  305                               310                       315                       320

Ile Leu Phe Asn Phe Val Arg Phe Met Glu Ser His Trp Phe Val Trp
   325                               330                       335

Val Thr Gln Met Ser His Ile Pro Met Asn Ile Asp Tyr Glu Lys Asn
   340                               345                       350

Gln Asp Trp Leu Ser Met Gln Leu Val Ala Thr Cys Asn Ile Glu Gln
   355                               360                       365

Ser Ala Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu
   370                               375                       380

His His Leu Phe Pro Thr Val Pro Arg His Asn Tyr Trp Arg Ala Ala
  385                               390                       395                       400

Pro Arg Val Arg Ala Leu Cys Glu Lys Tyr Gly Val Lys Tyr Gln Glu
   405                               410                       415

Lys Thr Leu Tyr Gly Ala Phe Ala Asp Ile Ile Arg Ser Leu Glu Lys
   420                               425                       430

Ser Gly Glu Leu Trp Leu Asp Ala Tyr Leu Asn Lys
   435                               440

```

<210> SEQ ID NO 16

<211> LENGTH: 444

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 16

```

Met Ala Pro Asp Pro Leu Ala Ala Glu Thr Ala Ala Gln Gly Leu Thr
  1           5           10           15

Pro Arg Tyr Phe Thr Trp Asp Glu Val Ala Gln Arg Ser Gly Cys Glu
  20           25           30

```

-continued

Glu Arg Trp Leu Val Ile Asp Arg Lys Val Tyr Asn Ile Ser Glu Phe
 35 40 45
 Thr Arg Arg His Pro Gly Gly Ser Arg Val Ile Ser His Tyr Ala Gly
 50 55 60
 Gln Asp Ala Thr Asp Pro Phe Val Ala Phe His Ile Asn Lys Gly Leu
 65 70 75 80
 Val Lys Lys Tyr Met Asn Ser Leu Leu Ile Gly Glu Leu Ser Pro Glu
 85 90 95
 Gln Pro Ser Phe Glu Pro Thr Lys Asn Lys Glu Leu Thr Asp Glu Phe
 100 105 110
 Arg Glu Leu Arg Ala Thr Val Glu Arg Met Gly Leu Met Lys Ala Asn
 115 120 125
 His Val Phe Phe Leu Leu Tyr Leu Leu His Ile Leu Leu Leu Asp Gly
 130 135 140
 Ala Ala Trp Leu Thr Leu Trp Val Phe Gly Thr Ser Phe Leu Pro Phe
 145 150 155 160
 Leu Leu Cys Ala Val Leu Leu Ser Ala Val Gln Ala Gln Ala Gly Trp
 165 170 175
 Leu Gln His Asp Phe Gly His Leu Ser Val Phe Ser Thr Ser Lys Trp
 180 185 190
 Asn His Leu Leu His His Phe Val Ile Gly His Leu Lys Gly Ala Pro
 195 200 205
 Ala Ser Trp Trp Asn His Met His Phe Gln His His Ala Lys Pro Asn
 210 215 220
 Cys Phe Arg Lys Asp Pro Asp Ile Asn Met His Pro Phe Phe Phe Ala
 225 230 235 240
 Leu Gly Lys Ile Leu Ser Val Glu Leu Gly Lys Gln Lys Lys Asn Tyr
 245 250 255
 Met Pro Tyr Asn His Gln His Lys Tyr Phe Phe Leu Ile Gly Pro Pro
 260 265 270
 Ala Leu Leu Pro Leu Tyr Phe Gln Trp Tyr Ile Phe Tyr Phe Val Ile
 275 280 285
 Gln Arg Lys Lys Trp Val Asp Leu Ala Trp Met Ile Thr Phe Tyr Val
 290 295 300
 Arg Phe Phe Leu Thr Tyr Val Pro Leu Leu Gly Leu Lys Ala Phe Leu
 305 310 315 320
 Gly Leu Phe Phe Ile Val Arg Phe Leu Glu Ser Asn Trp Phe Val Trp
 325 330 335
 Val Thr Gln Met Asn His Ile Pro Met His Ile Asp His Asp Arg Asn
 340 345 350
 Met Asp Trp Val Ser Thr Gln Leu Gln Ala Thr Cys Asn Val His Lys
 355 360 365
 Ser Ala Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu
 370 375 380
 His His Leu Phe Pro Thr Met Pro Arg His Asn Tyr His Lys Val Ala
 385 390 395 400
 Pro Leu Val Gln Ser Leu Cys Ala Lys His Gly Ile Glu Tyr Gln Ser
 405 410 415
 Lys Pro Leu Leu Ser Ala Phe Ala Asp Ile Ile His Ser Leu Lys Glu
 420 425 430

-continued

Ser Gly Gln Leu Trp Leu Asp Ala Tyr Leu His Gln
435 440

<210> SEQ ID NO 17
<211> LENGTH: 456
<212> TYPE: PRT
<213> ORGANISM: Pythium irregulare

<400> SEQUENCE: 17

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

His Asn Thr Ala Lys Ser Ala Trp Val Ile Ile Arg Gly Glu Val Tyr
20 25 30

Asp Val Thr Glu Trp Ala Asp Lys His Pro Gly Gly Ser Glu Leu Ile
35 40 45

Val Leu His Ser Gly Arg Glu Cys Thr Asp Thr Phe Tyr Ser Tyr His
50 55 60

Pro Phe Ser Asn Arg Ala Asp Lys Ile Leu Ala Lys Tyr Lys Ile Gly
65 70 75 80

Lys Leu Val Gly Gly Tyr Glu Phe Pro Val Phe Lys Pro Asp Ser Gly
85 90 95

Phe Tyr Lys Glu Cys Ser Glu Arg Val Ala Glu Tyr Phe Lys Thr Asn
100 105 110

Asn Leu Asp Pro Lys Ala Ala Phe Ala Gly Leu Trp Arg Met Val Phe
115 120 125

Val Phe Ala Val Ala Ala Leu Ala Tyr Met Gly Met Asn Glu Leu Ile
130 135 140

Pro Gly Asn Val Tyr Ala Gln Tyr Ala Trp Gly Val Val Phe Gly Val
145 150 155 160

Phe Gln Ala Leu Pro Leu Leu His Val Met His Asp Ser Ser His Ala
165 170 175

Ala Cys Ser Ser Ser Pro Ala Met Trp Gln Ile Ile Gly Arg Gly Val
180 185 190

Met Asp Trp Phe Ala Gly Ala Ser Met Val Ser Trp Leu Asn Gln His
195 200 205

Val Val Gly His His Ile Tyr Thr Asn Val Ala Gly Ala Asp Pro Asp
210 215 220

Leu Pro Val Asp Phe Glu Ser Asp Val Arg Arg Ile Val His Arg Gln
225 230 235 240

Val Leu Leu Pro Ile Tyr Lys Phe Gln His Ile Tyr Leu Pro Pro Leu
245 250 255

Tyr Gly Val Leu Gly Leu Lys Phe Arg Ile Gln Asp Val Phe Glu Thr
260 265 270

Phe Val Ser Leu Thr Asn Gly Pro Val Arg Val Asn Pro His Pro Val
275 280 285

Ser Asp Trp Val Gln Met Ile Phe Ala Lys Ala Phe Trp Thr Phe Tyr
290 295 300

Arg Ile Tyr Ile Pro Leu Val Trp Leu Lys Ile Thr Pro Ser Thr Phe
305 310 315 320

Trp Gly Val Phe Phe Leu Ala Glu Phe Thr Thr Gly Trp Tyr Leu Ala
325 330 335

Phe Asn Phe Gln Val Ser His Val Ser Thr Glu Cys Glu Tyr Pro Cys
340 345 350

-continued

Gly Asp Ala Pro Ser Ala Glu Val Gly Asp Glu Trp Ala Ile Ser Gln
 355 360 365

Val Lys Ser Ser Val Asp Tyr Ala His Gly Ser Pro Leu Ala Ala Phe
 370 375 380

Leu Cys Gly Ala Leu Asn Tyr Gln Val Thr His His Leu Tyr Pro Gly
 385 390 395 400

Ile Ser Gln Tyr His Tyr Pro Ala Ile Ala Pro Ile Ile Ile Asp Val
 405 410 415

Cys Lys Lys Tyr Asn Ile Lys Tyr Thr Val Leu Pro Thr Phe Thr Glu
 420 425 430

Ala Leu Leu Ala His Phe Lys His Leu Lys Asn Met Gly Glu Leu Gly
 435 440 445

Lys Pro Val Glu Ile His Met Gly
 450 455

<210> SEQ ID NO 18

<211> LENGTH: 439

<212> TYPE: PRT

<213> ORGANISM: Thraustochytrium sp.

<400> SEQUENCE: 18

Met Gly Lys Gly Ser Glu Gly Arg Ser Ala Ala Arg Glu Met Thr Ala
 1 5 10 15

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

Tyr Asp Ala Thr Asn Phe Lys His Pro Gly Gly Ser Ile Ile Asn Phe
 35 40 45

Leu Thr Glu Gly Glu Ala Gly Val Asp Ala Thr Gln Ala Tyr Arg Glu
 50 55 60

Phe His Gln Arg Ser Gly Lys Ala Asp Lys Tyr Leu Lys Ser Leu Pro
 65 70 75 80

Lys Leu Asp Ala Ser Lys Val Glu Ser Arg Phe Ser Ala Lys Glu Gln
 85 90 95

Ala Arg Arg Asp Ala Met Thr Arg Asp Tyr Ala Ala Phe Arg Glu Glu
 100 105 110

Leu Val Ala Glu Gly Tyr Phe Asp Pro Ser Ile Pro His Met Ile Tyr
 115 120 125

Arg Val Val Glu Ile Val Ala Leu Phe Ala Leu Ser Phe Trp Leu Met
 130 135 140

Ser Lys Ala Ser Pro Thr Ser Leu Val Leu Gly Val Val Met Asn Gly
 145 150 155 160

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

Ser Phe Thr Gly Val Ile Trp Leu Asp Asp Arg Met Cys Glu Phe Phe
 180 185 190

Tyr Gly Val Gly Cys Gly Met Ser Gly His Tyr Trp Lys Asn Gln His
 195 200 205

Ser Lys His His Ala Ala Pro Asn Arg Leu Glu His Asp Val Asp Leu
 210 215 220

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

Lys Pro Gly Ser Leu Leu Ala Leu Trp Leu Arg Val Gln Ala Tyr Leu

-continued

245					250					255					
Phe	Ala	Pro	Val	Ser	Cys	Leu	Leu	Ile	Gly	Leu	Gly	Trp	Thr	Leu	Tyr
			260					265					270		
Leu	His	Pro	Arg	Tyr	Met	Leu	Arg	Thr	Lys	Arg	His	Met	Glu	Phe	Val
			275					280					285		
Trp	Ile	Phe	Ala	Arg	Tyr	Ile	Gly	Trp	Phe	Ser	Leu	Met	Gly	Ala	Leu
			290					295					300		
Gly	Tyr	Ser	Pro	Gly	Thr	Ser	Val	Gly	Met	Tyr	Leu	Cys	Ser	Phe	Gly
			305					310					315		320
Leu	Gly	Cys	Ile	Tyr	Ile	Phe	Leu	Gln	Phe	Ala	Val	Ser	His	Thr	His
			325					330					335		
Leu	Pro	Val	Thr	Asn	Pro	Glu	Asp	Gln	Leu	His	Trp	Leu	Glu	Tyr	Ala
			340					345					350		
Ala	Asp	His	Thr	Val	Asn	Ile	Ser	Thr	Lys	Ser	Trp	Leu	Val	Thr	Trp
			355					360					365		
Trp	Met	Ser	Asn	Leu	Asn	Phe	Gln	Ile	Glu	His	His	Leu	Phe	Pro	Thr
			370					375					380		
Ala	Pro	Gln	Phe	Arg	Phe	Lys	Glu	Ile	Ser	Pro	Arg	Val	Glu	Ala	Leu
			385					390					395		400
Phe	Lys	Arg	His	Asn	Leu	Pro	Tyr	Tyr	Asp	Leu	Pro	Tyr	Thr	Ser	Ala
			405					410					415		
Val	Ser	Thr	Thr	Phe	Ala	Asn	Leu	Tyr	Ser	Val	Gly	His	Ser	Val	Gly
			420					425					430		
Ala	Asp	Thr	Lys	Lys	Gln	Asp									
			435												

<210> SEQ ID NO 19

<211> LENGTH: 446

<212> TYPE: PRT

<213> ORGANISM: Mortierella alpina

<400> SEQUENCE: 19

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

-continued

Ala Cys Ala Gln Val Gly Leu Asn Pro Leu His Asp Ala Ser His Phe
165 170 175

Ser Val Thr His Asn Pro Thr Val Trp Lys Ile Leu Gly Ala Thr His
180 185 190

Asp Phe Phe Asn Gly Ala Ser Tyr Leu Val Trp Met Tyr Gln His Met
195 200 205

Leu Gly His His Pro Tyr Thr Asn Ile Ala Gly Ala Asp Pro Asp Val
210 215 220

Ser Thr Phe Glu Pro Asp Val Arg Arg Ile Lys Pro Asn Gln Lys Trp
225 230 235 240

Phe Val Asn His Ile Asn Gln Asp Met Phe Val Pro Phe Leu Tyr Gly
245 250 255

Leu Leu Ala Phe Lys Val Arg Ile Gln Asp Ile Asn Ile Leu Tyr Phe
260 265 270

Val Lys Thr Asn Asp Ala Ile Arg Val Asn Pro Ile Ser Thr Trp His
275 280 285

Thr Val Met Phe Trp Gly Gly Lys Ala Phe Phe Val Trp Tyr Arg Leu
290 295 300

Ile Val Pro Leu Gln Tyr Leu Pro Leu Gly Lys Val Leu Leu Leu Phe
305 310 315 320

Thr Val Ala Asp Met Val Ser Ser Tyr Trp Leu Ala Leu Thr Phe Gln
325 330 335

Ala Asn His Val Val Glu Glu Val Gln Trp Pro Leu Pro Asp Glu Asn
340 345 350

Gly Ile Ile Gln Lys Asp Trp Ala Ala Met Gln Val Glu Thr Thr Gln
355 360 365

Asp Tyr Ala His Asp Ser His Leu Trp Thr Ser Ile Thr Gly Ser Leu
370 375 380

Asn Tyr Gln Ala Val His His Leu Phe Pro Asn Val Ser Gln His His
385 390 395 400

Tyr Pro Asp Ile Leu Ala Ile Ile Lys Asn Thr Cys Ser Glu Tyr Lys
405 410 415

Val Pro Tyr Leu Val Lys Asp Thr Phe Trp Gln Ala Phe Ala Ser His
420 425 430

Leu Glu His Leu Arg Val Leu Gly Leu Arg Pro Lys Glu Glu
435 440 445

<210> SEQ ID NO 20

<211> LENGTH: 447

<212> TYPE: PRT

<213> ORGANISM: Caenorhabditis elegans

<400> SEQUENCE: 20

Met Val Leu Arg Glu Gln Glu His Glu Pro Phe Phe Ile Lys Ile Asp
1 5 10 15

Gly Lys Trp Cys Gln Ile Asp Asp Ala Val Leu Arg Ser His Pro Gly
20 25 30

Gly Ser Ala Ile Thr Thr Tyr Lys Asn Met Asp Ala Thr Thr Val Phe
35 40 45

His Thr Phe His Thr Gly Ser Lys Glu Ala Tyr Gln Trp Leu Thr Glu
50 55 60

Leu Lys Lys Glu Cys Pro Thr Gln Glu Pro Glu Ile Pro Asp Ile Lys
65 70 75 80

-continued

```

Asp Asp Pro Ile Lys Gly Ile Asp Asp Val Asn Met Gly Thr Phe Asn
      85                      90                      95

Ile Ser Glu Lys Arg Ser Ala Gln Ile Asn Lys Ser Phe Thr Asp Leu
      100                    105                    110

Arg Met Arg Val Arg Ala Glu Gly Leu Met Asp Gly Ser Pro Leu Phe
      115                    120                    125

Tyr Ile Arg Lys Ile Leu Glu Thr Ile Phe Thr Ile Leu Phe Ala Phe
      130                    135                    140

Tyr Leu Gln Tyr His Thr Tyr Tyr Leu Pro Ser Ala Ile Leu Met Gly
      145                    150                    155                    160

Val Ala Trp Gln Gln Leu Gly Trp Leu Ile His Glu Phe Ala His His
      165                    170                    175

Gln Leu Phe Lys Asn Arg Tyr Tyr Asn Asp Leu Ala Ser Tyr Phe Val
      180                    185                    190

Gly Asn Phe Leu Gln Gly Phe Ser Ser Gly Gly Trp Lys Glu Gln His
      195                    200                    205

Asn Val His His Ala Ala Thr Asn Val Val Gly Arg Asp Gly Asp Leu
      210                    215                    220

Asp Leu Val Pro Phe Tyr Ala Thr Val Ala Glu His Leu Asn Asn Tyr
      225                    230                    235                    240

Ser Gln Asp Ser Trp Val Met Thr Leu Phe Arg Trp Gln His Val His
      245                    250                    255

Trp Thr Phe Met Leu Pro Phe Leu Arg Leu Ser Trp Leu Leu Gln Ser
      260                    265                    270

Ile Ile Phe Val Ser Gln Met Pro Thr His Tyr Tyr Asp Tyr Tyr Arg
      275                    280                    285

Asn Thr Ala Ile Tyr Glu Gln Val Gly Leu Ser Leu His Trp Ala Trp
      290                    295                    300

Ser Leu Gly Gln Leu Tyr Phe Leu Pro Asp Trp Ser Thr Arg Ile Met
      305                    310                    315                    320

Phe Phe Leu Val Ser His Leu Val Gly Gly Phe Leu Leu Ser His Val
      325                    330                    335

Val Thr Phe Asn His Tyr Ser Val Glu Lys Phe Ala Leu Ser Ser Asn
      340                    345                    350

Ile Met Ser Asn Tyr Ala Cys Leu Gln Ile Met Thr Thr Arg Asn Met
      355                    360                    365

Arg Pro Gly Arg Phe Ile Asp Trp Leu Trp Gly Gly Leu Asn Tyr Gln
      370                    375                    380

Ile Glu His His Leu Phe Pro Thr Met Pro Arg His Asn Leu Asn Thr
      385                    390                    395                    400

Val Met Pro Leu Val Lys Glu Phe Ala Ala Ala Asn Gly Leu Pro Tyr
      405                    410                    415

Met Val Asp Asp Tyr Phe Thr Gly Phe Trp Leu Glu Ile Glu Gln Phe
      420                    425                    430

Arg Asn Ile Ala Asn Val Ala Ala Lys Leu Thr Lys Lys Ile Ala
      435                    440                    445

```

<210> SEQ ID NO 21

<211> LENGTH: 444

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

-continued

<400> SEQUENCE: 21

Met Gly Lys Gly Gly Asn Gln Gly Glu Gly Ala Ala Glu Arg Glu Val
 1 5 10 15
 Ser Val Pro Thr Phe Ser Trp Glu Glu Ile Gln Lys His Asn Leu Arg
 20 25 30
 Thr Asp Arg Trp Leu Val Ile Asp Arg Lys Val Tyr Asn Ile Thr Lys
 35 40 45
 Trp Ser Ile Gln His Pro Gly Gly Gln Arg Val Ile Gly His Tyr Ala
 50 55 60
 Gly Glu Asp Ala Thr Asp Ala Phe Arg Ala Phe His Pro Asp Leu Glu
 65 70 75 80
 Phe Val Gly Lys Phe Leu Lys Pro Leu Leu Ile Gly Glu Leu Ala Pro
 85 90 95
 Glu Glu Pro Ser Gln Asp His Gly Lys Asn Ser Lys Ile Thr Glu Asp
 100 105 110
 Phe Arg Ala Leu Arg Lys Thr Ala Glu Asp Met Asn Leu Phe Lys Thr
 115 120 125
 Asn His Val Phe Phe Leu Leu Leu Leu Ala His Ile Ile Ala Leu Glu
 130 135 140
 Ser Ile Ala Trp Phe Thr Val Phe Tyr Phe Gly Asn Gly Trp Ile Pro
 145 150 155 160
 Thr Leu Ile Thr Ala Phe Val Leu Ala Thr Ser Gln Ala Gln Ala Gly
 165 170 175
 Trp Leu Gln His Asp Tyr Gly His Leu Ser Val Tyr Arg Lys Pro Lys
 180 185 190
 Trp Asn His Leu Val His Lys Phe Val Ile Gly His Leu Lys Gly Ala
 195 200 205
 Ser Ala Asn Trp Trp Asn His Arg His Phe Gln His His Ala Lys Pro
 210 215 220
 Asn Ile Phe His Lys Asp Pro Asp Val Asn Met Leu His Val Phe Val
 225 230 235 240
 Leu Gly Glu Trp Gln Pro Ile Glu Tyr Gly Lys Lys Lys Leu Lys Tyr
 245 250 255
 Leu Pro Tyr Asn His Gln His Glu Tyr Phe Phe Leu Ile Gly Pro Pro
 260 265 270
 Leu Leu Ile Pro Met Tyr Phe Gln Tyr Gln Ile Ile Met Thr Met Ile
 275 280 285
 Val His Lys Asn Trp Val Asp Leu Ala Trp Ala Val Ser Tyr Tyr Ile
 290 295 300
 Arg Phe Phe Ile Thr Tyr Ile Pro Phe Tyr Gly Ile Leu Gly Ala Leu
 305 310 315 320
 Leu Phe Leu Asn Phe Ile Arg Phe Leu Glu Ser His Trp Phe Val Trp
 325 330 335
 Val Thr Gln Met Asn His Ile Val Met Glu Ile Asp Gln Glu Ala Tyr
 340 345 350
 Arg Asp Trp Phe Ser Ser Gln Leu Thr Ala Thr Cys Asn Val Glu Gln
 355 360 365
 Ser Phe Phe Asn Asp Trp Phe Ser Gly His Leu Asn Phe Gln Ile Glu
 370 375 380
 His His Leu Phe Pro Thr Met Pro Arg His Asn Leu His Lys Ile Ala
 385 390 395 400

-continued

Pro Leu Val Lys Ser Leu Cys Ala Lys His Gly Ile Glu Tyr Gln Glu
 405 410 415

Lys Pro Leu Leu Arg Ala Leu Leu Asp Ile Ile Arg Ser Leu Lys Lys
 420 425 430

Ser Gly Lys Leu Trp Leu Asp Ala Tyr Leu His Lys
 435 440

<210> SEQ ID NO 22

<211> LENGTH: 444

<212> TYPE: PRT

<213> ORGANISM: Mus musculus

<400> SEQUENCE: 22

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

Pro Met Pro Thr Phe Arg Trp Glu Glu Ile Gln Lys His Asn Leu Arg
 20 25 30

Thr Asp Arg Trp Leu Val Ile Asp Arg Lys Val Tyr Asn Val Thr Lys
 35 40 45

Trp Ser Gln Arg His Pro Gly Gly His Arg Val Ile Gly His Tyr Ser
 50 55 60

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

Phe Val Gly Lys Phe Leu Lys Pro Leu Leu Ile Gly Glu Leu Ala Pro
 85 90 95

Glu Glu Pro Ser Leu Asp Arg Gly Lys Ser Ser Gln Ile Thr Glu Asp
 100 105 110

Phe Arg Ala Leu Lys Lys Thr Ala Glu Asp Met Asn Leu Phe Lys Thr
 115 120 125

Asn His Leu Phe Phe Phe Leu Leu Leu Ser His Ile Ile Val Met Glu
 130 135 140

Ser Leu Ala Trp Phe Ile Leu Ser Tyr Phe Gly Thr Gly Trp Ile Pro
 145 150 155 160

Thr Leu Val Thr Ala Phe Val Leu Ala Thr Ser Gln Ala Gln Ala Gly
 165 170 175

Trp Leu Gln His Asp Tyr Gly His Leu Ser Val Tyr Lys Lys Ser Ile
 180 185 190

Trp Asn His Val Val His Lys Phe Val Ile Gly His Leu Lys Gly Ala
 195 200 205

Ser Ala Asn Trp Trp Asn His Arg His Phe Gln His His Ala Lys Pro
 210 215 220

Asn Ile Phe His Lys Asp Pro Asp Ile Lys Ser Leu His Val Phe Val
 225 230 235 240

Leu Gly Glu Trp Gln Pro Leu Glu Tyr Gly Lys Lys Lys Leu Lys Tyr
 245 250 255

Leu Pro Tyr Asn His Gln His Glu Tyr Phe Phe Leu Ile Gly Pro Pro
 260 265 270

Leu Leu Ile Pro Met Tyr Phe Gln Tyr Gln Ile Ile Met Thr Met Ile
 275 280 285

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

Arg Phe Phe Tyr Thr Tyr Ile Pro Phe Tyr Gly Ile Leu Gly Ala Leu

-continued

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

<210> SEQ ID NO 23

<211> LENGTH: 459

<212> TYPE: PRT

<213> ORGANISM: Pythium irregulare

<400> SEQUENCE: 23

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

-continued

His Ser Ala Lys Asp Glu Gly Phe Ile Gly Asp Pro Asp Ile Asp Thr
 225 230 235 240
 Met Pro Leu Leu Ala Trp Ser Lys Glu Met Ala Arg Lys Ala Phe Glu
 245 250 255
 Ser Ala His Gly Pro Phe Phe Ile Arg Asn Gln Ala Phe Leu Tyr Phe
 260 265 270
 Pro Leu Leu Leu Leu Ala Arg Leu Ser Trp Leu Ala Gln Ser Phe Phe
 275 280 285
 Tyr Val Phe Thr Glu Phe Ser Phe Gly Ile Phe Asp Lys Val Glu Phe
 290 295 300
 Asp Gly Pro Glu Lys Ala Gly Leu Ile Val His Tyr Ile Trp Gln Leu
 305 310 315 320
 Ala Ile Pro Tyr Phe Cys Asn Met Ser Leu Phe Glu Gly Val Ala Tyr
 325 330 335
 Phe Leu Met Gly Gln Ala Ser Cys Gly Leu Leu Leu Ala Leu Val Phe
 340 345 350
 Ser Ile Gly His Asn Gly Met Ser Val Tyr Glu Arg Glu Thr Lys Pro
 355 360 365
 Asp Phe Trp Gln Leu Gln Val Thr Thr Thr Arg Asn Ile Arg Ala Ser
 370 375 380
 Val Phe Met Asp Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Asp His
 385 390 395 400
 His Leu Phe Pro Leu Val Pro Arg His Asn Leu Pro Lys Val Asn Val
 405 410 415
 Leu Ile Lys Ser Leu Cys Lys Glu Phe Asp Ile Pro Phe His Glu Thr
 420 425 430
 Gly Phe Trp Glu Gly Ile Tyr Glu Val Val Asp His Leu Ala Asp Ile
 435 440 445
 Ser Lys Glu Phe Ile Thr Glu Phe Pro Ala Met
 450 455

<210> SEQ ID NO 24

<211> LENGTH: 448

<212> TYPE: PRT

<213> ORGANISM: Borago officinalis

<400> SEQUENCE: 24

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

-continued

Tyr Gly Val Leu Phe Cys Glu Gly Val Leu Val His Leu Phe Ser Gly
 130 135 140
 Cys Leu Met Gly Phe Leu Trp Ile Gln Ser Gly Trp Ile Gly His Asp
 145 150 155 160
 Ala Gly His Tyr Met Val Val Ser Asp Ser Arg Leu Asn Lys Phe Met
 165 170 175
 Gly Ile Phe Ala Ala Asn Cys Leu Ser Gly Ile Ser Ile Gly Trp Trp
 180 185 190
 Lys Trp Asn His Asn Ala His His Ile Ala Cys Asn Ser Leu Glu Tyr
 195 200 205
 Asp Pro Asp Leu Gln Tyr Ile Pro Phe Leu Val Val Ser Ser Lys Phe
 210 215 220
 Phe Gly Ser Leu Thr Ser His Phe Tyr Glu Lys Arg Leu Thr Phe Asp
 225 230 235 240
 Ser Leu Ser Arg Phe Phe Val Ser Tyr Gln His Trp Thr Phe Tyr Pro
 245 250 255
 Ile Met Cys Ala Ala Arg Leu Asn Met Tyr Val Gln Ser Leu Ile Met
 260 265 270
 Leu Leu Thr Lys Arg Asn Val Ser Tyr Arg Ala Gln Glu Leu Leu Gly
 275 280 285
 Cys Leu Val Phe Ser Ile Trp Tyr Pro Leu Leu Val Ser Cys Leu Pro
 290 295 300
 Asn Trp Gly Glu Arg Ile Met Phe Val Ile Ala Ser Leu Ser Val Thr
 305 310 315 320
 Gly Met Gln Gln Val Gln Phe Ser Leu Asn His Phe Ser Ser Ser Val
 325 330 335
 Tyr Val Gly Lys Pro Lys Gly Asn Asn Trp Phe Glu Lys Gln Thr Asp
 340 345 350
 Gly Thr Leu Asp Ile Ser Cys Pro Pro Trp Met Asp Trp Phe His Gly
 355 360 365
 Gly Leu Gln Phe Gln Ile Glu His His Leu Phe Pro Lys Met Pro Arg
 370 375 380
 Cys Asn Leu Arg Lys Ile Ser Pro Tyr Val Ile Glu Leu Cys Lys Lys
 385 390 395 400
 His Asn Leu Pro Tyr Asn Tyr Ala Ser Phe Ser Lys Ala Asn Glu Met
 405 410 415
 Thr Leu Arg Thr Leu Arg Asn Thr Ala Leu Gln Ala Arg Asp Ile Thr
 420 425 430
 Lys Pro Leu Pro Lys Asn Leu Val Trp Glu Ala Leu His Thr His Gly
 435 440 445

<210> SEQ ID NO 25

<211> LENGTH: 446

<212> TYPE: PRT

<213> ORGANISM: Anemone leveillei

<400> SEQUENCE: 25

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

-continued

35					40					45					
Asn	Leu	Ala	Gly	Gln	Asp	Val	Thr	Asp	Ala	Phe	Val	Ala	Phe	His	Pro
50						55					60				
Gly	Ser	Ala	Trp	Lys	Asn	Leu	Asp	Lys	Phe	His	Ile	Gly	Tyr	Leu	Gln
65					70					75					80
Asp	Tyr	Val	Val	Ser	Asp	Val	Ser	Lys	Asp	Tyr	Arg	Lys	Leu	Val	Ser
				85					90					95	
Glu	Phe	Ser	Lys	Ala	Gly	Leu	Tyr	Glu	Lys	Lys	Gly	His	Gly	His	Leu
			100					105					110		
Ile	Arg	Leu	Leu	Val	Met	Ser	Leu	Val	Phe	Ile	Ala	Ser	Val	Ser	Gly
		115					120					125			
Val	Val	Leu	Ser	Asp	Lys	Thr	Ser	Val	His	Val	Gly	Ser	Ala	Val	Leu
	130					135					140				
Leu	Ala	Val	Ile	Trp	Met	Gln	Phe	Gly	Phe	Ile	Gly	His	Asp	Ser	Gly
145					150					155					160
His	Tyr	Asn	Ile	Met	Thr	Ser	Pro	Glu	Leu	Asn	Arg	Tyr	Met	Gln	Ile
			165					170						175	
Phe	Ser	Val	Asn	Val	Val	Ser	Gly	Val	Ser	Val	Gly	Trp	Trp	Lys	Arg
			180					185					190		
Tyr	His	Asn	Ala	His	His	Ile	Ala	Val	Asn	Ser	Leu	Glu	Tyr	Asp	Pro
		195					200					205			
Asp	Leu	Gln	Tyr	Val	Pro	Phe	Leu	Val	Val	Ser	Thr	Ala	Ile	Phe	Asp
	210					215					220				
Ser	Leu	Thr	Ser	His	Phe	Tyr	Arg	Lys	Lys	Met	Thr	Phe	Asp	Ala	Val
225					230					235					240
Ala	Arg	Phe	Leu	Val	Ser	Phe	Gln	His	Trp	Thr	Phe	Tyr	Pro	Leu	Met
				245					250					255	
Ala	Ile	Gly	Arg	Val	Ser	Phe	Leu	Ala	Gln	Ser	Ile	Gly	Val	Leu	Leu
			260					265					270		
Ser	Lys	Lys	Pro	Leu	Pro	Asp	Arg	His	Leu	Glu	Trp	Phe	Gly	Leu	Val
		275					280					285			
Val	Phe	Trp	Ala	Trp	Tyr	Ser	Leu	Leu	Ile	Ser	Cys	Leu	Pro	Asn	Trp
	290					295					300				
Trp	Glu	Arg	Val	Ile	Phe	Ile	Ala	Val	Asn	Phe	Ala	Val	Thr	Gly	Ile
305					310					315					320
Gln	His	Val	Gln	Phe	Cys	Leu	Asn	His	Tyr	Ser	Ala	Gln	Thr	Tyr	Ile
				325					330					335	
Gly	Ala	Pro	Cys	Ala	Asn	Asp	Trp	Phe	Glu	Lys	Gln	Thr	Lys	Gly	Ser
			340					345					350		
Ile	Asp	Ile	Ser	Cys	Ser	Pro	Trp	Thr	Asp	Trp	Phe	His	Gly	Gly	Leu
		355					360					365			
Gln	Phe	Gln	Ile	Glu	His	His	Leu	Phe	Pro	Arg	Met	Pro	Arg	Cys	Asn
					375						380				
Leu	Arg	Lys	Ile	Ser	Pro	Phe	Val	Lys	Glu	Leu	Cys	Arg	Lys	His	Asn
385					390					395					400
Leu	Val	Tyr	Thr	Ser	Val	Ser	Phe	Phe	Glu	Gly	Asn	Arg	Arg	Thr	Leu
				405					410					415	
Ala	Thr	Leu	Lys	Asn	Ala	Ala	Leu	Lys	Ala	Arg	Asp	Leu	Thr	Ser	Pro
				420				425					430		
Ile	Pro	Lys	Asn	Leu	Val	Trp	Glu	Ala	Val	His	Thr	His	Gly		
		435					440					445			

-continued

```

<210> SEQ ID NO 26
<211> LENGTH: 520
<212> TYPE: PRT
<213> ORGANISM: Ceratodon purpureus

<400> SEQUENCE: 26

Met Val Ser Gln Gly Gly Gly Leu Ser Gln Gly Ser Ile Glu Glu Asn
 1             5             10             15

Ile Asp Val Glu His Leu Ala Thr Met Pro Leu Val Ser Asp Phe Leu
 20             25             30

Asn Val Leu Gly Thr Thr Leu Gly Gln Trp Ser Leu Ser Thr Thr Phe
 35             40             45

Ala Phe Lys Arg Leu Thr Thr Lys Lys His Ser Ser Asp Ile Ser Val
 50             55             60

Glu Ala Gln Lys Glu Ser Val Ala Arg Gly Pro Val Glu Asn Ile Ser
 65             70             75             80

Gln Ser Val Ala Gln Pro Ile Arg Arg Arg Trp Val Gln Asp Lys Lys
 85             90             95

Pro Val Thr Tyr Ser Leu Lys Asp Val Ala Ser His Asp Met Pro Gln
100            105            110

Asp Cys Trp Ile Ile Ile Lys Glu Lys Val Tyr Asp Val Ser Thr Phe
115            120            125

Ala Glu Gln His Pro Gly Gly Thr Val Ile Asn Thr Tyr Phe Gly Arg
130            135            140

Asp Ala Thr Asp Val Phe Ser Thr Phe His Ala Ser Thr Ser Trp Lys
145            150            155            160

Ile Leu Gln Asn Phe Tyr Ile Gly Asn Leu Val Arg Glu Glu Pro Thr
165            170            175

Leu Glu Leu Leu Lys Glu Tyr Arg Glu Leu Arg Ala Leu Phe Leu Arg
180            185            190

Glu Gln Leu Phe Lys Ser Ser Lys Ser Tyr Tyr Leu Phe Lys Thr Leu
195            200            205

Ile Asn Val Ser Ile Val Ala Thr Ser Ile Ala Ile Ile Ser Leu Tyr
210            215            220

Lys Ser Tyr Arg Ala Val Leu Leu Ser Ala Ser Leu Met Gly Leu Phe
225            230            235            240

Ile Gln Gln Cys Gly Trp Leu Ser His Asp Phe Leu His His Gln Val
245            250            255

Phe Glu Thr Arg Trp Leu Asn Asp Val Val Gly Tyr Val Val Gly Asn
260            265            270

Val Val Leu Gly Phe Ser Val Ser Trp Trp Lys Thr Lys His Asn Leu
275            280            285

His His Ala Ala Pro Asn Glu Cys Asp Gln Lys Tyr Thr Pro Ile Asp
290            295            300

Glu Asp Ile Asp Thr Leu Pro Ile Ile Ala Trp Ser Lys Asp Leu Leu
305            310            315            320

Ala Thr Val Glu Ser Lys Thr Met Leu Arg Val Leu Gln Tyr Gln His
325            330            335

Leu Phe Phe Leu Val Leu Leu Thr Phe Ala Arg Ala Ser Trp Leu Phe
340            345            350

Trp Ser Ala Ala Phe Thr Leu Arg Pro Glu Leu Thr Leu Gly Glu Lys

```

-continued

355	360	365
Leu Leu Glu Arg Gly Thr Met Ala Leu His Tyr Ile Trp Phe Asn Ser		
370	375	380
Val Ala Phe Tyr Leu Leu Pro Gly Trp Lys Pro Val Val Trp Met Val		
385	390	395
Val Ser Glu Leu Met Ser Gly Phe Leu Leu Gly Tyr Val Phe Val Leu		
405	410	415
Ser His Asn Gly Met Glu Val Tyr Asn Thr Ser Lys Asp Phe Val Asn		
420	425	430
Ala Gln Ile Ala Ser Thr Arg Asp Ile Lys Ala Gly Val Phe Asn Asp		
435	440	445
Trp Phe Thr Gly Gly Leu Asn Arg Gln Ile Glu His His Leu Phe Pro		
450	455	460
Thr Met Pro Arg His Asn Leu Asn Lys Ile Ser Pro His Val Glu Thr		
465	470	475
Leu Cys Lys Lys His Gly Leu Val Tyr Glu Asp Val Ser Met Ala Ser		
485	490	495
Gly Thr Tyr Arg Val Leu Lys Thr Leu Lys Asp Val Ala Asp Ala Ala		
500	505	510
Ser His Gln Gln Leu Ala Ala Ser		
515	520	

<210> SEQ ID NO 27

<211> LENGTH: 525

<212> TYPE: PRT

<213> ORGANISM: Physcomitrella patens

<400> SEQUENCE: 27

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

-continued

Ala	Leu	Phe	Leu	Arg	Glu	Gln	Leu	Phe	Lys	Ser	Ser	Lys	Leu	Tyr	Tyr
195						200			205						
Val	Met	Lys	Leu	Leu	Thr	Asn	Val	Ala	Ile	Phe	Ala	Ala	Ser	Ile	Ala
210						215			220						
Ile	Ile	Cys	Trp	Ser	Lys	Thr	Ile	Ser	Ala	Val	Leu	Ala	Ser	Ala	Cys
225						230			235			240			
Met	Met	Ala	Leu	Cys	Phe	Gln	Gln	Cys	Gly	Trp	Leu	Ser	His	Asp	Phe
			245						250			255			
Leu	His	Asn	Gln	Val	Phe	Glu	Thr	Arg	Trp	Leu	Asn	Glu	Val	Val	Gly
			260						265			270			
Tyr	Val	Ile	Gly	Asn	Ala	Val	Leu	Gly	Phe	Ser	Thr	Gly	Trp	Trp	Lys
275						280						285			
Glu	Lys	His	Asn	Leu	His	His	Ala	Ala	Pro	Asn	Glu	Cys	Asp	Gln	Thr
290						295						300			
Tyr	Gln	Pro	Ile	Asp	Glu	Asp	Ile	Asp	Thr	Leu	Pro	Leu	Ile	Ala	Trp
305						310						315			
Ser	Lys	Asp	Ile	Leu	Ala	Thr	Val	Glu	Asn	Lys	Thr	Phe	Leu	Arg	Ile
			325						330			335			
Leu	Gln	Tyr	Gln	His	Leu	Phe	Phe	Met	Gly	Leu	Leu	Phe	Phe	Ala	Arg
			340						345			350			
Gly	Ser	Trp	Leu	Phe	Trp	Ser	Trp	Arg	Tyr	Thr	Ser	Thr	Ala	Val	Leu
355						360						365			
Ser	Pro	Val	Asp	Arg	Leu	Leu	Glu	Lys	Gly	Thr	Val	Leu	Phe	His	Tyr
370						375						380			
Phe	Trp	Phe	Val	Gly	Thr	Ala	Cys	Tyr	Leu	Leu	Pro	Gly	Trp	Lys	Pro
385						390						395			
Leu	Val	Trp	Met	Ala	Val	Thr	Glu	Leu	Met	Ser	Gly	Met	Leu	Leu	Gly
			405						410			415			
Phe	Val	Phe	Val	Leu	Ser	His	Asn	Gly	Met	Glu	Val	Tyr	Asn	Ser	Ser
			420						425			430			
Lys	Glu	Phe	Val	Ser	Ala	Gln	Ile	Val	Ser	Thr	Arg	Asp	Ile	Lys	Gly
435						440						445			
Asn	Ile	Phe	Asn	Asp	Trp	Phe	Thr	Gly	Gly	Leu	Asn	Arg	Gln	Ile	Glu
450						455						460			
His	His	Leu	Phe	Pro	Thr	Met	Pro	Arg	His	Asn	Leu	Asn	Lys	Ile	Ala
465						470						475			
Pro	Arg	Val	Glu	Val	Phe	Cys	Lys	Lys	His	Gly	Leu	Val	Tyr	Glu	Asp
			485						490			495			
Val	Ser	Ile	Ala	Thr	Gly	Thr	Cys	Lys	Val	Leu	Lys	Ala	Leu	Lys	Glu
			500						505			510			
Val	Ala	Glu	Ala	Ala	Ala	Glu	Gln	His	Ala	Thr	Thr	Ser			
515						520						525			

```
<210> SEQ ID NO 28
<211> LENGTH: 457
<212> TYPE: PRT
<213> ORGANISM: Mortierella alpina
```

<400> SEQUENCE: 28

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

-continued

Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Glu Phe Val Pro
 35 40 45
 Asp His Pro Gly Gly Ser Val Ile Leu Thr His Val Gly Lys Asp Gly
 50 55 60
 Thr Asp Val Phe Asp Thr Phe His Pro Glu Ala Ala Trp Glu Thr Leu
 65 70 75 80
 Ala Asn Phe Tyr Val Gly Asp Ile Asp Glu Ser Asp Arg Ala Ile Lys
 85 90 95
 Asn Asp Asp Phe Ala Ala Glu Val Arg Lys Leu Arg Thr Leu Phe Gln
 100 105 110
 Ser Leu Gly Tyr Tyr Asp Ser Ser Lys Ala Tyr Tyr Ala Phe Lys Val
 115 120 125
 Ser Phe Asn Leu Cys Ile Trp Gly Leu Ser Thr Phe Ile Val Ala Lys
 130 135 140
 Trp Gly Gln Thr Ser Thr Leu Ala Asn Val Leu Ser Ala Ala Leu Leu
 145 150 155 160
 Gly Leu Phe Trp Gln Gln Cys Gly Trp Leu Ala His Asp Phe Leu His
 165 170 175
 His Gln Val Phe Gln Asp Arg Phe Trp Gly Asp Leu Phe Gly Ala Phe
 180 185 190
 Leu Gly Gly Val Cys Gln Gly Phe Ser Ser Ser Trp Trp Lys Asp Lys
 195 200 205
 His Asn Thr His His Ala Ala Pro Asn Val His Gly Glu Asp Pro Asp
 210 215 220
 Ile Asp Thr His Pro Leu Leu Thr Trp Ser Glu His Ala Leu Glu Met
 225 230 235 240
 Phe Ser Asp Val Pro Asp Glu Glu Leu Thr Arg Met Trp Ser Arg Phe
 245 250 255
 Met Val Leu Asn Gln Thr Trp Phe Tyr Phe Pro Ile Leu Ser Phe Ala
 260 265 270
 Arg Leu Ser Trp Cys Leu Gln Ser Ile Met Phe Val Leu Pro Asn Gly
 275 280 285
 Gln Ala His Lys Pro Ser Gly Ala Arg Val Pro Ile Ser Leu Val Glu
 290 295 300
 Gln Leu Ser Leu Ala Met His Trp Thr Trp Tyr Leu Ala Thr Met Phe
 305 310 315 320
 Leu Phe Ile Lys Asp Pro Val Asn Met Ile Val Tyr Phe Leu Val Ser
 325 330 335
 Gln Ala Val Cys Gly Asn Leu Leu Ala Ile Val Phe Ser Leu Asn His
 340 345 350
 Asn Gly Met Pro Val Ile Ser Lys Glu Glu Ala Val Asp Met Asp Phe
 355 360 365
 Phe Thr Lys Gln Ile Ile Thr Gly Arg Asp Val His Pro Gly Leu Phe
 370 375 380
 Ala Asn Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Glu His His Leu
 385 390 395 400
 Phe Pro Ser Met Pro Arg His Asn Phe Ser Lys Ile Gln Pro Ala Val
 405 410 415
 Glu Thr Leu Cys Lys Lys Tyr Gly Val Arg Tyr His Thr Thr Gly Met
 420 425 430

-continued

Ile Glu Gly Thr Ala Glu Val Phe Ser Arg Leu Asn Glu Val Ser Lys
 435 440 445

Ala Ala Ser Lys Met Gly Lys Ala Gln
 450 455

<210> SEQ ID NO 29

<211> LENGTH: 443

<212> TYPE: PRT

<213> ORGANISM: *Caenorhabditis elegans*

<400> SEQUENCE: 29

Met Val Val Asp Lys Asn Ala Ser Gly Leu Arg Met Lys Val Asp Gly
 1 5 10 15

Lys Trp Leu Tyr Leu Ser Glu Glu Leu Val Lys Lys His Pro Gly Gly
 20 25 30

Ala Val Ile Glu Gln Tyr Arg Asn Ser Asp Ala Thr His Ile Phe His
 35 40 45

Ala Phe His Glu Gly Ser Ser Gln Ala Tyr Lys Gln Leu Asp Leu Leu
 50 55 60

Lys Lys His Gly Glu His Asp Glu Phe Leu Glu Lys Gln Leu Glu Lys
 65 70 75 80

Arg Leu Asp Lys Val Asp Ile Asn Val Ser Ala Tyr Asp Val Ser Val
 85 90 95

Ala Gln Glu Lys Lys Met Val Glu Ser Phe Glu Lys Leu Arg Gln Lys
 100 105 110

Leu His Asp Asp Gly Leu Met Lys Ala Asn Glu Thr Tyr Phe Leu Phe
 115 120 125

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

Tyr Leu Gly Trp Tyr Ile Thr Ser Ala Cys Leu Leu Ala Leu Ala Trp
 145 150 155 160

Gln Gln Phe Gly Trp Leu Thr His Glu Phe Cys His Gln Gln Pro Thr
 165 170 175

Lys Asn Arg Pro Leu Asn Asp Thr Ile Ser Leu Phe Phe Gly Asn Phe
 180 185 190

Leu Gln Gly Phe Ser Arg Asp Trp Trp Lys Asp Lys His Asn Thr His
 195 200 205

His Ala Ala Thr Asn Val Ile Asp His Asp Gly Asp Ile Asp Leu Ala
 210 215 220

Pro Leu Phe Ala Phe Ile Pro Gly Asp Leu Cys Lys Tyr Lys Ala Ser
 225 230 235 240

Phe Glu Lys Ala Ile Leu Lys Ile Val Pro Tyr Gln His Leu Tyr Phe
 245 250 255

Thr Ala Met Leu Pro Met Leu Arg Phe Ser Trp Thr Gly Gln Ser Val
 260 265 270

Gln Trp Val Phe Lys Glu Asn Gln Met Glu Tyr Lys Val Tyr Gln Arg
 275 280 285

Asn Ala Phe Trp Glu Gln Ala Thr Ile Val Gly His Trp Ala Trp Val
 290 295 300

Phe Tyr Gln Leu Phe Leu Leu Pro Thr Trp Pro Leu Arg Val Ala Tyr
 305 310 315 320

Phe Ile Ile Ser Gln Met Gly Gly Gly Leu Leu Ile Ala His Val Val
 325 330 335

-continued

Thr Phe Asn His Asn Ser Val Asp Lys Tyr Pro Ala Asn Ser Arg Ile
 340 345 350

Leu Asn Asn Phe Ala Ala Leu Gln Ile Leu Thr Thr Arg Asn Met Thr
 355 360 365

Pro Ser Pro Phe Ile Asp Trp Leu Trp Gly Gly Leu Asn Tyr Gln Ile
 370 375 380

Glu His His Leu Phe Pro Thr Met Pro Arg Cys Asn Leu Asn Ala Cys
 385 390 395 400

Val Lys Tyr Val Lys Glu Trp Cys Lys Glu Asn Asn Leu Pro Tyr Leu
 405 410 415

Val Asp Asp Tyr Phe Asp Gly Tyr Ala Met Asn Leu Gln Gln Leu Lys
 420 425 430

Asn Met Ala Glu His Ile Gln Ala Lys Ala Ala
 435 440

<210> SEQ ID NO 30

<211> LENGTH: 299

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 30

Met Glu His Phe Asp Ala Ser Leu Ser Thr Tyr Phe Lys Ala Leu Leu
 1 5 10 15

Gly Pro Arg Asp Thr Arg Val Lys Gly Trp Phe Leu Leu Asp Asn Tyr
 20 25 30

Ile Pro Thr Phe Ile Cys Ser Val Ile Tyr Leu Leu Ile Val Trp Leu
 35 40 45

Gly Pro Lys Tyr Met Arg Asn Lys Gln Pro Phe Ser Cys Arg Gly Ile
 50 55 60

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

Cys Glu Leu Val Thr Gly Val Trp Glu Gly Lys Tyr Asn Phe Phe Cys
 85 90 95

Gln Gly Thr Arg Thr Ala Gly Glu Ser Asp Met Lys Ile Ile Arg Val
 100 105 110

Leu Trp Trp Tyr Tyr Phe Ser Lys Leu Ile Glu Phe Met Asp Thr Phe
 115 120 125

Phe Phe Ile Leu Arg Lys Asn Asn His Gln Ile Thr Val Leu His Val
 130 135 140

Tyr His His Ala Ser Met Leu Asn Ile Trp Trp Phe Val Met Asn Trp
 145 150 155 160

Val Pro Cys Gly His Ser Tyr Phe Gly Ala Thr Leu Asn Ser Phe Ile
 165 170 175

His Val Leu Met Tyr Ser Tyr Tyr Gly Leu Ser Ser Val Pro Ser Met
 180 185 190

Arg Pro Tyr Leu Trp Trp Lys Lys Tyr Ile Thr Gln Gly Gln Leu Leu
 195 200 205

Gln Phe Val Leu Thr Ile Ile Gln Thr Ser Cys Gly Val Ile Trp Pro
 210 215 220

Cys Thr Phe Pro Leu Gly Trp Leu Tyr Phe Gln Ile Gly Tyr Met Ile
 225 230 235 240

Ser Leu Ile Ala Leu Phe Thr Asn Phe Tyr Ile Gln Thr Tyr Asn Lys

-continued

245					250					255					
Lys	Gly	Ala	Ser	Arg	Arg	Lys	Asp	His	Leu	Lys	Asp	His	Gln	Asn	Gly
			260					265					270		
Ser	Met	Ala	Ala	Val	Asn	Gly	His	Thr	Asn	Ser	Phe	Ser	Pro	Leu	Glu
		275					280					285			
Asn	Asn	Val	Lys	Pro	Arg	Lys	Leu	Arg	Lys	Asp					
	290					295									

<210> SEQ ID NO 31

<211> LENGTH: 290

<212> TYPE: PRT

<213> ORGANISM: Physcomitrella patens

<400> SEQUENCE: 31

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

-continued

<210> SEQ ID NO 32
 <211> LENGTH: 318
 <212> TYPE: PRT
 <213> ORGANISM: Mortierella alpina

<400> SEQUENCE: 32

```

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

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

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

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

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

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

Asn Phe Glu Arg Phe Glu Val Lys Thr Phe Ser Leu Leu His Asn Phe
100         105         110

Cys Leu Val Ser Ile Ser Ala Tyr Met Cys Gly Gly Ile Leu Tyr Glu
115         120         125

Ala Tyr Gln Ala Asn Tyr Gly Leu Phe Glu Asn Ala Ala Asp His Thr
130         135         140

Phe Lys Gly Leu Pro Met Ala Lys Met Ile Trp Leu Phe Tyr Phe Ser
145         150         155         160

Lys Ile Met Glu Phe Val Asp Thr Met Ile Met Val Leu Lys Lys Asn
165         170         175

Asn Arg Gln Ile Ser Phe Leu His Val Tyr His His Ser Ser Ile Phe
180         185         190

Thr Ile Trp Trp Leu Val Thr Phe Val Ala Pro Asn Gly Glu Ala Tyr
195         200         205

Phe Ser Ala Ala Leu Asn Ser Phe Ile His Val Ile Met Tyr Gly Tyr
210         215         220

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

Tyr Ile Thr Arg Ser Gln Met Thr Gln Phe Cys Met Met Ser Val Gln
245         250         255

Ser Ser Trp Asp Met Tyr Ala Met Lys Val Leu Gly Arg Pro Gly Tyr
260         265         270

Pro Phe Phe Ile Thr Ala Leu Leu Trp Phe Tyr Met Trp Thr Met Leu
275         280         285

Gly Leu Phe Tyr Asn Phe Tyr Arg Lys Asn Ala Lys Leu Ala Lys Gln
290         295         300

Ala Lys Ala Asp Ala Ala Lys Glu Lys Ala Arg Lys Leu Gln
305         310         315

```

<210> SEQ ID NO 33
 <211> LENGTH: 519
 <212> TYPE: PRT
 <213> ORGANISM: Thraustochytrium sp.

<400> SEQUENCE: 33

-continued

Met	Thr	Val	Gly	Tyr	Asp	Glu	Glu	Ile	Pro	Phe	Glu	Gln	Val	Arg	Ala	1	5	10	15
His	Asn	Lys	Pro	Asp	Asp	Ala	Trp	Cys	Ala	Ile	His	Gly	His	Val	Tyr	20	25	30	
Asp	Val	Thr	Lys	Phe	Ala	Ser	Val	His	Pro	Gly	Gly	Asp	Ile	Ile	Leu	35	40	45	
Leu	Ala	Ala	Gly	Lys	Glu	Ala	Thr	Val	Leu	Tyr	Glu	Thr	Tyr	His	Val	50	55	60	
Arg	Gly	Val	Ser	Asp	Ala	Val	Leu	Arg	Lys	Tyr	Arg	Ile	Gly	Lys	Leu	65	70	75	80
Pro	Asp	Gly	Gln	Gly	Gly	Ala	Asn	Glu	Lys	Glu	Lys	Arg	Thr	Leu	Ser	85	90	95	
Gly	Leu	Ser	Ser	Ala	Ser	Tyr	Tyr	Thr	Trp	Asn	Ser	Asp	Phe	Tyr	Arg	100	105	110	
Val	Met	Arg	Glu	Arg	Val	Val	Ala	Arg	Leu	Lys	Glu	Arg	Gly	Lys	Ala	115	120	125	
Arg	Arg	Gly	Gly	Tyr	Glu	Leu	Trp	Ile	Lys	Ala	Phe	Leu	Leu	Leu	Val	130	135	140	
Gly	Phe	Trp	Ser	Ser	Leu	Tyr	Trp	Met	Cys	Thr	Leu	Asp	Pro	Ser	Phe	145	150	155	160
Gly	Ala	Ile	Leu	Ala	Ala	Met	Ser	Leu	Gly	Val	Phe	Ala	Ala	Phe	Val	165	170	175	
Gly	Thr	Cys	Ile	Gln	His	Asp	Gly	Asn	His	Gly	Ala	Phe	Ala	Gln	Ser	180	185	190	
Arg	Trp	Val	Asn	Lys	Val	Ala	Gly	Trp	Thr	Leu	Asp	Met	Ile	Gly	Ala	195	200	205	
Ser	Gly	Met	Thr	Trp	Glu	Phe	Gln	His	Val	Leu	Gly	His	His	Pro	Tyr	210	215	220	
Thr	Asn	Leu	Ile	Glu	Glu	Glu	Asn	Gly	Leu	Gln	Lys	Val	Ser	Gly	Lys	225	230	235	240
Lys	Met	Asp	Thr	Lys	Leu	Ala	Asp	Gln	Glu	Ser	Asp	Pro	Asp	Val	Phe	245	250	255	
Ser	Thr	Tyr	Pro	Met	Met	Arg	Leu	His	Pro	Trp	His	Gln	Lys	Arg	Trp	260	265	270	
Tyr	His	Arg	Phe	Gln	His	Ile	Tyr	Gly	Pro	Phe	Ile	Phe	Gly	Phe	Met	275	280	285	
Thr	Ile	Asn	Lys	Val	Val	Thr	Gln	Asp	Val	Gly	Val	Val	Leu	Arg	Lys	290	295	300	
Arg	Leu	Phe	Gln	Ile	Asp	Ala	Glu	Cys	Arg	Tyr	Ala	Ser	Pro	Met	Tyr	305	310	315	320
Val	Ala	Arg	Phe	Trp	Ile	Met	Lys	Ala	Leu	Thr	Val	Leu	Tyr	Met	Val	325	330	335	
Ala	Leu	Pro	Cys	Tyr	Met	Gln	Gly	Pro	Trp	His	Gly	Leu	Lys	Leu	Phe	340	345	350	
Ala	Ile	Ala	His	Phe	Thr	Cys	Gly	Glu	Val	Leu	Ala	Thr	Met	Phe	Ile	355	360	365	
Val	Asn	His	Ile	Ile	Glu	Gly	Val	Ser	Tyr	Ala	Ser	Lys	Asp	Ala	Val	370	375	380	
Lys	Gly	Thr	Met	Ala	Pro	Pro	Lys	Thr	Met	His	Gly	Val	Thr	Pro	Met	385	390	395	400
Asn	Asn	Thr	Arg	Lys	Glu	Val	Glu	Ala	Glu	Ala	Ser	Lys	Ser	Gly	Ala				

-continued

405					410					415					
Val	Val	Lys	Ser	Val	Pro	Leu	Asp	Asp	Trp	Ala	Val	Val	Gln	Cys	Gln
				420					425					430	
Thr	Ser	Val	Asn	Trp	Ser	Val	Gly	Ser	Trp	Phe	Trp	Asn	His	Phe	Ser
				435					440					445	
Gly	Gly	Leu	Asn	His	Gln	Ile	Glu	His	His	Leu	Phe	Pro	Gly	Leu	Ser
				450					455					460	
His	Glu	Thr	Tyr	Tyr	His	Ile	Gln	Asp	Val	Phe	Gln	Ser	Thr	Cys	Ala
				465					470					475	480
Glu	Tyr	Gly	Val	Pro	Tyr	Gln	His	Glu	Pro	Ser	Leu	Trp	Thr	Ala	Tyr
				485					490					495	
Trp	Lys	Met	Leu	Glu	His	Leu	Arg	Gln	Leu	Gly	Asn	Glu	Glu	Thr	His
				500					505					510	
Glu	Ser	Trp	Gln	Arg	Ala	Ala									
				515											

<210> SEQ ID NO 34

<211> LENGTH: 541

<212> TYPE: PRT

<213> ORGANISM: Euglena gracilis

<400> SEQUENCE: 34

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

-continued

```

Leu Val Asn Arg Leu Ala Gly Trp Gly Met Asp Leu Ile Gly Ala Ser
      245                      250                      255
Ser Thr Val Trp Glu Tyr Gln His Val Ile Gly His His Gln Tyr Thr
      260                      265                      270
Asn Leu Val Ser Asp Thr Leu Phe Ser Leu Pro Glu Asn Asp Pro Asp
      275                      280                      285
Val Phe Ser Ser Tyr Pro Leu Met Arg Met His Pro Asp Thr Ala Trp
      290                      295                      300
Gln Pro His His Arg Phe Gln His Leu Phe Ala Phe Pro Leu Phe Ala
      305                      310                      315                      320
Leu Met Thr Ile Ser Lys Val Leu Thr Ser Asp Phe Ala Val Cys Leu
      325                      330                      335
Ser Met Lys Lys Gly Ser Ile Asp Cys Ser Ser Arg Leu Val Pro Leu
      340                      345                      350
Glu Gly Gln Leu Leu Phe Trp Gly Ala Lys Leu Ala Asn Phe Leu Leu
      355                      360                      365
Gln Ile Val Leu Pro Cys Tyr Leu His Gly Thr Ala Met Gly Leu Ala
      370                      375                      380
Leu Phe Ser Val Ala His Leu Val Ser Gly Glu Tyr Leu Ala Ile Cys
      385                      390                      395                      400
Phe Ile Ile Asn His Ile Ser Glu Ser Cys Glu Phe Met Asn Thr Ser
      405                      410                      415
Phe Gln Thr Ala Ala Arg Arg Thr Glu Met Leu Gln Ala Ala His Gln
      420                      425                      430
Ala Ala Glu Ala Lys Lys Val Lys Pro Thr Pro Pro Pro Asn Asp Trp
      435                      440                      445
Ala Val Thr Gln Val Gln Cys Cys Val Asn Trp Arg Ser Gly Gly Val
      450                      455                      460
Leu Ala Asn His Leu Ser Gly Gly Leu Asn His Gln Ile Glu His His
      465                      470                      475                      480
Leu Phe Pro Ser Ile Ser His Ala Asn Tyr Pro Thr Ile Ala Pro Val
      485                      490                      495
Val Lys Glu Val Cys Glu Glu Tyr Gly Leu Pro Tyr Lys Asn Tyr Val
      500                      505                      510
Thr Phe Trp Asp Ala Val Cys Gly Met Val Gln His Leu Arg Leu Met
      515                      520                      525
Gly Ala Pro Pro Val Pro Thr Asn Gly Asp Lys Lys Ser
      530                      535                      540

```

<210> SEQ ID NO 35

<211> LENGTH: 263

<212> TYPE: PRT

<213> ORGANISM: Isochrysis galbana

<400> SEQUENCE: 35

```

Met Ala Leu Ala Asn Asp Ala Gly Glu Arg Ile Trp Ala Ala Val Thr
  1                      5                      10                      15
Asp Pro Glu Ile Leu Ile Gly Thr Phe Ser Tyr Leu Leu Leu Lys Pro
      20                      25                      30
Leu Leu Arg Asn Ser Gly Leu Val Asp Glu Lys Lys Gly Ala Tyr Arg
      35                      40                      45
Thr Ser Met Ile Trp Tyr Asn Val Leu Leu Ala Leu Phe Ser Ala Leu
      50                      55                      60

```

-continued

Ser Phe Tyr Val Thr Ala Thr Ala Leu Gly Trp Asp Tyr Gly Thr Gly
 65 70 75 80
 Ala Trp Leu Arg Arg Gln Thr Gly Asp Thr Pro Gln Pro Leu Phe Gln
 85 90 95
 Cys Pro Ser Pro Val Trp Asp Ser Lys Leu Phe Thr Trp Thr Ala Lys
 100 105 110
 Ala Phe Tyr Tyr Ser Lys Tyr Val Glu Tyr Leu Asp Thr Ala Trp Leu
 115 120 125
 Val Leu Lys Gly Lys Arg Val Ser Phe Leu Gln Ala Phe His His Phe
 130 135 140
 Gly Ala Pro Trp Asp Val Tyr Leu Gly Ile Arg Leu His Asn Glu Gly
 145 150 155 160
 Val Trp Ile Phe Met Phe Phe Asn Ser Phe Ile His Thr Ile Met Tyr
 165 170 175
 Thr Tyr Tyr Gly Leu Thr Ala Ala Gly Tyr Lys Phe Lys Ala Lys Pro
 180 185 190
 Leu Ile Thr Ala Met Gln Ile Cys Gln Phe Val Gly Gly Phe Leu Leu
 195 200 205
 Val Trp Asp Tyr Ile Asn Val Pro Cys Phe Asn Ser Asp Lys Gly Lys
 210 215 220
 Leu Phe Ser Trp Ala Phe Asn Tyr Ala Tyr Val Gly Ser Val Phe Leu
 225 230 235 240
 Leu Phe Cys His Phe Phe Tyr Gln Asp Asn Leu Ala Thr Lys Lys Ser
 245 250 255
 Ala Lys Ala Gly Lys Gln Leu
 260

<210> SEQ ID NO 36
 <211> LENGTH: 419
 <212> TYPE: PRT
 <213> ORGANISM: *Euglena gracilis*

<400> SEQUENCE: 36

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

-continued

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

<210> SEQ ID NO 37

<211> LENGTH: 867

<212> TYPE: DNA

<213> ORGANISM: Caenorhabditis elegans

<400> SEQUENCE: 37

atggctcagc atccgctcgt tcaacggctt ctcgatgtca aattcgacac gaaacgattt	60
gtggctattg ctactcatgg gccaaagaat ttccctgacg cagaaggctc caagttcttt	120
gctgatcaact ttgatgttac tattcaggct tcaatcctgt acatggctcgt tgtgttcgga	180
acaaaatggt tcatgcgtaa tcgtcaacca ttccaattga ctattccact caacatctgg	240
aatttcaccc tcgcccatt ttccatcgca ggagctgtca aaatgacccc agagttcttt	300
ggaaccattg ccaacaaagg aattgtcgca tcctactgca aagtgtttga tttcacgaaa	360
ggagagaatg gatactgggt gtggctcttc atggcttcca aacttttcga acttgttgac	420
accatcttct tggttctccg taaacgtcca ctcatgttcc ttcactggta tcaccatatt	480

-continued

ctcaccatga tctacgcctg gtactctcat ccattgaccc caggattcaa cagatacgga	540
atztatctta actttgtcgt ccaagccttc atgtactctt actacttctt tcgctcgatg	600
aagattcgcg tgcaggatt catcgcccaa gctatcacat ctcttcaaat cgttcaattc	660
atcatctctt gcgcgttct tgctcatctt ggttatctca tgcacttcac caatgccaac	720
tgtgatttcg agccatcagt attcaagctc gcagttttca tggacacaac atacttggct	780
cttttcgtca acttcttctt ccaatcatat gttctccgcy gaggaaaaga caagtacaag	840
gcagtgccaa agaagaagaa caactaa	867

<210> SEQ ID NO 38
 <211> LENGTH: 1335
 <212> TYPE: DNA
 <213> ORGANISM: Danio rerio

<400> SEQUENCE: 38

atgggtggcg gaggacagca gacagaccga atcaccgaca ccaacggcag attcagcagc	60
tacacctggg aggaggtgca gaaacacacc aaacatggag atcagtgggg ggtgggggag	120
aggaagggtt ataacgtcag ccagtggttg aagagacacc ccggaggact gaggatcctc	180
ggacactatg ctggagaaga cgccacggag gcgttcactg cgtttcatcc aaaccttcag	240
ctggtgagga aatacctgaa gccgctgcta atcggagagc tggaggcgct tgaacccagt	300
caggaccggc agaaaaacgc tgctctcgtg gaggatttcc gagccctgcy tgagcgtctg	360
gaggctgaag gctgttttaa aacgcagccg ctgtttttcg ctctgcattt gggccacatt	420
ctgctcctgg aggccatcgc ttctcatgat gtgtggtatt tcggcaccgg ttggatcaac	480
acgctcatcg tcgctgttat tctggctact gcacagtcac aagctggatg gttgcagcat	540
gacttcggtc atctgtccgt gtttaaaacc tctggaatga atcatttggg gcacaaattt	600
gtcatcggac acctgaaggg agcgtctgcy ggctggtgga accatcggca cttccagcat	660
cacgctaaac ccaacatctt caagaaggac ccggacgtca acatgctgaa cgcctttgtg	720
gtgggaaacg tgcagccctg ggagtatggc gttaagaaga tcaagcatct gccctacaac	780
catcagcaca agtacttctt ctctcattgt cctccctgcy tcatccagct gtatttcag	840
ttccaaatct ttcacaatat gatcagtcac ggcatgtggg tggacctgct gtggtgtatc	900
agctactacg tccgatactt cctttgttac acgcagttct acggcgtctt ttgggctatt	960
atcctcttta atttgcctag gtttatggag agccactggg ttgtttgggt cacacagatg	1020
agccacatcc ccatgaacat tgactatgag aaaaatcagg actggctcag catgcagctg	1080
gtcgcgacct gtaacatcga gcagtctgcc ttcaacgact ggttcagcgg acacctcaac	1140
ttccagatcg agcatcatct ctttcccaca gtgcctcggc acaactactg gcgcgcccgt	1200
ccacgggtgc gagcgttggt tgagaaatac ggagtcaaat accaagagaa gacctgtac	1260
ggagcatttg cggatatcat taggtctttg gagaaatctg gcgagctctg gctggatgcy	1320
tatctcaaca aataa	1335

<210> SEQ ID NO 39
 <211> LENGTH: 31
 <212> TYPE: DNA
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

-continued

<400> SEQUENCE: 39

cccaagctta ctatgggtgg cggaggacag c

31

<210> SEQ ID NO 40

<211> LENGTH: 27

<212> TYPE: DNA

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 40

ccgctggagt tatttgttga gatacgc

27

<210> SEQ ID NO 41

<211> LENGTH: 27

<212> TYPE: DNA

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 41

gcgggtacca tggctcagca tccgctc

27

<210> SEQ ID NO 42

<211> LENGTH: 27

<212> TYPE: DNA

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 42

gcgggatcct tagttgttct tcttctt

27

<210> SEQ ID NO 43

<211> LENGTH: 6

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Conserved region

<400> SEQUENCE: 43

Asp His Pro Gly Gly Ser

1

5

<210> SEQ ID NO 44

<211> LENGTH: 7

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Conserved region

<400> SEQUENCE: 44

Trp Trp Lys Asp Lys His Asn

1

5

<210> SEQ ID NO 45

<211> LENGTH: 7

<212> TYPE: PRT

<213> ORGANISM: Artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Conserved region

<400> SEQUENCE: 45

-continued

Gln Ile Glu His His Leu Phe
1 5

<210> SEQ ID NO 46
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 46

tggtggaarc ayaarcayaa y 21

<210> SEQ ID NO 47
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (19)..(19)
<223> OTHER INFORMATION: n = any nucleotide

<400> SEQUENCE: 47

gcgagggatc caaggraana rrtgrtgytc 30

<210> SEQ ID NO 48
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Conserved region
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (4)..(4)
<223> OTHER INFORMATION: X = any amino acid

<400> SEQUENCE: 48

Phe Leu His Xaa Tyr His
1 5

<210> SEQ ID NO 49
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Conserved region
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (3)..(3)
<223> OTHER INFORMATION: X = any amino acid

<400> SEQUENCE: 49

Met Tyr Xaa Tyr Tyr Phe
1 5

<210> SEQ ID NO 50
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (14)..(14)
<223> OTHER INFORMATION: n = any nucleotide

-continued

```

<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (18)..(18)
<223> OTHER INFORMATION: n = any nucleotide
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (19)..(19)
<223> OTHER INFORMATION: n = any nucleotide
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (20)..(20)
<223> OTHER INFORMATION: n = any nucleotide

<400> SEQUENCE: 50

caggatcctt yytncatnnn tayca                25

<210> SEQ ID NO 51
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (18)..(18)
<223> OTHER INFORMATION: n = any nucleotide
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (19)..(19)
<223> OTHER INFORMATION: n = any nucleotide
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (20)..(20)
<223> OTHER INFORMATION: n = any nucleotide

<400> SEQUENCE: 51

gatctagara artartannn rtacat                26

<210> SEQ ID NO 52
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Conserved region

<400> SEQUENCE: 52

His Pro Gly Gly
1

<210> SEQ ID NO 53
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 53

agcacgacgs sarccacggc g                    21

<210> SEQ ID NO 54
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 54

```

-continued

gtggtgcayc abcacgtgct 20

<210> SEQ ID NO 55
 <211> LENGTH: 642
 <212> TYPE: DNA
 <213> ORGANISM: Pavlova salina
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (13)..(13)
 <223> OTHER INFORMATION: n = unknown
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (39)..(39)
 <223> OTHER INFORMATION: n = unknown
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (51)..(51)
 <223> OTHER INFORMATION: n = unknown
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (77)..(77)
 <223> OTHER INFORMATION: n = unknown
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (302)..(302)
 <223> OTHER INFORMATION: n = unknown
 <220> FEATURE:
 <221> NAME/KEY: misc_feature
 <222> LOCATION: (639)..(639)
 <223> OTHER INFORMATION: n = unknown

<400> SEQUENCE: 55

```

ggctgcgcaa ctnttggaag ggcgatcggg gcgggcctnt tcgttattac nccagctggc 60
gaaaggggga tgtgctncaa ggcgattaag ttgggtaacg ccagggtttc ccagtcacga 120
cgttgtaaaa cgacggccag tgaattgtaa tacgactcac tatagggcga attgggtacc 180
gggccccccc tcgagaagtc gggtcgcac tcgcggggcg acaagatctg gcagattggc 240
tttggcagtg ggttcaagtg caactcggcc gtgtggcagg cgaacaggag cgttgagcca 300
tntgagctcg actgacgagc tcggagctgc ggtacagaca ctgtcggcgg ctcgagaggg 360
ctgcgacttc agacgtgatc gggagattgt gcattggtgc gccgccgggc gcggcctgcc 420
gccccggcgc tgcacgtcat cgtcagtagt caccgtcggc atcagcgccc ggcccggtgt 480
tggtagctgg tagcgcaggc tgcgcagctg ccaacagccg ccgcccagag tgggtggtgg 540
gactccgggt gtcagtcaca ctcagtggcg gccgccggca gtaggccgtg actctgccgt 600
ggcgtagta tcagtggcag tcagctgctg tcgtcaatnt tt 642
  
```

<210> SEQ ID NO 56
 <211> LENGTH: 100
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 56

```

tgggttgagt actcgcccaa ccacacgacc aactgcgcgc cctcgtggtg gtgcgactgg 60
tggatgtctt acctcaacta ccagatcgag catcatctgt 100
  
```

<210> SEQ ID NO 57
 <211> LENGTH: 100
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:

-continued

<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 57

atagtgcagc ccgtgcttct cgaagagcgc cttgacgcgc ggcgcgatcg tcgggtggcg 60

gaattgcggc atggacggga acagatgatg ctcgatctgg 100

<210> SEQ ID NO 58

<211> LENGTH: 1612

<212> TYPE: DNA

<213> ORGANISM: Pavlova salina

<400> SEQUENCE: 58

gccttctgga cgactgtcat gccgcgcgc gatagctact cgtacgcgc ccgcgcgtcg 60

gcccagctgc acgaggtcga taccgccag gagcatgata agaaggagct cgtcatcggt 120

gaccgcgcgt acgacgtgac caactttgtg aagcgcacc cgggtggcaa gatcatcgca 180

taccaggttg gcacagatgc gacggacgcg tacaagcagt tccatgtgcg gtctgccaa 240

gcggacaaga tgctcaagtc gctgccttcg cgcgcggtgc acaagggcta ctgcgccgc 300

cgcgctgacc tcattgccga cttccaggag ttcaccaagc agctggaggc ggaggcatg 360

tttgagccgt cgctgcgcga cgtggcatac cgcctggcgg aggtgatcgc gatgcacgtg 420

gccggcgccg cgctcatctg gcacgggtac accttcgcgg gcattgccat gctcggcggt 480

gtgcaggggc gctcggcgtg gctcatgcac gagggcgggc actactcgct caccggcaac 540

attgcttttg accgtgccat ccaagtgcg tgctacggcc ttggctcggg catgtcgggc 600

gcgtgggtgg gcaaccagca caacaagc caccgcgcgc cgcagaagtt gcagcacgac 660

gtcgacctcg acaccctccc gctcgtcgcc ttccacgagc ggatagccgc caaggtgaag 720

agcccccgga tgaaggcgtg gcttagtatg caggcgaagc tcttcgcgc agtgaccacg 780

ctgctggctg cgctgggctg gcagctgtac ctgcacccgc gccatatgct gcgcaccaag 840

cactacgacg agctcgcgat gctcggcatt cgctacggcc ttgtcggcta cctcgcggcg 900

aactacggcg cgggggtacgt gctcgcgtgc tacctgctgt acgtgcagct cggcgccatg 960

tacatcttct gcaactttgc cgtgtcgcac acacacctgc cggttgtcga gcctaacgag 1020

cacgcaacgt ggggtggagta cgcgcgaac cacacgacca actgctcgc ctcgtgggtg 1080

tgcgactggt ggatgtcgtg cctcaactac cagatcgagc accacctcta cccgtccatg 1140

ccgcagttcc gccacccgaa gattgcgcgc cgggtgaagc agctcttcga gaagcacggc 1200

ctgcactacg acgtgcgtgg ctacttcgag gccatggcgg acacgtttgc caaccttgac 1260

aacgtcgcgc acgcgccgga gaagaagatg cagtgcgcgc gcgagtgcgc aacgccaagc 1320

gtccacccgc gagtcgccg tggtcctcct gccgatcgcg gcctgtctct ccagctgaca 1380

tctctgcttt gcttgcccat gatcgacgtc gcctccctct ctctccacac gatgtgcctg 1440

acgaatgacc tgcgggataa tcagcgctcg catgcccatg ccagcgccaa tggcagctgc 1500

tgccgcagcc gccaaagtgt aatccatgac acgtgctcc acgacgcgc cgccttccat 1560

cttgacaatc agcatggacg tagcatcatc agttcagtga ctaattcctt tc 1612

<210> SEQ ID NO 59

<211> LENGTH: 1278

<212> TYPE: DNA

<213> ORGANISM: Pavlova salina

-continued

<400> SEQUENCE: 59

```

atgccgccgc gcgatagcta ctcgtagcc gccccgccgt cggcccagct gcacgaggtc   60
gataccccgc aggagcatga taagaaggag ctcgtagccg gtgaccgccg gtacgacgtg   120
accaactttg tgaagcgcca cccgggtggc aagatcatcg cataccaggt tggcacagat   180
gcgacggacg cgtacaagca gttccatgtg cggctgtcca aggcggacaa gatgctcaag   240
tcgctgcctt cgcgcccggt gcacaagggc tactcgcccc gccgcgctga cctcattgcc   300
gacttccagg agttcaccaa gcagctggag gcggagggca tgtttgagcc gtcgctgccg   360
cacgtggcat accgcctggc ggaggtgatc gcgatgcacg tggccggcgc cgcgctcatc   420
tggcacgggt acaccttcgc gggcattgcc atgctcgcgg ttgtgcaggg ccgctgcggc   480
tggctcatgc acgagggcgg ccactactcg ctcacgggca acattgcttt tgaccgtgcc   540
atccaagtcg cgtgctacgg ccttggtgc ggcattgctg gcgctgggtg gcgcaaccag   600
cacaacaagc accacgcgac gccgcagaag ttgcagcacg acgtcgacct cgacaccctc   660
ccgctcgtcg ccttccacga gcggatagcc gccaaaggtg agagccccgc gatgaaggcg   720
tggcttagta tgcaggcgaa gctcttcgcg ccagtgaaca cgctgctggt cgcgctgggc   780
tggcagctgt acctgcaccc gcgccatatg ctgcgcacca agcactacga cgagctcgcg   840
atgctcgcca ttcgctacgg ccttgctggc tacctcgcgg cgaactacgg cgcgggggtac   900
gtgctcgctg gctacctgct gtacgtgcag ctgcgcgcca tgtacatctt ctgcaacttt   960
gccgtgtcgc acacacacct gccggttgc gagcctaacg agcacgcaac gtgggtggag   1020
tacgcccga accacacgac caactgctcg ccctcgtggt ggtgcgactg gtggatgtcg   1080
tacctcaact accagatcga gcaccacctc taccctgcca tgccgcagtt ccgccaccgc   1140
aagattgcgc cgcgggtgaa gcagctcttc gagaagcacg gcctgcacta cgacgtgcgt   1200
ggctacttcg aggccatggc ggacacgttt gccaaccttg acaacgtcgc gcacgcgccg   1260
gagaagaaga tgcagtga                                     1278

```

<210> SEQ ID NO 60

<211> LENGTH: 425

<212> TYPE: PRT

<213> ORGANISM: Pavlova salina

<400> SEQUENCE: 60

```

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

```

-continued

115					120					125						
Val	Ile	Ala	Met	His	Val	Ala	Gly	Ala	Ala	Leu	Ile	Trp	His	Gly	Tyr	
130					135					140						
Thr	Phe	Ala	Gly	Ile	Ala	Met	Leu	Gly	Val	Val	Gln	Gly	Arg	Cys	Gly	
145					150					155					160	
Trp	Leu	Met	His	Glu	Gly	Gly	His	Tyr	Ser	Leu	Thr	Gly	Asn	Ile	Ala	
				165					170					175		
Phe	Asp	Arg	Ala	Ile	Gln	Val	Ala	Cys	Tyr	Gly	Leu	Gly	Cys	Gly	Met	
			180					185					190			
Ser	Gly	Ala	Trp	Trp	Arg	Asn	Gln	His	Asn	Lys	His	His	Ala	Thr	Pro	
			195				200					205				
Gln	Lys	Leu	Gln	His	Asp	Val	Asp	Leu	Asp	Thr	Leu	Pro	Leu	Val	Ala	
		210				215					220					
Phe	His	Glu	Arg	Ile	Ala	Ala	Lys	Val	Lys	Ser	Pro	Ala	Met	Lys	Ala	
225					230					235					240	
Trp	Leu	Ser	Met	Gln	Ala	Lys	Leu	Phe	Ala	Pro	Val	Thr	Thr	Leu	Leu	
				245					250					255		
Val	Ala	Leu	Gly	Trp	Gln	Leu	Tyr	Leu	His	Pro	Arg	His	Met	Leu	Arg	
			260					265					270			
Thr	Lys	His	Tyr	Asp	Glu	Leu	Ala	Met	Leu	Gly	Ile	Arg	Tyr	Gly	Leu	
			275					280				285				
Val	Gly	Tyr	Leu	Ala	Ala	Asn	Tyr	Gly	Ala	Gly	Tyr	Val	Leu	Ala	Cys	
		290				295					300					
Tyr	Leu	Leu	Tyr	Val	Gln	Leu	Gly	Ala	Met	Tyr	Ile	Phe	Cys	Asn	Phe	
305					310					315					320	
Ala	Val	Ser	His	Thr	His	Leu	Pro	Val	Val	Glu	Pro	Asn	Glu	His	Ala	
				325					330					335		
Thr	Trp	Val	Glu	Tyr	Ala	Ala	Asn	His	Thr	Thr	Asn	Cys	Ser	Pro	Ser	
			340					345					350			
Trp	Trp	Cys	Asp	Trp	Trp	Met	Ser	Tyr	Leu	Asn	Tyr	Gln	Ile	Glu	His	
		355				360					365					
His	Leu	Tyr	Pro	Ser	Met	Pro	Gln	Phe	Arg	His	Pro	Lys	Ile	Ala	Pro	
		370				375					380					
Arg	Val	Lys	Gln	Leu	Phe	Glu	Lys	His	Gly	Leu	His	Tyr	Asp	Val	Arg	
385					390					395					400	
Gly	Tyr	Phe	Glu	Ala	Met	Ala	Asp	Thr	Phe	Ala	Asn	Leu	Asp	Asn	Val	
			405					410						415		
Ala	His	Ala	Pro	Glu	Lys	Lys	Met	Gln								
			420				425									

<210> SEQ ID NO 61
 <211> LENGTH: 8
 <212> TYPE: PRT
 <213> ORGANISM: Echium pitardii

<400> SEQUENCE: 61

Met Ala Asn Ala Ile Lys Lys Tyr
 1 5

<210> SEQ ID NO 62
 <211> LENGTH: 7
 <212> TYPE: PRT
 <213> ORGANISM: Echium pitardii

-continued

<400> SEQUENCE: 62

Glu Ala Leu Asn Thr His Gly
1 5

<210> SEQ ID NO 63

<211> LENGTH: 1347

<212> TYPE: DNA

<213> ORGANISM: Echium plantagineum

<400> SEQUENCE: 63

```

atggctaata caatcaagaa gtacattact gcagaggagc tgaagaagca tgataaagca    60
ggggatctct ggatctccat tcaaggaaaa atctatgatg ttccagattg gttgaaggac    120
catccagggtg ggaacttccc cttgctgagc cttgctggcc aagaggtaac tgatgcattt    180
gttgcatctc attctgttac aacttggaag cttcttgaaa aattcttcac tggttattac    240
cttaaagatt actctgtttc tgagggtgcc aaagattaca ggaagcttgt gtttgagttt    300
aataaaatgg gcttgtttga caaaaagggt catattgttc ttgtgactgt cttgtttata    360
gctatgttgt ttggtatgag tgtttatggg gttttgtttt gtgaggggtg tttggtacat    420
ttgcttgctg ggggggtgat gggttttgtc tggattcaga gtggttggat tggatcatgat    480
gctgggcatt atattgttat gcctgatgct aggcttaata agcttatggg tattgttgct    540
gccaattggt tatctggaat aagcattggt tggtggaagt ggaaccataa tgcacatcac    600
attgcctgta atagcctcga ttacgacccg gatttgcagt acattccgtt tcttgttgtg    660
tcgtccaagt tgtttagctc gctcacctct catttctatg aaaagaaact gacatttgac    720
tctttatcga gattctttgt aagccatcag cattggacgt tttacccggt tatgtgtatg    780
gctaggggta atatgtttgt gcagtctctg ataatgttgt tgactaagcg aaatgtgttc    840
tatagaagtc aagaactggt gggattggtg gtgttttgga tttgggtacc gtgcttgtt    900
tcttgcttgc ctaattgggg agaacgagta atgttcgttg ttgctagtct ctcggtgact    960
ggaatgcaac aagtgcagtt ctctttgaac catttctcgt cgagtgttta tgttggtcag    1020
cctaaggga acgattggtt cgagaacaa acatgtggga cgctcgacat ttcttgccct    1080
tcgtggatgg attggtttca tgggtgattg caattccaag ttgagcatca tttgttcct    1140
aagctgcccc gatgccacct tcggaaaatc tccccgttcg tgatggagtt atgcaagaag    1200
cataatttgt cttacaattg tgcattcttc tccgaggcca acaatatgac actcagaaca    1260
ttaagggaca cagcattgca agctcgcgat ttaaccaagc cgctcccca gaatttggt    1320
tggaagctc ttaatactca tgggtga    1347

```

<210> SEQ ID NO 64

<211> LENGTH: 448

<212> TYPE: PRT

<213> ORGANISM: Echium plantagineum

<400> SEQUENCE: 64

Met Ala Asn Ala Ile Lys Lys Tyr Ile Thr Ala Glu Glu Leu Lys Lys
1 5 10 15

His Asp Lys Ala Gly Asp Leu Trp Ile Ser Ile Gln Gly Lys Ile Tyr
20 25 30

Asp Val Ser Asp Trp Leu Lys Asp His Pro Gly Gly Asn Phe Pro Leu
35 40 45

-continued

Leu	Ser	Leu	Ala	Gly	Gln	Glu	Val	Thr	Asp	Ala	Phe	Val	Ala	Phe	His
50						55					60				
Ser	Gly	Thr	Thr	Trp	Lys	Leu	Leu	Glu	Lys	Phe	Phe	Thr	Gly	Tyr	Tyr
65					70					75					80
Leu	Lys	Asp	Tyr	Ser	Val	Ser	Glu	Val	Ser	Lys	Asp	Tyr	Arg	Lys	Leu
				85						90				95	
Val	Phe	Glu	Phe	Asn	Lys	Met	Gly	Leu	Phe	Asp	Lys	Lys	Gly	His	Ile
				100				105					110		
Val	Leu	Val	Thr	Val	Leu	Phe	Ile	Ala	Met	Leu	Phe	Gly	Met	Ser	Val
				115			120					125			
Tyr	Gly	Val	Leu	Phe	Cys	Glu	Gly	Val	Leu	Val	His	Leu	Leu	Ala	Gly
	130					135					140				
Gly	Leu	Met	Gly	Phe	Val	Trp	Ile	Gln	Ser	Gly	Trp	Ile	Gly	His	Asp
145					150					155					160
Ala	Gly	His	Tyr	Ile	Val	Met	Pro	Asp	Ala	Arg	Leu	Asn	Lys	Leu	Met
				165					170					175	
Gly	Ile	Val	Ala	Ala	Asn	Cys	Leu	Ser	Gly	Ile	Ser	Ile	Gly	Trp	Trp
				180				185					190		
Lys	Trp	Asn	His	Asn	Ala	His	His	Ile	Ala	Cys	Asn	Ser	Leu	Asp	Tyr
		195					200					205			
Asp	Pro	Asp	Leu	Gln	Tyr	Ile	Pro	Phe	Leu	Val	Val	Ser	Ser	Lys	Leu
	210					215					220				
Phe	Ser	Ser	Leu	Thr	Ser	His	Phe	Tyr	Glu	Lys	Lys	Leu	Thr	Phe	Asp
225					230					235					240
Ser	Leu	Ser	Arg	Phe	Phe	Val	Ser	His	Gln	His	Trp	Thr	Phe	Tyr	Pro
				245					250					255	
Val	Met	Cys	Met	Ala	Arg	Val	Asn	Met	Phe	Val	Gln	Ser	Leu	Ile	Met
			260					265					270		
Leu	Leu	Thr	Lys	Arg	Asn	Val	Phe	Tyr	Arg	Ser	Gln	Glu	Leu	Leu	Gly
		275					280					285			
Leu	Val	Val	Phe	Trp	Ile	Trp	Tyr	Pro	Leu	Leu	Val	Ser	Cys	Leu	Pro
	290					295					300				
Asn	Trp	Gly	Glu	Arg	Val	Met	Phe	Val	Val	Ala	Ser	Leu	Ser	Val	Thr
305					310					315					320
Gly	Met	Gln	Gln	Val	Gln	Phe	Ser	Leu	Asn	His	Phe	Ser	Ser	Ser	Val
				325					330					335	
Tyr	Val	Gly	Gln	Pro	Lys	Gly	Asn	Asp	Trp	Phe	Glu	Lys	Gln	Thr	Cys
		340						345					350		
Gly	Thr	Leu	Asp	Ile	Ser	Cys	Pro	Ser	Trp	Met	Asp	Trp	Phe	His	Gly
		355					360					365			
Gly	Leu	Gln	Phe	Gln	Val	Glu	His	His	Leu	Phe	Pro	Lys	Leu	Pro	Arg
	370					375					380				
Cys	His	Leu	Arg	Lys	Ile	Ser	Pro	Phe	Val	Met	Glu	Leu	Cys	Lys	Lys
385					390					395					400
His	Asn	Leu	Ser	Tyr	Asn	Cys	Ala	Ser	Phe	Ser	Glu	Ala	Asn	Asn	Met
				405					410					415	
Thr	Leu	Arg	Thr	Leu	Arg	Asp	Thr	Ala	Leu	Gln	Ala	Arg	Asp	Leu	Thr
				420				425					430		
Lys	Pro	Leu	Pro	Lys	Asn	Leu	Val	Trp	Glu	Ala	Leu	Asn	Thr	His	Gly
		435					440					445			

-continued

```

<210> SEQ ID NO 65
<211> LENGTH: 448
<212> TYPE: PRT
<213> ORGANISM: Echium gentianoides

<400> SEQUENCE: 65

Met Ala Asn Ala Ile Lys Lys Tyr Ile Thr Ala Glu Glu Leu Lys Lys
 1           5           10           15
His Asp Lys Glu Gly Asp Leu Trp Ile Ser Ile Gln Gly Lys Val Tyr
 20           25           30
Asp Val Ser Asp Trp Leu Lys Asp His Pro Gly Gly Lys Phe Pro Leu
 35           40           45
Leu Ser Leu Ala Gly Gln Glu Val Thr Asp Ala Phe Val Ala Phe His
 50           55           60
Ser Gly Ser Thr Trp Lys Phe Leu Asp Ser Phe Phe Thr Gly Tyr Tyr
 65           70           75           80
Leu Lys Asp Tyr Ser Val Ser Glu Val Ser Lys Asp Tyr Arg Lys Leu
 85           90           95
Val Phe Glu Phe Asn Lys Met Gly Leu Phe Asp Lys Lys Gly His Ile
100          105          110
Val Leu Val Thr Val Leu Phe Ile Ala Met Met Phe Ala Met Ser Val
115          120          125
Tyr Gly Val Leu Phe Cys Glu Gly Val Leu Val His Leu Leu Ala Gly
130          135          140
Gly Leu Met Gly Phe Val Trp Ile Gln Ser Gly Trp Ile Gly His Asp
145          150          155          160
Ala Gly His Tyr Ile Val Met Pro Asn Pro Arg Leu Asn Lys Leu Met
165          170          175
Gly Ile Val Ala Gly Asn Cys Leu Ser Gly Ile Ser Ile Gly Trp Trp
180          185          190
Lys Trp Asn His Asn Ala His His Ile Ala Cys Asn Ser Leu Asp Tyr
195          200          205
Asp Pro Asp Leu Gln Tyr Ile Pro Phe Leu Val Val Ser Ser Lys Leu
210          215          220
Phe Ser Ser Leu Thr Ser His Phe Tyr Glu Lys Lys Leu Thr Phe Asp
225          230          235          240
Ser Leu Ser Arg Phe Phe Val Ser His Gln His Trp Thr Phe Tyr Pro
245          250          255
Val Met Cys Ser Ala Arg Val Asn Met Phe Val Gln Ser Leu Ile Met
260          265          270
Leu Leu Thr Lys Arg Asn Val Phe Tyr Arg Ser Gln Glu Leu Leu Gly
275          280          285
Leu Val Val Phe Trp Ile Trp Tyr Pro Leu Leu Val Ser Cys Leu Pro
290          295          300
Asn Trp Gly Glu Arg Ile Met Phe Val Val Ala Ser Leu Ser Val Thr
305          310          315          320
Gly Met Gln Gln Val Gln Phe Ser Leu Asn His Phe Ser Ala Ser Val
325          330          335
Tyr Val Gly Gln Pro Lys Gly Asn Asp Trp Phe Glu Lys Gln Thr Cys
340          345          350
Gly Thr Leu Asp Ile Ser Cys Pro Ser Trp Met Asp Trp Phe His Gly
355          360          365

```

-continued

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

<210> SEQ ID NO 66
 <211> LENGTH: 448
 <212> TYPE: PRT
 <213> ORGANISM: Echium pitardii

<400> SEQUENCE: 66

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

-continued

```

Leu Val Val Phe Trp Ile Trp Tyr Pro Leu Leu Val Ser Cys Leu Pro
 290                295                300

Asn Trp Gly Glu Arg Ile Met Phe Val Val Ala Ser Leu Ser Val Thr
 305                310                315                320

Gly Leu Gln Gln Val Gln Phe Ser Leu Asn His Phe Ala Ala Ser Val
      325                330                335

Tyr Val Gly Gln Pro Lys Gly Ile Asp Trp Phe Glu Lys Gln Thr Cys
      340                345                350

Gly Thr Leu Asp Ile Ser Cys Pro Ser Trp Met Asp Trp Phe His Gly
      355                360                365

Gly Leu Gln Phe Gln Val Glu His His Leu Phe Pro Lys Leu Pro Arg
      370                375                380

Cys His Leu Arg Lys Ile Ser Pro Phe Val Met Glu Leu Cys Lys Lys
 385                390                395                400

His Asn Leu Ser Tyr Asn Cys Ala Ser Phe Ser Gln Ala Asn Glu Met
      405                410                415

Thr Leu Arg Thr Leu Arg Asp Thr Ala Leu Gln Ala Arg Asp Leu Thr
      420                425                430

Lys Pro Leu Pro Lys Asn Leu Val Trp Glu Ala Leu Asn Thr His Gly
      435                440                445

```

<210> SEQ ID NO 67

<211> LENGTH: 448

<212> TYPE: PRT

<213> ORGANISM: Borago officinalis

<400> SEQUENCE: 67

```

Met Ala Ala Gln Ile Lys Lys Tyr Ile Thr Ser Asp Glu Leu Lys Asn
 1                5                10                15

His Asp Lys Pro Gly Asp Leu Trp Ile Ser Ile Gln Gly Lys Ala Tyr
      20                25                30

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

Lys Ser Leu Ala Gly Gln Glu Val Thr Asp Ala Phe Val Ala Phe His
      50                55                60

Pro Ala Ser Thr Trp Lys Asn Leu Asp Lys Phe Phe Thr Gly Tyr Tyr
      65                70                75                80

Leu Lys Asp Tyr Ser Val Ser Glu Val Ser Lys Asp Tyr Arg Lys Leu
      85                90                95

Val Phe Glu Phe Ser Lys Met Gly Leu Tyr Asp Lys Lys Gly His Ile
      100               105                110

Met Phe Ala Thr Leu Cys Phe Ile Ala Met Leu Phe Ala Met Ser Val
      115               120                125

Tyr Gly Val Leu Phe Cys Glu Gly Val Leu Val His Leu Phe Ser Gly
      130               135                140

Cys Leu Met Gly Phe Leu Trp Ile Gln Ser Gly Trp Ile Gly His Asp
      145               150                155                160

Ala Gly His Tyr Met Val Val Ser Asp Ser Arg Leu Asn Lys Phe Met
      165               170                175

Gly Ile Phe Ala Ala Asn Cys Leu Ser Gly Ile Ser Ile Gly Trp Trp
      180               185                190

Lys Trp Asn His Asn Ala His His Ile Ala Cys Asn Ser Leu Glu Tyr

```

-continued

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

<210> SEQ ID NO 68

<211> LENGTH: 446

<212> TYPE: PRT

<213> ORGANISM: Borago officinalis

<400> SEQUENCE: 68

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

-continued

```

Ile Tyr Cys Leu Ser Phe Ile Ala Leu Leu Leu Cys Gly Cys Val Tyr
    115                120                125

Gly Val Leu Cys Ser Asn Ser Leu Trp Val His Met Leu Ser Gly Ala
    130                135                140

Met Leu Gly Met Cys Phe Ile Gln Ala Ala Tyr Leu Gly His Asp Ser
    145                150                155                160

Gly His Tyr Thr Met Met Ser Ser Lys Gly Tyr Asn Lys Phe Ala Gln
    165                170                175

Val Leu Asn Gly Asn Cys Leu Thr Gly Ile Ser Ile Ala Trp Trp Lys
    180                185                190

Trp Thr His Asn Ala His His Ile Ala Cys Asn Ser Leu Asp Tyr Asp
    195                200                205

Pro Asp Leu Gln His Leu Pro Val Phe Ala Val Pro Ser Ser Phe Phe
    210                215                220

Lys Ser Leu Thr Ser Arg Phe Tyr Gly Arg Glu Leu Thr Phe Asp Gly
    225                230                235                240

Leu Ser Arg Phe Leu Val Ser Tyr Gln His Phe Thr Ile Tyr Leu Val
    245                250                255

Met Ile Phe Gly Arg Ile Asn Leu Tyr Val Gln Thr Phe Leu Leu Leu
    260                265                270

Phe Ser Thr Arg Lys Val Pro Asp Arg Ala Leu Asn Ile Ile Gly Ile
    275                280                285

Leu Val Tyr Trp Thr Trp Phe Pro Tyr Leu Val Ser Cys Leu Pro Asn
    290                295                300

Trp Asn Glu Arg Val Leu Phe Val Leu Thr Cys Phe Ser Val Thr Ala
    305                310                315                320

Leu Gln His Ile Gln Phe Thr Leu Asn His Phe Ala Ala Asp Val Tyr
    325                330                335

Val Gly Pro Pro Thr Gly Thr Asn Trp Phe Glu Lys Gln Ala Ala Gly
    340                345                350

Thr Ile Asp Ile Ser Cys Ser Ser Trp Met Asp Trp Phe Phe Gly Gly
    355                360                365

Leu Gln Phe Gln Leu Glu His His Leu Phe Pro Arg Met Pro Arg Cys
    370                375                380

Gln Leu Arg Asn Ile Ser Pro Ile Val Gln Asp Tyr Cys Lys Lys His
    385                390                395                400

Asn Leu Pro Tyr Arg Ser Leu Ser Phe Phe Asp Ala Asn Val Ala Thr
    405                410                415

Ile Lys Thr Leu Arg Thr Ala Ala Leu Gln Ala Arg Asp Leu Thr Val
    420                425                430

Val Pro Gln Asn Leu Leu Trp Glu Ala Phe Asn Thr His Gly
    435                440                445

```

<210> SEQ ID NO 69

<211> LENGTH: 458

<212> TYPE: PRT

<213> ORGANISM: Helianthus annuus

<400> SEQUENCE: 69

```

Met Val Ser Pro Ser Ile Glu Val Leu Asn Ser Ile Ala Asp Gly Lys
  1          5          10          15

Lys Tyr Ile Thr Ser Lys Glu Leu Lys Lys His Asn Asn Pro Asn Asp
    20          25          30

```

-continued

Leu Trp Ile Ser Ile Leu Gly Lys Val Tyr Asn Val Thr Glu Trp Ala
 35 40 45
 Lys Glu His Pro Gly Gly Asp Ala Pro Leu Ile Asn Leu Ala Gly Gln
 50 55 60
 Asp Val Thr Asp Ala Phe Ile Ala Phe His Pro Gly Thr Ala Trp Lys
 65 70 75 80
 His Leu Asp Lys Leu Phe Thr Gly Tyr His Leu Lys Asp Tyr Gln Val
 85 90 95
 Ser Asp Ile Ser Arg Asp Tyr Arg Lys Leu Ala Ser Glu Phe Ala Lys
 100 105 110
 Ala Gly Met Phe Glu Lys Lys Gly His Gly Val Ile Tyr Ser Leu Cys
 115 120 125
 Phe Val Ser Leu Leu Leu Ser Ala Cys Val Tyr Gly Val Leu Tyr Ser
 130 135 140
 Gly Ser Phe Trp Ile His Met Leu Ser Gly Ala Ile Leu Gly Leu Ala
 145 150 155 160
 Trp Met Gln Ile Ala Tyr Leu Gly His Asp Ala Gly His Tyr Gln Met
 165 170 175
 Met Ala Thr Arg Gly Trp Asn Lys Phe Ala Gly Ile Phe Ile Gly Asn
 180 185 190
 Cys Ile Thr Gly Ile Ser Ile Ala Trp Trp Lys Trp Thr His Asn Ala
 195 200 205
 His His Ile Ala Cys Asn Ser Leu Asp Tyr Asp Pro Asp Leu Gln His
 210 215 220
 Leu Pro Met Leu Ala Val Ser Ser Lys Leu Phe Asn Ser Ile Thr Ser
 225 230 235 240
 Val Phe Tyr Gly Arg Gln Leu Thr Phe Asp Pro Leu Ala Arg Phe Phe
 245 250 255
 Val Ser Tyr Gln His Tyr Leu Tyr Tyr Pro Ile Met Cys Val Ala Arg
 260 265 270
 Val Asn Leu Tyr Leu Gln Thr Ile Leu Leu Leu Ile Ser Lys Arg Lys
 275 280 285
 Ile Pro Asp Arg Gly Leu Asn Ile Leu Gly Thr Leu Ile Phe Trp Thr
 290 295 300
 Trp Phe Pro Leu Leu Val Ser Arg Leu Pro Asn Trp Pro Glu Arg Val
 305 310 315 320
 Ala Phe Val Leu Val Ser Phe Cys Val Thr Gly Ile Gln His Ile Gln
 325 330 335
 Phe Thr Leu Asn His Phe Ser Gly Asp Val Tyr Val Gly Pro Pro Lys
 340 345 350
 Gly Asp Asn Trp Phe Glu Lys Gln Thr Arg Gly Thr Ile Asp Ile Ala
 355 360 365
 Cys Ser Ser Trp Met Asp Trp Phe Phe Gly Gly Leu Gln Phe Gln Leu
 370 375 380
 Glu His His Leu Phe Pro Arg Leu Pro Arg Cys His Leu Arg Ser Ile
 385 390 395 400
 Ser Pro Ile Cys Arg Glu Leu Cys Lys Lys Tyr Asn Leu Pro Tyr Val
 405 410 415
 Ser Leu Ser Phe Tyr Asp Ala Asn Val Thr Thr Leu Lys Thr Leu Arg
 420 425 430

-continued

 Thr Ala Ala Leu Gln Ala Arg Asp Leu Thr Asn Pro Ala Pro Gln Asn
 435 440 445

 Leu Ala Trp Glu Ala Phe Asn Thr His Gly
 450 455

<210> SEQ ID NO 70

<211> LENGTH: 449

<212> TYPE: PRT

<213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 70

 Met Ala Asp Gln Thr Lys Lys Arg Tyr Val Thr Ser Glu Asp Leu Lys
 1 5 10 15

 Lys His Asn Lys Pro Gly Asp Leu Trp Ile Ser Ile Gln Gly Lys Val
 20 25 30

 Tyr Asp Val Ser Asp Trp Val Lys Ser His Pro Gly Gly Glu Ala Ala
 35 40 45

 Ile Leu Asn Leu Ala Gly Gln Asp Val Thr Asp Ala Phe Ile Ala Tyr
 50 55 60

 His Pro Gly Thr Ala Trp His His Leu Glu Lys Leu His Asn Gly Tyr
 65 70 75 80

 His Val Arg Asp His His Val Ser Asp Val Ser Arg Asp Tyr Arg Arg
 85 90 95

 Leu Ala Ala Glu Phe Ser Lys Arg Gly Leu Phe Asp Lys Lys Gly His
 100 105 110

 Val Thr Leu Tyr Thr Leu Thr Cys Val Gly Val Met Leu Ala Ala Val
 115 120 125

 Leu Tyr Gly Val Leu Ala Cys Thr Ser Ile Trp Ala His Leu Ile Ser
 130 135 140

 Ala Val Leu Leu Gly Leu Leu Trp Ile Gln Ser Ala Tyr Val Gly His
 145 150 155 160

 Asp Ser Gly His Tyr Thr Val Thr Ser Thr Lys Pro Cys Asn Lys Leu
 165 170 175

 Ile Gln Leu Leu Ser Gly Asn Cys Leu Thr Gly Ile Ser Ile Ala Trp
 180 185 190

 Trp Lys Trp Thr His Asn Ala His His Ile Ala Cys Asn Ser Leu Asp
 195 200 205

 His Asp Pro Asp Leu Gln His Ile Pro Ile Phe Ala Val Ser Thr Lys
 210 215 220

 Phe Phe Asn Ser Met Thr Ser Arg Phe Tyr Gly Arg Lys Leu Thr Phe
 225 230 235 240

 Asp Pro Leu Ala Arg Phe Leu Ile Ser Tyr Gln His Trp Thr Phe Tyr
 245 250 255

 Pro Val Met Cys Val Gly Arg Ile Asn Leu Phe Ile Gln Thr Phe Leu
 260 265 270

 Leu Leu Phe Ser Lys Arg His Val Pro Asp Arg Ala Leu Asn Ile Ala
 275 280 285

 Gly Ile Leu Val Phe Trp Thr Trp Phe Pro Leu Leu Val Ser Phe Leu
 290 295 300

 Pro Asn Trp Gln Glu Arg Phe Ile Phe Val Phe Val Ser Phe Ala Val
 305 310 315 320

 Thr Ala Ile Gln His Val Gln Phe Cys Leu Asn His Phe Ala Ala Asp
 325 330 335

-continued

Val Tyr Thr Gly Pro Pro Asn Gly Asn Asp Trp Phe Glu Lys Gln Thr
 340 345 350

Ala Gly Thr Leu Asp Ile Ser Cys Arg Ser Phe Met Asp Trp Phe Phe
 355 360 365

Gly Gly Leu Gln Phe Gln Leu Glu His His Leu Phe Pro Arg Leu Pro
 370 375 380

Arg Cys His Leu Arg Thr Val Ser Pro Val Val Lys Glu Leu Cys Lys
 385 390 395 400

Lys His Asn Leu Pro Tyr Arg Ser Leu Ser Trp Trp Glu Ala Asn Val
 405 410 415

Trp Thr Ile Arg Thr Leu Lys Asn Ala Ala Ile Gln Ala Arg Asp Ala
 420 425 430

Thr Asn Pro Val Leu Lys Asn Leu Leu Trp Glu Ala Val Asn Thr His
 435 440 445

Gly

<210> SEQ ID NO 71
 <211> LENGTH: 449
 <212> TYPE: PRT
 <213> ORGANISM: Arabidopsis thaliana

<400> SEQUENCE: 71

Met Ala Glu Glu Thr Glu Lys Lys Tyr Ile Thr Asn Glu Asp Leu Lys
 1 5 10 15

Lys His Asn Lys Ser Gly Asp Leu Trp Ile Ala Ile Gln Gly Lys Val
 20 25 30

Tyr Asn Val Ser Asp Trp Ile Lys Thr His Pro Gly Gly Asp Thr Val
 35 40 45

Ile Leu Asn Leu Val Gly Gln Asp Val Thr Asp Ala Phe Ile Ala Phe
 50 55 60

His Pro Gly Thr Ala Trp His His Leu Asp His Leu Phe Thr Gly Tyr
 65 70 75 80

His Ile Arg Asp Phe Gln Val Ser Glu Val Ser Arg Asp Tyr Arg Arg
 85 90 95

Met Ala Ala Glu Phe Arg Lys Leu Gly Leu Phe Glu Asn Lys Gly His
 100 105 110

Val Thr Leu Tyr Thr Leu Ala Phe Val Ala Ala Met Phe Leu Gly Val
 115 120 125

Leu Tyr Gly Val Leu Ala Cys Thr Ser Val Phe Ala His Gln Ile Ala
 130 135 140

Ala Ala Leu Leu Gly Leu Leu Trp Ile Gln Ser Ala Tyr Ile Gly His
 145 150 155 160

Asp Ser Gly His Tyr Val Ile Met Ser Asn Lys Ser Tyr Asn Arg Phe
 165 170 175

Ala Gln Leu Leu Ser Gly Asn Cys Leu Thr Gly Ile Ser Ile Ala Trp
 180 185 190

Trp Lys Trp Thr His Asn Ala His His Leu Ala Cys Asn Ser Leu Asp
 195 200 205

Tyr Asp Pro Asp Leu Gln His Ile Pro Val Phe Ala Val Ser Thr Lys
 210 215 220

Phe Phe Ser Ser Leu Thr Ser Arg Phe Tyr Asp Arg Lys Leu Thr Phe
 225 230 235 240

-continued

```

Asp Pro Val Ala Arg Phe Leu Val Ser Tyr Gln His Phe Thr Tyr Tyr
      245                      250                      255

Pro Val Met Cys Phe Gly Arg Ile Asn Leu Phe Ile Gln Thr Phe Leu
      260                      265                      270

Leu Leu Phe Ser Lys Arg Glu Val Pro Asp Arg Ala Leu Asn Phe Ala
      275                      280                      285

Gly Ile Leu Val Phe Trp Thr Trp Phe Pro Leu Leu Val Ser Cys Leu
      290                      295                      300

Pro Asn Trp Pro Glu Arg Phe Phe Phe Val Phe Thr Ser Phe Thr Val
      305                      310                      315                      320

Thr Ala Leu Gln His Ile Gln Phe Thr Leu Asn His Phe Ala Ala Asp
      325                      330                      335

Val Tyr Val Gly Pro Pro Thr Gly Ser Asp Trp Phe Glu Lys Gln Ala
      340                      345                      350

Ala Gly Thr Ile Asp Ile Ser Cys Arg Ser Tyr Met Asp Trp Phe Phe
      355                      360                      365

Gly Gly Leu Gln Phe Gln Leu Glu His His Leu Phe Pro Arg Leu Pro
      370                      375                      380

Arg Cys His Leu Arg Lys Val Ser Pro Val Val Gln Glu Leu Cys Lys
      385                      390                      395                      400

Lys His Asn Leu Pro Tyr Arg Ser Met Ser Trp Phe Glu Ala Asn Val
      405                      410                      415

Leu Thr Ile Asn Thr Leu Lys Thr Ala Ala Tyr Gln Ala Arg Asp Val
      420                      425                      430

Ala Asn Pro Val Val Lys Asn Leu Val Trp Glu Ala Leu Asn Thr His
      435                      440                      445

```

Gly

```

<210> SEQ ID NO 72
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Conserved motif

```

<400> SEQUENCE: 72

```

His Pro Gly Gly
1

```

```

<210> SEQ ID NO 73
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Conserved motif
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (2)..(2)
<223> OTHER INFORMATION: X = any amino acid
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (3)..(3)
<223> OTHER INFORMATION: X = any amino acid

```

<400> SEQUENCE: 73

```

Gln Xaa Xaa His His
1                5

```

-continued

```

<210> SEQ ID NO 74
<211> LENGTH: 271
<212> TYPE: PRT
<213> ORGANISM: Thraustochytrium sp.

<400> SEQUENCE: 74
Met Met Glu Pro Leu Asp Arg Tyr Arg Ala Leu Ala Glu Leu Ala Ala
 1             5             10             15
Arg Tyr Ala Ser Ser Ala Ala Phe Lys Trp Gln Val Thr Tyr Asp Ala
      20             25             30
Lys Asp Ser Phe Val Gly Pro Leu Gly Ile Arg Glu Pro Leu Gly Leu
      35             40             45
Leu Val Gly Ser Val Val Leu Tyr Leu Ser Leu Leu Ala Val Val Tyr
      50             55             60
Ala Leu Arg Asn Tyr Leu Gly Gly Leu Met Ala Leu Arg Ser Val His
      65             70             75             80
Asn Leu Gly Leu Cys Leu Phe Ser Gly Ala Val Trp Ile Tyr Thr Ser
      85             90             95
Tyr Leu Met Ile Gln Asp Gly His Phe Arg Ser Leu Glu Ala Ala Thr
      100            105            110
Cys Glu Pro Leu Lys His Pro His Phe Gln Leu Ile Ser Leu Leu Phe
      115            120            125
Ala Leu Ser Lys Ile Trp Glu Trp Phe Asp Thr Val Leu Leu Ile Val
      130            135            140
Lys Gly Asn Lys Leu Arg Phe Leu His Val Leu His His Ala Thr Thr
      145            150            155            160
Phe Trp Leu Tyr Ala Ile Asp His Ile Phe Leu Ser Ser Ile Lys Tyr
      165            170            175
Gly Val Ala Val Asn Ala Phe Ile His Thr Val Met Tyr Ala His Tyr
      180            185            190
Phe Arg Pro Phe Pro Lys Gly Leu Arg Pro Leu Ile Thr Gln Leu Gln
      195            200            205
Ile Val Gln Phe Ile Phe Ser Ile Gly Ile His Thr Ala Ile Tyr Trp
      210            215            220
His Tyr Asp Cys Glu Pro Leu Val His Thr His Phe Trp Glu Tyr Val
      225            230            235            240
Thr Pro Tyr Leu Phe Val Val Pro Phe Leu Ile Leu Phe Phe Asn Phe
      245            250            255
Tyr Leu Gln Gln Tyr Val Leu Ala Pro Ala Lys Thr Lys Lys Ala
      260            265            270

```

```

<210> SEQ ID NO 75
<211> LENGTH: 266
<212> TYPE: PRT
<213> ORGANISM: Dabio rerio

```

```

<400> SEQUENCE: 75
Met Ser Val Leu Ala Leu Gln Glu Tyr Glu Phe Glu Arg Gln Phe Asn
 1             5             10             15
Glu Asp Glu Ala Ile Arg Trp Met Gln Glu Asn Trp Lys Lys Ser Phe
      20             25             30
Leu Phe Ser Ala Leu Tyr Ala Ala Cys Ile Leu Gly Gly Arg His Val
      35             40             45

```

-continued

```

Met Lys Gln Arg Glu Lys Phe Glu Leu Arg Lys Pro Leu Val Leu Trp
 50                      55                      60

Ser Leu Thr Leu Ala Ala Phe Ser Ile Phe Gly Ala Ile Arg Thr Gly
 65                      70                      75                      80

Gly Tyr Met Val Asn Ile Leu Met Thr Lys Gly Leu Lys Gln Ser Val
                      85                      90                      95

Cys Asp Gln Ser Phe Tyr Asn Gly Pro Val Ser Lys Phe Trp Ala Tyr
          100                      105                      110

Ala Phe Val Leu Ser Lys Ala Pro Glu Leu Gly Asp Thr Leu Phe Ile
          115                      120                      125

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

Thr Val Leu Leu Tyr Ser Trp Tyr Ser Tyr Lys Asp Met Val Ala Gly
          145                      150                      155                      160

Gly Gly Trp Phe Met Thr Met Asn Tyr Leu Val His Ala Val Met Tyr
          165                      170                      175

Ser Tyr Tyr Ala Leu Arg Ala Ala Gly Phe Lys Ile Ser Arg Lys Phe
          180                      185                      190

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

Val Val Asn Tyr Leu Val Tyr Leu Trp Met Gln Gln Gly Gln Glu Cys
          210                      215                      220

Pro Ser His Val Gln Asn Ile Val Trp Ser Ser Leu Met Tyr Leu Ser
          225                      230                      235                      240

Tyr Phe Val Leu Phe Cys Gln Phe Phe Phe Glu Ala Tyr Ile Thr Lys
          245                      250                      255

Arg Lys Ser Asn Ala Ala Lys Lys Ser Gln
          260                      265

```

<210> SEQ ID NO 76

<211> LENGTH: 320

<212> TYPE: PRT

<213> ORGANISM: Pavlova lutheri

<400> SEQUENCE: 76

```

His Glu Ala Ser Cys Arg Ile Arg His Glu Ala Ala Leu Trp Ser Trp
 1                      5                      10                      15

Leu Pro Thr Tyr Asp Glu Phe Val Asp Gly Leu Ser Phe Val Asp Arg
          20                      25                      30

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

Glu Trp Ala Ala Ile Ser Val Asp Lys His Met Ser Phe Phe Ser Asp
          50                      55                      60

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

Ala Leu Tyr Leu Val Met Ile Val Val Gly Pro Met Ile Met Ala Asn
          85                      90                      95

Arg Pro Pro Leu Pro Val Asn Gly Leu Ala Cys Ala Trp Asn Trp Phe
          100                      105                      110

Leu Ala Ala Phe Ser Thr Phe Gly Val Ala Cys Thr Trp His Cys Ile
          115                      120                      125

Phe Thr Arg Leu Arg Ser Arg Gly Phe Glu Ser Thr Thr Cys Gly Ser
          130                      135                      140

```

-continued

Ala Met Phe Met Ser Gln Gly Tyr Val Gly Leu Ala Met Leu Leu Phe
 145 150 155 160

Ile Tyr Ser Lys Leu Phe Glu Leu Ile Asp Thr Phe Phe Leu Ile Ala
 165 170 175

Lys Lys Ala Asp Val Ile Phe Leu His Trp Tyr His His Val Thr Val
 180 185 190

Leu Leu Tyr Cys Trp His Ser His Ser Val Arg Ile Pro Ser Gly Ile
 195 200 205

Trp Phe Ala Ala Met Asn Tyr Phe Val His Ala Ile Met Tyr Ser Tyr
 210 215 220

Phe Ala Met Thr Gln Met Gly Pro Arg Tyr Arg Lys Leu Val Arg Pro
 225 230 235 240

Tyr Ala Arg Leu Ile Thr Thr Leu Gln Ile Ser Gln Met Phe Val Gly
 245 250 255

Leu Ile Val Asn Gly Ser Ile Ile Tyr Phe Thr Ser Leu Gly His Ala
 260 265 270

Cys Lys Ser Ser Lys Thr Asn Thr Ile Leu Ser Trp Leu Met Tyr Leu
 275 280 285

Ser Tyr Phe Val Leu Phe Gly Leu Leu Tyr Leu Arg Asn Tyr Ile Leu
 290 295 300

Gly Thr His Gly Lys Pro Ala Gly Lys Arg Ala Lys Gly Lys Ala Glu
 305 310 315 320

<210> SEQ ID NO 77

<211> LENGTH: 282

<212> TYPE: PRT

<213> ORGANISM: Danio rerio

<400> SEQUENCE: 77

Met Glu Thr Phe Ser His Arg Val Asn Ser Tyr Ile Asp Ser Trp Met
 1 5 10 15

Gly Pro Arg Asp Leu Arg Val Thr Gly Trp Phe Leu Leu Asp Asp Tyr
 20 25 30

Ile Pro Thr Phe Ile Phe Thr Val Met Tyr Leu Leu Ile Val Trp Met
 35 40 45

Gly Pro Lys Tyr Met Lys Asn Arg Gln Ala Tyr Ser Cys Arg Ala Leu
 50 55 60

Leu Val Pro Tyr Asn Leu Cys Leu Thr Leu Leu Ser Leu Tyr Met Phe
 65 70 75 80

Tyr Glu Leu Val Met Ser Val Tyr Gln Gly Gly Tyr Asn Phe Phe Cys
 85 90 95

Gln Asn Thr His Ser Gly Gly Asp Ala Asp Asn Arg Met Met Asn Val
 100 105 110

Leu Trp Trp Tyr Tyr Phe Ser Lys Leu Ile Glu Phe Met Asp Thr Phe
 115 120 125

Phe Phe Ile Leu Arg Lys Asn Asn His Gln Ile Thr Phe Leu His Val
 130 135 140

Tyr His His Ala Thr Met Leu Asn Ile Trp Trp Phe Val Met Asn Trp
 145 150 155 160

Val Pro Cys Gly His Ser Tyr Phe Gly Ala Thr Phe Asn Ser Phe Ile
 165 170 175

His Val Leu Met Tyr Ser Tyr Tyr Gly Leu Ser Ala Val Pro Ala Leu

-continued

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

<210> SEQ ID NO 78

<211> LENGTH: 396

<212> TYPE: PRT

<213> ORGANISM: Pavlova lutheri

<400> SEQUENCE: 78

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

-continued

Glu Cys Ala Gly Ala Ala Met Lys Thr Asn Ile Ser Val Leu Ala Pro
 260 265 270

Leu Ile Leu Pro Val Ser Glu Gln Val Arg Phe Leu Ala Asn Tyr Val
 275 280 285

Ala Arg Lys Trp Leu Arg Met Lys Gly Val Lys Gly Tyr Val Pro Asp
 290 295 300

Phe Thr Thr Ala Val Gln His Phe Cys Ile His Thr Gly Gly Arg Ala
 305 310 315 320

Val Leu Asp Ala Leu Gln Ala Asn Leu Ser Leu Ser Asp Tyr Tyr Leu
 325 330 335

Glu Pro Ser Arg Tyr Ser Leu Trp Arg Trp Gly Asn Val Ser Ser Ala
 340 345 350

Ser Val Trp Tyr Glu Leu Asp Trp Leu Glu Lys Ser Gly Arg Ile Arg
 355 360 365

Arg Gly Asp Lys Val Trp Gln Ile Gly Phe Gly Ser Gly Phe Lys Cys
 370 375 380

Asn Ser Ala Val Trp Arg Ala Cys Arg Ala Met Pro
 385 390 395

<210> SEQ ID NO 79
 <211> LENGTH: 315
 <212> TYPE: DNA
 <213> ORGANISM: Heterocapsa niei

<400> SEQUENCE: 79

```

gattcaggat cctttttgca taggtaccac tccaccatgt ttccgatctg gtaggggtgt      60
gtgcgtggtc cagcctctt cacttggaaca agtgccgtcg ggccaagtgc cgtcggggcca    120
agtgcgctcg ggccaaggaa agcactccag cgctcacaac cacctcaccc cccctccccg    180
ccccccgctt cgttttcgct tgctttcagg tggatggggg cccgctgggt gcctggaggc    240
cagtcgtatt tttgtgcgac catcaattcc accgtgcatg ttgtcatgta cgcctattac    300
ttttctagat caatc                                          315

```

<210> SEQ ID NO 80
 <211> LENGTH: 100
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Protein encoded by SEQ ID NO:79 which comprises
 a stop codon between residues 17 and 18 suggesting the presence of
 an intron

<400> SEQUENCE: 80

Asp Ser Gly Ser Phe Leu His Arg Tyr His Ser Thr Met Phe Pro Ile
 1 5 10 15

Trp Gly Cys Val Arg Gly Pro Arg Leu Phe Thr Trp Thr Ser Ala Val
 20 25 30

Gly Pro Ser Ala Val Gly Pro Ser Ala Val Gly Pro Arg Lys Ala Leu
 35 40 45

Gln Arg Ser Gln Pro Pro His Pro Pro Leu Pro Pro Pro Ala Ser Phe
 50 55 60

Ser Leu Ala Phe Arg Trp Met Gly Ala Arg Trp Val Pro Gly Gly Gln
 65 70 75 80

Ser Tyr Phe Cys Ala Thr Ile Asn Ser Thr Val His Val Val Met Tyr
 85 90 95

-continued

Ala Tyr Tyr Phe
100

<210> SEQ ID NO 81
<211> LENGTH: 46
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 81

gctacgcccc gggatcctcg aggctggcgc aacgcaatta atgtga 46

<210> SEQ ID NO 82
<211> LENGTH: 84
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 82

cacaggaaac agcttgacat cgattaccgg caattgtacg gcggccgcta cggatatacct 60

cgctcgagct cgcccggggt agct 84

<210> SEQ ID NO 83
<211> LENGTH: 46
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 83

agcacatcga tgaaggagat ataccatgg ctaatgcaat caagaa 46

<210> SEQ ID NO 84
<211> LENGTH: 36
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 84

acgatgcggc cgctcaacca tgagtattaa gagctt 36

<210> SEQ ID NO 85
<211> LENGTH: 294
<212> TYPE: PRT
<213> ORGANISM: Pavlova salina

<400> SEQUENCE: 85

Met Pro Thr Trp Gly Glu Phe Val Ala Gly Leu Thr Tyr Val Glu Arg
1 5 10 15

Gln Gln Met Ser Glu Glu Leu Val Arg Ala Asn Lys Leu Pro Leu Ser
20 25 30

Leu Ile Pro Glu Val Asp Phe Phe Thr Ile Ala Ser Val Tyr Val Gly
35 40 45

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

Ile Thr Leu Gly Pro Gln Leu Met Ala Arg Arg Pro Pro Leu Pro Ile
65 70 75 80

-continued

```

Asn Thr Leu Ala Cys Leu Trp Asn Phe Ala Leu Ser Leu Phe Ser Phe
      85                      90                      95
Val Gly Met Ile Val Thr Trp Thr Thr Ile Gly Glu Arg Leu Trp Lys
      100                      105                      110
Asn Gly Ile Glu Asp Thr Val Cys Gly His Pro Ile Phe Met Gly Tyr
      115                      120                      125
Gly Trp Ile Gly Tyr Val Met Leu Ala Phe Ile Trp Ser Lys Leu Phe
      130                      135                      140
Glu Leu Ile Asp Thr Val Phe Leu Val Ala Lys Lys Ala Asp Val Ile
      145                      150                      155                      160
Phe Leu His Trp Tyr His His Val Thr Val Leu Leu Tyr Cys Trp His
      165                      170                      175
Ser Tyr Ala Val Arg Ile Pro Ser Gly Ile Trp Phe Ala Ala Met Asn
      180                      185                      190
Tyr Phe Val His Ala Ile Met Tyr Ala Tyr Phe Gly Met Thr Gln Ile
      195                      200                      205
Gly Pro Arg Gln Arg Lys Leu Val Arg Pro Tyr Ala Arg Leu Ile Thr
      210                      215                      220
Thr Phe Gln Leu Ser Gln Met Gly Val Gly Leu Ala Val Asn Gly Leu
      225                      230                      235                      240
Ile Ile Arg Tyr Pro Ser Ile Gly His His Cys His Ser Asn Lys Thr
      245                      250                      255
Asn Thr Ile Leu Ser Trp Ile Met Tyr Ala Ser Tyr Phe Val Leu Phe
      260                      265                      270
Ala Ala Leu Tyr Val Lys Asn Tyr Ile Phe Ser Lys Leu Lys Ser Pro
      275                      280                      285
Lys Arg Lys Lys Val Glu
      290

```

```

<210> SEQ ID NO 86
<211> LENGTH: 276
<212> TYPE: PRT
<213> ORGANISM: Pavlova salina

<400> SEQUENCE: 86

```

```

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

```

-continued

130				135				140							
His	Trp	Tyr	His	His	Val	Thr	Val	Leu	Leu	Tyr	Cys	Trp	His	Ser	Tyr
145					150					155				160	
Ala	Val	Arg	Ile	Pro	Ser	Gly	Ile	Trp	Phe	Ala	Ala	Met	Asn	Tyr	Phe
			165					170						175	
Val	His	Ala	Ile	Met	Tyr	Ala	Tyr	Phe	Gly	Met	Thr	Gln	Ile	Gly	Pro
			180					185					190		
Arg	Gln	Arg	Lys	Leu	Val	Arg	Pro	Tyr	Ala	Arg	Leu	Ile	Thr	Thr	Phe
			195				200					205			
Gln	Leu	Ser	Gln	Met	Gly	Val	Gly	Leu	Ala	Val	Asn	Gly	Leu	Ile	Ile
	210					215					220				
Arg	Tyr	Pro	Ser	Ile	Gly	His	His	Cys	His	Ser	Asn	Lys	Thr	Asn	Thr
	225				230					235					240
Ile	Leu	Ser	Trp	Ile	Met	Tyr	Ala	Ser	Tyr	Phe	Val	Leu	Phe	Ala	Ala
			245						250					255	
Leu	Tyr	Val	Lys	Asn	Tyr	Ile	Phe	Ser	Lys	Leu	Lys	Ser	Pro	Lys	Arg
			260				265						270		
Lys	Lys	Val	Glu												
			275												

<210> SEQ ID NO 87
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Oligonucleotide primer

<400> SEQUENCE: 87

cgctctagaa ctagtggatc

20

<210> SEQ ID NO 88
 <211> LENGTH: 223
 <212> TYPE: PRT
 <213> ORGANISM: Melosira sp.

<400> SEQUENCE: 88

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

-continued

Thr	Gly	Lys	Ser	Leu	Pro	Ile	Trp	Trp	Lys	Ser	Ser	Leu	Thr	Met	Met
145					150				155					160	
Gln	Met	Ile	Gln	Phe	Val	Thr	Met	Met	Ser	Gln	Ala	Ser	Tyr	Leu	Leu
			165					170						175	
Val	Thr	Asn	Cys	Glu	Lys	Thr	Ser	Arg	Gly	Val	Val	Ile	Ala	Tyr	Phe
		180						185						190	
Val	Tyr	Ile	Phe	Thr	Leu	Leu	Val	Leu	Phe	Ala	Gln	Phe	Phe	Arg	Ala
		195				200					205				
Ser	Tyr	Met	Lys	Pro	Lys	Gly	Lys	Lys	Ala	Lys	Met	Lys	Lys	Lys	Val
	210					215				220					

<210> SEQ ID NO 89

<211> LENGTH: 683

<212> TYPE: DNA

<213> ORGANISM: Melosira sp.

<400> SEQUENCE: 89

```

acgatcttca agtcaaacgc cgtccctgcc ctggatccat accccatcaa attcgtttac    60
aatgtgtccc agatcatgat gtgcgcgtac atgacgatcg aggcaggcct ggtggcctac    120
cgcaagtggct atactgtcat gccatgcaac gactacaaca ccaacaaccc ccctgtcggg    180
aacctgctgt ggctgtttta catttccaaa gtttgggact tttgggacac catctttatc    240
gtgattggca aaaagtggaa gcagctgagc ttcttgacag tgtaccacca caccaccatc    300
tttttgttct actggctcaa ctgcgatgtc aactacgacg gagatattta tctgacgatt    360
ctgttgaacg gcttcattca caccgcatg tacacttatt acttcgtttg catgcacacg    420
aaggtgcccc agactggaaa gtcgttgccc atttggtgga aatccagtct caccatgatg    480
caaatgatcc aattcgtcac catgatgagc caggcttcgt acttgctcgt gacgaactgc    540
gaaaagacca gtcgggggggt cgttattgct tactttgtgt acattttcac tctactogtc    600
ttatttgctc agttcttcct agcatcttac atgaagccca agggaaagaa ggcgaaaatg    660
aagaaggtat aagctgctgg cat                                     683

```

1-221. (canceled)**222.** A transgenic seed comprising a LC-PUFA, wherein the LC-PUFA is selected from the group consisting of:

- i) EPA,
- ii) DPA,
- iii) DHA,
- iv) EPA and DPA, and
- v) EPA, DHA, and DPA.

223. The seed of claim 222, derived from an isogenic non-transgenic seed which produces LA and/or ALA.**224.** The seed of claim 223, wherein the isogenic non-transgenic seed comprises a higher concentration of ALA than LA in its fatty acids.**225.** The seed of claim 223, wherein the isogenic non-transgenic seed comprises at least about 13% ALA of the total fatty acids (w/w).**226.** The seed of claim 222, wherein the total fatty acid in the oil of the seed comprises at least 9% C20 fatty acids (w/w).**227.** The seed of claim 222, wherein the seed is derived from an oilseed plant.**228.** The seed of claim 227, wherein the oilseed plant is oilseed rape, maize, sunflower, soybean, sorghum, flax, sugar beet, cotton, peanut, poppy, mustard, castor bean, sesame, or safflower.**229.** The seed of claim 222, wherein the seed has a germination rate which is substantially the same as that of an isogenic non-transgenic seed.**230.** The seed of claim 222 which comprises DPA which is synthesized in the seed and wherein the total fatty acid of the seed comprises at least 0.1% DPA (w/w).**231.** The seed of claim 222 which comprises DHA which is synthesized in the seed.**232.** The seed of claim 222 which comprises at least one ω 3 C20 LC-PUFA which is synthesized in the seed and wherein the total fatty acids of the seed comprises at least 2.5% ω 3 C20 LC-PUFA (w/w).**233.** The seed of claim 222 wherein the efficiency of conversion of ALA to EPA in the seed is at least 14.6%.

234. The seed of claim 222 which comprises ω 3 polyunsaturated fatty acids that are the products of Δ 6-desaturation of ALA and/or the products of 9 elongation of ALA, which products are synthesized in the seed, and wherein the efficiency of conversion of ALA to said products in the seed is at least 22%.

235. The seed of claim 222 which comprises DPA which is synthesized from EPA in the seed and wherein the efficiency of conversion of EPA to DPA in the seed is at least 5%.

236. The seed of claim 222, wherein at least 25% (w/w) of the LC-PUFA in the seed form part of triacylglycerols.

237. A method of preparing a feedstuff, the method comprising admixing a transgenic seed according to claim 222, or an extract comprising a LC-PUFA obtained from said seed, with a suitable carrier.

238. The method of claim 237, wherein the feedstuff is for consumption by a mammal or a fish.

239. A method of producing a transgenic seed according to claim 222, the method comprising

- i) introducing into a progenitor cell of a seed one or more polynucleotides which encode at least two enzymes each of which is a 5/6 bifunctional desaturase, 5 desaturase, 6 desaturase, 5/6 bifunctional elongase, 5 elongase, 6 elongase, 4 desaturase, 9 elongase, or 8 desaturase, wherein the one or more polynucleotides are operably linked to one or more promoters that are capable of directing expression of said polynucleotides in the cell, thereby producing a recombinant progenitor cell,

- ii) culturing said recombinant progenitor cell to produce a plant which comprises said transgenic seed, and

- iii) recovering the seed from the plant so produced.

240. A method of producing a transgenic seed comprising cultivating a transgenic plant which produces the transgenic seed of claim 222, and harvesting said transgenic seed from the plant.

241. A transgenic plant which comprises the transgenic seed of claim 222.

242. An extract of the transgenic seed of claim 222, wherein said extract comprises an increased level of the LC-PUFA in its fatty acid relative to a corresponding extract from an isogenic non-transformed plant.

243. A method of producing oil comprising at least one LC-PUFA, the method comprising obtaining the transgenic plant of claim 241, and extracting oil from said plant.

244. A method of producing oil comprising at least one LC-PUFA, the method comprising obtaining the seed of claim 222, and extracting oil from said seed.

245. A composition comprising the seed of claim 222, or an extract thereof comprising LC-PUFA, and a suitable carrier.

246. A method of treating a condition which would benefit from a LC-PUFA, comprising administering to a subject a transgenic seed according to claim 222 in an amount effective to treat the condition.

247. A method of treating a condition which would benefit from a LC-PUFA, comprising administering to a subject a composition according to claim 245 in an amount effective to treat the condition.

* * * * *