

HIGH PALMITIC OIL COMPOSITIONS AND USES THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This international application claims benefit to U.S. Provisional Application No. 63/549,843, filed on February 5, 2024, which is incorporated herein by reference in its entirety.

INCORPORATION BY REFERENCE TO A SEQUENCE LISTING

[0002] The instant application contains a sequence listing Sequence Listing XML which has been submitted electronically in XML file format and is hereby incorporated by reference in its entirety. Said Sequence Listing XML, created on February 3, 2025, is named 009919_00098_WO_SL.xml and is 142,039 bytes in size.

BACKGROUND

[0003] Oleaginous microorganisms have the ability to convert carbon substrates into oils, including triacylglycerols (TAG) or lipids, and accumulate these oils intracellularly. Some microorganisms can accumulate lipids in amounts of up to 89% dry weight. Thus, oleaginous microorganisms, including microalgae, bacteria, fungi, and yeasts, can serve as an ideal source for biobased oil production. Genetic and non-genetic modification techniques can allow for the production of oils with particular fatty acid profiles.

[0004] Oils containing significant amounts of palmitic acid are valuable across many industries, including foods, personal care, home care, cosmetics, and biofuel. However, the primary oil responsible for providing high levels of palmitic acid is palm oil derived from oil palms (*Elaeis guineensis*). Because palm oil is semisolid at room temperature, they are routinely fractionated into a liquid fraction (palm olein) and a solid fraction (palm stearin). It is generally known that palm olein contains higher levels of oleic and linoleic acids compared to palm oil. The demand for palm oil and the development of this industry has led to a significant reduction of biodiversity in tropical locations where palms are grown. Thus, a need exists for sustainable sources of palm oil.

INCORPORATION BY REFERENCE

[0005] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

SUMMARY

[0006] In some aspects, the present disclosure provides an oil comprising a TAG component, wherein the TAG component has a fatty acid content comprising at least 30% or more of palmitic (C16:0) acid.

[0007] In some aspects, the present disclosure provides an oil, obtained through non-genetic engineering means, comprising TAG components having a fatty acid content comprising 30% or more of palmitic (C16:0) acid, 45% or more of oleic (C18:1) acid and 8% or less of linoleic (C18:2) acid.

[0008] In some aspects, the present disclosure provides an oil derived from a microalgal cell comprising a TAG component, wherein the TAG component has a fatty acid content comprising 30% or more of (C16:0) acid, wherein the TAG component has a low linoleic content.

[0009] In some aspects, the present disclosures provide an oil derived from a microalgal cell, wherein the oil provides an alternative to palm oil-derived olein.

[0010] In some aspects, the present disclosure provides an oil derived from a microalgal cell comprising a TAG component, wherein the TAG component has at least 30% C16:0 fatty acids and at least 40% C18:1 fatty acids, and wherein the oil is produced through classical strain improvement of a microalgal cell.

[0011] In some aspects, the present disclosure provides an oil derived from a microalgal cell comprising a TAG component, wherein the TAG component has a fatty acid content comprising 30% or more C16:0 fatty acid, wherein the oil further comprises less than 1% tripalmitin.

[0012] In some aspects, the present disclosure provides an oil comprising a TAG component having a fatty acid content comprising at least 30% C16:0 fatty acids, less than about 1% diacylglycerols, and less than about 1% monoacylglycerols.

[0013] In some aspects, the present disclosure provides an oil that further comprises sterols, such as ergosterol.

[0014] In some aspects, the present disclosure provides an oil, obtained through non-genetic engineering means, comprising TAG components having a fatty acid content comprising 49% or more of palmitic (C16:0) acid, about 36% or less of oleic (C18:1) acid, and 8% or less of linoleic (C18:2) acid.

[0015] In some aspects, the present disclosure provides an oil obtained through non-genetic engineering means, comprising TAG components having a POP (palmitic-oleic-

palmitic) TAG population comprising 45% or more of all TAG species, while PPP (tripalmitin) comprises less than 2.7%.

[0016] In some aspects, the present disclosure provides a combination of screening modalities for *Prototheca moriformis*, including mutagenesis with agents including UV light (254 nm) and 4-nitroquinoline-1-oxide (4-NQO), followed by repeated cycles of growth in a lipid production medium at temperatures up to about 38°C followed by a brief (up to about 4 min) incubation in near lethal temperatures (up to about 68 °C). After incubation at lethal temperatures, the cells are allowed to recover for about 1 to 5 days (e.g., up to about 3 days) in a low glucose (i.e., a media containing up to 5 g/L) growth media, followed by plating the cells to growth medium containing plates. Individual clones are then selected for growth in a lipid production medium, followed by screening for fatty acid profile of their resulting lipid (oil). The fatty acid profile is screened by generating fatty acid methyl esters (FAMEs), interrogating those FAMEs by gas chromatography utilizing flame ionization detection (GC/FID), and analyzing the triacylglycerol (TAG) composition of the oils by Liquid-Chromatography/Time of Flight-Mass Spectrometry (LC/TOF-MS) equipped with an Atmospheric Pressure Chemical Ionization (APCI) source. Using such a method, a person of skill in the art can produce a microalgal cell that can make an oil with the desired TAG composition.

[0017] In some aspects, the present disclosure provides a means to produce valuable structured fats through the biotransformation of a sugar, wherein the fats have a similar composition or components to the palm mid fraction, but in the absence of the capital equipment, exhaustive fractionation, and wasteful processes.

[0018] In one embodiment, an isolated microalgal cell is provided such as a *Prototheca* sp. cell obtained through classical microalgal cell improvement, wherein the microalgal cell produces an oil characterized as comprising: (a) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 17% of the TAG component, or of at least 23%, or of at least 29%, or of at least 32%, or of at least 35%, or at least 45%, and (b) a tripalmitin (PPP) content less than 0.3%, or less than about 0.6%, or less than about 1.0%, or less than at least 1.2%, or less than about 1.5%, or less than about 2.7% of the TAG component.

[0019] In another embodiment, the isolated microalgal cell has a POP content that is at least about 45%, or at least about 50% or at least about 55%, or at least about 60%, and a PPP content that is less than about 2.7% of the TAG component.

[0020] In another embodiment, the microalgal oil comprising the TAG component, has a TAG component of (a) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 17% of the TAG component, or at least 23%, or at least 29%, or at least 32%, or at least 35%, or at least 45%, and (b) a tripalmitin (PPP) content less than 0.3% of the TAG component, or less than about 0.6%, or less than about 1.0%, or less than at least 1.2%, or less than about 1.5%, or less than about 2.7% of the TAG, and wherein the microalgal oil is produced by a microalgal *Prototheca* cell obtained through classical strain improvement.

[0021] In one aspect, the microalgal oils produced by the microalgal cells described are from a cell that has had no exogenous nucleic acid introduced. In another aspect, the oil-producing microalgal cell obtained by the classical means described herein can be further recombinantly engineered with an exogenous nucleic acid that is not native to the parent microalgal cell that produced the strain.

[0022] In a further embodiment, the microalgal oil has a TAG component that comprises: (a) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 45%, and (b) a tripalmitin (PPP) content of less than about 2.7% of the TAG. Alternatively, a contemplated microalgal oil can have (a) the POP content of at least 35% and the PPP content of less than about 1.5%, or (b) the POP content of at least 45% and the PPP content of less than about 2.7%. The microalgal oils described above and herein can further comprise at least about 30% C16:0 fatty acids, or at least about 35% C16:0 fatty acids, or at least about 40% C16:0 fatty acids, or at least about 45% C16:0 fatty acids, or at least about 50% C16:0 fatty acids.

[0023] Also contemplated is a method of obtaining a microalgal cell that produces the oil of any of claims 4 to 8 comprising the steps of: (a) mutagenizing microalgal cells with at least one or more agents, which promote errors in DNA excision repair known (e.g., ultraviolet light (e.g., 254 nm), 4-nitroquinoline 1-oxide (4-NQO), methyl methanesulfonate (MMS), Ni, Ar, or polyaromatic hydrocarbons while culturing the microalgal cells at about 28 °C to about 38 °C in a vegetative or a lipid production media; (b) enriching the mutagenized microalgal cells by exposing the mutagenized cells to physical conditions (e.g., an elevated temperature of about 50 to about 68°C or by using density fractionation on sucrose or percoll gradients to isolate specific subpopulations based on density) or a chemical compound (e.g., herbicides, fungicides, or inhibitors of lipid biosynthesis, such as tebutam or cofenstrole); (c) growing the mutagenized, enriched microalgal cells in a lipid production medium; and (d) harvesting the microalgal oil from the microalgal cells and interrogating the microalgal oil using a FAME analysis and a triacylglyceride (TAG) analysis to determine the component of the screened

microalgal cells; and wherein the TAG component of the screened microalgal cells comprises (e) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 17% of the TAG component, or of at least 23%, or of at least 29%, or of at least 32%, or of at least 35%, or at least 45%, and (f) a tripalmitin (PPP) content less than 0.3%, or less than about 0.6%, or less than about 1.0%, or less than at least 1.2%, or less than about 1.5%, or less than about 2.7% of the TAG component. In a preferred embodiment, the method uses a *Prototheca sp.* cell.

[0024] Another embodiment is a method of obtaining a microalgal oil of a microalgal cell as described herein by: (a) mutagenizing cells with agents that introduce mutations through promoting errors in DNA excision repair known to include, but not limited to, UV light (254 nm), 4-nitroquinoline 1-oxide (4-NQO), methyl methanesulfonate (MMS), Ni, Ar, or polyaromatic hydrocarbons culturing the microalgal cells at about 28 °C to about 38 °C in vegetative or lipid production media and (b) Enriching said cell populations using physical change (e.g., elevated temperatures of about 50, about 60 or up to about 68°C or fractionation on sucrose or percoll gradients to isolate specific subpopulations based on density) or chemical compound (e.g., herbicides, fungicides or inhibitors of lipid biosynthesis such as tebutam or cofenstrole) and (c) plating aforementioned mutagenized and enrich populations to solid vegetative medium, (d) growing said individuals in a lipid production medium, (e) harvesting the microalgal cells from said lipid production medium, (f) obtaining the microalgal oil from the cultured microalgal cells and interrogating the oil via FAME analysis as well as TAG analysis.

[0025] This Summary is provided to introduce certain concepts in a simplified form that are further described in the Detailed Description. The intention of this Summary is not to identify key or necessary features of the claimed subject matter, nor is it intended to be used to limit, in any way, the claimed subject matter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0026] The novel features of the TAG component of microalgal oils, microalgal oils, and the microalgae producing them, as well as the methods of obtaining said unpurified oils, are set forth with particularity in the appended claims. A better understanding of the features and advantages of the methods, compositions, and components will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the compositions, components, and methods are utilized, and the accompanying drawings of which:

[0027] FIG. 1A and 1B show a sequence alignment of a plastidic 23 S rDNA between a UTEX 1533; SEQ ID NO:30, designated as *Prototheca wickerhamii*, and a *Prototheca moriformis* strain, designated UTEX 1435; SEQ ID NO:31.

[0028] FIG. 2A and 2B show an outline of the mutagenesis trait selection, high throughput, and automated screening steps of the improvement process described herein.

[0029] FIG. 3 shows an overview of the workflow and strategy utilized in classical strain improvement for the generation of mutants derived from UTEX 1533 in the current disclosure.

[0030] FIG. 4 shows TIC chromatograms of extracted triacylglycerols derived from classically improved mutant isolates (a-c/CHK100 mutant isolates 1-3) derived from the classically improved base strain CHK100(d).

[0031] FIG. 5 is a mass spectrum of POP/PPO triglyceride present in mutant isolate (a)/CHK100 mutant isolate 1 in FIG. 4.

[0032] FIG. 6 is a mass spectrum of POP/PPO triglyceride present in mutant isolate (b)/CHK100 mutant isolate 2 in FIG. 4.

[0033] FIG. 7 is a mass spectrum of POP/PPO triglyceride present in mutant isolate (c)/CHK100 mutant isolate 3 in FIG. 4.

[0034] FIG. 8 is a mass spectrum of POP/PPO triglyceride present in the CHK100 parental strain in FIG. 4.

[0035] FIG. 9 is a mass spectrum of POP/PPO triglyceride of red palm oil.

[0036] FIG. 10 is a mass spectrum of 1,3-palmitin-2-olein analytical standard.

[0037] FIG. 11 is a mass spectrum of 1,2-palmitin-3-olein analytical standard.

[0038] FIG. 12 is a graph illustrating the triglyceride profiles of some of the higher palmitate strains of the present disclosure compared to red palm oil.

[0039] FIG. 13 is a graph illustrating the triglyceride profiles of mutant isolate (b) in FIG. 4 compared to red palm oil.

[0040] FIG. 14 illustrates a comparison between the levels of palmitic acid (C16:0) and unsaturated fatty acids (sum of C18:1 and C18:2 fatty acids) found in a novel oil of the present disclosure compared to the fatty acid levels found in palm olein.

[0041] FIG. 15 shows the TAG profile of the oil produced by the high-palmitic acid algal strains of the present disclosure compared to the TAG profiles of palm olein.

[0042] FIG. 16 shows FAME (fatty acid methyl ester) profiles of algal strain 268U derived mutant BLK08 F04 run in lipid production at 28 °C versus 32 °C.

[0043] FIG. 17 shows TAG profiles of algal strain 268U derived mutant BLK08 F04 run in lipid production at 28 °C versus 32 °C with PPO and POP regioisomers indicated.

[0044] FIG. 18 shows levels of PPO and POP regioisomers and PPP in classically improved algal strains of the present disclosure compared with red palm oil and palm olein mid-fraction

DETAILED DESCRIPTION

[0045] Provided herein are microalgal oil compositions having high-palmitic acid content, high POP TAG content, and low PPP content, as well as methods of making thereof, formulations, and applications thereof. The microalgal oils having such a ratio of high-palmitic acid content and POP TAG content, and having a low PPP content, can then be a ready source substitute for the fractions obtained from palm oil that involves numerous fractionation steps and many days to obtain the mid-fractions of interest.

[0046] The principal species of palm under global cultivation is *Elaeis guineensis* or various cultivars and varieties derived from it. One source of palm oil is the palm fruit, which is comprised of a fleshy outer mesocarp that surrounds the hard palm kernel or endocarp. The oil of the palm fruit differs markedly from the oil of the hard palm kernel, termed “palm kernel oil” (“PKO”). PKO is characterized as being high in C12:0 fatty acids (lauric, >45%) and C14:0 (myristic, > 15%), and with relatively low C16:0, palmitic (8%), and C18:1, oleic (15%). The palm oil from the palm fruit is virtually devoid of C12:0 and C14:0 fatty acids, being comprised primarily of palmitic acid (40-46%) and oleic acid (37-45%) and some C18:2 fatty acids (linoleic, 8-12%).

[0047] Palm oil is used ‘as is’ for cooking and other applications, including personal care, but a significant amount of commercial palm oil is fractionated to create more valuable components. (Jin et al., “Characteristics of palm mid-fractions produced from different fractionation paths and their potential usages,” *Int. J. Food Prop.* 21: 58-69, 2018).

[0048] The first fractionation of refined palm oil, a so-called dry fractionation process because it does not use solvents, yields two fractions: a palm olein fraction and a palm stearin fraction. The palm olein fraction is a liquid at room temperature and is enriched in oleic and linoleic fatty acids; the palm stearin fraction is a solid at room temperature and is highly enriched in palmitate (50-60%).

[0049] Palm olein can be further dry fractionated to a “palm mid-fraction A” (“PMF-A”) component. During the fractionation process, palm olein is cooled to roughly 16°C, with the rate depending on the specific manufacturer and their process. This is followed by a further

decrease in temperature to 12-13°C and holding at this temperature for a period of time to allow for further crystallization of the solid PMF-A fraction. The PMF-A fraction is enriched in palmitate, particularly POP (40%). After this fractionation step, the liquid component at these low temperatures is termed “super olein.”

[0050] Palm stearin can be further fractionated to produce a palm mid-fraction B (PMF-B) component employing a wet fractionation technique utilizing acetone (ratios of 1:4-8 wt:wt of solvent:oil) and temperatures ranging from 17-25 °C with holds at the indicated temperature of 8-48 hr. The resulting solid fraction produced by the further fractionation is termed “super stearin,” and enriched in C18:1, while the PMF-B is further enriched in palmitate and the POP fraction (45% POP).

[0051] The PMF-A and PMF-B fractions can be blended and wet fractionated yet again using acetone (ratios of 1:10, wt:wt of solvent:oil) and at a very low temperature (e.g., about 4 °C) combined with a hold time of up to 24 hr, to produce a third, highly enriched POP fraction (60-65% POP), termed PMF-C.

[0052] The resulting high POP fractions obtained by these methods from palm oil are highly valued for their use in the confectionery industry, where they give chocolates their characteristic sharp melt profile (*see e.g., BAILEY’S INDUSTRIAL OILS AND FAT PRODUCTS*. 2005. 6th Ed. Shahidi, Fereidoon, Ed. John Wiley and Sons, Inc., Hoboken, NJ. Volume 4: “Edible oil and fat products: Products and Applications,” 4: 159-173).

[0053] It will not be lost to the average layperson that the processes involved in the generation of these PMF components are indeed quite involved and costly from a capital (cost of equipment), energy (chilling and alternately heating large quantities of oil) and hazardous chemicals (large quantities of acetone) perspective.

[0054] Provided herein are examples of a simpler process, whereby fermentable sugars are converted into triglyceride oils in microalgae with POP compositions equal or superior to those found in PMF derived from a more complicated and more costly process currently employed in the manufacture of these materials.

[0055] Having a more commercially available form of these structured oils for use from a non-GMO organism, given cocoa butter and palm oil availability, will result in economic savings in production and will not be as prone to climate changes or plant viruses and blights.

[0056] Oil compositions provided herein can be produced by a microorganism that is genetically modified or non-genetically modified. Non-genetically modified or non-genetically engineered microorganisms or non-genetically modified organisms (non-GMO) can be

produced using strategies such as those described herein. The non-GMO microorganisms obtained by the methods described herein can be further recombinantly modified if desired.

[0057] Genetic and non-genetic modification techniques can allow for the production of oils having particular phenotypes. While genetic engineering techniques can tend to be more targeted to phenotypes elicited in a host oleaginous microbe. Classical strain improvement or other non-genetic engineering techniques can also be employed to enhance phenotypes. Enhancement of phenotypes can include the elaboration of a particular fatty acid profile (e.g., high-palmitic acid content), yield on carbon, volumetric oil accumulation (e.g., grams oil/L culture), oil productivity (e.g., g oil/L culture day), and oil as a percent dry cell weight (DCW) as a measure of strain performance.

[0058] A further benefit of classical strain improvement techniques, when used as the sole means to alter or improve strain phenotype and performance, can be realized from both a regulatory and business/marketing perspective. From a regulatory perspective, non-genetically engineered microbes may be exempt from regulatory oversight by entities such as the U.S. EPA's Toxic Substances Control Act (TSCA) and the requirement to file a Microbial Commercial Activity Notice (MCAN) when the material is to be used in chemical (non-food) applications. Such dispensation extends to other geographies as well, such as Brazil, for example, where such microbes are exempt from filing a Strain Dossier with the Brazilian regulatory body (the National Technical Commission of Biosafety, Ministry of Science, Technology, Innovation, and Communications or CTNBio) that oversees industrial microbes. The avoidance of such regulatory oversight can save millions of dollars in development costs. From a marketing and consumer branding perspective, the raw materials produced by such non-genetically engineered means can meet GMO-free and organic labeling standards, as well as brands' and consumers' desires for "clean labeling".

[0059] As used herein, the term "classical strain improvement" or "classical strain optimization" refers to methods of random or semi-random mutagenesis of microbes to create non-naturally occurring strains with improved properties. These methods include, but are not limited to, mutagenesis of a population to create genetic variants, random selection or screening of a surviving population to identify an improved strain, and identification of improved strains by assaying fermentation broth for products. Classical strain improvement methods include exposure to UV radiation (e.g., 254 nm), chemical mutagens (e.g., 4-nitroquinoline 1-oxide, which forms bulky quinolone purine adducts but also produces mutagenic superoxide radicals (see e.g., Downes, D. et al., "Characterization of the Mutagenic Spectrum of 4-Nitroquinoline 1-Oxide (4-NQO) in *Aspergillus nidulans* by Whole Genome Sequencing," *G3, Genes*,

Genomics and Genetics 1-10, 2014) and/or selective or enrichment agents (see e.g., Tebutam, a microtubule inhibitor and potent inhibitor of fatty acyl-ACP thioesterase, FATA as discussed in Johnen et al., "Inhibition of acyl-ACP thioesterase as site of action of the commercial herbicides cumyluron, oxaziclomefone, bromobutide, methyldymron and tebutam," *Pest Manag Sci.* 78: 3620-3629, 2022).

[0060] The latter class of agents is deployed after a cell population is mutagenized as a means to enrich or differentiate truly mutagenized cells from those that are not, by looking for phenotypes that lie outside those of the 'wild type' or non-mutagenized control population. A person in the art would understand that such control populations, or 'mock controls', are always assessed alongside the mutagenized population when using enrichment agents as a means to assess the degree of divergence of the mutagenized population from the control population. Selective or enrichment agents can be both physical (e.g., tolerance to extremes in temperature, the buoyancy of the mutagenized cells in percoll, sucrose, or other gradients, the intensity with which cells take up neutral lipid dyes such as Nile Red or 4,4-difluoro-4-bora-3a,4a-diaza-s-indacene-BODIPY, and a means to sort highly fluorescent populations from less fluorescent ones) or chemical (e.g., the resistance of mutagenized populations to tolerate inhibition by pharmacological agents, herbicides, salts). Classical strain improvement methods do not include recombinant genetic engineering methods targeted to one or more genomic regions, e.g., via homologous recombination, random insertion, markerless transformation, or CRISPR-Cas9-based technologies.

[0061] As used herein, the term "microbial oil" refers to an oil produced or extracted from a microorganism (microbe), e.g., an oleaginous, single-celled, eukaryotic, or prokaryotic microorganism, including but not limited to, microalgae, yeast, bacteria, and fungi. Preferably, the microorganism is a microalgal organism. More preferably, the microalgal organism is a *Prototheca* strain and has been classically enhanced via exposure to chemical mutagens, compounds, temperature, and/or UV light.

[0062] As used herein, the term "triacylglycerol", "triglyceride", or "TAG" are synonymous and refer to esters between glycerol and three fatty acids. Generally, fatty acids of TAGs have chain lengths of 6 carbon atoms or more.

[0063] Because there are three possible positions on the glycerol backbone to which fatty acids can esterify, referred to as *sn-1*, -2, and -3, the number of possible TAG molecules in an oil is equal to the number of fatty acid species present in the oil raised to the 3rd power. By example, an oil comprised of seven different fatty acids will have 343 possible TAG species.

[0064] Regioisomers of TAGs are triacylglycerol molecules that have the same fatty acid composition and molecular weight but whose fatty acid moieties are arrayed in a different order, e.g., POP contains palmitate at the *sn-1* and *sn-3* positions and oleate at *sn-2*, while PPO contains palmitate both at *sn-1* and *sn-2* and oleate at *sn-3*. PPP, with palmitate at *sn-1*, -2, and -3, has no regioisomer.

[0065] As used herein, the term "TAG purity", "molecular purity", or "oil purity" refers to the number of molecular species that make up an oil composition on an absolute basis or present in amounts above a certain threshold. The fewer the number of TAG species in an oil, the greater the "purity" of the oil.

[0066] Properties of oils and their modes of characterization, processing, characteristics, and compositions are well known (*see e.g.*, THE LIPID HANDBOOK. 2007. Gunstone, F.D., Harwood, J.L. and Dijkstra, A.J Eds. CRC Press Taylor and Francis Group).

[0067] As used here, the terms "oil" or "lipid" are synonymous and refer to all the TAGs comprising the oil and lipid, accounting for >98% of the molecules in the lipid or oil; the remaining components are comprised of free fatty acids, sterols, color bodies, metals, and phospholipids. This is largely the composition of oil or lipid as it exists in the microalgal cell prior to being separated from the cell.

[0068] The term "fats" as used here, refers to oils or lipids that are solid at room temperature, about 20-25 °C.

[0069] A microalgal oil can be further categorized based on the degree to which it has been purified and separated from the other cellular debris or non-lipid biomass. This separation can be affected using mechanical means, i.e., by expeller pressing, or by chemical means, i.e., by solvent extraction and bead milling, where solvents include combinations of hydrocarbons such as hexane, ethanol, or propane used alone or separately.

[0070] Solvents can also include aqueous solvents comprised of water and salts such as ammonium sulfate, magnesium sulfate, sodium chloride, etc.

[0071] Once the oil is separated from the non-lipid biomass either by mechanical pressing or solvent extraction, the resulting oil is termed "crude oil."

[0072] In the case of solvent-extracted oils, the solvent-oil mixture is termed "micellae" and, depending on the solvent type, the solvent is removed, which can be further purified via passage through a centrifuge or filter press to remove "fines" or small particles of non-lipid biomass.

[0073] The composition of the “crude oil” is comprised of TAGs, free fatty acids, sterols, color bodies, metals, and phospholipids.

[0074] Crude oils can be further processed by refining. Refining oil serves to remove free fatty acids and can be achieved by adding NaOH or “caustic” to create soaps, which are removed by centrifugation.

[0075] Refined oils are further processed by degumming, which includes neutralization of residual caustic with citric acid and water washing, followed by centrifugation to remove the precipitated “gums” or phospholipid fraction of the crude oil.

[0076] The refined, degummed oil can next be bleached, using earthen bleaching clays to remove residual free fatty acids, metals including Ca, PO₄, Fe, Cl, and K, among others, and color bodies that can affect oil quality and appearance.

[0077] Refined, degummed, and bleached oils are finally purified via deodorization, where steam can be used to strip the oil under a vacuum of the volatile components and off-putting flavors. The final oil is termed an RBD oil (i.e., refined, bleached and deodorized oil).

[0078] In the present disclosure, the fatty acid composition or FAME profile and the TAG composition or profile as determined from an intact cell are equivalent to what is determined from an RBD oil for two reasons. First, the analytical methods used to determine FAME and TAG profiles, whether from whole cells or RBD oil affect a certain purification of the oil away from the cellular debris. Second, the oils within the cells are already highly pure materials, >98% TAGs.

[0079] As used herein, the terms “diacylglycerol”, “diglyceride”, and “DAG” refer to glycerides having two fatty acids esterified to a glycerol molecule.

[0080] As used herein, the terms “monoacylglycerol”, “monoglyceride”, and “MAG” refer to glycerides having one fatty acid esterified to a molecule of glycerol.

[0081] As used herein, the terms “fatty acid profile” or “fatty acid content” refer to a fatty acid composition of an oil, e.g., an oil produced by a cell provided herein or a derivative thereof. Derivatives of an oil produced by a cell provided herein include a refined, bleached, and deodorized oil. Fatty acid profiles can be determined by subjecting an oil to transesterification to generate fatty acid methyl esters (FAMES) and subsequently quantitating fatty acid type by Gas-Chromatography (GC) equipped with a Flame-Ionization Detector (GC/FID).

[0082] As used herein, the term “sterol profile” refers to a sterol composition of an oil (e.g., an oil produced by a cell provided herein or a derivative thereof). Derivatives of an oil produced by a cell provided herein include a refined, bleached, and deodorized oil.

[0083] As used herein, "oleic content", "oleic acid content", "oleate content", and "olein content" refer to the percentage amount of oleic acid in the fatty acid profile of a substance (e.g., a TAG oil).

[0084] As used herein, the term "C18:1 content" refers to the percentage amount of a C18:1 fatty acid (e.g., oleic acid) in the fatty acid profile of a substance (e.g., a microbial oil).

[0085] As used herein, "linoleic content" and "linoleic acid content" refer to the percentage amount of linoleic acid in the fatty acid profile of a substance (e.g., a TAG oil)

[0086] As used herein, the term "C18:2 content" refers to the percentage amount of a C18:2 fatty acid (e.g., linoleic acid) in the fatty acid profile of a substance (e.g., a microbial oil).

[0087] As used herein, "palmitate content", "palmitic content", and "palmitic acid content" refer to the percentage amount of palmitic acid in the fatty acid profile of a substance (e.g., a microalgal oil).

[0088] As used herein, the term "C16:0 content" refers to the percentage amount of C16:0 fatty acid (e.g., palmitic acid) in the fatty acid profile of a substance (e.g., a microbial oil).

[0089] As used herein, "PPP", "tripalmitin", "tripalmitate", and "tripalmitoylglycerol" refer to a triacylglycerol having palmitic acid or palmitate as the only fatty acid on the glycerol backbone. The PPP content or PPP level is generally discussed as a percentage of the TAG component.

[0090] The term "POP" or "palmitic-oleic-palmitic" can refer to a triacylglycerol having palmitic acid or palmitate at the *sn-1* and *sn-3* positions on the glycerol backbone and oleate at the *sn-2* position on the glycerol backbone. The POP content or POP level is generally discussed as a percentage of the TAG component.

[0091] The term "PPO" or "palmitatic-palmitic-oleate" can refer to a triacylglycerol having palmitic acid or palmitate at the *sn-1* and *sn-2* positions on the glycerol backbone and oleate at the *sn-3* position on the glycerol backbone. The PPO content or PPO level is generally discussed as a percentage of the TAG component.

[0092] As used herein, the term "high palmitate" or "high palmitic" can refer to greater than 30% palmitic acid, or greater than 35% palmitic acid, greater than 40% palmitic acid, greater than 45% palmitic acid, greater than 50% palmitic acid, greater than 55% palmitic acid, or greater than 60% palmitic acid.

[0093] As used herein, the term "low linoleic content" or "low linoleic acid content" can refer to lower than 10%, lower than 9%, lower than 8%, lower than 7%, lower than 6%,

lower than 5%, lower than 4%, lower than 3%, lower than 2% or lower than 1% linoleic acid of the total fatty acid content.

[0094] As used herein, the term “low MAG” or “low monoacylglycerol” can refer to lower than 5%, lower than 4%, lower than 3%, lower than 2%, or lower than 1% of MAG of the total lipid content of an oil.

[0095] As used herein, the term “low DAG” or “low diacylglycerol” can refer to lower than 5%, lower than 4%, lower than 3%, lower than 2%, or lower than 1% of DAG of the total lipid content of an oil.

[0096] As used herein, “sequence identity” refers to a percentage of identical nucleotide or amino acid residues between two sequences being compared after an optimal alignment of sequences. An optimal alignment of sequences may be produced manually or by means of computer programs that use a sequence alignment algorithm (e.g., Clustal W, T-coffee, COBALT, BestFit, FASTA, BLASTP, BLASTN, and TFASTA). Sequence identity can be calculated by determining the number of identical positions between the two sequences being compared, dividing this number by the number of positions compared, and multiplying the result obtained by 100 to obtain the sequence identity between the two sequences; this calculation can be done for DNA, RNA, and protein sequences.

[0097] As used herein, the term “about” refers to $\pm 10\%$ of the value provided.

[0098] As used herein, cells are grown in a growth media suitable for the cells being used during classic enrichment and selection strategies. For example, in one embodiment, *Prototheca* microalgal cells from a specific *Prototheca* strain can use a vegetative growth media comprising NaH_2PO_4 , K_2HPO_4 , citric acid monohydrate, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$, dextrose, $(\text{NH}_4)_2\text{SO}_4$, Sigma antifoam 204 at final concentrations of 13.6 mM, 11.4 mM, 5 mM, 0.15 mM, 40 g/L, 7.5 mM and 0.23 g/L, respectively. The vegetative growth media also can comprise vitamins such as thiamine HCl, D-pantothenic acid, biotin, cyanocobalamin, riboflavin, and pyridoxine HCl at final concentrations of $2.96\text{E-}06$, $2.24\text{E-}07$, $6.53\text{E-}09$, $8.30\text{E-}11$, $1.33\text{E-}08$, and $1.27\text{E-}08$ molar, respectively. The vegetative growth media can further comprise one or more of the following micronutrients: H_3BO_3 , $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, $\text{NaMoO}_4 \cdot 2\text{H}_2\text{O}$, $\text{Ni}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$, citric acid monohydrate, $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ and $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ at final concentrations of $2.96\text{E-}05$, $1.22\text{E-}05\text{M}$, $1.46\text{E-}05$, $3.97\text{E-}07$, $2.75\text{E-}07$, $1.95\text{E-}04$, $3.96\text{E-}07$, and $5.4\text{E-}06$ molar, respectively. The *Prototheca* cells in the log phase of growth in the vegetative medium are then subjected to mutagenesis by means of chemicals or UV light (FIG. 3, Step 1). The exemplified *Prototheca* cells are then sub-cultured into a lipid production medium where the cells are then subjected to selection/enrichment strategies (FIG. 3, Step 2).

A suitable lipid production medium can be used to culture the cells. For *Prototheca* cells, a suitable culture media comprises the same vitamins and micronutrients at the same concentrations as used in a vegetative medium. Macronutrients in the lipid production medium, relative to the vegetative medium, can be adjusted to include 25 mM citric acid monohydrate, 50 g/L dextrose, and 1.5 mM (NH₄)₂SO₄. Microalgal cells can then be plated on a suitable solid vegetative medium containing 0.8% agarose to obtain clonal isolates (FIG. 3, Step 3). Once suitable clonal isolates are obtained, the isolates are interrogated in a suitable lipid production medium in a 96-well format (FIG. 3, Step 4). Using glucose consumption as a surrogate for oil production, high glucose consuming microalgal cell strains can be validated by culturing in a suitable lipid production medium in a tube or shake flask format (FIG. 3, Step 5). Clonal isolates that are successfully validated are then subcultured for multiple generations to stabilize mutations (FIG. 3, Step 6), followed by purification of clonal isolates (FIG. 3, Step 7). The clonal isolates that are stable can undergo subsequent re-interrogation in a lipid production medium (FIG. 3, Step 8). Clones deemed to be phenotypically stable (e.g., clones having %cv <5% for profile, glucose consumption, and oil titer) after the second interrogation round are ready for validation in fermentation (FIG. 3, Step 9), while clones that still demonstrate variability in their phenotype (FIG. 3, Step 8), are passaged once more (FIG. 3, Step 6) to generate stable clonal lines. Unless indicated otherwise, microalgal *Prototheca* cells were grown at about 28 °C with shaking at 200 rpm for cultures grown in tubes or shake flasks while blocks were grown with shaking at 900 rpm.

[0099] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present teachings, some exemplary methods and materials are described herein.

Microbial Cells and Oils Produced Therefrom

[0100] An oil provided herein is obtained from a non-genetically modified microorganism (microbe), for example, oleaginous microalgae, yeast, or bacteria. In some embodiments, the non-genetically modified microorganism is a microalgal cell. In some embodiments, the microalgal cell is a non-genetically modified *Prototheca* sp. strain. The non-genetically modified *Prototheca* sp. strain can be produced by one or more classical strain improvement strategies described herein.

[0101] In some embodiments, a cell provided herein does not comprise an exogenous gene or exogenous nucleotides that encode for an exogenous protein or gene. For example, a cell provided herein does not comprise an exogenous gene in a lipid biosynthetic pathway. In some embodiments, a cell provided herein does not comprise a genetic disruption of one or more alleles of an endogenous gene in a lipid biosynthetic pathway. In some embodiments, a cell provided herein does not comprise a heterologous insertion within a genomic region that encodes for an endogenous gene in a lipid biosynthetic pathway. Non-limiting examples of genes involved in lipid biosynthesis include acyl-ACP thioesterase (FAT), delta-12 fatty acid desaturase (FAD), ketoacyl-ACP synthase (KAS), stearoyl-ACP desaturase (SAD), lysophosphatidic acid acyltransferase (LPAAT), ketoacyl-CoA reductase (KCR), hydroxyacyl-CoA dehydratase (HACD), and enoyl-CoA reductase (ECR).

[0102] In some embodiments, a cell provided herein does not comprise an exogenous acyl-ACP thioesterase gene. In some embodiments, a cell provided herein does not comprise a genetic disruption of one or more alleles of an endogenous acyl-ACP thioesterase gene. In some embodiments, a cell provided herein does not comprise a heterologous insertion within a genomic region that encodes for an endogenous acyl-ACP thioesterase gene. In some embodiments, the endogenous acyl-ACP thioesterase gene is FATA (fatty-acyl-ACP A).

[0103] In some embodiments, a cell provided herein does not comprise an exogenous fatty acid desaturase gene. In some embodiments, a cell provided herein does not comprise a genetic disruption of one or more alleles of an endogenous fatty acid desaturase gene. In some embodiments, a cell provided herein does not comprise a heterologous insertion within a genomic region that encodes for an endogenous fatty acid desaturase gene. In some embodiments, the endogenous fatty acid desaturase gene is a delta-12 fatty acid desaturase. In some embodiments, the endogenous fatty acid desaturase gene is FAD2.

[0104] In some embodiments, a cell provided herein does not comprise an exogenous ketoacyl-ACP synthase gene. In some embodiments, a cell provided herein does not comprise a genetic disruption of one or more alleles of an endogenous ketoacyl-ACP synthase gene. In some embodiments, a cell provided herein does not comprise a heterologous insertion within a genomic region that encodes for an endogenous ketoacyl-ACP synthase gene. In some embodiments, the endogenous ketoacyl-ACP synthase gene is KASI, KASII, or KASIII.

[0105] In some embodiments, a cell provided herein does not comprise an exogenous stearoyl-ACP desaturase gene. In some embodiments, a cell provided herein does not comprise a genetic disruption of one or more alleles of an endogenous stearoyl-ACP desaturase gene. In some embodiments, a cell provided herein does not comprise a heterologous insertion within a

genomic region that encodes for an endogenous stearyl-ACP desaturase gene. In some embodiments, the endogenous stearyl-ACP desaturase gene is SAD2.

[0106] In some embodiments, a cell provided herein does not comprise an exogenous lysophosphatidic acid acyltransferase gene. In some embodiments, a cell provided herein does not comprise a genetic disruption of one or more alleles of an endogenous lysophosphatidic acid acyltransferase gene. In some embodiments, a cell provided herein does not comprise a heterologous insertion within a genomic region that encodes for an endogenous lysophosphatidic acid acyltransferase gene.

[0107] In some embodiments, a cell provided herein does not comprise a genetic disruption of one or more alleles within 1.5 kb of an endogenous V-type proton ATPase catalytic subunit A isoform 1 gene or 6S genomic region. In some embodiments, a cell provided herein does not comprise a heterologous insertion within 1.5 kb of an endogenous V-type proton ATPase catalytic subunit A isoform 1 gene or 6S genomic region.

[0108] In some embodiments, a cell provided herein does not comprise a genetic disruption of one or more alleles within 1.5 kb of an endogenous DAO1B gene. In some embodiments, a cell provided herein does not comprise a heterologous insertion within 1.5 kb of an endogenous DAO1B gene.

[0109] In some embodiments, a cell provided herein does not comprise a genetic disruption of one or more alleles within 1.5 kb of an endogenous Thi4 gene. In some embodiments, a cell provided herein does not comprise a heterologous insertion within 1.5 kb of an endogenous Thi4 gene.

[0110] Accordingly, a cell provided herein comprises uninterrupted sequences of endogenous genes or genomic regions, including FAD2, FATA1, KASII, SAD2, V-type proton ATPase catalytic subunit A isoform 1 (6S), DAO1B, and Thi4 described herein.

[0111] In some embodiments, a cell provided herein comprises at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity to any one of SEQ ID NO:1-26 (shown in Table 11 of the present disclosure). In some embodiments, a cell provided herein comprises any one of SEQ ID NO:1-26 (genomic regions of CHK22 and CHK100). In some embodiments, a cell provided herein comprises SEQ ID NO:1-26 (genomic regions of CHK22 and CHK100).

[0112] In some embodiments, a cell provided herein comprises at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity to any one of SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 18, 19, 21, 23, and 25 (genomic regions of CHK22 and CHK100). In some embodiments, a cell provided herein comprises any one of SEQ ID NO:1, 3, 5, 7, 9, 11,

13, 15, 18, 19, 21, 23, and 25 (genomic regions of CHK22 and CHK100). In some embodiments, a cell provided herein comprises SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 18, 19, 21, 23, and 25 (genomic regions of CHK22 and CHK100).

[0113] In some embodiments, an oil provided herein is produced by microalgae. In some embodiments, the microalgae is a species of a genus selected from the group consisting of: *Chlorella sp.*, *Pseudochlorella sp.*, *Heterochlorella sp.*, *Prototheca sp.*, *Arthrospira sp.*, *Euglena sp.*, *Nannochloropsis sp.*, *Phaeodactylum sp.*, *Chlamydomonas sp.*, *Scenedesmus sp.*, *Ostreococcus sp.*, *Selenastrum sp.*, *Haematococcus sp.*, *Nitzschia*, *Dunaliella*, *Navicula sp.*, *Trebouxia sp.*, *Pseudotrebouxia sp.*, *Vavicula sp.*, *Bracteococcus sp.*, *Gomophonema sp.*, *Watanabea sp.*, *Botryococcus sp.*, *Tetraselmis sp.*, and *Isochryis sp.* In some embodiments, the microalgae is *Prototheca sp.* In some embodiments, the microalgae are a *P. moriformis* strain. In some embodiments, the microalgae are a *P. wickerhamii* strain.

[0114] In some embodiments, the microalgal cell described herein for use in manufacturing microbial oil is derived from a UTEX 1435 base strain. In some embodiments, a cell provided herein can be derived from a UTEX 1533 base strain. In some embodiments, a cell provided herein is derived from a base strain having a 23S ribosomal DNA sequence that is at least 80%, at least 85%, at least 90%, at least 95%, or 100% identity to SEQ ID NO:30 or SEQ ID NO:31. In some embodiments, a cell provided herein has a 23S ribosomal DNA sequence that is at least 80%, at least 85%, at least 90%, at least 95%, or 100% identity to SEQ ID NO:30 or SEQ ID NO:31.

[0115] In some embodiments, an oil provided herein can be produced by oleaginous yeast. In some embodiments, the oleaginous yeast is a species of a genus selected from the group consisting of: *Candida sp.*, *Cryptococcus sp.*, *Debaromyces sp.*, *Endomycopsis sp.*, *Geotrichum sp.*, *Hyphopichia sp.*, *Lipomyces sp.*, *Pichia sp.*, *Rodosporidium sp.*, *Rhodotorula sp.*, *Sporobolomyces sp.*, *Starmerella sp.*, *Torulaspora sp.*, *Trichosporon sp.*, *Wickerhamomyces sp.*, *Yarrowia sp.*, and *Zygoascus sp.*

[0116] In some embodiments, an oil provided herein is obtained from or produced by oleaginous bacteria. In some embodiments, the oleaginous bacteria is a species selected from the group consisting of: *Flavimonas oryzae*, *Pseudomonas aeruginosa*, *Morococcus sp.*, *Rhodobacter sphaeroides*, *Rhodococcus opacus*, *Rhodococcus erythropolis*, *Streptomyces jeddahensis*, *Ochrobactrum sp.*, *Arthrobacter sp.*, *Nocardia sp.*, *Mycobacteria sp.*, *Gordonia sp.*, *Catenisphaera sp.*, and *Dietzia sp.*

[0117] Further provided herein are bioreactors comprising a non-naturally occurring microorganism. For example, these bioreactors comprise an oleaginous, non-naturally occurring microorganism and an oil produced by the microorganism.

[0118] While in many embodiments, an oil provided herein is obtained from a non-genetically modified microorganism or a classically-improved microorganism, in other embodiments, an oil provided herein is obtained from a genetically modified microorganism, for example, oleaginous microalgae, yeast, or bacteria. In some embodiments, an oil provided herein is obtained from a classically-improved microorganism that is then genetically modified to produce a genetically modified microorganism. In some embodiments, the genetically modified microorganism is a genetically modified *Prototheca sp.* strain.

Classical Strain Improvement.

[0119] “Classical strain improvement” and “classical microalgal strain improvement” or “classical microalgal cell improvement” are the terms used herein to characterize the strategy of selecting for organisms having desired phenotypes, e.g., high palmitic oil production. Classical strain improvement (also called “mutation breeding”) involves exposing organisms to chemicals or radiation to generate mutants with desirable traits. These classical strain improvement methods introduce random or semi-random mutations, which can thereby allow the selection of strains exhibiting desirable traits as a result of random mutagenesis. Several iterations of mutagenesis and selection can be performed with one or more mutagens to obtain a strain having desirable phenotypes. Ultraviolet (UV) light can be used to introduce random mutations within a microorganism's nuclear genome. Chemical mutagens include compounds that inhibit or disrupt the biosynthetic processes of a microorganism, e.g., antibiotics, antifungals, or carcinogens. Non-limiting examples of chemical mutagens include ICR-191, ethyl methanesulfonate (EMS), and 4-nitroquinoline- 1-oxide (4-NQO). Non-limiting examples of chemical mutagens include acridine mutagens, amino acid analogs, fatty acid biosynthesis inhibitors, cholesterol biosynthetic inhibitors, mTOR inhibitors, and membrane solubilizing agents. Combinations of chemical mutagens can also be used simultaneously (or serially) to induce mutagenesis. Following mutagenesis, selective or enrichment agents can be used to select or enrich for strains of interest. Non-limiting examples of enrichment agents include L-canavanine, cerulenin, triparanol, clomiphene, clomiphene citrate, clotrimazole, terfenadine, fluphenazine, AZD8055, BASF 13-338, cafenstrole, clomiphene, PF-042110, and phenethyl alcohol, as well as the use of a combination of enrichment agents.

[0120] By “an isolated cell” is meant a microalgal cell or other microorganism as set forth herein that has been isolated from its native environment and grown and cloned such that all the clones are the same.

[0121] Methods provided herein include classical strain improvement methods to improve strain productivity, carbon yield, and palmitic acid content. The glucose consumption rate of the microbial cells can be a highly predictive indicator of lipid titer and is known in the art. As such, the glucose consumption rate can be used as an enrichment tool in the mutant selection process. The methods provided herein can include one or more of determining a total lipid titer of the cell, determining a fatty acid profile of the oil, or determining a C16:0 (e.g., palmitic acid) content of the oil. Further, the methods provided herein can include assaying cell media to determine glucose consumption rate, total lipid titer, a fatty acid profile, a C16:0 (e.g., palmitic acid) content, and TAG profile.

High Palmitic Oils of the Disclosure.

[0122] The high palmitic oils of the microbial strains, e.g., microalgal strains, described herein are obtained from isolated clonal cells that have been manipulated using classical strain improvement strategies as discussed herein. Exemplary strains are *Prototheca* microalgal strains such as *P. wickerhamii* or *P. moriformis*.

[0123] The complexity and physical properties of an oil can be evaluated by the fatty acid and the TAG profile of the TAG component of a microalgal oil. The fatty acid profile is a measure of fatty acid composition and can be determined by subjecting an oil to transesterification to generate fatty acid methyl esters (FAMES) and subsequently quantitating fatty acid type by Gas-Chromatography equipped with a Flame-Ionization Detector (GC/FID). Accordingly, fatty acid content can be determined by GC/FID. Since TAGs comprise three fatty acids arrayed along the glycerol backbone in the TAG molecule, the number of possible distinct regioisomers of TAGs can be defined by the number of fatty acid species in the oil raised to the third power. The TAG profile provides relative amounts of various TAG species in an oil, which can be determined by subjecting the oil to TAG fractionation using Liquid-Chromatography/Time of Flight-Mass Spectrometry (LC/TOF-MS) equipped with an Atmospheric Pressure Chemical Ionization (APCI) source.

[0124] In some embodiments, an oil provided herein has a high palmitic acid content. For example, the fatty acid content of the TAG component of an oil provided herein can be high in palmitic acid, comparable to the palmitic acid content in palm olein derived from palm oil.

[0125] In some embodiments, an oil provided herein has a high palmitic acid content. For example, the fatty acid content of the TAG component of an oil provided herein can be high in palmitic acid, comparable to the palmitic acid content in palm olein derived from palm oil.

[0126] In some embodiments, an oil provided herein comprises saturated fatty acid content of at least 30%, at least 40%, at least 50%, or at least 60%. For example, an oil provided herein has a C16:0 content of at least about 30%, at least about 31%, at least about 32%, at least about 33%, at least about 34%, at least about 35%, at least about 36%, at least about 37%, at least about 38%, at least about 39%, at least about 40%, at least about 41%, at least about 42%, at least about 43%, at least about 44%, at least about 45%, at least about 46%, at least about 47%, at least about 48%, at least about 49%, at least about 50%, at least about 51%, at least about 52%, at least about 53%, at least about 54%, at least about 55%, at least about 56%, at least about 57%, at least about 58%, at least about 59%, at least about 60%, or more.

[0127] In some embodiments, an oil provided herein has a fatty acid content comprising at least 30%, at least 40%, at least 50%, or at least 60%, or more of C16:0 fatty acid. In some embodiments, the C16:0 fatty acid comprises palmitic acid.

[0128] In some embodiments, an oil provided herein has a fatty acid content of the TAG component comprising at least 40%, at least 45%, or at least 50%, or more of a C18:1 fatty acid. In some embodiments, the C18:1 fatty acid comprises oleic acid.

[0129] In some embodiments, the oil provided herein has a fatty acid content of the TAG component comprising less than 8%, less than 7%, less than 6%, less than 5%, or less of a C18:2 fatty acid. In some embodiments, the C18:2 fatty acid comprises linoleic acid.

[0130] In some embodiments, an oil provided herein comprises less than 5%, less than 4%, less than 3%, less than 2%, less than 1 %, less than 0.5%, less than 0.1%, or less tripalmitin.

[0131] In some embodiments, an oil provided herein comprises polyunsaturated fatty acids. In some embodiments, the polyunsaturated fatty acid comprises C18:2. In some embodiments, the oil has a C18:2 content from about 3% to about 9%, from about 4% to about 8%, or from about 5% to about 8%. In some embodiments, an oil provided herein can have a C18:2 content of less than 8%, less than 7%, less than 6%, or lower.

[0132] In some embodiments, the fatty acid content of the TAG component can comprise from about 35% to about 60% C16:0 fatty acids.

[0133] In addition to the fatty acid profile, an oil can be further evaluated by the sterol profile or composition. Sterol composition can be determined by mass spectrometry, for

example, gas chromatography-mass spectrometry (GC-MS), liquid chromatography-mass spectrometry (LC-MS), tandem mass spectrometry (MS/MS), and coupled liquid and gas chromatography with subsequent flame ionization detection (LC-GC-FID). The concentration of the different sterols present in an oil can be expressed as mg sterol/100 grams of oil. Non-limiting examples of sterols include ergosterol, ergost-8(14)-en-3-ol, (3 β), 5.Xi.-Ergost-7-en-3 β -, 9,19-cyclolanost-24-en-3-ol, (3B), and 9,19-cyclolanostan-3-ol, 24-methylene-, (3B). In some embodiments, ergosterol may be present in an amount of at least about 30mg/100 g oil. In some embodiments, 5.Xi.-Ergost-7-en-3 β -ol may be present in an amount of at least about 25 mg/100 g oil. In some embodiments, 9,19-Cyclolanost-24-en-3-ol, (3B)- may be present in an amount of at least 15 mg/100 g oil. In some embodiments, 19-cyclolanostan-3-ol, 24-methylene-, (3B) may be present in an amount of at least about 25 mg/100g oil.

[0134] In some embodiments, an oil provided herein can comprise less than 5%, less than 4%, less than 3%, less than 2%, less than 1%, less than 0.5%, less than 0.1%, or less of monoacylglycerols (MAG).

[0135] In some embodiments, an oil provided herein can comprise less than 5%, less than 4%, less than 3%, less than 2%, less than 1%, less than 0.5%, less than 0.1%, or less of diacylglycerols (DAG).

[0136] The microalgae-derived oil of the present disclosure can be useful as an alternative to palm oil in a variety of applications that utilize palmitic acid. A microalgae-derived oil of the present disclosure can deliver palmitic acid like palm olein and palm mid-fraction, the palm mid-fraction is valued for its high levels of the structuring fat, POP. However, the microalgae-derived oil of the present disclosure can also be advantageous over palm oil and its various fractions due to the reduced C18:2 fatty acid levels. The microalgae-derived oil of the present disclosure can comprise a lower level of C18:2 in comparison to palm oil. Because C18:2 fatty acids can be prone to oxidation, an oil comprising high levels of C18:2 can be more prone to oxidative degradation. Therefore, the oil of the present disclosure can be advantageous in food and nutrition product production that require long-shelf lives. The microalgae-derived oil of the present disclosure can also be advantageous over palm oil due to the reduced levels of MAG and DAG. MAG and DAG can be more sensitive to oxidation than TAG. MAG and DAG can also act as precursors for the formation of glycidyl esters, harmful process contaminants, during the thermal process of refining vegetable oils and fats.

[0137] For at least these reasons, the oil of the present disclosure can be advantageous in the production of food and nutrition products requiring the highest level of food safety.

Food, Nutrition and Home and Personal Care Applications

[0138] The oils described herein can be used or formulated with one or more excipients for a variety of food applications, including but not limited to, food products, e.g., food-grade oils, cooking oil, frying oils, coatings, salad dressings, spreads, frozen desserts, pharmaceuticals, nutritional supplements, nutraceuticals, meal replacements, infant formulas, beverages, flavoring agents, and food additives.

[0139] Further, the oils of the present disclosure may be used in personal and home care as alternatives to palm oils and palm oil fractions which are commonly used as feedstock to prepare surfactants and derivatives.

EXAMPLES**Example 1. Generation of a Classically Improved Microalgal Strain Capable of Producing a Triglyceride Oil that is Highly Enriched in Palmitic Acid.***Strain Acquisition, Isolation, and Characterization*

[0140] A *Prototheca wickerhamii* strain (UTEX 1533) was obtained from the University of Texas at Austin Culture Collection of Algae (UTEX). Analysis of the 23S ribosomal DNA (rDNA) sequence of UTEX 1533 (SEQ ID NO:30) suggests that UTEX 1533 is very closely related to a *Prototheca moriformis* strain, UTEX 1435 (SEQ ID NO:31). The 23S rDNA sequence of UTEX 1533 has 100% sequence identity with the published 23S rDNA sequence from UTEX 1435 (FIG. 1). The UTEX 1533 strain acquired from UTEX was verified as axenic through repeated passages on solid media containing various antibiotic and antifungal agents, followed by passage on antibiotic-free solid media to confirm a single cell type and a homogeneous morphology. This parent strain was given the designation “Strain 0” and is also referred to herein as CHK22.

First Round of Mutagenesis.

[0141] The classical strain improvement of CHK22 was undertaken to improve its productivity, yield on carbon, and to increase its palmitic acid content. The strain lineage of mutants generated in the examples are shown in FIG.2A and 2B, while an overview of the process flow for mutagenesis, trait selection, and high-throughput, automated screening steps of the improvement process is outlined in FIG. 3. As shown in FIG. 2A and 2B, algal cells were cultured in a vegetative growth medium where indicated, comprised of the following components including macronutrients comprised of NaH₂PO₄, K₂HPO₄, citric acid monohydrate, MgSO₄•7H₂O, CuCl₂•2H₂O, dextrose, (NH₄)₂SO₄, Sigma antifoam 204, at final concentrations of 13.6 mM, 11.4 mM, 5 mM, 0.15 mM, 40 g/L, 7.5 mM and 0.23 g/L,

respectively. Vitamins were present as follows in the vegetative growth media: Thiamine HCl, D-pantothenic Acid, Biotin, Cyanocobalamin, Riboflavin, and Pyridoxine HCL at final concentrations of 2.96E-06, 2.24E-07, 6.53E-09, 8.30E-11, 1.33E-08 and 1.27E-08 molar, respectively. Micronutrients were present in the vegetative growth media used as follows: H₃BO₃, ZnSO₄•7H₂O, MnSO₄•H₂O, NaMoO₄•2H₂O, Ni(NO₃)₂•6H₂O, citric acid monohydrate, CuSO₄•5H₂O and FeSO₄•7H₂O at final concentrations of 2.96E-05, 1.22E-05M, 1.46E-05, 3.97E-07, 2.75E-07, 1.95E-04, 3.96E-07, and 5.4E-06 molar, respectively. Cells in the log phase of growth in the vegetative medium, were subjected to mutagenesis by means of chemicals or UV light (FIG. 3, Step 1). Cells were then sub-cultured into the lipid production medium, where they were subjected to selection/enrichment strategies (FIG. 3, Step 2).

[0142] The lipid production medium was comprised of the same vitamins and micronutrients at the same concentrations as used in the vegetative medium. Macronutrients in the lipid production medium, relative to the vegetative medium, were adjusted to include 25 mM citric acid monohydrate, 50 g/L dextrose, and 1.5 mM (NH₄)₂SO₄. The microalgal strains were plated to a solid medium vegetative medium containing 0.8% agarose to obtain clonal isolates (FIG. 3, Step 3), followed by the interrogation of these isolates in the lipid production medium in a 96-well format (FIG. 3, Step 4).

[0143] Using glucose consumption as a surrogate for oil production, high glucose-consuming strains can be validated in a lipid production medium in a tube or shake flask format (FIG. 3, Step 5). Isolates that successfully were validated were sub-cultured for multiple generations to stabilize mutations (FIG. 3, Step 6), followed by purification of clonal isolates (FIG. 3, Step 7) and subsequent re-interrogation in the lipid production medium (FIG. 3, Step 8). Clones deemed to be phenotypically stable (% cv <5% for profile, glucose consumption and oil titer) were then validated via fermentation (FIG. 3, Step 9), while clones that still exhibited variability in (FIG. 3, Step 8) were passaged once more (FIG. 3, Step 6) to generate stable lines. Unless indicated otherwise, cells were grown at about 28 °C with shaking at 200 rpm for cultures grown in tubes or shake flasks, while blocks were grown with shaking at 900 rpm.

[0144] As detailed in FIG. 3, CHK22 was chemically mutagenized with 44 μM 4-nitroquinoline 1-oxide (4-NQO) or subjected to sham mutagenesis with the addition of only the mutagen solvent, dimethyl sulfoxide (DMSO). The mutagen was inactivated with the addition of sodium thiosulfate which was then removed with repeated washing in with water. The cells were then allowed to recover in limited a sugar growth media. The limited sugar

growth media is the same as the vegetative media, but contains 5 g/L glucose. The mutagenized and the mock-mutagenized populations were then each independently cultured in the lipid production media at 38 °C for five days. The optimal lipid production temperature for CHK22 is typically between 28-32 °C, so the higher-than-optimal temperature was applied as a stressor, with the goal of providing an advantage to those mutants with a predisposition toward exhibiting higher palmitate levels. A similar phenomenon has been observed in both fungal model systems and in some commercially relevant seed oil-producing higher plant species following chronic exposure to higher temperatures during growth or cultivation, respectively. At the end of the 5-day, 38 °C lipid culture, the cells were collected and then exposed next to 65 °C for 4 min. Based on range-finding experiments conducted in the same format, exposure to temperatures in this range for this duration of time typically kills >99% of the cell population. After heat exposure, the cells were recovered in limited sugar growth media for three days and then diluted and plated. Mutant clones were then picked and assessed for glucose consumption rate and fatty acid profile in a 72 hrs, 96-well block-based lipid production assay (FIG. 2A). The first three rows in Table 1 depict values for glucose consumption rate and fatty acid profile for the CHK22 parental strain, tested in triplicate. The last three columns in Table 1 depict CHK22 mutants. As shown in Table 1, CHK22 mutant 1 exhibited nearly 8% more palmitate than the CHK22 parent, albeit with a significant reduction in glucose consumption rate, a hallmark of lower lipid titer.

Table 1: Fatty acid profiles of select CHK22 mutant strains exhibiting higher glucose consumption in a 96-well block lipid assay and increased palmitate levels.

	Sample					
	CHK22	CHK22	CHK22	CHK22 mutant 1	CHK22 mutant 2	CHK22 mutant 3
Glucose consumption Rate (g/L*D)	7.00	7.17	7.18	5.26	7.11	6.90
C14:0	1.51	1.58	1.59	5.38	1.68	1.53
C16:0	27.92	28.09	28.61	35.86	31.02	30.60
C16:1	0.84	0.76	0.85	1.38	1.00	0.92
C18:0	3.16	3.38	3.29	2.16	2.70	3.00
C18:1	56.73	56.11	55.89	43.45	53.39	54.78
C18:2	8.19	8.23	8.02	9.40	8.61	7.57

	Sample					
	CHK22	CHK22	CHK22	CHK22 mutant 1	CHK22 mutant 2	CHK22 mutant 3
C18:3 alpha	0.59	0.61	0.63	0.76	0.54	0.56
C20:0	0.33	0.35	0.31	0.26	0.26	0.26
C20:1	0.10	0.13	0.10	0.11	0.11	0.10

[0145] As shown in FIG. 2A, mutant strains were grown in lipid production medium (0.5 mL) with shaking (900 rpm) for 72 hours at 28 °C at which point glucose consumption was measured. Isolates with high glucose consumption levels were further interrogated for fatty acid composition with lead strains shown in Table 1.

[0146] To further adapt the CHK22 mutant 1 to its altered fatty acid composition and improve its lipid productivity, CHK22 mutant 1 was serially passaged in the absence of any selection and then plated to solid media to yield single colonies. These sub-clones were screened in a lipid production assay to assess the phenotypic stability of the lineage (e.g., %cv <5%) for both glucose consumption rate and fatty acid profile. After one cycle of serial passaging, the mutant was still phenotypically unstable. Several sub-clones that retained the higher palmitate level phenotype, along with increased rates of glucose consumption, were taken through additional rounds of serial passaging and stability assessment. After a total of five rounds of serial passaging and stability assessment, a stable strain, designated CHK100, was identified. CHK100 exhibited about 5% more palmitate and had an oil content, dry cell weight, and lipid titer greater than the CHK22 parental strain (FIG. 2A).

Table 2: Tube based assays on CHK100, a classically improved derivative of strain CHK22, showing increased C16:0 content and improved lipid titer.

	Sample							
	CHK100 (1)	CHK100 (2)	CHK100 (3)	CHK100 (4)	CHK22 (1)	CHK22 (2)	CHK22 (3)	CHK22 (4)
Lipid% w/w	78.6	78.0	75.5	77.1	77.1	76.1	76.0	74.7
Avg. Lipid %w/w	77.3				76.2			
DCW (g/L)	11.4	11.3	11.6	11.6	8.8	9.1	9.4	9.6
Avg. DCW (g/L)	11.5				9.2			
Lipid Titer (g/L)	9.0	8.8	8.8	8.9	6.8	6.9	7.1	7.2
NLB (g/L)	2.4	2.5	2.8	2.7	2.0	2.2	2.3	2.4
PCP	3.7	3.5	3.1	3.4	3.5	3.2	3.2	3.0
C14:0	1.22	1.22	1.18	1.18	1.39	1.39	1.36	1.36
C14:1	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.02
C15:0	0.03	0.03	0.03	0.04	0.03	0.03	0.04	0.03
C16:0	32.25	32.25	32.07	32.15	27.59	27.62	27.39	27.37
C16:1	1.36	1.36	1.40	1.38	0.85	0.85	0.85	0.85
C18:0	2.03	2.02	1.99	1.99	3.27	3.27	3.27	3.29
C18:1	55.65	55.61	55.89	55.91	57.92	57.86	58.28	58.34
C18:2	6.36	6.37	6.27	6.25	7.45	7.44	7.30	7.30
C18:3 alpha	0.50	0.53	0.57	0.55	0.64	0.63	0.68	0.67
C20:0	0.19	0.19	0.18	0.19	0.33	0.32	0.31	0.32
C20:1	0.05	0.05	0.05	0.05	0.07	0.07	0.07	0.07
C21:0	0.05	0.05	0.05	0.04	0.06	0.06	0.06	0.06

[0147] As shown in FIG. 2A, all strains were run under standard lipid production conditions in which strains were grown in duplicate in 10 mL of lipid production medium. Cultures were grown with shaking (200 rpm) for 121 hours at 28 °C at which point 1 mL of biomass was removed, applied to a polycarbonate filter, washed with an equal volume of Milli Q water, and placed in a tared glass vial at -80 °C for 30 minutes. Vials containing filters and frozen biomass were lyophilized to dryness overnight, their weights recorded, and filters with

dried biomass were subjected to direct transesterification followed by GC/FID to quantitate FAMES.

Mutagenesis of CHK100

[0148] CHK100 cells were recovered from a 24-hour lipid production culture and chemically mutagenized for 30 min at 32 °C with 22 µM 4-nitroquinoline 1-oxide (4-NQO) or subjected to sham mutagenesis with the addition of only the mutagen solvent, DMSO. The mutagen was inactivated with the addition of sodium thiosulfate which was then removed with repeated washing with water and media. The cells were then allowed to recover in limited sugar growth media. The mutagenized and the mock-mutagenized populations were then each independently cultured, first in seed media for 24 h at 38 °C and then in lipid production media for three days at 38 °C. At the end of the 3 days in the 38 °C lipid culture, the cells were collected and then exposed next to 65 °C for 1, 2, 3, or 4 min. After heat exposure, the surviving cells were recovered in limited sugar growth media for three days and then diluted and plated. Viable cells were only recovered from the 1 min at 65 °C exposure condition. Mutant clones were then picked and assessed for glucose consumption rate and fatty acid profile in a 72 h, 96-well block-based lipid production assay (FIG. 2A and 2B). Two of the mutants identified (Table 3; highlighted in bold font and underlined) exhibited 4-4.5% more palmitate than the CHK100 parent while maintaining a reasonable glucose consumption rate, which is typically a hallmark of a strain having an acceptable lipid titer.

Table 3: Fatty acid profiles of select mutant strains exhibit higher glucose consumption rate (g/L*d, grams per liter per day cultured) in a 96-well block lipid assay and increased palmitate levels.

CHK100 Controls	Gluc. Cons. Rt.	C14:0	C14:1	C15:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3 alpha	C20:0	C20:1
1	7.75	1.50	0.04	0.04	33.69	1.21	2.38	53.05	6.87	0.54	0.21	0.10
2	8.99	1.37	0.04	0.05	33.55	1.26	2.32	53.68	6.61	0.49	0.20	0.10
3	8.41	1.41	0.02	0.05	33.26	1.17	2.42	53.53	6.72	0.50	0.21	0.09
4	-	1.43	0.03	0.04	33.10	1.19	2.43	53.33	6.78	0.50	0.21	0.10
5	9.23	1.40	0.00	0.04	33.02	1.31	2.24	54.35	6.58	0.49	0.17	0.04
6	8.01	1.45	0.03	0.04	32.97	1.16	2.39	53.99	6.77	0.50	0.20	0.08
7	8.03	1.43	0.02	0.04	32.78	1.14	2.41	54.15	6.78	0.50	0.21	0.08
8	8.28	1.36	0.03	0.04	32.67	1.18	2.44	54.14	6.87	0.53	0.20	0.10
9	9.30	1.31	0.03	0.03	32.57	1.20	2.27	54.87	6.54	0.49	0.18	0.08

CHK100 Controls	Gluc. Cons. Rt.	C14:0	C14:1	C15:0	C16:0	C16:1	C18:0	C18:1	C18:2	C18:3 alpha	C20:0	C20:1
10	8.66	1.34	0.03	0.05	32.53	1.17	2.43	54.43	6.69	0.53	0.21	0.09
11	8.26	1.36	0.02	0.04	32.32	1.11	2.39	54.90	6.67	0.49	0.20	0.10
12	8.31	1.37	0.02	0.04	32.26	1.11	2.40	54.81	6.79	0.50	0.21	0.10
13	8.30	1.35	0.02	0.04	32.16	1.09	2.40	55.09	6.71	0.50	0.21	0.09
14	8.37	1.34	0.02	0.04	32.08	1.13	2.41	55.01	6.77	0.51	0.20	0.09
15	8.29	1.34	0.03	0.03	32.02	1.10	2.39	55.23	6.71	0.50	0.20	0.10
16	8.67	1.33	0.03	0.03	31.96	1.11	2.36	55.25	6.73	0.51	0.20	0.09
17	8.30	1.31	0.03	0.04	31.72	1.08	2.37	55.50	6.72	0.51	0.20	0.09
18	8.75	1.27	0.03	0.04	31.43	1.10	2.32	55.93	6.72	0.51	0.20	0.09
Avg.	8.47	1.37	0.03	0.04	32.56	1.16	2.38	54.51	6.72	0.51	0.20	0.10
CHK100 mutant isolate 2	8.83	1.68	0.04	0.04	<u>37.27</u>	1.14	2.44	49.48	6.72	0.47	0.22	0.08
CHK100 mutant isolate 3	8.82	1.60	0.03	0.04	<u>36.51</u>	1.24	2.24	51.04	6.21	0.45	0.20	0.08

[0149] CHK100 mutant strains were grown in the lipid production medium (0.5 mL) described above with shaking (900 rpm) for 72 hours at which point glucose consumption was measured. Isolates with high glucose consumption levels were further interrogated for fatty acid composition with lead strains shown in Table 3 here compared to aggregated CHK100 parental control data.

[0150] These mutants were re-screened in a 5-day bioreactor tube-based lipid production assay to quantitatively assess their fatty acid profile, lipid productivity, and triacylglycerol (TAG) distribution, with a particular focus on the level of valuable palmitic-oleic-palmitic (POP)-containing TAGs. POP is valuable as an alternative to traditional POP obtained from palm oil, which requires a great deal of industrial processing to obtain and can be subject to supply issues given climate change.

[0151] As shown in Table 4, one of the mutants, "...Blk07;F02" or "CHK100 mutant isolate 2", continues to exhibit ~4.5% more palmitate than the CHK100 parental strain. In addition, the lipid titer and oil content are modestly improved compared to the parent. An overview of the lineage history of these classically improved high palmitate strains is illustrated in FIG. 2A, with oleate (C18:1) and palmitate (C16:0) levels reported for each strain in the lineage. The method of mutagenesis utilized at each step is presented above each arrow, and the enrichment strategy utilized in conjunction with each mutagenesis strategy is presented

below each arrow. “...Blk07;F02” is otherwise referred to as “CHK mutant isolate 2” in the present disclosure.

Table 4: Tube-based assays evaluating CHK100 mutants showing increased C16:0 levels and improved lipid titer. “Dry cell weight” (“DCW”) was determined by applying 1 mL of culture broth to a pre-weighed glass filter sitting atop a sintered glass funnel under vacuum and washing with 5 mL of water. Filters were then lyophilized to dryness and reweighed to determine the DCW of each sample. “Non lipid biomass” (“NLB”) was determined by subtracting the lipid (g/L) from the DCW. “Per cell production” (“PCP”) was determined by dividing the NLB by the lipid titer.

	Sample											
	CHK100 mutant isolate 2 Rep A (1)	CHK100 mutant isolate 2 Rep A (2)	CHK100 mutant isolate 2 Rep B (1)	CHK100 mutant isolate 2 Rep B (2)	CHK100 mutant isolate 1 Rep A (1)	CHK100 mutant isolate 1 Rep A (2)	CHK100 mutant isolate 1 Rep B (1)	CHK100 mutant isolate 1 Rep B (2)	CHK100 control Rep A (1)	CHK100 control Rep A (2)	CHK100 control Rep B (1)	CHK100 control Rep B (2)
%w/w	70.0	71.3	73.0	75.8	70.9	74.1	69.1	68.9	70.7	70.9	69.2	70.7
Avg. % w/w	72.5				70.8				70.4			
Biomass wt (mg)	16.5	16.2	16.1	15.7	16.1	15.7	16.2	15.9	16.3	16.6	16.1	16.2
Vol. applied to filter (mL)	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
DCW (g/L)	16.2	16.2	16.1	15.7	16.1	15.7	16.2	15.9	16.3	16.6	16.1	16.2
Avg. DCS (g/L)	16.1				16.0				16.3			
Lipid Titer (g/L)	11.6	11.6	11.8	11.9	11.4	11.6	11.2	11.0	11.5	11.8	11.1	11.5
Avg. Lipid Titer (g/L)	11.7				11.3				11.5			
NLB (g/L)	1.9	4.6	4.3	3.8	4.7	4.1	5.0	4.9	4.8	4.8	5.0	4.7
PCP	2.3	2.5	2.7	3.1	2.4	2.9	2.2	2.2	2.4	2.4	2.2	2.4
C14:0	1.42	1.43	1.45	1.40	1.38	1.36	1.41	1.42	1.18	1.17	1.17	1.17
C14:1	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
C15:0	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.04	0.04	0.03
C16:0	36.56	36.63	36.75	36.54	35.40	35.38	35.52	35.65	32.18	32.13	32.15	32.14
C16:1	1.37	1.37	1.37	1.34	1.40	1.38	1.41	1.41	1.32	1.31	1.30	1.30
C18:0	1.96	1.95	1.97	2.01	1.99	2.02	1.98	1.99	2.06	2.08	2.07	2.07
C18:1	51.88	51.86	51.74	52.02	53.21	53.29	53.00	52.87	56.20	56.32	56.32	56.32
C18:2	5.81	5.81	5.80	5.77	5.69	5.64	5.73	5.72	6.00	6.00	6.03	6.01
C18:3 alpha	0.37	0.37	0.37	0.37	0.37	0.36	0.37	0.37	0.40	0.40	0.40	0.40
C20:0	0.14	0.14	.015	.015	0.15	0.16	0.15	0.14	0.15	0.15	0.16	0.15
C20:1	0.03	0.03	0.04	0.04	0.04	0.05	0.04	0.02	0.05	0.05	0.05	0.04

[0152] All strains included in Table 4 were run under standard lipid production conditions in which strains were grown in duplicate in 10 mL of the lipid production medium. Cultures were grown with shaking (200 rpm) for 121 hours at 28 °C at which point 1 mL of biomass was removed, applied to a polycarbonate filter, washed with an equal volume of Milli Q water and placed in a tared glass vial at -80 °C for 30 minutes. Vials containing filters and frozen biomass were lyophilized to dryness overnight, their weights recorded and filters with dried biomass were subjected to direct transesterification followed by GC/FID to quantitate FAMES. Abbreviations included in Table 4 include: DCW (dried cell weight), NLB (non-lipid biomass), and PCP (per-cell production, defined as [lipid titer (g/L)]/NLB).

[0153] FIG. 2B shows the continued progression of the mutant isolate "...Blk07;F02", also referred to herein as strain "CHK mutant isolate 2." Serial passaging of the isolate and screening in the lipid assays, as described above, resulted in a stable strain, CHK224, having a palmitate level of 37% in the isolated crude oil as determined by FAME analysis. A separate stability screen resulting from SSID 1447 produced subclone "C06" (FIG. 2B) and the intermediate strain CHK243u, which, upon additional serial passage and interrogation of isolates in lipid production, resulted in subclone "A10." Subclone A10 is characterized as having a palmitate level of 39% in the isolated oil. Subclone A10, upon another round of passaging and interrogation in lipid production assays (SSID 1515), resulted in a branch in the lineage. One branch in the lineage gave rise to clone "E04", which, upon additional passaging and stability assays in lipid production medium, resulted in the stable strain, CHK263, which is characterized as having 40% palmitate in the isolated crude oil.

[0154] The second branch point, as illustrated in FIG. 2B from the SSID 1515 clone, gave rise to the unstable strain "sub-clone E07", which was subsequently designated CHK256u. A lipid stage culture of CHK256u was subsequently exposed to 2.8 µM 4-NQ at 30 °C for 30 minutes and allowed to recover for 24 hrs in fresh vegetative medium at 30 °C. The mutant population of the exposed CHK256u cells was then exposed during culture in lipid production medium to 100 µM tebutam (herbicide) at 35 °C for 72 h. Surviving cells were recovered and subjected to screening. The first clone from the screening of SSID1642 resulted in a mutant, "clone F07", with a palmitate level of 42% in the isolated crude oil; this clone is designated CHK 268u. CHK 268u was subsequently subjected to serial passage and further interrogation using lipid production assays (SSID 1710). After screening of the SSID 1710

cells, from this screen was cloned “Blk08,F04,” which has 49.9% palmitate when cultivated in lipid production medium at 32 °C in the isolated microalgal oil.

LCMS Analytical Method

[0155] Triglyceride (TAG) profile analysis was conducted on CHK100 and the CHK100 mutant strains using liquid chromatography-mass spectrometry (LCMS). The TAGs were extracted from *P. moriformis* dried biomass into a solution of 3:1 toluene/2-propanol (v/v) by means of a mechanical disruption using ceramic beads and vigorous agitation. The filtered extracts were then injected into an Agilent 1290 Infinity II UHPLC system coupled to a 6470B triple quadrupole mass spectrometer and APCI ionization source according to the parameters described by Co, E.D. et al. (Co, E.D., Koutchekinia, M., Carney, J.R. et al., “Matching the Functionality of Single-Cell Algal Oils with Different Molecular Compositions,” *J. Am. Oil Chem. Soc.* 91: 533–48, 2014) and their Total Ion Chromatograms (TIC) are shown in FIG. 4. FIG. 4 illustrates, through the LCMS data shown, that TAG designations are non-regiospecific.

LCMS Determination of Regioisomers

[0156] Chromatographic separation of regioisomers is not achieved by the LCMS method described, however, a qualitative assessment regarding the regiospecificity of the predominant isomer can be made based on the abundance of diacylglycerol (DAG) ions resulting from fragmentation of the TAG. In LC-APCI-MS analyses of TAGs, the loss of the fatty acid at the *sn-1* and *sn-3* positions are energetically more favorable compared to loss at the *sn-2* position, resulting in a greater abundance of the diacylglycerol ions formed from the former fragments. For example, in the case of 1,2-palmitin-3-olein, (PPO), a near 1:1 ratio of [OP]⁺ DAG ion (*m/z* 577) and [PP]⁺ DAG ion (*m/z* 551) can be observed as the contribution of the 1-palmitin, 3-olein DAG species is a relatively rare event (also the [OP]⁺ DAG ion species (*m/z* 577). In the case of 1,3-palmitin-2-olein (POP) a greater abundance of the [OP]⁺ DAG ion (*m/z* 577) can be observed, as elimination of palmitate at either *sn-1* or *sn-3* results in [OP]⁺ DAG ion (*m/z* 577) while the corresponding [PP]⁺ DAG ion (*m/z* 551) can only arise via the rare loss of olein at the *sn-2* position. For AAB and ABA type triglycerides, a linear calibration curve between [AA]⁺/[AB]⁺ ratios and mixtures of AAB and ABA pure regioisomers with known compositions can be constructed. The method used is described in, e.g., Byrdwell (Byrdwell WC., “The updated bottom-up solution applied to mass spectrometry of soybean oil in a dietary supplement gelcap,” *Anal. Bioanal. Chem.* 407(17): 5143-60, 2015); using this

method, the precise composition of each regioisomer can be calculated based on this linear relationship and simplified by means of a two-point calibration curve using only the endpoints of the calibration curve. This can be expressed by the following equation:

$$\%ABA = \frac{\left(\frac{[AA] +}{[AB] +} \right)_{AAB} - \left(\frac{[AA] +}{[AB] +} \right)_{Obs}}{\left(\frac{[AA] +}{[AB] +} \right)_{AAB} - \left(\frac{[AA] +}{[AB] +} \right)_{ABA}}$$

[0157] Mass spectra and DAG ion ratios for the higher palmitate-producing microalgal strains obtained using the procedure depicted in FIG. 3, along with the CHK100 control and palm oil, are shown in FIGs. 5-8. FIGs. 9-12 depict the mass spectra and DAG ion ratios of native palm oil, and 1,3-palmitin-2-olein and 1,2-palmitin-3-olein, respectively, purchased from Larodan Research Chemicals. In Table 5, the [OP]⁺ and [PP]⁺ DAG ion abundance and calculated the %POP in the TAG component of the unpurified oil for each sample are shown (including their biological replicates designated rep A and rep B, and instrument injection replicates designated r001 and r002). In FIG. 9, the red palm oil analyzed was Nutiva® Nurture Vitality purchased from Berkeley Bowl, Berkeley, California. In FIGs. 10 and 11, the 1,3-palmitin-2-olein and 1,2-palmitin-3-olein analytical standards were both purchased from Larodan Research Chemicals.

Table 5: Calculated [OP]⁺/[PP]⁺ DAG ion abundance and calculated %POP based on their ratio compared to pure POP and PPO regioisomer standards from palm oil versus triglycerides extracted from various classically improved *P. moriformis* isolates using the screening method depicted in FIG. 2A and 2B.

Sample	[OP] ⁺ : [PP] ⁺	Abundance m/z 551 [PP] ⁺	Abundance m/z 577 [OP] ⁺	[PP] ⁺ : [OP] ⁺	%POP of POP/PPO peak in TIC	TIC POP/PPO Area %	% POP in triglyceride oil
CHK100 mutant isolate 1 Rep A (1)	3.98	130173.26	517672.62	0.25	97.77	19.45	18.82
CHK100 mutant isolate 1 Rep A (2)	4.13	121581.72	502444.75	0.24	98.62	19.11	18.85
CHK100 mutant isolate 1 Rep B (1)	3.90	124613.99	486305.69	0.26	95.83	19.37	18.56
CHK100 mutant isolate 1 Rep B (2)	4.30	116280.24	500479.19	0.23	100.50	19.45	19.45
CHK100 mutant isolate 2 Rep A (1)	4.07	198769.05	809166.75	0.25	97.90	25.43	24.90
CHK100 mutant isolate 2	3.89	194987.83	759016.38	0.26	95.71	24.92	23.85

Sample	[OP] ⁺ :[PP] ⁺	Abundance m/z 551 [PP] ⁺	Abundance m/z 577 [OP] ⁺	[PP] ⁺ :[OP] ⁺	%POP of POP/PPO peak in TIC	TIC POP/PPO Area %	% POP in triglyceride oil
Rep A (2)							
CHK100 mutant isolate 2 Rep B (1)	4.11	1920223.77	782385.88	0.24	98.39	24.79	24.39
CHK100 mutant isolate 2 Rep B (2)	4.19	195791.72	820783.44	0.24	99.29	25.39	25.21
CHK100 mutant isolate 3 Rep A (1)	3.95	157399.91	621418.69	0.25	96.41	23.41	22.57
CHK100 mutant isolate 3 Rep A (2)	4.02	154569.34	622093.69	0.25	97.35	23.58	22.96
CHK100 mutant isolate 3 Rep B (1)	3.96	218106.58	864771.00	0.25	96.62	23.19	22.41
CHK100 mutant isolate 3 Rep B (2)	3.90	226937.56	884796.38	0.26	95.79	23.54	22.55
CHK100 Rep A (1)	4.06	145233.64	589042.88	0.25	97.73	19.03	18.60
CHK100 Rep A (2)	3.91	148040.62	579523.50	0.26	95.99	18.79	18.04
CHK100 Rep B (1)	4.13	139082.73	574351.00	0.24	98.58	19.26	18.99
CHK100 Rep B (2)	3.99	140938.38	562031.62	0.25	96.90	19.21	18.62
Palm oil (1)	3.53	128347.91	453653.34	0.28	90.63	25.70	23.29
Palm oil (2)	3.42	139511.14	477637.72	0.29	88.64	25.64	22.78
1,3- palmitin-2- olein analytical std (1)	4.26	303582.72	1292106.50	0.23	99.99	100.00	99.99
1,3-palmitin-2- olein analytical std (2)	4.26	324865.12	1383256.50	0.23	100.01	100.00	100.01
1,2-palmitin-3- olein analytical std (1)	1.35	793718.38	1073788.75	0.74	1.57	100.00	1.57
1,2-palmitin-3- olein analytical std (2)	1.32	850459.00	1126033.50	0.76	-1.57	100.00	-1.57

[0158] The complete triglyceride profile (area %) of each classically improved derivative of CHK100, the parental CHK100 control, and red palm oil are shown in Table 6. Further, as illustrated in Table 6, Sample CHK100 mutant isolate 2 (RepA), has higher POP and lower PPO compared to Red Palm Oil. Tripalmitin levels are also higher in red palm oil.

Triglyceride designations for all other TAGs are non-regiospecific. The acronyms used in Table 6 and elsewhere are as follows:

LLL	linoleic-linoleic-linoleic
LLP	linoleic-linoleic-palmitic
LOL	linoleic-oleic-linoleic
MLP	myristic-linoleic-palmitic
MOL	myristic-oleic-linoleic
MOP	myristic-oleic-palmitic
OOA	oleic-oleic-arachidic
OOB	oleic-oleic-behenic
OOL	oleic-oleic-linoleic
OOLn	oleic-oleic-linolenic
OOM	oleic-oleic-myristic
OOO	oleic-oleic-oleic
OOP	oleic-oleic-palmitic
OOS	oleic-oleic-stearate
PCP	palmitic-capric-palmitic
PLP	palmitic-linoleic-palmitic
PLS	palmitic-linoleic-stearic
PMP	palmitic-myristic-palmitic
POA	palmitic-oleic-arachidic
POB	palmitic-oleic-behenic
POL	palmitic-oleic-linoleic
POP	palmitic-oleic-palmitic
POLn	palmitic-oleic-linolenic
POS	palmitic-oleic-stearic
PPA	palmitic-palmitic-arachidic
PPO	palmitic-palmitic-oleic
PPP	palmitic-palmitic-palmitic
PPS	palmitic-palmitic-stearic
PSS	palmitic-stearic-stearic
SOL	palmitic-oleic-linoleic
SOS	stearic-oleic-stearic
SSG	stearic-stearic-gondolic (C18:1n-11)

SSID Strain/Sample Identifier

TAG triacylglycerols or triglyceride and are non-regiospecific unless indicated

Table 6: Triglyceride profiles of the higher palmitate strains compared to Nutiva® Nurture Vitality Red Palm Oil.

TAG (non-regiospecific unless indicated)	ECN	Sample				
		CHK100 mutant isolate 1 (Rep A)	CHK100 mutant isolate 2 (Rep A)	CHK100 mutant isolate 3 (Rep A)	CHK100 parental strain / control (Rep A)	Palm Oil
		Area Sum %	Area Sum %	Area Sum %	Area Sum %	Area Sum %
LLL	42	-	-	-	-	0.19
LOL	44	0.48	0.32	0.33	0.51	0.92
OOLn	44	0.27	0.15	0.24	0.30	-
LLP	44	0.71	0.71	0.66	0.80	2.92
MOL (+POLn)	44	0.35	0.95	0.45	0.96	-
POLn	44	0.54	-	-	-	-
MLP	44	0.19	0.28		0.33	0.24
OOL	46	2.95	2.10	2.27	3.12	2.58
POPo + PoLS	46	0.93	1.03	1.09	1.28	-
POL	46	5.61	6.26	6.09	6.68	11.12
OOM	46	2.86	3.33	3.26	3.07	-
PLP	46	1.75	2.88	2.74	2.16	7.43
MOP	46	1.91	2.09	1.93	1.49	0.89
PMP	-	-	-	-	-	-
OOO	48	14.71	10.26	11.37	14.70	5.70
SOL	48	0.40	0.23	0.24		1.32
OOP	48	36.53	35.51	36.96	37.62	21.76
PLS	48	0.31	0.30	0.34	0.20	1.41
POP regoisomer	48	18.82	24.90	22.57	18.60	23.29
PPO regoisomer	48	0.63	0.53	0.84	0.43	2.41
PPP	48	0.49	0.81	0.64	0.41	4.08
OOS	50	3.64	2.31	2.35	2.62	3.66
POS	50	3.36	2.57	2.38	2.24	5.59
PPS	50	-	-	-	-	0.91
OOA	52	0.21	0.16	0.14	0.12	0.24
SOS + POA	52	0.30	0.21	0.18	0.20	0.60
PSS + PPA	52	-	-	-	-	-
OOB	54	-	-	-	-	-
POB	54	-	-	-	-	-
SSG	54	-	-	-	-	-
POLg	56	-	-	-	-	-

TAG (non-regiospecific unless indicated)		Sample				
		CHK100 mutant isolate 1 (Rep A)	CHK100 mutant isolate 2 (Rep A)	CHK100 mutant isolate 3 (Rep A)	CHK100 parental strain / control (Rep A)	Palm Oil
		Area Sum %	Area Sum %	Area Sum %	Area Sum %	Area Sum %
AAS	58	-	-	-	-	-
SUM		97.95	97.89	97.53	97.84	97.26
577/551 [OP] ⁺ :[PP] ⁺		3.98	4.07	3.95	4.06	3.53
Calculated %POP in oil		18.82	24.90	22.57	18.60	23.29

[0159] TAG designations are non-regiospecific, unless indicated, as is the case with POP and PPO. FIGs. 12 and 13 are graphical representations of the same data shown in Table 6. “ECN” in Table 6 denotes the equivalent carbon number. Sample CHK100 mutant isolate 2 Rep A specifically shows a higher POP content compared to palm oil. The red palm oil used as the analytical standard in FIGs. 12 and 13 was Nutiva® Nurture Vitality Red Palm Oil. Palm oil has a higher occurrence of PPO compared to the CHK100 control and CHK100 mutant isolates as well as higher tripalmitin.

[0160] Further, FAME analysis was performed on sample CHK100 mutant isolate 2 Rep A (see Table 4) as well as the analytical standard red palm oil (Nutiva® Nurture Vitality Red Palm Oil) and show the palmitate levels of the oil derived from CHK100 mutant isolate 2 of the present disclosure.

Table 7: FAME Analysis of Red Palm Oil.

<i>Fatty Acids</i>	<i>Sample Name</i>	
	Red Palm Oil Nutiva® Nurture Vitality (1)	Red Palm Oil Nutiva® Nurture Vitality (2)
C8:0	0.03	0.03
C10:0	0.03	0.03
C12:0	0.40	0.40
C14:0	1.00	1.00
C15:0	0.00	0.00
C16:0	42.87	42.79
C18:0	4.84	4.85
C18:1 n-9	39.03	39.11
C18:2 n-6	10.39	10.38
C18:3 alpha	0.28	0.28
C20:0	0.36	0.36

Example 2. Fatty acid profile.

[0161] As shown in FIG. 14, the fatty acid profile of the CHK100 mutant isolate 2, a high-palmitic acid strain obtained through classical strain optimization of CHK100 using the methods described herein, was analyzed using GC-FID (gas chromatography flame ionization detection). The results reveal levels of both palmitic and unsaturated fatty acids, calculated by adding the concentrations of oleic acid and linoleic acids (abbreviated as 18:1 and 18:2, respectively in FIG. 14) found in the samples. These levels of palmitic acid and unsaturated fatty acids of the microalgal oils described herein were found to be equivalent to those reported in the literature for palm olein (Sampaio et al., (2011) Steam deacidification of palm oil. *Food and Bioproducts Processing*. 89, 383-90).

Example 3. Triacylglycerol profile.

[0162] FIG. 15 illustrates the triacylglycerol (TAG) profile of an exemplary high-palmitic acid strain of the present disclosure, obtained through classical strain optimization of CHK100, has been analyzed using LC-MS (liquid chromatography – gas spectrometry). FIG. 18 compares the levels of PPP, PPO regioisomer and POP regioisomer of the various strains

depicted in FIG. 2A and 2B, or which the strain depicted in FIG. 15 provides the results for a different strain.

[0163] FIG. 15 shows the levels of 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) in the classically improved algae strain, CHK100 mutant isolate 2, is equivalent to the level of POP level found in palm olein reported in the literature (Sampaio et al., 2011). The analysis also showed that the algae strain contains less than 1% tripalmitin (referred to as PPP in FIG. 15).

Example 4. Diacylglycerol and Monoacylglycerol profiles.

[0164] Both palm oil and palm fractions contain significant levels of partial glycerides (Sampaio et al., 2011 & Sasue et al., (2015) “Synthesis of palm oil high in diacylglycerol through direct esterification.” *The Malaysian J. of Analytical Sciences*. 19(1): 222-29), particularly diacylglycerols (DAG) and monoacylglycerols (MAG). According to the literature, palm oil and its fractions typically have about 2-6% diacylglycerols and 2-5% monoacylglycerols (Sasue et al., 2015). The algae oil described herein, with a fatty acid and triacylglycerol (TAG) profile similar to palm olein, contains only trace amounts (e.g., <2%) of the total MAG and DAG combined.

[0165] As documented in existing literature, the presence of partial glycerides such as DAG and MAG can adversely affect the shelf-life stability of products, given their heightened susceptibility to oxidative degradation compared to triacylglycerols (TAG). The high-palmitic oil resulting from the classical strain optimization of CHK100 also contains less than 6% of linoleic acid (expressed as % of total TAG), while palm oil and palm olein contain typically levels of linoleic acid ranging from 9-12%, which also makes palm oil and palm olein more sensitive to oxidative degradation.

[0166] Additionally, the occurrence of DAG in oil has been identified as a precursor for the formation of process contaminants, including glycidyl esters, that generated when one takes a microalgal oil through a deodorization process (Destailats et al. (2012) Glycidyl esters in refined palm (*Elaeis guineensis*) oil and related fractions. Part I: Formation mechanism. *Food Chemistry*, 131: 1391-98).

[0167] Consequently, owing to the reduced MAG and DAG levels, the high-palmitic oil resulting from the classical strain optimization of CHK100 increases the oxidative stability of the oil and is seen to decrease the residual levels of glycidyl esters, making the resulting crude oil from the strain significantly more useful than what can be obtained from existing strains. A more oxidatively stable oil increases to the number of applications for use of the oil in comestibles for human and animal health and nutrition.

Example 5. Characterization of High Palmitic Acid Triacylglycerol Oils Produced by Classically Improved Microalgal Strains.

[0168] The sterol compositions of high palmitic refined, bleached, and deodorized (RBD) triacylglycerol oils of the present disclosure were determined using gas chromatography combined with mass spectrometry and flame ionization detection. Samples were prepared using the technique described in the German standardized method F-III for sterols in fats and oils from Aitzetmüller, et al. (1998) (“Analysis of sterol content and composition in fats and oils by capillary-gas liquid chromatography using an internal standard. Comments on the German sterol method,” *Fett/Lipid* 100: 429-435), which is incorporated herein by reference in its entirety. After addition of an internal standard, the oils were saponified. The unsaponifiable fractions were isolated using solid-phase extraction. Trimethylsilyl (TMS) derivatives of the unsaponifiable fraction were injected on an Agilent 8890 gas chromatograph equipped with a 5977B single quadrupole EI-MS and FID (flame ionization detection) detector. A capillary flow technology (CFT) splitter was used to split the sample between the MS and FID detectors following separation using a DB-5MS Ultra Inert column (60 m length×0.25 mm inner diameter X 0.25 µm film thickness). Mass spectra data, in conjunction with retention time comparison with analytical standards, if available, were used to identify each of the sterol species. Quantification was based on the GC-FID detected response of the sterol TMS ethers compared to that of the internal standard.

[0169] The resulting sterol profiles are shown in Table 8. Amounts are shown as mg/100 g of the oil and as approximate percentages of total detected sterols. Minor unidentified peaks were not accounted for. Not all sterols were identifiable and are listed as unknown. Ergosterol was the main sterol present in both oils. Various other sterols were present in the RBD oils in significant amounts, including 5.xi.-ergost-7-en-3(3-ol, (3β) and 9,19-cyclolanostan-3-ol, 24-methylene-(3B)-,O-TMS.

Table 8: Sterol Profile of RBD Oil Strains - CHK22 and CHK100

	Sample	
	RBD0061 CHK22 (mg/100g)	RBD0062 CHK100 (mg/100g)
Unknown1	4.78	21.36
Ergosterol, TMS derivative	80.02	35.12

	Sample	
	RBD0061 CHK22 (mg/100g)	RBD0062 CHK100 (mg/100g)
Ergost-8(14)-en-3-ol, (3 β)-, TMS derivative	10.87	5.14
Ergosta-5,8-dien-3-ol, (3 β)-	-	-
5.Xi.-Ergost-7-en-3 β -ol, TMS derivative	46.21	28.73
9,19-cyclolanost-24-en-3-ol, (3B)-, TMS derivative	14.10	18.60
9,19-cyclolanostan-3-ol, 24-methylene-(3B)-,O-TMS	17.71	29.52
Unknown2	5.14	8.47
TOTAL	178.83	146.93

Example 6: Fatty acid and TAG profile of clone “Blk08,F04”.

[0170] As shown in FIG. 16 and Table 9, the fatty acid profile of the “Blk08,F04” high-palmitic acid strain, obtained through classical strain optimization of CHK 268u (FIG. 2B), resulted in a strain producing almost 50% palmitate when grown at 32 °C and 47% when grown at 28 °C for the same length of time.

[0171] Table 9 shows the results of two biological replicates of clone “Blk08,F04” when run in a 5-day tube assay at either 32 or 28 °C. As can be seen, the lipid titer, DCW, and palmitate levels in the obtained oil are higher when grown at 32 °C, but the lipid % and per cell production (PCP;[Lipid g/L]/[DCW-lipid g/L]) are lower.

[0172] FIG. 17 shows the TAG profiles of the lipids isolated from the same samples of clone “BLK08,F04” grown at 32 or 28 °C in which the isolate makes 45 or 40% POP, respectively.

Table 9: Clone “Blk08,F04” run at 32 or 28 °C.

	Sample	
	“Blk08,F04” 28°C	“Blk08,F04” 32°C
Lipid as % DCW	80.1	77.7
DCW g/L	11.5	12.2
PCP	4.02	3.51

	Sample	
	“Blk08,F04” 28°C	“Blk08,F04” 32°C
Lipid g/L	9.2	9.5
C14:0	1.72	2.26
C16:0	46.90	49.87
C16:1	1.47	1.04
C18:0	2.03	3.04
C18:1	40.30	35.81
C18:2	6.13	6.29
C18:3	0.43	0.37
C20:0	0.19	0.28

[0173] Table 10 and FIG. 18 show levels of POP and PPO regioisomers and PPP levels in classically improved strains, which are the subject of this disclosure, as shown in FIG.2, as well as commercially available oils, including palm mid fraction (PMF) and red palm oil. The PMF is valued for its high POP content and is a key component in structuring fats generated through multiple stages of fractionation of palm oil fractions. POP, stearate-oleate-stearate (SOS), and palmitic-oleic-stearic (POS) are the principal structuring fats in cocoa butter. PMF and algal fats described herein can be used as blend stocks along with cocoa butters from different geographies, whose content of the three principle structuring fats can differ, or with other cocoa butter extenders or equivalents used to affect the physical properties of these important structuring fats. As can be seen, clone “BLK08;F04” shows POP levels significantly higher than commercially available palm oil and approximately 70% of what is seen in commercially available palm mid fraction (PMF). Importantly, the levels of PPO, and PPP, which can adversely affect the melt profile and mouthfeel of confections, is markedly reduced in microalgal-derived oils described herein.

[0174] Close inspection of Table 10 and FIG. 18 demonstrate the utility of the microalgal cells of the genus *Prototheca* for generating fats of superior composition relative to traditional incumbent oils, such as red palm oil. Note that the levels of PPP and PPO, two TAG species which can disrupt and interfere with the melt profile and mouthfeel of important structuring fats, such as cocoa butter, are elevated in palm oil and in the palm mid fraction (PMF, designated as “AAK PMF” in FIG. 18 and below in Table 10) compared to the crude

microalgal oils of the enhanced strains described herein. Note the inherently elevated levels of both PPO and PPP in red palm oil and PPO in particular in PMF.

[0175] All microalgal oils of the present disclosure show PPP and PPO levels well below those for palm oil (e.g., clone BLK08;FO4 shows PPP levels only 65% that seen in red palm oil and PPO levels at just 16% of what is seen in red palm oil, while at the same time possessing POP levels over 1.9 times higher) and PPO levels between 4.7- and 12.3-fold lower than what is seen in AAK PMF.

Table 10. Levels of POP and PPO regioisomers and PPP levels in classically improved strains, which are the subject of this disclosure, and commercially available oils, also shown in FIG.18.

Strain/sample	POP regioisomer	PPO regioisomer	PPP
AAK PMF	59.77	4.68	2.10
SSID1731-13 32 °C BLK08;F04	45.15	0.38	2.68
SSID1731-13 28 °C BLK08;F04	40.02	0.46	1.55
CHK268u 32 °C	35.89	0.89	1.46
CHK268u 28 °C	32.51	0.58	0.92
CHK263	28.48	0.89	0.90
CHK256u	26.69	1.00	1.19
CHK224	23.20	0.39	0.52
CHK100	17.79	0.56	0.28

Strain/sample	POP regioisomer	PPO regioisomer	PPP
Red Palm Oil	23.29	2.41	4.08

The sequences of Table 11 represent exemplary insertion loci in *Prototheca* that can be used to assess the absence of recombinantly engineered introductions of heterologous genes into the parent strain. Given that certain governments regulate the use of products and labeling of products produced by or from genetically manufactured organisms (GMOs), the developed *Prototheca* strains obtained in this application can use these exemplary loci to demonstrate that they were not engineered. However, the strains can be recombinantly engineered if so desired.

Table 11: shows SEQ ID NOS: 1-26, 30-31 as referenced herein.

SEQ ID NO.	Source features	Sequence
1	Moltype = DNA length = 4218 Location/Qualifiers 1..4218 moltype = other DNA organism = synthetic construct	cggcgcgctg cttegcgtgc cgggtgcagc aatcagatcc aagtctgacg acttgcgcgc acgcgccgga tcc tteaatt ccaaagtgtc gtccgcgtgc gcttcttcgc cttegtcctc ttgaacatcc agcgacgcaa ggcaggggc g ctggggcgct ggcgtcccg accggcctcg gcgcacgcgg ctgaaattgc cgtatgcgcg aatgtagtgc c gtccgccc acctctcaat taagttttc agcgcgtggt tgggaatgat ctgcgctcat gggggcgaag aagggg ttca gaggtgcttt attgttact gactgggcgt accagcattc gtgcattgact gattatacat acaaaagtac agc tcgett c aatgcctgc gattctact cccgagcgag cactctctc accgtcgggt tgcttccac gaccacgcc g gtaagagggt ctgtggctc gcgcccctcg cgagcgcatc ttccagcca cgtctgtatg atttgcgt cata cgtctg gcccgctgac ccaaaatga cgggatcctg cataatatcg cccgaaatgg gatccaggca ttcgtcag ga ggcgtagcc ccgcgggaga tgcgggtccc gccgcattgg aaaggtgtag aggggggtgaa tccccattt catgaaatgc ttgtcaacg atggtgcga ttcgtgcaa gtgaatatgg ggtcacgcgg tggacgaacg cgg agggggc atgaccgaat ctaggctcgc attctcaga tcaattcatg ccggcggtcc ggggtttgcg cgtcgcg caa ggctacgtct cctagccgc tgcgcaccac gcgtgcgacg cggaggccat ctccggagc aacgaccat g gattgtcttagcgatcgac gaatgagtgc tagtgagtgc tacgtcgcac ccagtcgctc gcaggagaag gcg gcagctg ccgagcttcg gcttaccagt cgtgactcgt atgtgatcag gaatcattgg cattgttagc attataatc ggcttcgcg ctgcgtatgg gcattggcaat gtctcatgca gtcgactta gtaaccaat ttgggtggc caggtc cggg cgaccgggct ccgtgtgcc gggcaccacc tctgccagg agtagcaggg ccgccctctc gtccga cgt tggccactg aataccgtgg ctccagccc tacatgatgg gctgcctagt cgggcgggac gcgcaactgc ccgcgcgac tgggggctgg tctgaatcct tcaggcgggt gttaccgcag aaagaaaggg tgccgatttc aaa gcagacc catgtgccgg gcctgtggc ctgtgtggc gcctatgtag tccccccc taccacaatt gtcgccag tt tgcgactcc ataaactcaa aacagcaget tctgagctgc gctgttcaag aacacctctg gggttgtc accc

SEQ ID NO.	Source features	Sequence
		gcgagg tcgacgccc gcatggctat caagacgaac aggcagcctg tggagaagcc tccgttcacg atcggg acgc tgcgcaaggc catccccgcg cactgtttcg agcgctcggc gcttcgtacg agcatgtacc tggcctttga catcgccgtc atgtccctcg tctacgtcg gtcgacgtac atcgaccctg caccgggtcc tacgtgggtc aagta cggca tcatgtggcc gctctactgg ttctccagg tgtgtttgag ggttttggt gcccgattg aggtcctggt gg cgcgcatg gaggagaagg cgctgtccc gctgaccccc ccgccgaccc tccccacc ttccaggggc ctt cggcac ggggtgtctgg gtgtgcgcgc acgagtgcgg tcaccaggcc ttctctcca gccaggccat caacgac ggc gtgggcctgg tgtccacag cctgtcgtg gtgcctact actcctggaa gcactcgcac cgcgccacc actcaacac ggggtgcctg gacaaggacg aggtgtttgt gccgccgac cgtgcgggtg cgcacgaggg c ctggagtgg gaggagtggc tgcccatccg catgggcaag gtctgtgca ccttgacct gggtggccg ctgt acctca tgtcaacgt cgctccccg cttaccgcg gcttcgcaa caatttgac cgtgtgtcg ccatcttcag caagcgcgag cgcacgagg tggcatctc cgacctcgc ttgtggcgg tgcacagcg gtcacgctg ctg ggccgca ccatgggctg ggctggctg gtaagacct acgtgtgccc ctacatgac gtgaacatgt ggctggt gct catcacgtg ctccagcaca cgcacccggc cgtgccgac tacttcgaga aggactggga ctgctacgc ggcgccatgg ccaccgtga ccgtccatg ggccccccct tcatggacag catctgcac cacatctccg aca cccactg gctgcaccac ctcttcagca ccatccgca ctaccagcc gaggaggcct cgcgcgccat ccggcc catc ctgggcaagt actaccaac cgacagccgc tgggtcggcc gcgccctgtg ggaggactgg cgcgactg cc gctacgtct ccccgacgcg ccgaggacg actccgcgt ctggttcac aagtgcgcg gctcgcgga ggacgcagaa caacgtgccc gccgtgctt tgcacgcgc gactccggcg cttcgtggt ggacacccca taa agaaacc ctcaattctg ttgtggaag acacggtgta cccccacca ccaactgca cctctattat tggattatt gacgcgggag tgggcgtgt accctacaac gtagctctc tagtttcag ctgctccca ccatgtaaa tcatgc tag aatagtgcgt ggttatgtga gaggtatagt gtgtctgagc agacggggcg ggtgcatgt cgtgtgtgtg a tctttggt caaggcgtcg tcgacgtgac gtgcccgatc atgagagcaa taccgcgtc aaagccgacg catag cttt actecgcaat ccaaacgact gtcgtcgta tttttggat atctattta aagagcgagc acagcgccgg gc atgggcct gaaaggcctc gcggcgtgc tctgtgtggg ggccgcgagc gctgtgggca tcgcggcagt gca ccaggcg cagacggagg aacgcatggt gcgtgcgcaa tataagatac atgtattgt gtctgtacta taatgcta ga gcatcaccag gggcttagtc atcgacctg ctttggatc tacagaaatt gcacaagggc gtctccggg at gaggagat gtaccagtc aagctggagc ggcttcgagc caagcaggag cgcggcgcat gacgacctat cca catgtga caaaggggtc tgggtcgac gaatcgaaa accagtcagg agacagcggg gtccatgagt tggcc cgctc gcagcttcag cgcttgagc atcgcgcat tgtccacgat ccaaggctcc cgcagcgct cctgggac c cctgtgcgga gaagcgtgtg ccgagagggc tgtgagggcg gaggtggag gagaggctgc gaatcattgc gtgcctagg cgtctcgga gtagtagcac accacgcgc ggtgcttca ttctattgt tccgttgctg taatgca tgc gagcaatct gcactaaaag ctgcgatcac cacctcggcg ttgtgtct ctagtaccg tcacctgacc ac gcgcacgg aagtatggcg gtcggcgacg gcctgtgca cgtggcagt gcatccacg tggccagga gca ggtctcg cagtgcaac accgcgcct cgtctcgcc ggggtccatc cagtccatac cccgcggcg gccctc gagc cggacggcca gcagccagg ccgggggtcc gcgccgacgc ccggcgcccc cgcgcgcag ctgg ggtcgg ccagggaag cactggacg gccggccgg gcgagcccc cgcgcgctg ggacccacc gca

SEQ ID NO.	Source features	Sequence
		gcaccac cgcctgctcc aggtccacgt cgtcggcgg caggcccgcg tgcggcgct gcgccagcgt ccag ggcagg cggggccggg cgtgcatg
2	Moltype = DNA length = 4217 Location/Qualifiers 1..4217 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	cggcgcgtg cttegcgtgc cgggtgcagc aatcagatcc aagtctgacg acttgcgcgc acgcgccgga tcc ttcaatt ccaaagtgtc gtccgcgtgc gcttcttcgc cttegtctc ttgaacatcc agcgacgcaa gcgcaggggc g ctgggcggct ggcgccccga accggcctcg gcgcacggcg ctgaaattgc cgaatgcgc aatgtagtgc c gtccgccc acctctcaat taagttttc agcgcgtggt tgggaatgat ctgcgctcat gggggcgaag aagggg ttca gaggtgctt attgttact gactgggcgt accagcattc gtgcatgact gattatacat acaaaagtac agc tcgttc aatgccctgc gattcctact cccgagcgag cactcctc accgtcgggt tgcctccac gaccagcc g gtaagagggt ctgtggctc gcgccctcg cgagcgcatc ttccagcca cgtctgatg attttgcgt cata cgtctg gcccgctgac cccaaatga cgggacctg cataatatc cccgaatgg gatccaggca ttctgacg ga ggcgtagcc ccgcgggaga tgcgggtccc gccgcattgg aaaggtgtag aggggggaa tccccatt catgaaatgc ttggtcaacg atggtgcgca ttctgcaaa gtgaatatgg ggtcacgcgg tggacgaacg cgg agggggc atgaccgaat ctaggctcgc attcctcaga tcattcatg ccggcgctcc ggggtttgcg cgtcgcg caa ggctacgtct cctagccgc tgcgcaccac gcgtgcgacg cggaggccat ctccggagc aacgacct g gattgtctta gcgatcgac gaatgagtgc tagtgagtcg tacgtcgac ccagtcgctc gcaggagaag gc ggcagctg ccgagctcg gcttaccagt cgtgactcgt atgtgatcag gaatcattgg cattgtagc attataatt c ggctccgcg ctgcgtatgg gcatggcaat gtctcatgca gtcgatctta gtcaaccaat ttgggtgac caggt ccggg cgaccgggct cgtgtcgcc gggcaccacc tctgccagg agtagcaggg ccgccctc gtcgg acgt tggccactg aataccgtgg ctteagccc tacatgatgg gctgcctagt cgggcgggac gcgcaactg c ccgcgcgac tgggggtggt tctgaatcct tcaggcgggt gttaccgag aaagaaaggg tcccatttc aa agcagacc catgtgccg gccctgtggt ctgtgttggc gcctatgtag tcaccccc tcaccaatt gtcgcca gtt tgcgactcc ataaactcaa aacagcagct tctgagctgc gctgtcaag aacacctc ggggttgc acc cgcgagg tcgacgcca gcatggctat caagacgaac aggcagcctg tggagaagcc tccgtcacg atcgg gacgc tgcgaaggc catccccgc cactgttgc agcgtcggc gcttcgtac agcatgacc tggccttg a catcgcggtc atgtccctg tctacgtcg gtcgacgtac atcgacctg caccggtgcc tacgtgggc aag tacggca tcagtgtgccc gctctactg ttctccagg tgtgttgag ggttttggg gcccgattg aggtcctgt g gcgcgcatg gaggagaagg cgctgtccc gctgaccccc ccgccgaccc tccccacct tccaggcgc ctt cggcacg ggtgtctggg tgtgcgcgca cagtgcggt caccaggcct ttctccag ccaggccatc aacgac ggcg tggcctggt gtccacagc ctgctgctgg tgcctacta ctctggaag cactcgacc gccgccacca ctccaacacg ggtgcctg acaaggacga ggtgttgg ccgccgacc gtcgggtg gcacgagggc c tggagtggg aggagtgggt gccatccgc atgggcaagg tctggtcac cttgacctg gctggccgc tgta cctcat gttcaacg gcctccgcc ctaccgcg ctteggcaac cacttgacc cgtggtgcc catcttcagc aagcgcgagc gcatcgaggt ggtcatctc gacctcgct tgggtgggt gtcagcggg ctcagcgtgc tgg gccgac catgggctgg gctggctgg tcaagacct cgtggtgcc tacatgatg tgaacatgt gctggtg ctc atcagctgc tccagcac gcacccggcc ctgccgact acttcgagaa ggactgggac tggctacgcg gcgcatg gcacgtgac cgtccatgg gccggccct catggacagc atctgcacc acatctcca cac ccagtg ctgcaccacc tctcagcac catccgcac taccagccg aggaggcctc cgccgccatc cggccc

SEQ ID NO.	Source features	Sequence
		atcc tgggcaagta ctaccaatcc gacagccgct gggtcggccg cggcctgtgg gaggactggc gcgactgc cg ctacgtcgtc ccgacgcgc ccgaggacga ctcgcgcgc ttgttccaca agtgagcgcg cctgcgcgag gacgcagaac aacgctgccg ccgtgtcttt tgcacgcgcg actccggcgc ttgcgtgggtg gcacccccat aaa gaaaccc tcaattctgt ttgtggaaga caggtgttac cccacccac ccacctgcac ctctattatt ggtattattg acgcgggagt gggtgttga cctacaacg tagcttctct agtttcagc tggctcccac cattgtaaat tcatgctaga atagtgcgtg gttatgtgag aggtatagtg tgtctgagca gacggggcgg gatgca tgtc gtggtgtgta tctttgctc aaggcgtcgt cgactgacg tccccgatca tgagagcaat accgcgtca aagccgacgc atagccttta ctcgcaatc caaacgactg tcgctcgtat tttttgata tctattttaa agagcgag ca cagcggcggg catgggcctg aaaggcctc cggccgtgct cgtggtgggg gccgcgagcg cgtggggc at cgccgcagtg caccaggcgc agacggagga acgcatggtg cgtgcgcaat ataagatata tttattgtg t ctgtactat aatgctagag catcaccagg ggcttagtca tcgcacctgc ttgtgctt acagaaattg cacaagg gcg tctccggga ttaggatag taccagtc aactggagcg gcttcgagcc aagcaggagc gcggcgca tg acgacctat cacatgtgac aaaggggtct gggtcgtacg aatcgacaaa ccagtcagga gacagcgggg tccatgagtt ggccgcctc cagcttcagc gctttgagca tcgcggcatt gtccacgac caaggctccc gcagc gcctc cttgggacct ctgtcgggag aagcgtgtgc cgagagggtc gtgaggggcg aggtcggagg agaggc tgcg aatcattgcg ctgcctaggc gtctcccgag tagtagcaca ccacgcgcg gtgcttcat ttcattgtg cc gttctgt aatgcatgcg agcaatctg cactaaaagc tgcgatcacc acctcggcgt ttttggctc tagtaccg ct cacctgacca cgcgcacgga agtatggcg tcggcgacgg cctgctgcac gtggcagtc agatccact ggccaggag caggtctgc agctgcaaca ccgcgccctc gtgctgcgg gggtccatcc agtccatcc cc gcggcgcc cctcgagcc ggacggccag cagccagggc cgggggtccg cggcgacgcc cggcgggccc ggcgcgcagc tgggtcggc caggaagagc cactggacgg ccggccggcg cagccccgc gccggcgtg g gaccacccg cagcaccacc gcctgtcca ggtccacgct gctcggcggc agggccgcgt gcggcgctg cgccagcgtc caggcgagc gggtccgggc gtgcatg
3	Moltype = DNA length = 4227 Location/Qualifiers 1..4227 moltype = other DNA organism = synthetic construct	cggctcgtg cttgcgtgc cgggtgcagc gatcagatcc aagtctgacg acttgtgtg atgtactgtg tctttg agt ccagagcgcc ggccgcacgc gcttctccc cttctcttc ctctgaaca tccagcagtg caagtgcagg g cgctggcg gctggcgcc cgaaccggcc tcggcgacg cggtgaaat tgccaatgtc ggcaatgtag tgc cgctcgg cccatccctc gatcaagttt ttacgcgct ggttggggat gatctgcgt catgggaaga taaaaggg gt tctgaggtg ttgattgta cttaccgga agtaactata ttcatacatg actgatccca caaaaaaca aagct cactt caaagaaccg cgcatgtcta cteccagca atcacttcgc tcaccgtcgg gttgctccc acgacaacgc cggtagaggg gtgggtggc tcgcgacct cgggggcaca tcttccagc catgtctgta taatctcacg ctcata cgtc tggcccgctg acccaaaaat gacgggatcc tgcattgat cgcccgagat ggggtccagg cattcctctg gaggcgtcag cctcgggga gatgcgggc ccaccgcat ggaaaggcac aaaggggtg aatccccat tt catgaaat ttttggtag cgatggtgc cactcgtgc caatgaatat ggggtcacgc ggtggacgaa cgcgg agggg gcctggccga atctaggctt gatttctca gatcacttgc tgcggcggt ccgggggttg cgcgtcgcg aacgtccgt ctcctagcc gctcgcacc gcgcgtgcga cgcgaaggtc atttccaga acaacgacca tgg cttctct tagcgatgc tgaatgact gctagtgtg cgtacgctc acccagtcgc tcgaggaga acgcggc aac tggcagctt cggcttgcca gtcgtgact gtatgtgac aggaatcatt ggcattgga gcattataat tcgg

SEQ ID NO.	Source features	Sequence
		cttccg cgctgtttat ggcatggca atgtctcatg cagtcgacct tagtaacca attctgggtg gccagctccg ggcgaccggg ctccgtgtcg ccgggcacca cctctgcca tgagtaacag ggccgcctc tectcccgac gtt ggccac tgaataccgt gtctggggc cctacatgat gggtgccta gtcgggcggg acgcgcaact gccgcg gcaa tctgggacgt ggtctgaatc ctccaggcgg gttccccga gaaagaaagg gtgccgattt caaagcaga g ccatgtgccg ggccctgtgg cctgtgttg cgcctatgta gtcaccccc ctcaccaat tctgccagt ttgc gcaatc cataaactca aaactgcagc ttctgagctg cgctgttcaa gaacacctct ggggtttgct caccgcga g gtcgacgcc agcatggcta tcaagacgaa caggcagcct gtggagaagc ctccgttcac gatcgggacg ctgcgaagg ccatccccgc gactgttc gagcgctcgg cgcttcgtag cagcatgtac ctggccttg acatc gcggt catgtccctg ctctacgtc cgctgacgta catcgacct gcgccgtgc ctacgtgggt caagtatggc gtcatgtgc cgctactg gttcttcag gtgtgtgtga ggggtgtgtg tgcctgatac gaggtctgg tggcgc gcat gggggagaag gcgcctgtcc cgctgacccc ccggtacc ctccggcacc ctccagggc gccttcgg ca cgggtgtctg ggtgtgcgcg cagagtgcg gccaccaggc ctttctcc agccaggcca tcaacgacgg cgtgggcctg gtgtccaca gcctgtgtg ggtgccctac tactctgga agcactcgca ccgcgccac cact ccaaca cggggtgcct ggacaaggac gaggtgttg tggcccgca ccgcgagtg gcgcacagg gcctg gagtg ggaggagtgg ctgccatcc gcctgggcaa ggtgctgtc acctgaccc tgggtggcc gctgtac ctc atgtcaacg tcgcctcgc gccgtaccgc cgcttcgcca accacttga ccctgtgtg ccatcttca gc aagcgca gcgcacgag gtgtcatct ccgacctggc gctggtggcg gtgtcagcg ggtcagcgt gctg ggcgc accatgggt gggcctggt ggtcaagacc tacgtgtgc cctacgtat cgtgaacatg tggctgt gc tcatcagct gtccagcacc gcacccgg cgctccgca ctactegag aaggactgg actggtgcg cggcgccatg gccaccgtgg accgtccat gggcccgccc tcatggaca acatctgca ccatctcc gac accacg tctgcacca cctctcagc accatccgc actaccacgc cgaggaggcc tcgccgcca tcaggc ccat cctgggcaag tactaccgt ccgacagcc cggttcggc gcgcacctg gggaggactg gcgcgact gc cgctacgtc tcccgacgc gcccgaggac gactccgcg tctgttcca caagtgagtg agtgagtcg tactcagcg cgctgcgcg gggtgcgga acgcccgcg cgctgtgtt ttgcacgcg cgaactcgt gctt cgcggg tggcaccac attgaaaaa acctcaatc tttgttga agacacggtg taccaccaac caccacc tg cactctatt attgtatta ttgacgggg agcgggcgtt gtacttaca acgtagcgc tctgtttc agctgg ctcc caccattgta aattctgtc aaaatagtc gtggttatgt gagaggtatg gtgtaacagg gcgtcagta tgt tggttt cgtgctgac tcgggcacaa ggcgtcgtc agtgacgtg ccctgatga gagcaatacc gcgtca aag ccgacgcatg gcctttact cgactccaa acgactgtc cctgtattt tcggatatct atttttaag agcga gcaca gcgccgggca tggcctgaa aggcctcgc gccgtgctg tgggtggggc cgcgagcgc tgggg catcg cggcagtga ccaggcgag acggaggaac gcatggtgag tgcgcatcag aagatgcatg tctgttg tc tgtactataa tctagagca tcaccaggg ctagtcac gcacctgct tggctattc agaaattgca caag ggcgtc ctccgggatg aggatagta ccagtcgag ctggagcggc ttcgagcaa cgaggagcgc ggcgc atgac gacctacca catcgacaa aggggtctg gtcgtacgac aaaccagta ggaggcggc gggtec atga gctggcccgc tcgagcttc agcgttcga gcatcgccg gttgtccgc atcaagcct ccgcagcg c ctcttgggg ccctgtgcg gagaggcggg tggcagagg gctgtgagg cgaggctgg aggacaggc t gcgaatcgtc gcgtgccta ggcgtctgc gaggagaagc acaccacgc ccgcgctt catttcattg tgc

SEQ ID NO.	Source features	Sequence
		ccgttgcc tgtaatgcat gcgagcgatc ctgcactaaa gctgcgatcg ccacctcagc gttgttcgct tctagcacc g ctcacctgac cacgcgcacg gaagagtggc ggtcggcgac ggcttgcgct acgtggcagt gcaggtccac gtggccgagg agcaggtctc gcagctgcag caccgcgccc tctgtctcgc cgggggccat ccagtccatg cc ccgcgccg ggccctcgag ccggacggcc agcagccagg gccgggggtc cgcgccgacg ccgcggcgcc ccgccgcga gctggggctc gccaggaaga gccactgggc ggccggcccg ggcgagcccc gcgccggcg c gggaccacc cgcagcacca ccgcctggc caggtcgacg tctctcgcg gcaggcccg gtgcggcg c tgcgccagcg tccagggcag gcggggcgcg gcgtgca
4	Moltype = DNA length = 4227 Location/Qualifiers 1..4227 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	cggtctgctg ctttgcgtgc cgggtgcagc gatcagatcc aagtctgacg acttgcgtg atgtactgtg tctttg agt ccagagcgcc ggccgcacgc gcttctccc cttctcttc ctctgaaca tccagcgatg caagtgcagg g cgctggcg gctggcgcc cgaaccggcc tggcgcacg cggctgaaat tgccaatgct ggcaatgtag tgc cgctcgg cccatccctc gatcaagttt ttacgcgctg gttggggat gatctgcgct catgggaaga taaaagg gt tctgaggtgg ttgattgta cttaccgga agtactcata ttacatag actgatccca caaaaaaca aagct cactt caaagaaccg cgcattgcta ctecccgca atcacttgc tcaccgtcg gttgctccc acgacaacgc cggtgagagg gtcggtggc tcgcgacct cgcgggcaca tcttccag catgtctgta taatctcag ctcata cgtc tggcccgctg acccaaaaat gacgggacgc tgcattgat cgcccgagat ggggtccagg cattctctg gaggcgctag cctgcggga gatgccggc ccaccgcat ggaaaggcac aaagggggtg aatccccat tt catgaaat tgttggtcag cgtggtgcg cactcgtcg caatgaatat ggggtcacgc ggtggacgaa cgcgg agggg gcctggccga atctaggtt gattctca gatcatttc tgcggcggt cgggggttg cgcgtcgcg aacgctccgt ctccctagcc gctcgcacc gcggtgcga cgcgaaggc atttccaga acaacgacca tgg ctgtct tagcgatcg tgaatgact gctagttagt cgtacgctg accagtcg tcgcaggaga acgcggc aac tgcgagctt cgcttgcca gtcgtgact gtatgtgac aggaatcatt ggcatgta gcatataat tgg cttcg cgtgtttat ggcatgga atgtctcat cgtcgacct tagtcaacca attctgggtg gccagctccg ggcgaccggg ctccgtgctg ccgggcacca cctctgcca tgagtaacag ggccgcctc tctcccgac gtt ggccac tgaataccgt gtttggggc cctacatgat gggtgccta gtcggcggg acgcgcaact gccgc gcaa tctgggacgt ggtctgaat ctccaggcg gttccccga gaaagaaagg gtgccgattt caaagcaga g ccatgtccg ggccctgtg cctgtgttg cgcctatgta gtcaccccc ctcaccaat tgcgccagt ttgc gcaat cataaactca aaactgcagc ttctgagctg cgtgttcaa gaacacctt ggggtttgct caccgcga g gtcgacgcc agcatggta tcaagacgaa caggcagcct gtggagaagc ctcggtcac gatgggacg ctgcgaagg ccatccccc gactgttc gagegctcg cgttcctag cagcatgtac ttggccttg acatc gcggt catgtccctg ctctacgctg cgtcgacgta catcgacct gcgcgggtg ctacgtgggt caagtatgc gtcatgtggc cgtctactg gttctccag gtgtgtgta ggggtgtgtg tggcgatc gaggtcttg tggcg gcat gggggagaag gcgcctgtcc cgtgacccc ccgggtacc ctcggcgac cttccagggc gcctcgg ca cgggtgtctg ggtgtcgcg cagagtgcg gccaccaggc ctttctcc agccaggcca tcaacagcg cgtgggctg gtttcaca gcctgctgt ggtgcctac tactcctgga agactcgca cgcgcgac cact ccaaca cgggggtcct ggacaaggac gaggtgttg tggcccgca ccgcgagtg gcgcacagg gcctg gagtg ggaggagtgg ctgccatcc gcattggcaa ggtgctgtg acctgaccc tgggtggcc gctgtac ctc atgttcaac tgcctcgcg gccgtaccg cgttcgcca accatttga cccgtggtg ccatcttca gc

SEQ ID NO.	Source features	Sequence
		aagcgcga gcgcacgcag gtggtcatct ccgacctggc gctggtggcg gtgctcagcg ggctcagcgt gctg ggccgc accatgggct gggcctggct ggtaagacc tacgtgtgct cctacctgat cgtgaacatg ttgctcgt gc tcatacgt gctccagcac acgcacccgg cgtgccgca ctacttcgag aaggactggg actggtgcg cggcgccatg gccaccgtgg accgtccat ggcccgcgc tcatggaca acatctgca ccacatcc gac accacg tctgcacca cctctcagc accatcccgc actaccacgc cgaggaggcc tcgcgcgca tcaggc ccat cctgggcaag tactaccagt ccgacagccg ctgggtcggc cgcgccctgt gggaggactg gcgcgact gc cgtacgtc tcccgagcg gcccgaggac gactccgcgc tctggttcca caagtgagt agtgagtgc tactcagcg cgcctgcgc gggtatcgga acgcccgcgc cgcctgtct ttgcacgc cgactccgtc gctt cgcggg tggcaccccc attgaaaaa acctcaattc tttgtgga agacacgggt taccaccaac caccacc tg caccttatt attgtatta ttgacgggg agcgggctt gtactetaca acgtagcgc tctggtttc agctgg ctcc caccattgta aattctgct aaaatagtc gtggttatgt gagaggatg gtgaacagg gcgtcagta tgt tggttt cgtgctgac tgggcacaa ggcgtcgtc agtgacgtg cccgtgatga gagcaatacc gcgtca aag ccgacgcag gctttact cgcactcaa acgactgtc ctcgtattt tcgatatct attttaag agcga gcaca gcggcgggca tggcctgaa aggcctcgc gccgtgctc tgggtggggc cgcgagcgc tgggg catcg cggcagtgca ccagcgagc acggaggaa gcattgtgag tgcgcatc aagatgcag tctgttg tc tctactata tctagagca tcaccaggg cttagtcac gcacctgct tggcattac agaaattgca caag ggcgtc ctccgggatg aggatagta ccagctcaag ctggagcggc tcgagccaa cgaggagcgc ggccg atgac gacctacca catcgacaa aggggtctg gtcgtacgac aaaccagta ggaggcggc gggtc atga gctggcccgc tcgagctc agcgttcga gcacgcggc gttgtcgc atcaaacct cccgcagcg c ctcctgggg cccctgtgc gagaggcggg tgccgagagg gctgtaggg cggaggctgg aggacaggg t gcgaatgct gcgctgcta ggctgtgc gaggagaagc acaccacgc ccgccgctt catttcattg tgc ccgttc tgaatgcat gcgagcgc ctgcactaaa gctgcgac ccacctcagc gttgtcgc tctagacc g ctcacctgac cagcgcagc gaagagtggc ggtcggcgc gccctgctc acgtggcagt gcaggtccac gtggccgagg agcaggtc gcagtcgag caccgcgcc tctgtcgc cgggtccat ccagtcagc cc ccgcccgc ggcctcgc cgggacggc agcagccagg gccgggggc cgcgccgag cccggcgcc cccccgcga gctggggtc gccaggaaga gccactggc gccgggggc ggcgagccc gcgcccgc c gggaccacc cgcagacca ccgctggtc caggtcgac tgcctggc gcaggcccgc gtcggcgc c tgcgacgcg tcagggcag cggggggc gcgtgca
5	Moltype = DNA length = 4067 Location/Qualifiers 1..4067 moltype = other DNA organism = synthetic construct	gcaatggcg tcggtacagg gtctgcgtc gtgctgggt cctctccta cgtgcacaa gggagcgc cgg ccagtc agcgcgtcca caacctccc tcgtcacaca cacacctgc gaaccaggcc gccatttgc tcttga gca tgccttgc catgtccggg ttcccatca tctgttgc gttctgggc tcagcttct gctccagcac acca tcctgt cgtcgaaga gaaggagaca tgtgtacatt attggtgta gggcgtgaa tcggccattt tttaatgat cacgtcatg ccaatagac cggcacataa cgacgttcaa acccccgc aagcccgga caaccacc cc tccacac cccacacaa agaaccgc accgcttacc ttcccacga ggtaggcct tctgtgcga aaaccg gcct cggtgatgaa tgcagccc ttctgacga gcgtgccc ggccaacac ctttttgc gcgtctc a ggcttggg gccctctgg gcttgggtc cgccatgac tgcgcgac agagaacgt tctgtgtaa aggag cgccc ggctgcgcaa tatatatata ggcatgcaa cacagccaa cctactcgg gagccgtcc caccacc

SEQ ID NO.	Source features	Sequence
		cc aagtcgctg ccttgacggc atactgctgc agaagcttca tgagaatgat gccgaacaag aggggcacga ggaccaatc cgggacatcc ttgtcgataa tgatctcgtg agtccccatc gtccgccga cgtccgggg agcc cgccga tgcctcaagac gagagggccc tggaccagga ggggctggcc cgggcgggca ctggcgtcga aggt gcgccc gtcgttcgcc tgcagtcta tggcacaata caagtctct gacgggggtg gtttctccc gtgcgggca g gcaacagagg tattcaccct ggctatgggg agatcgccga tcgagctggg ataagagata cggccccgc c aaggatcgc tcactctggt ctgagccga cagtattct ggcaagcaat gacaactgt caggaccgga ccgtg ccata tattctcac ctgcccgc aaaacctaac aatttgggag tcactgtgcc actgagttc actggtagct g aatggagtc gctgctccac taaacgaatt gtcagcaccg ccagccggcc gaggaccga gtcatacgga ggg tagtagc gcgccatggc accgaccagc ctgcttgcca gtactggcg ctctccgct tctctgtgt cctctgcgc g ctccagcgcg tgcgtttt cgggtggatc tgcggctcg ggcgaccgc agcgccgct gccatgcag c gccgtgct tccgaacagt ggcggtcagg gccgcaccg cggtagccgt ccgtccgaa cccgccaaag ag tttggga gcagcttgag ctctgcaaga tggcggagga caagcgatc ttctggagg agcaccggtg cgtgg aggtc cggggtgac cggccgtgc attcaacgta atcaatcga tgatgatcag aggacacgaa gcttgggtg g cgggtggcag aaacactgt cattgaagg gcataggat gcgttcttc acctctcatt tctcttct gaatcc ctcc ctgctcactc ttctctcc tccttccgt tcacgagca ttcggggcaa cgaggtgggc cctcgcagc g gctgacgat cacggcggtg gccaacatcc tgcaggaggc ggcgggcaac cagcggtgg ccatgtgggg c cggagctcg gagggttgc cgacggacce ggagctgcag gagcggggtc tcactttgt gatgacgcg atgc agatcc agatgtaccg ctaccgcgc tggggcgacc tgatgcaggt ggagacctgg ttccagacgg cgggca agct gggcgcgag cgcgagtgg tctgcgca caagctgacc ggcgagcgcg tggcgcgcc cactc gagc tgggtcatga tcaacatcc cagcgccgg ccgtgccga tggcgagct cgtccgcgc aagtcggc ct tcttcgcgc cagccgcgc cgctggcg tcccgccgc ggtcacgct gccaagctgc ccaacatgc gagccggcg ccgtgcgc ggcaccgca gtcgcgcgc cgcaccgaca tggacatgaa cgggcacgtg aacaacgtg cctacctgc ctggtgcctg gagccgtgc ccgagcagct ctccagcag taccacctct acca gatgga gatgacttc aaggccaggt gccacgggg cgacgtatc tcctccagg ccgagcagat cccgcc ccag gagcgctca cgcacaacgg cgcggccgc aacctctct gttctgcca tagcattctg cgcgccga ga ccgagctcgt ccgcgcgca accacatgtt cggccccc cagcgcgcc gccccaagc cgcacaag gc gagccactga ggacagggtg gttggctgga tggggaaacg ctggtcgcgg gattcgatcc tctgtttat a tcctcctg gaagcacacc cagactctg aagaagaaaa cgtgcacaca cacaaccaa ccggccgaat attgtctc ttatccggg tccaagagag actgcgatgc cccctcaat cagatcctc ctccctgcc ctcaat ctt ccctgctgc ctgcgccgc ggtgcgccgt ctgccgcc agtcagtcac tctgcacag gccctgtg c gcagtctc ctgtaccct taccgtct tcatctgc gagggccct attgaatgta ttggtgct gtgtggcca a cggggtgct ggcgcgcc cgtcgggca gtgctggcg acttggcg aagccgattg ttctctgta ag ccacgcgc ttgctctt gggaagagaa ggggggggta ctgaatgat gaggaggaga aggagggta ttgg tattat ctgagttggg gaggcaggga gattggaaa atgtaagtgg cagcagggc aaggagaaat gtgagca tgt gcatggtgat gtcgttggtc gaggacgac ctgcacgcgt gtatctgatg tagaatacgg caatcacct agt ctacatc tatacctct ccgtataacg cctttccaa atgcctccc gtttctctc tattctgat ccacatgat acc ctggcac tattcaagg gctggacatt tcaagaaggt ttgcgtatc gaagaaggat tgggttgag aggtggccg

SEQ ID NO.	Source features	Sequence
		a tgaagtggg gtcaagetgc gtggagccct ctgcacggat ttctatggta atctgcgtcc acgtcatcag tagc cgtacg cctctgggc gtcgccgtgc tctcagccg ccgcttcacc accagccagg agccccatac gcacgc gctc ccagacggcg cgcgcgacgg ggacgatcat gagcagcgcg atctggggcg tgcgtgcga gctgagc cgc acaaccacgt tgcctggcc gcgcgtggtg tactggcagt tggtcgagac catggtgtcg gacttgaagt aggagccgtc aaagcgcggg aggcgcgtct cggccgagac gttgccgccg atgaccaggc tctggcgtt ct ggcgcggg tcgggccct gccactgcgc ggagccgcc aggagcaggc gcgactcgc acggccgcgc a gcttgacgt cggccgccag cgcgtgcc ggctcgtagg ccgacccat cttgacgaag aggcggccga tgc ctgcgcg gacctgacc gcgtcggaca ggcgaggcg ctcgtccagc ttgacccga gcgcgcgtt gccgg gcttg aacggcagcg agaactctc gccagcttg gagccgagca ggcccagcgc cagcttctgc cgttgc cgc gcgcgcgag gcgcgagtc acgcgcagcg ttagag
6	Moltype = DNA length = 4069 Location/Qualifiers 1..4069 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	gcaatggcg tcggtacagg gtctgcgtcc gtgtgggt cctctccta cgtgcacaa gggagcgcgc cgg ccagtc agcgcgtcca caactcccc tcgtcacaca cacactgcg gaaccaggcc gccatttgc tegtga gca tgccttgc catgtccggg ttcccatca tctgttgag gttcttggc tccagcttct gctccagcac acca tctgt cgtcgaaga gaaggagaca tgtgtacatt attggtgta gggcgctgaa tcggccattt ttataaatga tcacgtcat gccaatagac gcggcacata acgacgtca aacccccgc aaagcgcgg acaacccat cc ctccacac cccccacaca aagaaccgc caccgttac cttgccacg aggtaggcct tctgtgcgc aaaacc ggcc tcggtgatga atgcagccc gttctgacg agcgcgtccc gggccaacac gctctttgc tgcgtctct c aggcttggg ggcctcttg ggttgggtg ccgcatgat ctgcgcgcat cagagaaagc ttgctgtaa aaagg agcgc ccggtgcgc aatatata taggcatgcc aacacagccc aacctactc gggagccct cccaccac cc ccaagtcgc tgcctgacg gcatactgt gcagaagct catgagaatg atgccgaaca agaggggcac gaggaccaa tccggacat cttgtcgt aatgatctg tgagtccca tcgtccccc gacgctccg ggag cccgc gatctcaag acgagaggc cctcgaccag gaggggctg cccgggcggg cactggcgc gaag gtgcgc ccgtcgtc cctgcagtc tatgccaca aacaagtct ctgacgggt gcgttgcgc ccgtgcggg c aggcaacaga ggtattacc ctgtgatgg ggagatcgc gatcagctg ggataagaga tacgttccc c gcaaggatc gctcactcgt gtctgagccg gacagtcat ctggcaagca atgacaact gtcaggacc gaccg tgcca tatattctc acctagcgc gcaaaccta acaatttgg agtactgtg cactaggtt cactggtag c tgaatggag tcgtgtcc actaaacgaa ttgtcagcac cgcagccgg ccgaggacc gactcatagc gag ggtagta gcgcgccatg gcaccgacca gcctgttgc cagtactggc gtctctccg cttctgtg gtctctgc g cgtccagcg cgtgcgtt tccggtgat catgcgtcc gtggcgacc gcagcggcg cgtccatgc a gcgcgctg cttccgaaca gtggcggtca gggccgcacc cgcggtagc gtccgtccg aacccgccca ag agtttgg gagcagctt agccctgcaa gatggcggag gacaagcgca tctcttga ggagaccgg tgcgt ggagg tccgggctg accggccgtc gcattcaacg taatcaatc catgatgac agaggacacg aagtctt gt ggcgtggcc agaactcgc tccattgcaa gggcatagg atcggtctc tcacctca ttctctt ctga atccct cctgtctac tcttctct cctcttccc gttacgcag cattcgggc aacgaggtg gccctcga gcggctgac atcacggcg tggccaacat cctgcaggag gcggcgggca accacgcgt ggccatgtg ggccggagct cggagggtt cgcgacggac ccggagctgc aggagggcg tctcatctt gtgatgcgc gc atgcagat ccagatgtac cgtacccgc gctggggcg cctgatgcag gtggagacct ggtccagac ggcg

SEQ ID NO.	Source features	Sequence
		ggcaag ctgggcgcgc agcgcgagtg ggtgctgcgc gacaagctga ccggcgaggc gctgggcgcgc gcc acctcga gctgggtcat gatcaacatc cgcacgcgcc ggccgtgccg catgccggag ctctccgcgc tcaagt cggc cttcttcgcg cgcgagccgc cgcgcctggc gctgccgcc cgggtcacgc gtgccaagct gcccaaca tc gcgacgccgg cgcgcctgcg cgggcaccgc caggtcgcgc gccgcaccga catggacatg aacgggca cg tgaacaacgt ggctacatg gcctggtgcc tggaggccgt gcccgagcac gtcttcagcg actaccact c taccagatg gagatcgact tcaaggccga gtgccacgc ggcgacgca tctctccca ggccgagcag atcc cgcccc aggaggcgct cagcacaac ggccgcggcc gcaaccctc ctgcttcgc catagcatc tgcgcg ccga gaccgagtc gtccgcgcgc gaaccacatg gtggccccc atcgacgcgc ccgccccaa gccgcc caag gcgagccact gaggacaggg tggttggtg gatggggaaa cgtggtgcg gggttcgat cctgtgct t atactctcc tgaagcaca cccacgactc tgaagaaga aacgtgcaca cacacaacc aaccggccga a tatttgctt cttatcccc ggtccaagag agactcgat gccccctca atcagcatc tctctctgc cgttcaa tc ttcctgctt gcctgcgcc cgggtgcgc gctgcccc cagtcagtc actcctgcac agggccctg tgc gcagtgc tctgtacc ttaccgctc ctccattct gcaggcccc ctattgaatg tattegtgc ctgtgtgcc a agcgggctg ctgggcgcgc cgcgctcggg cagtgcctgg cgactttggc ggaagccgat tgtttctg taag ccacgc gcttctgct ttgggaagag aagggggggg tactgaatgg atgaggagga gaaggagggg tattgg tatt atctgagtg gggaggcagg gagagttgga aaatgtaagt ggacgcagcg gcaaggagaa tggtagca t gtgcatggtg atgtcgttg tcgaggacga tctgcacgc gtgtatctga tgaagaatc ggcaatcacc ctagt ctaca tctatacct ctccgtataa cgcccttcc aaatgccctc cgtttctct cttattctg atccatga tgacc ctgc actattcaa gggtggaca tttaagaag gttgctgtat ctgaagaagg attggttg agaggtggcc gatgaaagt gggtcaagct gcgtggagcc ctctgcacgg attctatgg taactcgt ccacgtcacc agtag ccgta cgcctctgc gcgtgccgt gctcctcagc cccgcttca ccaccagcca ggagccccat acgcacgc gc tcccagacgg cgcgcgcgc ggggacgac atgagcagcg cgatctgggg cgtgtcgtgc gagctgagc c gcacaaccac gttgccctgg ccgcgcgtgg tgtactggca gttgtcag accatgtgt cgacttgaa gt aggagccg tcaaagcgc ggaggcgcgt ctgcccgag acgttgccgc cgatgaccag gctctgcgc ttct ggcgcg ggtcgcggcc ctgccactgc cgcgagccgc ccaggagcag gcgcgactgc tcacggccgc gca gcttgac gtcggccgc agcgcctgc ccggctcgt ggccgacccc atctgacga agaggcggcc gatcg ctgc cggacctga ccgcgtcgga caggcgcagg cgctctcca gcttgacccc gagcgcgcgc ttgccgg gct tgaacggcag cgagaactc tcgccagct tggagccgag caggccagc gccagctct gccgcttc c gcgcgcgc aggcgcgagt cgacgcgag cgtgtagag
7	Moltype = DNA length = 4083 Location/Qualifiers 1..4083 moltype = other DNA organism = synthetic construct	gcaatggcgc tctgtgcagg gtctgtgccc ggtctccctc tctacgatg cacaaggcaa cgcctttgct gactca gcgc gtccacaacc tccctcgtc acacacacct gcggaaccag gccgccatt tctgcttga gcatgccttg catcatgtcc gggttgcca tcatatggt gaggttctg ggtccagtt tctgtccag cagccatcc tctggtc ga agaggaggag acatgtaagc gttgttgga tgaagggtgc taaattggcc attatttta atgacacgc tct gccaat agacgcggca cattgcatg ttaaaactc cgtcaaagcc ggggacaacc acatccctcc acacccc aca cacaagaac ccgccaccgc tcacctgcc cagaggtag gccttctgt gcgcaaaccc ggctcgggtg atgaatgct gcccgctct gacgagcgt gccggggcca acacgctct ttgctgcgc tctcaggt tgggg gcctc cttgggcttg gggggcccca tgatctgcgc gcaccagaga gaagtcattg gtgaaaagga gcgcccc

SEQ ID NO.	Source features	Sequence
		get gcgctatata tatatatata tgctcaggcc aataaagtc aacctcactg gggagccccg tcccaccacc cc caagtcgc gtacctgac ggcatactgc tgcagaagct tcatgaggat gatgccgaac aagaggggca caagg acceca atcccgaca tccttgctga taatgatctc gtgagtcctc atctccgcc cgacgtgctg gggagccccg c cgatgctcaa gacgagaggg cctcgcacca ggaggggctg gcccgggcgg gcactggcgt cgaaggtgc g cccgtcgttc gctgcagtc ccatgacaca aaacaagtct tctgacgggg tgcgtttgct cccgtgcggg cag gcaacag aggtattcaa cctggctcat ggcagatcgg ccatcgagct gggataagag atacggtccc gcgcg aggat cgtctatct ggtctgagcc ggacagtcct tctggcaagc aatgacaact tggcaggacc ggaccgtgc c atatattct cacctagcgc ggcaaacct aacaatttgg aagtcactgt gccactgagc tgcactgta gctga atgga gtcgctgct cactaatcga attgtcaga ccgccagccg gccgaggacc cgagtcatac cgagggtga gt agcgcgcat ggcaccgacc agcctgctt gccgtactgg cgtctcttcc gttctctgt gctctctac gcg ctccggc gctgctgct ttccgtgga tcatcgctc cgtggcgac cgacgggcc gctgcccag cagcgc cgct gttccgaac agtggctgc agggccgac ccgcagtagc cgtccgtccg gaaccgccc aagatttt g ggagcagctt gagccctgca agatggcgga ggacaagcgc atctctctgt aggagcaccg gtgcgcggag gtccggggct gaccggccgt cgcattcaac gtaataatc gcatgatgat cacaggacgc gacgtcttgg tggc ggtggc cagggacact gccattgca caggcatagg aatgcgttcc ttctatttc tcatgtttt gagccctcc cttctactc ttctctcc tctccctc taccgcaga ttcgtggcaa cgaggtgggc cctcgcagc ggctga cgat cagggcggtg gccaacatcc tgcaggaggc ggcgggcaac cagcggttgg ccatgtgggg tggag ctcg gaggggttgc cgacggacc ggagctgcag gagggggcc tcatcttct gatgacgcgc atgcagatc c aaatgtaccg ctaccgcgc tggggcgacc tcatgcaggt ggagacctgg ttccagacgg cgggcaagct a ggcgcgag cgcgagtggt tctgcgcga caagctgacc ggcgaggcgc tggcgcggc cactccagc t gggtcatga tcaacatcc cagcgccgg ccgtgccga tggccagct cgtccgctc aagtggtct tctc gcgcg cgagccgccc cgcctggcgc tggcgcccac ggtcacgcgc gccaagtgc ccaacatgc gacgc cggcg ccgctgcgc ggcaccgcca ggtcgcgcgc cgcaccgaca tggacatgaa cggccactgt aacaa cgttg cctactggc ctggtgctg gaggcgtgc ccgagcagct ctccagcag taccacctc accagatgg a gatcacttc aagccgaggt gccacgcggg cgactcctc tctccagg ccgagcagat cccgccccag gagggcgtca cgcacaagg cgccggcgc aacctctct gttctgcca cagcattctg cgcgcgaga cc gagctctt ccgcgcgc accacctgt cgccccct cgacgcgccc gccgccaag cgccaaggc ga gccactga gcagaggaga ggttggttgg tagttgcggg atggctggtc gccgtcgat cctgtctgt ctattg tctc ctctgcaca agccccca cgactccgaa gaagaagaag aaacgcgc caccacaaac ccaaccg gcc gaattttgc ttcttatcc cgggtccaag agagacggcg atgccccct caatcagct cctctcct gc cgctcaa tcttctctg ttcatgctc ccgcgagagg ctgtctgcgc gcccgctcag tcatccccg tgcaga cgcc tctgtctgg tctctgta tctttaccg ctctttcat tctgcaggc cccctgttga atgtattgt tgcct gtgtg gccaagcgc ctgctggggc cgccggcgtc gggcggttct cggcgactct ggcggaagcc ggttgtt ctt ctgtaagcca cgcgttctt gtttttgaa aagagggggg ttactgaat ggaggaggag caggataatt g gtagtatct gattgttgg ggaggcagga agagttggga aatggaagt gcacgatggg caaggagaat ggtg agcatg cgcagggtga tctgttggc cgagaacgat cctgcactcc cgttaagtag ataccgcaat caggttga tc tacatcteta acttctcat ttaagccct ttccatgc cctcgcttt gcatctaca cctttgatc cacatgat

SEQ ID NO.	Source features	Sequence
		tg atgacctgg cactatctta agggctggac attcaagaa ggtttgcgta tctgaagaag ggctggtttg aag gggtcgc cgatgagagt gggaacaagc tgcgaggatg cctctgcgcg aattcctaaa ggagtaatct gcgtcc acgt catcagtagc cgtaggcctc tgcgacgtcg cegtgtctct cggccgcgcg ttccgcgcg cccagcagc c ccagacgcac gcgtcccgac acggcgcgcg cgacggggac gatcatgagc agcgcgatct ggggcgtgt c gtgcgagctg agccggacga ccacgttccc ctggccgcgc gtggtgtact ggagtttg cgagaccatg g tgtctgact tgaagtagga gccgtcaaag cgcggggaggc gcgtctcggc cgagacgttg ccgcccagca cca ggctctg gcgttctga cgcgggtcgc ggccctgcca ctgcgcggag cctcccagga gcaggcgcgc ctctg cgcgg ctgcgcagct tgaggtcggc cgccagcgcc gtgcccggct cgtaggccga ccccatcttg acgaaga ggc ggccgaccgc cgcgcggagc ttgaccgcgt cggacaggcg caggcgctgc tccagcttga ccccgag cgc cgcgttgccg ggcttgaacg gcagcgagaa ctctcgcgc agcttgaggc cgagcaggcc cagcgcca gc ttctgccgct tgcgcgcgc gcgcaggcgc gagtgcagc gcagcgtgta gag
8	Moltype = DNA length = 4083 Location/Qualifiers 1..4083 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	gcaatggcgc tctgtcagg gtctgtgtcc ggtctccctc tctacgatg cacaaggcaa cgccttctgt gactea gcgc gtccacaacc tccctcgtc acacacacct gcggaaccag gccgccatt tctgtcttga gcattccttg catcatgtcc gggttgccc tcatatcgtt gaggttcttg ggtccagtt tctgtccag cagccatcc tctcgttc ga agaggaggag acatgtaagc gttgttggga tgagggttgc taaattggcc attatttita atgatcacgc tctg gccaat agacgcggca cattgcgatg tcaaaactcc cgtcaaagcc gcggacaacc acatccctcc acacccc aca cacaagaac ccgccaccgc tcacttgcgc cagcaggtag gccttctgt gcgcaaaccc ggccctcgtg atgaatgcgt gcccttctc gacgagcgt gcccgggcca acacgtctt ttctgcgc tcctcaggct tgggg gcctc cttggcctg gggggcccca tgatctgcgc gcaccagaga gaagtcatg gtgaaaagga gcgccc gct gcgtatata tatatatata tctcaggcc aataaagtc aacctactg gggagccccg tcccaccacc cc caagtgc gtacctgac ggcatactgc tgcagaagct tcatgaggat gatgccgaac aagaggggca caagg acceca atcccgaca tcttgcga taatgatctc gtgagtcccc atgtccgcg cgacgtcgc gggagcccc c cgtatcctaa gacgagaggc cctcgcaca ggaggggctg gcccgggcgc gcactggcgt cgaaggtgc g cccgtcgttc gctgcagtc cgtatgacaca aaacaagtct tctgacgggg tgcgttctc cccgtcggg cag gcaacag aggtattcaa cctggctatg gcgagatcgg cgtatcagct gggataagag atacggtccc gcgcg aggat cgtctactc ggttgagcc ggacagtc tctggcaagc aatgacaact tgccaggacc ggaccgtgc c atatattct cacctagcgc ggcaaaacct aacaatttgg aagtcactgt gccactgagc tgcactgta gctga atgga gtcgtctc cactaatga attgtcagca ccgccagccg gccgaggacc cgagtcatag cgagggtga gt agcgcgcat ggcaccgacc agcctgcttg cccgtactgg cgtctctcc gcttctctg gctctctac gcg ctccggc gcgtgcgctt ttccgtgga tcatgcgtc cgtggcgac cgcagcgcc cgtgccatg cagcgc cgct gttccgaac agtggtgtc agggccgcac ccgcagtagc cgtccgtccg gaaccgcgc aagatttt g ggagcagctt gagccctgca agatggcgga ggacaagcgc atcttcttg agggagaccg gtgcgcggag gtccggggct gaccggcgt cgcattcaac gtaatcaate gcatgatgat cacaggacgc gacgttcttg tgg cgggtgc cagggacat gccattgca caggcatagg aatgcgttcc ttctatttc taagtcttct gagccctcc ctcttactc ttctctcc tctccctc tcacgcagca tctgtggcaa cgaggtgggc cctcgcagc ggctga cgat cagggcggtg gccacaatcc tgcaggaggc ggccgggcaac cagcggttg ccatgtgggg tcggag ctcg gagggtttc cgacggacc ggagctgcag gagggggcc tcatcttct gatgacgcgc atgcagatc

SEQ ID NO.	Source features	Sequence
		c aaatgtaccg ctacccgcgc tggggcgacc tgatgcaggt ggagacctgg ttccagacgg cgggcaagct a ggcgcgcag cgcgagtggg tgctgcgcga caagctgacc ggcgaggcgc tgggcgcggc cacctccagc t gggtcatga tcaacatccg cagcgcgcgg cegtgcgcga tggcagct cgtccgcgtc aagtggcct tcttc gcgcg cgagccgcgc cgctggcgc tggcggccac ggtcacgcgc gccaagctgc ccaacatcgc gacgc cggcg ccgtgcgcgc ggaccgccca ggtcgcgcgc cgcaccgaca tggacatgaa cggccacgtg aacaa cgttg cctacctggc ctggtgcctg gaggccgtgc ccgagcacgt ctcagcgac taccacctct accagatgg a gatcacttc aaggccgagt gccacgcggg cgacgtcctc tctcccagg ccgagcagat ccgccccag gaggcgctca cgcacaacgg cgccggcgc aaccctctct gttctgcca cagcattctg cgcgcgcgaga cc gagctcgt ccgcgcgcgc accacctggt cggccccctg cgacgcgcgc gccgccaagc cgcccaaggc ga gccactga gcagaggaga ggggtggctgg tagttgcggg atggctggtc gcccgctcat cctgctgtg ctattg tctc ctctgcaca agcccaccca cgactccgaa gaagaagaag aaaacgcgca cacacacaac ccaaccg gcc gaatatftgc tctctatcc cgggtccaag agagacggcg atgccccct caatcagct cctctcct gc cgtccaa tcttccctgc tgcagcgc ccgcgagagg ctgtctgcgc gcccgctcag tcactcccc tgcaga cgcc tcgtgctcgg tgctctgta tctttaccg ctcctttcat tctgcaggc cccctgttga atgtattctg tgcct gtgtg gccaaagcgc ctgctggggc cgccgccgtc gggcggtgct cggcgactct ggcggaagcc ggttgtt ctt ctgtaagcca cgcgcttct gcttttgaa aagagggggg tttactgaat ggaggaggag caggataatt g gtagtatct gattgttgg ggaggcagga agagttggga aatggaagt gcacgatggg caaggagaat ggtg agcatg cgcagggtga tgcgttggc cgagaacgat cctgcactcc cgtaagtaga atacggcaat cactgtga tc tacatcteta acttctcat taaagccct ttccacatgc cctcgcgtt gcatcctaca cctttgac cactgat tg atgacctgg cactatctta agggctggac attcaagaa ggtttgcgta tctgaagaag ggttggttg aag gggtcgc cgatgagagt gggaacaag tgcgaggatg cctctgcgcg aattcctaaa ggagtaact cgtcc acgt catcagtagc cgtaggctc tgcgacgtc cgtgctct cggccgcgc ttcgccgcg cccagcagc c ccagacgcac gcgtcccg acggcgcgc cgacggggac gatcatgagc agcgcgatct ggggcgtgt c gtgcgagctg agccggacga ccacttgcc ctggccgcgc gtggtgtact ggcagttgc cgagaccatg g tgtctgact tgaagtagga gccgtcaaa cgcgggaggg gcgtctcggc cgagacgtt cgcggacga cca ggctctg gcgcttctga cgcgggtcgc ggccctgcca ctgcgcggag cctcccagga gcaggcgcgc ctcgt cgcgg ctgcgcagct tgaggtcggc cgccagcgc gtgcccggct cgtaggccga ccccatctg acgaaga ggc ggccgaccgc cgcgcggagc ttgaccgct cggacaggcg caggcgtcgc tccagcttga ccccgag cgc cgcgttgccg ggcttgaac gcagcgagaa ctctcgcgc agcttgagc cgagcaggcc cagcgcca gc ttctgccgt tgcgcgcgc gcgcaggcgc gactgcgc gcagcgtgta gag
9	Moltype = DNA length = 4768 Location/Qualifiers 1..4768 moltype = other DNA organism = synthetic construct	gcctgcctcg ccaactgcga gttctgcgc cgcggggggc gcgagggccg cgacggcgc cgggggttca ccgctcgtg cgaggagctc tgggagcggc ggcgaacaga gttgtgcac ctcccggagc aggagtttcg cg gcatcatg ctcgtcatca tgcctgccg ctctgacaag tcgaaaaaag gcgtcgccaa gtacgcaagc ggggg tgggg tgggttggg gtggtggcga tgcgatgaaa atggtggtgt tgggtgtcc gattggcgtg tcggaatttt gtgaactgag cggcactctg cggtctgcc ggccatgcc tagtccagtt ttcgatgtgc gctatcatgt accac ggccg atgcgaatct caccctttct tctcgcgc aggtatggtc acctgagcga cggcctgttg cacttggtgc tcacaggcg gtgcagccgc ctccagttacc tcaagttct gtcaggatg tcgacatcg gccttgaggc gggc

SEQ ID NO.	Source features	Sequence
		gggcag cacgggtcct acatccaggt ctgcctgcc cacgcggtgc acatcgaagc ggtacgagcg tggatc ggga ggcacttgac agtcgcgtcg ccgaacaagg gacgggtacc ctgaggagct tgttgaccg ctgattcac t cctgaccte cttttctct ttccaaaaca ggttgccag gagagccact ggaacgtcga tgggagctc attc aaagcc gcacgataaa cgcgagctc caccgaggag tgatgatgt gttgcccga ggcgtggagg gctgac gtgc gcgaaactag ctgggggccc ccattccgc ctttaaacg cctgctcct ggtccccgac cgtgctggc ggccctgac aattgcca ttccgttta ttcttgaca atgagcgct catccagtg ccaccgcca tcccaa att gttcctcag aacctcag atacctcct tcaaaactgct ccaagagt gacgagtact tgtaattga tgcga ccgct gttcacaatg tagtcggcat gcttggtga gcgcgttga aactgctca gcccgctggc ctgccatgg acgtcccctg atcgcttat gccaccgca tggaccgca cagccagtt ttggattca atagcaacga atacg catcg ataattctg gacattgcat gcctaccgc gtgaattgct gtcccaaac taagcattat catgctcgg tc acgcgac ctggatccgg ggtactgga ccgctggtg agagcgctgc cgtcgattg gtggcaagca agatt gcgca ggtggcgaa gggagagacc aaaaccggag gctggaagcg ggcacaacat cgtattattg cgtatag tag agcagtggca gtcgattc gagtccgca acggtctcg caagctcgt acgctcag taggagatag gggaccactg ccctgccag aatggtcgc accctgtccc tggcgccc cgctgcaac acgagtgcg tat ccagca ggggtgtc gcctcaacc gccccatgt tggcgccg gctcgatcag gtgcgtgag gggggt ttgg tggggccg ccttgggcc cgtgctggc gtgcggacgt gggggccgg gtagtgatc agcaggg gtt gcatgcaat gcctataccg gcgattgaat agcgatgaac gggataggt tgcgctact ccattgccc g cgacccgt ttctgccc cagcgtggt cggcgagct gcgaagcgg accccacca gcgattgtg atca ccggaa tggcggtgc ctccgtgtt ggcaacgat tcgagacct ttacgacaag ctctggaag gaacgagc gg cgtggacctg attccaggt gcgtaggcc ttggatgaat gcgtctaggt tgcgagtgta ctggccagga ag cagcaggc ttggggttg gtgttctg ttctgtaat ttgagtttc attataagt tctgacggt ctgtttcga aa acatgcaa caactccaca cacacact ccttcaact gactctgag gttgacatc tccgagtcc cgaccaa gtt tgcggcgcag atcaccggt tctccgtga gactgcgtg gacaagaaga acgcgcggcg gtacgacga c gcgtctgt acgcatggt ggcctccaag aaggccctgc gccaggcagg cctggagaag gacaagtgc ccgagggcta cggggcgct gacaagacgc gcacgggct gctggtggc tgggcatgg cgggctgac ggtttccag gacggcgtca aggcgtggt ggagaaggc tacaagaaga tgagccctt cttatccc tac gccatca ccaacatggg ctccgcgtg gtgggcatc accagggtt catggcccc aactactcg tctccac agc ctgcgcgac tccaactac catttgtga cgggccaac cacatccga agggcgacgc ggacgtcat g gtcgtggcg gcaccgagc ctccatcgt cccgtgggccc tggcggtt tgtggcctgc cgcgcgtgt c cacgcgca cgacgagccc aagcgcgca gccggccgtg ggacgaggc cgcgacggt ttgtatggg c gaggggcg gccgtgctg tcatggagt gctggagca gcgcagaagc gtggcgac catcctggg ga gtacctg gcgggccat gacctgcgac gcgaccaca tgacggacc gcaccccgag ggcctggcg tga gcacctg catcgctg gcgtcgagg acgcggcgt ctcgccgac gaggtcaact acgtcaacg gcacg ccacc tccacctgg tggcgacaa ggccgaggt gcgcgggtca agtcggtt tggcagatg aagggtat ca agatgaacgc caccaagagt atgacgggc actgcctggg cggcgccggc ggcattggag ccgtcgca c gctcatggc atccgaccg gctgggtgca cccaccatc aaccacgaca accccatcg cgaggtcat g gcctggacg tcgtgcca cgcaaggcc cagcagaca tcaactgc catctcaac tcttcggct ttggc

SEQ ID NO.	Source features	Sequence
		gggca caactccgtc gtcgcctttg cgccttccg cgagtaggtg aagcgagcgt gctttgctga ggaggagag gc ggggtgcgag cgctctggcc gtgcgcgcga tactctccc gcatgagcag actctcctg ccacgcccgga atctactgt caacgagcaa ctgtgtgtt gtccgtggc caatctatt atttccga ctgtggcctg actctgttg gctgtgcaag caccacagga actctttgt tgagcggggg gttatcgtag agagggttg cgaatgggc accg atcgga aggccttga aacgtggcgg tcattggccg acatgattc cgtggattca ctgatgcggc attgcttacc attcatattg aaaagacagc ctgcaacct acaatttgat ggagcaaata catatatata cggggccgac gtgtca gatg gccgttgcgc tcttgagctc cacggcccga cacagctgct ctagatcttc ctccagcggc aggcagccg c gctggccgtc agcgtacgtg cgcaccgcca gggtcgagct cccggcctct tccggcccga tgaccgcat g acggggact ttggcctct cggcgttgcg aacgagcttg ccatggagg cgcggccggg cgcgagctcc acg cgcaggc cgcgcgcgcg catggcggct gccaccgtct ccatgtacgg ccgcaccgcg tccgtcgtgg gcag cagtcg cacctgctcc gggggccagcc agagcgggaa ggcgcccgcg tagtgctcga tcaggatgcc gaaga agcgc tccagcgagc cagcagggc gcggtgcacc atgatgggc gctggcggct gccgtcggcc gccgc gtact cgagcccga cgcctcggc aggttgaagt ccagctgcac cgtggagcac tgccactgc gcccgag cgc gtctcggatc ttgacgtcga tcttcgggcc gtaaaaggcg ccgcccccga cgtccaccgt gaaggcccag ccctgtcgtat gaattgtgt tgggaggggg ggttcgagg ttggatgaca ccacaataat atagtcagcg acac gcgtgc tgcacgataa ttaccagtac agactcggcg atgataacac caccctttc atcaaccgaa atgagccgtt ccctgttgac aaacacacac gcacacacac cttgtcgtt agcgtttag cgagcgcgcg ctcggccctg gcc cagtcct cgtcggacc gaccgactc tccggccgcg tggagaggtt gacctgaac cggcacgcgt cgaag ccaaa ggcgcggaag acgcgtccg tgaggtccag caggcccgg atctcggccg cgtatcgtc cggcag gcag aagatgtggg cgtcgtcctg cgtgaagccg gcacgcgga agaggccgtg catgtgccc gaggcgt cgt agcggtagag agtcccagc tccgccagc gcagcggcag gtcgggtag ctgacgggccc tggccgcg ta gacgtga
10	Moltype = DNA length = 4768 Location/Qualifiers 1..4768 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	gcctgcctcg ccaactgcga gtctgcgcg cgcggcgggc gcgagggccg cgacggcgcg ccggggttca ccgctcgtt cgaggagctc tgggagcggc ggcgaacaga gttgtgcac ctcccggagc aggagtttc cg gcatcatg ctcgtcatca tgcctgccg ctctgacaag tcgaaaaaag gcgtcgccaa gtacgcaagc ggggg tgggg tgggttggg gtggtggcga tgcgatgaaa atggtggtg ttgtggtcc gattggcgtg tcggaattt gtgaactgag cggcactctg cggctctgcc ggccatgcc tagtccagt ttgatgtgc gctatcatg accac gccgc atgcgaatct caccctttt tctcgcgcc aggtatggtc acctgagcga cggcctgttg cacttggtgc tcacaggcg gtgcagccg ctcagtagc tcaagttct gtcaggatg tcgacatcg gccttgaggc gggc gggcag cagggtctt acatccaggt cttgcctgcc cagcgggtgc acatcgaagc ggtacgagcg tggatc ggga ggcacttgac agtcgcgtcg ccgaacaagg gacgggtacc ctgaggagct tgttgaccg ctgattcac t ccctgacctc cctttctct ttccaaaaca ggttgccag gagagccact ggaacgtcga tggggagctc attc aaagcc gcacgataa cgcgcagctc caccgaggag tgatcatgt gtttccccga ggcgtggagg gctgac gtgc gcgaaactag ctgggggccc ccattccgc cctttaaag cctgcctct ggtccccgac cgtgctggc ggccctgate aattcgtcca ttccgttta ttcttgaca atgagcgct catccagtg ccaccgccc tcccga att gttcctctca aaactctcag atacctcct tcaactgct ccaagagtg cagagtagt tgtaattta tgcga ccgct gttcacaatg tagtcggcat gcttgtgtga gcgcgttcga aactcgtca gcccgtcggc ctgccatgg

SEQ ID NO.	Source features	Sequence
		acgtcccctg atcgctttat gccaccgcga tggaccgcga cagccaggt ttcggattca atagcaacga atacg catcg ataattctgt gacattgcat gcctcaccgc gtgaattgct gtcccaaag taagcattat catggctcgg tc acgcgacg ctggatccgg ggatcctgga ccgctggtgg agagcgctgc cgtcggattg gtggcaagca agatt gcgca ggttggcgaa gggagagacc aaaaccggag gctggaagcg ggcacaacat cgtattattg cgtatag tag agcagtggca gtcgcatttc gaggtccgca acggatctcg caagctcgt acgtcacag taggagatag gggaccactg cccctgccag aatggtcgcg accctgtccc tcgccggccc cgctcgcaac acgcagtgcg tat ccagcaa gcgggttgc gccttaacc gcccctatgt tggcgtccgg gctcgatcag gtgcgctgag gggggt ttgg tgggcccgcg cctctgggcc cgtgtcgcc gtgcggacgt ggggcccggg gtagtgatc agcaggg gtt gcatgcaat gcctataccg gcgattgaat agcgatgaac gggatacgtg tgcgctact ccattcccat g cgaccccggt ttctgtccgc cagcgtggt cgcccgagct gcgaagcggg accccacca gcgcattgtg atca ccggaa tggcggtggc ctccgtgtt ggcaacgatg tcgagacct ttacgacaag ctctggaag gaacgagc gg cgtggacctg atttccaggt gcgtaggctc ttggatgaat gcgttaggt tgcgaggtga ctggccaggga ag cagcaggc ttgggtttg gtgttctgat ttctgtaat ttgaggttc attataagt tctgacgtt cttgttcca aa acatgcaa caactccaca cacacacct ccttcaact gagtctgcag gtttgacat tccgagtcc cgaccaa gtt tgcggcgagc atcaccggtc tctccgtgga ggactgcgtg gacaagaaga acgcgcggcg gtacgacga c gcgctgtcgt acgcgatggt ggcctccaag aaggccctgc gccaggcagg cctggagaag gacaagtgcc ccgagggcta cggggcgctg gacaagacgc gcacgggctg gctggtcggc tcgggcatgg gcgggctgac ggtttccag gacggcgta aggcgtggt ggagaaggcg tacaagaaga tgagccctt cttcatccc tac gccatca ccaacatggg ctccgcgtg gtgggcatcg accagggtt catggcccc aactactcg tctccac agc ctgcgcgacg tccaactacg catttgtga cgcggccaac cacatccgca agggcgacgc ggacgtcat g gtcgtcggcg gcaccgaggc ctccatcgtg cccgtgggcc tggcggtt tgtgctcgc cgcgcgtgt c cagcgcaa cgacgagccc aagcgcgcga gccggccgtg ggacgaggcg cgcgacggt ttgtatggg c gaggcgcg gccgtgctg tcatggagtc gctggagcac gcgcagaagc gtggcgcgac catcctgggc ga gtacctgg gcggcgccat gacctgcgac gcgcaccaca tgacggacce gcaccccgag ggcctggcg tga gcacctg catcgcctg gcgtcagg acgcggcgct ctcgccgac gaggtaact acgtcaacg gcacg ccacc tccacctgg tggcgacaa ggccgaggtg cgcgcgtca agtcggtt ttggacatg aagggtat ca agatgaacgc caccaagagt atgatcgggc actgcctggg cgcgcggcg ggcattggagg ccgtcgcga c gctcatggc atccgacccg gctgggtgca cccaccatc aaccacgaca accccatcg cgaggtcgat g gcctggacg tcgtgccaa cgccaaggcc cagcacgaca tcaacgtcg catctcaac tcctcggct ttggc gggca caactccgtc gtcgctttg cgccctccg cgagtaggtg aagcgagcgt gctttgctga ggaggagg gc ggggtgcgag cgtctggcc gtgcgcgcga tacttcccc gcatgagcag actcctcgt ccacgccga atctactgt caacgagcaa ctgtgtttt tgcctggc caatctatt atttccga ctgtggcgt actctgtttg gctgtgcaag caccacgga actctttt tgagcggggg gttatcgtag agaggggtt cgaaatgggc accg atcgga aggccttgca aacgtggcg tcatggcg acatgattc cgtggattca ctgatcggc attgcttacc attcatattg aaaagacgc cttgcaacct acaatttgat ggagcaaata catatatata cggggccgac gtgtca gatg gccgttgcg tcttgagtc cacggcccga cacagctgt ctagatctc ctcagcggc aggcagccg c gctggccgtc agcgtacgt gcacccgca ggtcgagct cccggcctt tccgggccga tgaccgcat g

SEQ ID NO.	Source features	Sequence
		acggggact ttggccctct cggcgttgcg aacgagcttg cccatggagg cgccgccggg cgcgagctcc acg cgcaggc cgcgcgcgcg catggcggt gccaccgtct ccatgtacgg ccgcaccgcg tccgtcgtgg gcag cagtcg cacctgctcc gggggccagc agagcgggaa ggcgcccgcg tagtgctcga tcaggatgcc gaaga agcgc tcacgcgagc cgagcagggc gcggtgcacc atgatgggcc gctggcgggt gccgtcggcc gccgc gtact cgagcccgaa ccgtcgggc aggttgaagt ccagctgcac cgtggagcac tgccactgc gcccgag cgc gtctcggatc ttgacgtcga tctcgggcc gtaaaaggcg ccgcccgca cgtccaccgt gaagcccgag ccctgtcgat gaatgattgt tgggaggggg ggttcgagg ttggatgaca ccacaataat atagtcagcg acac gcgtgc tgcacgataa ttaccagtac agactcggcg atgataacac caccctttc atcaaccgaa atgagccgtt ccctgttgac aaacacacac gcacacacac cttgtcgttg agcgttcag cgagcgccgc ctcggccctg gcc cagtcct cgtcggaccc gaccgactcg tccggccgcg tggagagggt gacctgaac cggcacgcgt cgaag ccaaa ggcgcggaag acgcgtccg tgaggtccag caccggcccg atctcggccg cgatctctc cggcag gcag aagatgtggg cgtcgtcctg cgtgaagccg cgcacgcgga agaggccgtg catggtgccg gacgcgt cgt agcggtagac agtcccagc tcccccagc gcagcggcag gtcgcgtag ctgacgggcc tggccgcg ta gacgctga
11	Moltype = DNA length = 4792 Location/Qualifiers 1..4782 moltype = other DNA organism = synthetic construct	gcctgcctcg ccaactgcga gttctcgcg cgcggcgggc gcgagggcgc gcagggcgcg gcgggggtta ccgcctcggc cgaggagctc tgggagcggc ggcgagcaga gttgtgcac ctcccggagc aggagttgc ag gcatcatg ctcgtcatca tgcctgccg ctctgacaag tcgaaaaagg gcgtcgccaa gtacgcaagc acagg gggtg ggggtggggtg ggggtgggtg cggtgcgatg aaaatggtgg tgttgggtg ttcacattgg cgtgtcggaa g tttgtgaag tgaagtcac tctcggctc tcccggcgca tgcctagtc agtttctgat gtgcgtatc gtgtac cagc ccgcgtgcga atcccacct ttctctccg cccaggtat ggtcacctga gcgacggcct gctgcacctg gtgctcatca ggcgggtcag ccgcctcag tacctcaagt ttctgctcag gatgtgcac atcggcctcg aggc gggcgg gcagcacggg tctacatcc aggtcctgcc cggccacgcg gtgcacatcg aagcggtacg agcgt ggatc gggaggcact tgacagtcgc gtcgcccgc aagcacggg cgccctgagg agctgtgtg gccgctg att cacttccctg gcctccctt tctcttcca aaccagggtg gccaggagag ccaactggaac gtcgatggg ag ctattca aagccgcag atcaacgcgc agctccaccg cggagtgtc gacgtgttg ccagaggcgt ggagg gtga cgagcgcaa actagctgg ggcceccatc ccgcccgtga aacgcctgcc tctggggccc tggccggt gc tggcggccct gaccgattcg tccattcgt ttattctt gacaatgagc gcctcatccc agtgcctcat ccag tgcca ccgcccaccc ccaaatgtt cctctcaaac ctctcagata ccccttcaa actgtccca agagtgcacg agtatttga atgttatag accgctgttc acaatgtagt cggcatgctt gtatgagcct gttcgaaact cgtcagc cc gtcggcctgc ccatggacct cccctgatc cttcacgccc accgcatgga ctgcgacacg ccagttttg g attcaatat caaaaaatac gttcggtaaa ttctgtaacg ttgcatgcct caccgcgtga attgtgtcc caaacgtaa g catcatcgtg gctcggtcac gcgatcctgg atccggggat cctagaccgc tggtagagag cgctgccgtc g gattggtgg caagtaagat tgcgcagggt ggcgaaggga gagacaaaa ccggaggctg gaagcgggca ca acatcgtg ttattgctga tagtagagca gtggcagtcg catttcagg tccgaacgg atctcgcaag ctcgtac gc tcacagtagg agaaagggga cactgcccc tgcagaatg gtcgcgacc tctcctcgc cggccccgcc tgcaacacgc agtgcgtatc cggcaagcgg gctgtcgcct tcaaccgcc ccatgttggc gtccgggctc gat cagggtc gctgaggggg gtttgggtg cccgcgcctc tggggccgtg tggccgtgc ggacgtgggg ccctg

SEQ ID NO.	Source features	Sequence
		ggcag tggatcagca gggtttgcgt gcaaatgcct ataccggcga ttgaatagcg atgaacggga tacggttgc g ctactccat gcccatgcga ccccgttct gtcgccagc cgtggcgcc cgggctgcga agcgggaccc c accagcgc attgtgatca ccggaatggg cgtggcctcc gtgtttggca acgatgtcga gacctttac aacaag cttc tggaaagAAC gagcggcgtg gacctgattt ccaggtgcgt aggtccttgg atcatgctg ctaggttggg aggcggctgg cgaggaagca gcagccttgg ggtttggtgt tccgatttct ggcaattga ggttcattg tgagat tcta tgcgggtctg ttccgaaac atgcaacaac tccacacaca cacactctc tccaccaact ctgcaggttt ga catctccg agttcccgac caagtttgcg gcgcagatca ccggcttctc cgtggaggac tgcgttgaca agaaga acgc gcggcgttac gacgacgcgc tgcgtacgc gatgttgccc tccaagaagg ccctgcgcca ggcggga ctg gagaaggaca agtccccga gggctacgga gcgctggata agacgcgcgc gggcgtgctg gtcggctc gg gcatggggcg gctgacgctc tccaggacg gcgtcaaggc gctggtggag aagggtaca agaagatga g ccccttctc atccccacg ccatcacaa catgggctcc gcgctgtgg gcatcgacca gggcttcag ggg cccaact actcgtctc cacggcctgc gcgacctca actacgctt tgtgaacgcg gcccaaccaca tccgaa ggg cgacgcggac gtcattgctg tgggcggcac cgaggcctcc atcgtcccc tgggctggg cggtttgt g gcctgccgcg cgctgtccac gcgcaacgac gagcccaagc gcgcgagccg gccgtgggac gagggccg cg acggcttctg gatggcgag ggcggcgccg tctgtgtcat ggagtcgctg gacgacgcgc agaagcgcg g gcgcaccac ctgggcgagt acctgggggg cgccatgacc tgcgacgcgc accacatgac ggaccgcac cccaggggc tgggctgag cactgcatc cgctgggcgc tcgaggacgc cggcgtctc cccagcagg t caactacg caacgcgcac gccacctca ccctggggg cgacaaggcc gaggtgcgcg cggtcaagtc ggt ctttgc gacatgaagg gcatcaagat gaacgccacc aagtcacga tgggcaactg cctggcgcc gccgg cggca tggaggccgt cgccacgctc atggccatcc gcaccggctg ggtgcacccc accatcaacc acgaca ccc catcgccgag gtcgacggcc tggacgtct cgccaacgcc aaggccacgc acaaaatcaa cgtcgccat c tccaactct tgggttccg cgggcacaac tccgtctcgt ctttgcgc cttccgcgag tagcgggagc ga gcgcgctt ggctgaggag ggagcgggg tgcgagccct ttgctgcgc gcgatactc cccgcacga gca gactcca cgcgctgaa tctattgtc aacgagcaac cgtgtgttt gtcgttgccc attcttata ttctccgac t gtggcgta ctctgttg ctgtgcaagc acccaggaa ctctttgtc gagcggggg tgcgtagag agggtc cgcg aaacaggcac cgtgcgag cctcgtgtg gtcattgtc gacatgatt cgggagatt actgatcgg cattgcttac cattcattt aaaagacagc ttgcaaccta caattcgaag gagcgaatac atatataac ggggccc ggcg tgcagatcg ccgttctc ctgagctc acggcccgac acagctgctc gatgacttc tccagcgga ggcagccgc ctggccgtc gcgtacgtc gcaccgccag ggctgagctc ccggcctct cggggccgat g accgccatg acggggacct tggccctc ggcgttgcg atgagcttc ccatggaggc cccgccgggc gcg agctcca cgcgcaggcc gcgcgcgcgc atggcgtctc ccaccgtctc catgtaggc gcgaccgct ccgtc gtggg cagcaggcgc acctgctcc gggccagcca gagcgggaag gcgcccgcgt agtgcctg catgga tgccg aagaagcgt ccagcgagcc gagcaggcg cggtgaatca tgatgggccc ctggcggctg ccgtc ggcg ccgctgact gagcccaac cgtcgggca aattgaagtc cagctgcacc gtggagcact gccattgc g cccgagcgc tctggtatc tgacgtgat cttcgggccc taaaaggcgc cgcccccgc gtcaccgtg a aggccacgc cctgctgat aatgattct ggttggagg gcggtttt cgaggtctg atgatgata ggtatgc gcc acaataatc tacactcgc ccagtcagc acacgcgtc tgcacaatc ctactactat agactcggg a

SEQ ID NO.	Source features	Sequence
		tgatgatac cacccttttc atcgaacgag atgggccatt cctactgac aaacacacac gcacctgtgc gttcagc gct tcggcgagcg ccgcctcggc cctggcccag tctcgtcgg acccgaccga ctcgtccggc cgcggtggag a ggttgacctc gaaccggcgc gcgtcgaagc caaaggcgcg gaagacgcgc tccgtgaggt ccagcacgg c ccggatctcg gccgcgatct gctccggcag gcagaagacg tgggcgtcgt cctgcgtgaa gccgcgcacg cggaaagaggc cgtgcatggt gccggagcgc tcgtagcgggt agaccgtgcc cagctccgcc cagcgcagcg g cagggtcgcg gtagctgacg ggccgggccc cgtagacgct ga
12	Moltype = DNA length = 4793 Location/Qualifiers 1..4793 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	gcctgcctcg ccaactgcga gttctgcgcg cgcggcgggc gcgagggccc gcacggcgcg gcgggggtta ccgcctcggc cgaggagctc tgggagcggc ggcgagcaga gtttgcac ctcccggagc aggagttgc ag gcatcatg ctcgtcatca tgcctgccg ctctgacaag tcgaaaaagg gcgtcgccaa gtacgcaagc acagg gggtg ggggtggggtg ggggtggttg cgggtcgatg aaaatggtgg tgtttggtg ttcacattgg cgtgcgga g tttgtgaag tgagcttcac tctgcggctc tcccggcgca tgcctagtcc agtttccgat gtgcgtatc gtgtac cacg ccgcgtgcga atcccacct ttctctccg ccccgaggtat ggtaacctga gcgacggcct gctgcacctg gtgctcatca ggccgtgcag ccgcctccag tacctcaagt ttctgctcag gatgtcgac atcgccctcg aggc gggcgg gcagcacggg tctacatcc aggtcctgcc cggccacgcg gtgcacatcg aagcggtagc agcgt ggatc gggaggcact tgacagtgc gtcgcccgac aaggcacggg cgcctgagg agctgttgg gccgctg att cactccctg gcctccctt ttctttcca aaccagggtg gccaggagag ccaactggaac gtcgatgggg ag ctcattca aagccgcacg atcaacgcgc agctccaccg cggagtgate gacgtgttg ccagaggcgt ggagg gtga cgagcgcgaa actagctggg ggcccattc ccgcccttga aagcctgcc tctggggccc tggccggt gc tggcgccct gaccgattcg tccattccgt ttattcttt gacaatgagc gcctcatccc agtgcctcat ccag tgcca cggccatcc ccaattgtt cctctcaaac ctctcagata ccccttca aactgetccc aagagtgcac gagtattgt aatgttatac gaccgctgt cacaatgtag tcggcatgct tgtatgacc tgttgaaac tgcctcag cc cgtcggcctg cccatggacc tccctgate gcttcacgcc caccgcatgg actgcgacac gccagtttg g gattcaata tcaaaaaata cgcttcggta attctgtaac gttgcatgcc tcaccgcgtg aattgctgtc ccaaacgta a gcatcatcgt ggctcggta cgcgacctg gatccgggga tcttagaccg ctggtggaga gcgtgccgt cg gattggtg gcaagtaaga ttgcgcaggt tggcgaagg agagacaaa accggagggt ggaagcgggc aca acatcgt attattcgt atagtagagc agtggcagtc gcattcgag gtccgcaacg gatcgcgaa gctcgtac g ctacagtag gagaaagggg accactgcc ctgccagaat ggctgcgacc ctctccctcg ccggccccgc ctgcaaacg cagtgcgtat ccggcaagcg ggctgtgcc ttaaccgcc cccatgttg cgtccgggt cgt cagggtg cgctgagggg ggtttgtgt gcccgcgcct ctgggccgt gtcggccgtg cggacgtgg gccctg ggca gtggatcagc agggtttgcg tgcaaatgcc tataccggcg attgaatagc gatgaacgg atacggttgc gctcactcca tgcccatgcg acccgttct tctcggcag ccgtggctgc ccgggctgcg aagcgggacc cc accagcg cattgtgate accggaatgg gcgtggcctc cgtgtttggc aacgatgtcg agaccttta caacaag ctt ctggaaggaa cgagcggcgt ggacctgatt tccaggtgcg taggtccttg gatgcacgc tctaggttgg g aggcggctg gcgagggaagc agcaggcttg gggtttgggtg ttccgatttc tggaatttg aggtttcatt gtgagat tct atcggttctt gttcgaaaa catgcaaca ctccacacac acacactcct ctcaccaac tctgcaggtt tga catctcc gagttccga ccaagtttgc ggcgagatc accggcttct ccgtggagga ctgctggac aagaaga acg cgcggcggtg cgacgacgcg ctgctgtacg cgtatgtggc ctccaagaag gcctgcgcg aggcggga

SEQ ID NO.	Source features	Sequence
		ct ggagaaggac aagtccccg agggctacgg agcgctggat aagacgcgcg cgggctgtct ggtcggctc g ggcattggcg ggcgtacggt ctccaggac ggcgtcaagg cgctggtgga gaagggtac aagaatga gcccccttt catccctac gccatcacca acatgggctc cgcgctggtg ggcacgacc agggcttcat gggg cccaac tactcgtct ccacggcctg cgcgacctcc aactacgctt ttgtgaacgc ggccaaccac atccgcaa gg gcgacgcgga cgtcatggtc gtgggcggca ccgaggcctc catcgtgccc gtgggcttg gcggtttgt ggcctgccgc gcgctgtcca cgcgcaacga cgaagccaag cgcgcgagcc ggccgtggga cgaaggccgc gacggcttcg tgatgggca gggcgcggcc gtgctggtca tggagtcgt ggagcacgcg cagaagcgcg gcgcgacct cctgggcgag tacctggggg gcgccatgac ctgcgacgcg caccacatga cggaccgcga c cccaggggc ctgggctga gcacctgcat ccgctggcg ctcgaggacg ccggcgcttc gcccgacgag gt caactacg tcaacgcgca cgccacctcc accctggtgg gcgacaaggc cgaagtgcgc gcggtcaagt cggt ctttgg cgacatgaag ggcataaga tgaagccac caagtcctat atcgggcaat gcctggcgcg cgcggg cggc atggaggccg tcgccacgct catggccatc cgcaccggct ggggtgaccc caccatcaac cagacaa cc ccatcgccga ggtcgacggc ctggacgtcg tcgccaacgc caaggccag caaaaatca acgtcgccat ctcaactcc ttggttcg gcgggcacaa ctccgtctc gcctttgcgc cttccgcga gtaggcgag cga gcgcgct tggctgagga gggaggcggg gtgcgagccc ttggtctcg cgcgatactc tccccgcag agca gactcc acgcgcctga atctactgt caacgagcaa ccgtgtgtt tgcctgtgc cattctatt atttccga ct gtggcgt actctgttg gctgtcaag caccacagga actctttgt cgaaggggg gtgctgtaga gagggtc cgc gaaacaggca ccgatcgca gcctcgtgtt ggtcattgt cgacatgatt cccggagatt cactgatcg g cattgctta ccattcatt gaaaagacag ctgcaacct acaatcgaa ggagcgaata catatatata cggggcc ggc gtgtagatc gccgttcgc tttgagtc cacggcccga cacagctgt cgtgacttc ctccagggc a ggcagccgc gctggccgc ggcgtacgtg cgcaccgcca gggtcgagct cccggcctct tccggccga tg accgcat gacggggacc ttggccctct cgcgcttgc gatgagctt ccatggagg ccccgccgg gcgc agctcc acgcgcaggc cgcgcgcgcg catggctct gccaccgtct ccatgtagg ccgcaccgcg tccgtc gtgg gcagcagcg cactgtctc ggggccagcc agagcgggaa ggcggccgcg tagtgctga tcagga tgcc gaagaagcg tccagcgagc cgaagcggc gcggtgaatc atgatgggc gctggcggt gccgtcg gcg gccgcgtact cgaagccaaa ccgctgggc aaattgaagt ccagctgcac cgtggagcag tgccattgc gcccagcgc gtcctggatc ttgacgtga tttcgggc gtaaaaggcg ccgcccgcg cgtccaccgt ga aggcccag ccctgtcat gaattgtcg tggtagag ggcggatttt tcgaggtctg gatgatgata cggtagc gc cacaataata ctacactgc gccagtcagc gacacgcgtg ctgcacaatg cctactata tagactcgt ga tgatgata ccacctttt catcgaacga gatgggcat tcctactga caaacacaca cgcacctgt cgttcagc gc ttcggcgagc gccgcctcg ccctggccca gtctcgtcg gaccgaccg actcgtccgg ccgctggag aggttgacct cgaaccggcg cgcgtcgaag ccaaaggcg ggaagacgcg ctccgtgag tccagcagg cccggatctc gggccgcatc tgcctggca ggcagaagac gtgggcgtcg tcctcgtga agccgcgac gc ggaagagg ccgtgcatg tccgggagcg ctcgtagcgg tagaccgtgc ccagctccgc ccagcgagc ggc aggtcgc ggtagctgac gggccgggccc gcgtagacgc tga
13	Moltype = DNA length = 4205 Location/Qualifiers 1..4205	aaaccgatg atgccagcg cggagctcag catgctccc agcatcgctg gataggacga attggcagg ta ccactga agaattgggag gcaggtgctg ctgattacga gtgtgaaaa gaaagggggc agagagccgt cctc

SEQ ID NO.	Source features	Sequence
	moltype = other DNA organism = synthetic construct	agatcc gaccactatg caggtagccg ctgccccgtg cccgcctggc tgaatttga cacatgccca tcaaggca gg caggeatttc tgtgcatgca ccaaacccat aatcttcgaa aacacacage atgtaccaac gcacgcgtaa aa gttgggggt gctgccagtg cgtcatgccca ggcatgatgt gctcctgcac atcgcccatg atctctcca tegtctcg gg cgtctccggc gcctgggtccg ggagccgttc cgccagatac ccagacgccca cctccgacct cacgggggtac tttcagcgc actgtcggta gtcgacgac gcgtccacca tggagtagcc gagggcgcgg aactggcgtg acg gagggag aacagggggg gttgatgac acacgccagt ctcaaacgc atgtaagacc cgttgatta ttagtat aat catgcattac tagtgggtg agcggcaggc ataaggcaca ccgacgtcga tggcacgagc aactcccga tcatatttc tattgtcta acgccaagcc ggaccacc ccgcatgctg tactacagcg cacgcaccgc ttcgtg atcc accgggtgaa ctagtctc gacggaaaca tctggttcgg gcctcctgct tgcactccg cccatgccga caacctttct gctgttaca cgaccacaa tgcaacgcga cagaccgtg tgggactgat cgttctactg cacct gcatg caattgtcac aagcgttac tcaattgta ttggtttt ttctgggagc agttgtcga cggccgcgt cc cgcaggca gcgatgacgt gtgcgtggcc tgggtgttc gtcgaaaggc cagcaacct aaatgcagg cgatc cggag attgggatct gatccgagt tggaccagat cggcccgat gcggcacggg aactgcatg actcggcg cg gaaccagct ttcgtaaat ccagattggt gtcgatacc tggatttgc atcagcgaaa caagactca gca gcgagcg tatttggcg gcgtgctacc aggggtgcat acattgccca ttctgtctg gaccgttta ctggcgag a ggggtgagtg atgggggttg caggcatcga aacgcgcgtg catggtgtgc gtgtctgtt tcggtgcac gaa ttcaata gtcggatgg cgacggtaga attgggtgtg gcgtcgcgt gcatgcctcg cccctcggg tgcattg acc gggactggaa tccccctcg cgaccatct gtaacgct ccgactctc cgaccgcgc caggatagac tctgttcaa ccaatgaca atggcgtctg ccgtcacct tgcgtgcgc cctccccgc gcggtgctgc cgcgc cgggt cggcgcgtg cctcgcgtc cctgggtgtg cgcgcggctg ccagcaggc cccgctgggc gttccgc cct cgtgacagc cccctcccc gtgtctact ccaagctgga caagcagac gcctgacgc cagagcgc t ggagctggtg cagagcatgg ggcagttgc ggaggagagg gtgctgccc tegtgcaccc cgtggacaag ctgtggcagc cgcaggactt ttgcccgc cccgagtcgc ccgactcga gcatcaggtg gcggagctgc gcg cgcgcg caaggacctg cccgacgagt acttgtgtg gctggtgggg gacatgatca cggaggaggc gctgc cgacc tacatggcca tgcacaac gctggacggc gtgcgcgac acagggcgc gccgaccac cgtgg gcgc gctggacgc gcatgggtg gccgaggaga accggcacgg cgacctgtg aacaagtaact gctggct gac ggggcgcgtc aacatgcggg ccgtggaggt gacctcaac aacctgatca agagcggcat gaaccgc ag acggacaaca acccttattt ggggttcgtc tacacctct tccaggagcg gccaccaag tacagccag g caacaccgc gcgcttgcg gccgagcag gcgacaagaa cctgagcaag atctcgggc tgatgccag cg acgagggc cggcacgaga tcgctacac gcgcatcgtg gacgagttct tccgcctcga ccccgagggc gcc gtcgccg cctacgcca catgatgcgc aagcagatca ccatgccgc gcacctcatg gacgacatgg gccac ggca ggccaaccg ggccgcaacc tctcggcga cttctccgc gtcgccgaga agatcgacgt ctacgac gcc gaggactact gccgacctt ggagcacctc aacgcgcgt ggaaggtgga cagcggcag gtcagcgg cc agggccgcgc ggaccaggag tacgtctgtg gcctgcccc gcgttcgg aaactcgcg agaagaccg c cgccaagcgc aagcgcgtc gcgcaggcc cgtgcctt tctggatct cgggcgcga gatcatggtc t agggagcga cagtggtgc tgcggggtg gcgggagtg gacgcctc tctctctct ctgtttgaa cgga acaat ggccacccc cgctacgcgc cagcatcga gcaacgaaga aaaccccc atgataggtt gcggtg

SEQ ID NO.	Source features	Sequence
		getg cgggatata gatccggccg cacatcaaag ggcccctccg ccagagaaga agctccttc ccagcaga ct cctctgctg ccaaaacact tctctgtcca cagcaacacc aaaggatgaa cagatcaact tgcgtctccg cgt agcttcc tcggctagcg tgcttgcaac aggtccctgc actattatct tctgcttcc ctcgtgaatta tgcggcagcg gagcgctcgc tctggcagcg gctccttcgc gccgccctcg ctgacgaggt gtacagtcga tgaatggtec tggg cgaaga acgaggggaat ttgtgggtaa aacaagcacc gtctctcagg ccccgccgca gtggccgta aagtcca aga ccgtgaccag gcagcgacgc gcgtccgtgt gcgggccctg cctggcggct cgccgtgcca ggctcag ag cagctccctc aggtcgctt ggacggcctc tgcaggccg gtgagggcct gcaggagcgc ctgagcgtg gcagtggcgg tcgtatccgg gtgcggcgtc accgcctcg actgccatc cgaagagcca tcgtcgtcgt ctc ccagact ggacgcgctc gccatcgcc ccgccatgc agaacacgcc tctgcgtcg ccgtgcgcag cgcg gcgcgc agcgctcgc gcgccaggcg cgcgtccagc aggcccatc ccagctccc gccgcgaggg ccg gccagcg cctcaaagcc gccggcgcg gcggccgcca ccacgagcg cagcgggcg cgcgcgcct cc agggtcgg cggccgggc acggcgctc gcgccagcg gccagcgcc gcccccgta gaacgcagt c gctccatc gaccgcacaa agtgggagac ccggtggctc gtggggtcgt agggccgcac gcggcgcgcc g cggccctcgt tgggtgcgc cgtgctgctc cgcagcccg caatcgtaac gctcccttc cgtccacgc tcgcg gcggc cagctggcc cagccagct ggaggccag gcgcagctc gcgggacct gctccagaaa ggggc cgtcg ccagcatgc acagcgcca ggggagcgtc aggtcgaacc gcaggcgctc cagcgtggcg acgcg cggag gctgcccgg gagcgccag ggacgagcgg gcgaltgggt ggccggcgcc gaataaccag ggtg cactcc gtcagcagcc tctccgacc tcacacccc ctgccttct agctggcgcc accggggcg gatggcc gca ccggcgctgt ccagcagcc cgttcacaca gccacgtgca gctcgatcc ggcgacgtcc gtctc
14	Moltype = DNA length = 4205 Location/Qualifiers 1..4205 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	aaaccgatg atgccagcg cggagctcag catgtcccg agcatcgctg gataggacga attggcaggg ta ccactga agaattggag gcaggtgctg ctgattacga gtgtggaaaa gaaaggggcg agagagccgt cctc agatcc gaccactatg caggtagcgc ctgcccgtg cccgcctggc tgaattatga cacatgcca tcaaggca gg caggcattt tgctcatgca ccaaacccat aatcttcgaa aacacacagc atgtaccaac gcacgcgtaa aa gttgggt gctgccagt cgtcatgcca ggcatgatgt gctcctgcac atcgccatg atctctcca tctctcg gg cgtctccggc gcctggctcg ggagccgtc cgccagatac ccagacgcca cctccagct caggggtac tttcgagcg actgtcgga gtcagcgc gcgtccacca tggagtagcc gagggcgccg aactggcgtg acg gagggag aacagggggg gttgatgac acagccagt ctcacaacgc atgtaagacc cgttgatta tgagtat aat catgcattac tagtgggtg agcgccagcg ataaggcaca ccgacgtcga tggcagagc aactcccga tcatattcc tattgtcta acgccaagcc ggaccacc cgcagctcg tactacagcg cagcaccgc ttcgtg atcc accgggtgaa cgtagtcctc gacggaaca tctggtcgg gcctctgct tgcactccg cccatgccga caaccttct gctgttaca cgaccacaa tgcaacgca cagaccgtg tgggactgat cgttactg cacct gcatg caattgtac aagcgttac tcaattgta ttgtttgt ttctgggagc agttgctga ccgcccgct cc cgcagga gcgatgacgt gtgcgtggcc tgggtgttc gtcgaaaggc cagcaacct aatcgcagg cgatc cggag attgggatct gatccgagt tggaccagat ccgcccgat gcggcacggg aactgcatg actcgcg cg gaaccagct ttcgtaaat ccagattgt gtccgatacc tggattgcc atcagcga aaagactca gca gcgagcg tatttgccg gcgtgctacc aggtgtgcat acattgccca ttctgtctg gaccgttta ctggcgag a gggtagtgg atggggtgg caggcatcga aacgcgcgt catggtgtgc gtgtctgtt tcggctgcac gaa

SEQ ID NO.	Source features	Sequence
		ttcaata gtcggatggg cgacggtaga attgggtgtg gcgctcgcgt gcattgcctcg ccccgctggg tgcattg acc gggactggaa tccccctcg cgaccattct gctaacgctc ccgactctcc cgaccgcgcg caggatagac tctgttcaa ccaatcgaca atggcgtctg ccgtcacctt tgcgtgcgcc cctccccgcg gcgcggctgc cgcgc cgggt cgccgcgctg cctcgcgtcc cctgggtgtg gcgcgggtcg ccagcgaggc cccgctgggc gttccgc cct cgggtgcagc cccctcccc gtgtctact ccaagctgga caagcagcac cgcctgacgc ccgagcgcc t ggagctggtg cagagcatgg ggcagtttc ggaggagagg gtgctgccc tgctgcacc cgtggacaag ctgtggcagc cgcaggactt ttgcccgc cccgagtcgc ccgacttga ggtacagtg gcggagctgc gcg cgcgcgc caaggacctg cccgacgagt accttgggt gctggtgggg gacatgata cggaggaggc gctgc cgacc tacatggcca tgctcaacac gctggagggc gtgcgcgacg acacgggcgc ggccgaccac ccgtgg gcgc gctggacgcg gcagtggtg gccaggaga accggcacgg cgacctgctg aacaagtact gctggct gac gggggcgctc aacatgcggg cgtggagggt gaccatcaac aacctgata agagcgcat gaaccgc ag acggacaaca accctattt gggttcgtc tacactctc tccaggagcg gcgaccaag tacagccacg g caacaccgc gcgccttgc gccgagcag gcgacaaga cctgagcaag atctggggc tgatgccag cg acgagggc cggcacgaga tcgcctacac gcgcatctg gcagagtct tccgctcga ccccgagggc gcc gtgcgcg cctacgcaa catgatcgc aagcagata ccatgcccgc gcacctatg gacgacatg gccac ggca ggccaaccgc ggccgaacc tcttcgcca ctttcgcgc gtgcgcgaga agatgcagct ctacgac gcc gaggactact gccgactct ggagcacctc aacgcgcgt ggaaggtgga cagcgccag gtcagcgg cc agggcgccgc ggaccaggag tacgtcctg gcctgcccc gcgcttcgg aaactcgcc agaagaccg c cgccaagcg aagcgcgtc gcgcaggcc cgtgccttc tctggtatc cggggcgga gatcatgtc t agggagcga cagtggtgc tgcggggctg gcgggagtg gacgcccctc tcgtctctc ctgtctgaa cgga acaatc ggccacccc cgtacgcgc cagcatcga gcaacgaaga aaacccccg atgataggt gcggtg gctg ccgggataa gatccggccg cacatcaag ggcccctcg ccagagaaga agctcttcc ccagcaga ct cctctgctg ccaaacact tctgttcca cagcaaac aaaggatga catatcaat tgcgtctcc cgt agcttc tcggtagcg tgcttgaac aggtccctgc actattatc tctgcttc ctctgaata tgcgcaggc gagcgtcgc tctgagcgc gctcttcgc gcccccctc ctgacgagt gtacatgaa tgaatgtcc tggg cgaaga acgagggaat ttgtgggtaa aacaagcgc gtctctcagg ccccgcgga gtcgctgta aagtcca aga ccgtgaccag gcagcgcgc gcgtccgtg gcgggcccgc cctggcggt cgcgctgcca ggctcgag ag cagctccctc aggtgcctt ggacggcctc tgcgagggcg gtgagggct gcaggagcgc ctcgagcgtg gcagtggcgg tcgtatccgg gtgcgggct accgctcgc actcgccatc cgaagagcca tcgtcgtcgt ctc ccagact ggagcgcgc gccatcgcc ccgcccgtgc agaacagcc tctgcgtcg ccgtgcgcag cgcg gcgcgc agcgtcgc gcgccagggc cgcgtccagc agggccatc ccagctccc gccgcgaggc ccg gccagcg cctcaagcc gccggcgcg gcggccgcca ccacgaggc cagcgcggcg cgcgcgccct cc aggtcgcg cggcggggc acggcgtcgc gcgccaggc ggccagcgcc ccccccgta gaacgcagt c gctccatc gaccgcaca agtgggagac ccgtgggtc gtggggctgt agggccgcac gcggcgggcg g cgggcccgc tgggtgcgc cgtgctgctc gcagccccg caatcgtaac gctccctgc cgtccacgc tcgcg gcggc cagctggcc cagccagct ggagggccg gcgcatgct gcgcggacct gctccagaaa ggggc cgtcg cccagcatgc acagcgcca ggggagcgtc aggtcgaacc gcaggcgctc cagctggcg acgcg

SEQ ID NO.	Source features	Sequence
		cggag gctcgcccg ggcgcggcag ggacgagcgg gcgattgggt ggccggcggc gaataaccag ggtg cactcc gtcagcagcc tectccgacc tcacaccccc ctgcctcttg agctggcggc accggggcgc gatggcc gca ccggcgctgt ccagcacgcc cgttccaca gccacgtgca gctcgatgcc ggcgacgtcc gtctc
15	Moltype = DNA length = 4184 Location/Qualifiers 1..4184 moltype = other DNA organism = synthetic construct	ggagctcagc atgtcgccag catcgctgga taagacgaat tggcagggta ccacctgaag aatgggaggg ag gtgttgc gattacgagt gtgtaaaaga aaggggtaga gagccgtcct cagatccgac cactatgcag gtagcc gtc gcccatgccc gcctggctga atattgacac atgccatca aggcaggcag gcatttctgt gcattgcacca agcccacaat cttccacaac acacagcatg taccaacgca cgcgtaaaag ttgggggtgt gccagtgcgt catg ccaggc atgatgtgt cctgcacatc ggccatgac tectccatcgt tctcgggtgt ctcggcgcc ttgtccggg a gccgttccgc cagataccca gacgccacct ccgacctcac ggggtacttt tcgagcgact gtcggtatgc ga cgatcgcg tccacatgg agtagccgag gcgcgggaac tggcgtgacg gagggaggag agggaggaga g agggggggg ggatgatcac acgccagtct cacaacgcat gtaagaccgg ttgtattatg agtacaatca tgcac tacta gttgatgag cggcaggcat aaggcacacc gacgtcgatg gcacgagcaa ctcccgcatc atatttct a ttgtctcac gccagccgg tcaccatccg catgctcgta ttacagcgca cgcaccgctt cgtgatccac cgg gtgaacg tagtctcga cggaaacatc ttgttcgggc ctcgtgctgg cactccctcc catgccgaca acctttt gc tgtaccacg acccagatg caacgcgaca cgaccgggtg gactgatcgg ttactgcac ctgcatgcaa t tgtcacaag cgcatactcc aatcgatcc gtttgatttc tgtgaaaact cgtcgaccg ccccggtccc gcaggca gcg atgacgtgtg cgtggcctgg gtgttcgtc gaaaggccag caaccccaaa tcgacggcga tccggagatt gggatctgat ccgagttgga ccagatcccc ccgatcggc accgggaactg catcgactcg gcgcgggaacc ca gcttctgt aaatgccaga ttgtgtccg ataccttgat ttgccatcag cgaacaaga cttcagcagc gacgtatt t ggcgggcgtg ctaccagggt tgcatacatt gcccatttct gtctggaccg ctttactggc gcagagggtg agtt gatggg gttggcaggc atcgaaacgc gcgtgcatgg tgtgtgtgtc tgtttcggc tgcacaatc aatagtcgga tgggcgacgg tagaattggg tgtgtgctc gcgtgcatgc ctcgccccgt cgggtgtcat gaccgggact gga atcccc ctcgcgaccc tcttgtaac gctcccgact ctcccgccc gcgcgaggat agactctagt tcaaccaa tc gacaatggcg tctgccgca ctttgcgtg cgcctctccc cgcggcgcgg tcgcccgccc gggcgcgcgc gtgcctcgc gtccctggt ggtgcgcgc gtcgccagcg aggcctcgct gggcgtgccg ccctcggtgc ag cgccctc ccccggtgct tactccaagc tggacaagca acaccgctg acgcccagac gcctggagct ggtg cagagc atgggtcagt ttgcggagga gaggggtgct cccgtgctgc acccgtgga caagctgtgg cagccg cagg actttttgcc cgaccggag tcgccgact tcgaggacca ggtggcggag ctgcgcgcgc gcgccaag ga cctggcgac gactactttg ttgtgctggt gggcgacatg atcacggagg aggcgctgcc gacttacatg gccatgctca acacttgga cgtgtgctgc gacgacacgg gcgcggctga ccaccgtgg gcgcgctgga c gcggcagtg ggtggccgag gagaaccggc acggcgacct gctgaacaag tactgttggc tgacggggcg cg tcaacatg cgggcccgtg aggtgacct caacaacctg atcaagagcg gcataaccc gcagacggac aaca acctt acttgggctt cgtctacacc tcttccagg agcgcgcgac caagtacagc cagggcaaca ccgcgcg cct ggcggccgag cagggcgaca agggcctgag caagatctgc gggctgatcg ccagcgacga gggccgg cac gagatgcct acacgcgcat cgtggacgag ttcttcgcc tcgacccga gggcgccgtc gccgcctacg ccaacatgat gcgcaagcag atccatgc ccgcgacct catggacgac atgggccag gcgagggcaa c ccggggcgc aacctcttg ccgacttctc cgccgtcgcc gagaagatcg acgtctacga cggcaggac tact

SEQ ID NO.	Source features	Sequence
		gccgca tctggagca cctcaacgcg cgctggaagg tggacgagcg ccaggtcagc ggccaggccg ccgc ggacca ggagtacgtt ctgggcctgc cccagcgctt ccggaactc gccgagaaga ccgccgcaaa gcgca agcgc gtcgcgcgca ggcccgtcgc cttctcctgg atctccggac gcgagattat ggtctaggga ggtacgag cg cgcgcgaggg attggtggga gtgggacgcg ctcgtcgtc cttctatc tgaagggaag attggccacc c cgctccacg cgccacgc atgagcaacga agaaaacccc ccgatgatag gttgcagtgg ctgccgagat ata gatccgg ctgcacgtca aagggccctt cggccagaga agaagctctt ttccagcgac cgcagactcc ttctgcc aaa aacactctc tctgtccaca gcaacaccaa tggatggaca gatcaacttg tctctcgc tagcttctc ggc tagcgtg cttgcaacag gtcctgcac tattactc ctgcttct ctgaattatg cggcaggcga gcgctcgtc ttgcgagcgc tcttcgcgc cgcctcgt gatcgagtgt acagtcaatg aatggtcctg ggcgagaac gagg gaattt gtgggcgaga gagcatcgtc tctcaggccc cagcgcagtg gccgttaag tcaagaccg tgaccag gca gcgcacgcg tccgtgtcg ggccctgcct ggcggtcgg cgcgccagcg tcgagagcag ctccctca gg tcgcttga cagcctcgc gaggccgtg agggctcga ggagcgcctc gagcgtggca gtggcggtg tatccgggtc gccggtcact gactgcgact cccatccga agagccatcg tcgtcgtc ccagactgga agc gtcgc atcgccccc cgatacaga aaacgcctt cgcctcgcg tgcgcagcgc ggcgcgacc aggg ccggt ccagcagacc catcgccagc tcccgccgc gagggccggc cagcgcctca aagccgccgg cgcg ggcggc cgcccccgc gaggccagcg cggcgcgcg ccctccagg gtcggcgccc gggtcacggc gtc gcgcgc caggcgccca gcgcgcgcc ccgcaggacg cagtgcctt ccatcgaccg cacaagtgg gag accggt ggctcgtggg gtcgtagggc cgcacgcggc ggcgcgcggg cctcgtgaac gtgcgttgc cattc cgag cctcgcgtc gtcactcc cttgccatc cagcgtcgc cgccagcc tggccagc cagctgg agg cccaggagca tggcgcgcg gacctgtcc agaaaggggc cgttgcagc catgcacagt ggccagg gga gcgtcaggtg gaaccgcagg cgctccagcg tggcgacgc cggaaggctc cccgggagcg gccagg gacg cgtggacgat tgggtggcg catcgaatg ctggagcga ctccgtcagc ggctccgcc gacgcc cac cccctgcct cttgagctgg cggcaccggg ccgcgatggc cgcaccggcg ctgtccagca cgcgcgtt c cacagccag tgcagctcga tccggcgac gtctgtctc atca
16	Moltype = DNA length = 4202 Location/Qualifiers 1..4202 moltype = unassigned DNA note = <i>Prototheca</i> <i>wickerhamii</i> organism = unidentified	ggagctcagc atgtgccca gcacgcggc ataagacgaa ttggcagggt accacctgaa gaatggagg ca ggtgtgtg tgattatgag tgtgtaaaag aaaggggtag agagccgtcc tcagatccga ctactatgca gtagcc gct cgcccatgcc cgcctggctg aatattgatg catgccatc aaggcaggca ggcatcttg tgcacgcacc a agcccacaa tcttcacaa cacacagcat gtaccaacgc acgcgtaaaa gttgggtg tgcagtgcg tcatg ccagg catgatgtc tctgcacat ccgccatgat ctctccatc gtctcgggtg ttccggcg cttgctcggg agccgttcg ccagatacc agacgccacc tccacctca cggggtactt ttcgagcgtc tccggtagt cgac gategc gtccaccatg gtagtaccga ggcgccggaa ctggcgtgac ggaggaggga gagggaggag aga gagggggg gggggggggg ggggatgatt acagccagc ctcacaacgc atgcaagacc cgtttgatta tgagt acaat catgactac tagatgatg agcgccaggc ataaggcaca ccgacttga tggcatgagc aactccc ca tcatatttc tattgtctc acgccaagcc ggaccatc cgcagctca tattacagc cagcaccgc ttctg gatec accgggtgaa cgtagtctc gacggaaaca tctggctcgg gcctcgtgt ggactccct cccatgcg a caacctttt gctgtacca cagccacga tgcaacgca cagaccggg tgggactgat cggttactg ca cctgcatg caattgtac aagcgcatc tccaatcga tccgtttgat ttctgtgaaa actcgtcga ccgcccg

SEQ ID NO.	Source features	Sequence
		t cccgcaggca gcgatgacgt gtgcgtgacc tgggtgttcc gtcgaaaggc cagcaacccc aaatcgagg c gatccggag attgggatct gatccgagct tggaccagat cccccacgat gcggcacggg aactgcatcg actc ggcgcg gaaccacgct ttcgtaaatg ccagattggt gtccgatacc ttgattgcc atcagcgaaa caagacttca gcagcgagcg tatttgccgg gcgtgtacc aggggtgcat acattgccca ttctgtctg gaccgcttta ccggc gcaga gggtagattg atgggggtgg caggcatcga aacgcgcgtg catggtgtgt gtgtctgtt tgggtgcac aattcaata gtcggatggg cgacggtaga attgggtgtt gcgctcgcgt gcattgcctc ccccgctggg tgtc atgacc gggactggaa tccccctcg cgacctctct gtaacgctc ccgacttcc cggcgcgcg caggata gac tctagtcaa ccaatcgaca atggcgtctg ccgtcacctt tgcgtgcgcc cctcccccg gcgcggtgc c gcgccgggt cggcgcgctg cctcgcgtcc cctggtggtg cagccgctcg ccagcgaggc cccgctgggc gt gccgccct cgggtcagcg cccctcccc gtggttact ccaagctgga caagcaaac cgcctgacgc ccga gcgcct ggagctggtg cagagcatgg gtcagttgc ggaggagagg gtgtccccg tctgacacc cgtgga caag ctgtggcagc cgaggactt cctccccgac cccgagtcgc ccgacttca ggaccaggtg gcggagct gc gcgcgcgcg caaggacgtg cccgacgagt actttgtggt gctggtgggc gacatgatca cggaggaggc gctcccgacc tacatggcca tctcaaac cttgacggt gtgcgcgacg acacgggcgc ggctgaccac cc gtgggcgc gctggacgcg gcagtgggtg gccgaggaga accggcacgg cgacctctg aacaagtact gtt ggctgac ggggcgcgac aacatgcggg ccgtggaggt gaccatcaac aacctgatca agagcggcat gaac ccgcag acggacaaca acccttactt gggcttctg tacacctct tccaggagcg cgcgaccaag tacagcca cg gcaacaccgc gcgcctggcg gccgagcag gcgacaagg cctgagcaag atctcggggc tgatcgcc ag cgacgaggcg cggcacgaga tcgctacac gcgcatcgtg gacgagttt tccgctcga ccccgaggg c gccgtcgccg cctacgcaa catgatgcgc aagcagatca ccatgcccgc gcacctatg gacgacatgg g ccacggcg ggccaacccg gggcgcaacc tcttcgccg ctttcgccg gtcgccgaga agatcgacgt cta cgacgcc gaggactact gccgcatct ggagcacctc aacgcgcgtt ggaagggtga cagcgccag gtca gcggcc agggcgccgc ggaccaggag tacgttctgg gcctgcccc gcgcttcgg aaactcgcc agaag accgc cgccaagcg aagcgcgtg cgcgcaggcc cgtgccttc tcttgatct ccggacgca gattatg gtc tagggaggtg cagcgcgcg cgagggtg gtgggaggtg gacgcgctg tgcctctt ctattctgaa gggaagattg gccacccgc tccacgccc acgcatcgag caacgaagaa aacccccga tgataggtg ca gtggctgc cgagatatag atccggctgc acgtcaaagg gccctcggc cagagaagaa gctctttcc cagcg accgc agactcctt tgcacaaaac actctctct gtccacagca acaccaatgg atggacagat caacatgtgt cttcgctag cttctcggc tagcgtgctt gcaacaggtc cctgcactat tatctctctg ctttctctg aattatgcg g caggcgagcg ctcgctctt cagcgctcc ttcgcgccg cctcgtgat cgagtgtaca gtcaatgaat ggt cctgggc gaagaacgag ggaattgtg ggcgagagag catcgtctct caggccccg cgcagtggcc gttaa agtcc aagaccgtga ccaggcagcg catcgcgtcc gtgtcgggc cctgcctggc ggctcgcgcc gccagg ctcg agagcagtc cctcaggtcg ccttgacag cctcgcgag gccggtgagg gtctcagga gcgcctcg ag cgtggcagtg gcggtgtat ccgggtcgcc ggtcatcgac tgcgactcg catccgaaga gccatcgctg t cgtctcca gactggaagc gctcgccatc ggccccgccc atacagaaaa cgcctctcg ctcgccgtgc gca gcgcggc gcgcaccagg gccgcgtcca gcagacccat cgcagctcc cggccgagc ggccggccag cg cctcaaag ccgcccggcg gggcgggccc cgccgcccag gccagcgcg cgcgcgccc ctcagggtc g

SEQ ID NO.	Source features	Sequence
		gcgcccggg tcacggcgtc gcgcgccag gcggccagcg ccgcccccg caggacgcag tgcgcctcca t cgaccgcac aaagtgggag acccgggtgc tcgtggggtc gtagggccgc acgcggcggc gcgcgggect c gtgaacgtg cgtttgcat tccgcagccc tgcggtcgtc aactccctt gccgatccac gtcgcggcg gccag cctgg ccagggccag ctggaggccc aggagcatgc cggcgcggac ctgtccaga aaggggccgt tgcgc agcat gcacagtggc caggggagcg tcaggtgaa ccgcaggcgc tccagcgtgg cgacgcggcg aggtc cgccc gggagcggc agggacgcgt ggacgattgg gtggcgggca tcgaagtgt ggagcgcact ccgtca ggcg cctccgccga cggcgaccc cctgcctct tagctggcg gcaccgggccc gcgatggcg caccggc gct gtccagcag cccgttcca cagccagtg cagctcgatg ccggcgacgt ctgttccat ca
17	Moltype = DNA length = 4448 Location/Qualifiers 1..4448 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	acgaattgtt gcttgaagg gtgctgaatg gcaccagata ctgggtcggg ggcattgtct caggcgtcat tttag gttc tgccaataat aatccatga ggacagcaag ccattgaca gccactgac tcgaatttta gagcaaccta c gagaccaca ccatcgtgc acgcacctg aggtgtttt tggatagtg cagccgagcg catggccctg cgcc gtcctc ggtagaggga gggcagcaca gggccctctc ttctttttg actggccaag ctggcacgtc ttgggtgc c gggatcaggt gagcttctg cggcgggggc gagcagggcc accaggtcga ggacctccct ctccaggcag ccaatggtgt ggagcaggcc tctatctgg tgatagccg cttgtctgc ctggctgctg gcaccggtcc tcttg gatgg cggaggtggc gcaagaacac ccgtgtgtc cagttcatc aacagctctt gcgcagtgc gcacaccg ca aggacgcagg tactttaaa agacacaagc gtgacatgtg cgtgatcatt gcaacggctg ccgcagagga cttaccacag ctcagcaagc agtgttgaga cggccttggc caccctatg agtcacgca gcagctttc gtgct ggaag gccaggacct cgtttcttg ctgggtctcc acggctcgc gtcgggattc ctggaggag gagggagc cc aggggagagt ggtgttgta tgacgacgc tgtgcgaga ggcatgaga tctcatgag tgcagctct t tgcagccag gtttgcctgt cgcacagact tgagaggcac caccaaaagt ctactatcca aataatgctc cagagc atca gcaatgttg ccacctcag taccaacgt cccacctcg tctgcgacc tgcgtgccc gccccgcgc ggcctcggca gcagggagc gcgccgcgc gccctcggcg gccctcgct gcgctcag gcctcacgt t cccttcca gcgcgccag cgaggcctgg acgtccacca ggtgcttga ccgagcaag ttgcgtctc gcag cccggc caggttggc accagctct ccatgtgga gtacgtgtg gccgtggaga tcttctgga agctcgg tg atggcctgct cgcgtgcgc gtgcgacgca aagtggacgc ctggtgaggc cgcgccttg ccgtcagg g agggaccccc gcccgcccg gccctcggcg cggcgccgc gccgcccag gccctcggc ccgcccgc gc gccgcccgtgc gcttcagcc gcacgggcgc tgggtagtcc tgcgagaaga gctgggcat ctctgctc cgcgcgac gctccgcgc gcgatcggc tcccgggtcg cgcgtctct ctggatgag gccgccagcc gccgccact ctgctcagc gcgccgact cgcgtccgc ctgcgccgc gctgcgccc ccagcgcct gg ccttttg ccgctcgt gcgctcgt gatgtccat accaggtcca tgaggtctg ctgcccgg ctgagcc act gttcgtccg ggcggccaag aggagcatga gggaggactc ctggtccagg gtcctgacgt ggtcgggc t ctgggagcg gccagcatc tctggtctg ccgaccgag gccgcctcca actggtctc cagcagccg a gtcgcgcc gacctggca gaggaagaca ggtgagggg gtatgaattg tacagaaca ccacagcct tgt ctaggca gaatccctc cagtcagtc ttacctgga tgacggcctg cgaacagtg tccagcgacc ctcgtg ccg ccgtcttc cgcacgttc ttccagcac cgtgatggcg cgagccagcg ccgcacgctg gcgtgcgt tcgccgatct gaggacatc ggggaactct gatcagtcta aaccccttg cgcgttagtg ttccatct ttgcag accg gtgagagccg actgtgtg cgccacccc cacaccact cctccagac caattctgc accttttg c

SEQ ID NO.	Source features	Sequence
		gaaggcatc ggctcggcc tgcagagagg acagcagtgc ccagccgctg ggggttgccg gatgcacgt ca gcttgtt tccagaagga gttgctcct gagccttca ttctagect cgataacctc caaagecgt ctaattgtgg aggggggtcg aatttaaaag ctggaatgt tggctcgtgc gtctggaaca agcccagact tgtgctcac tgggaa aagg accatcagct ccaaaaaact tgccgctcaa accgcgtacc tctgcttcg cgcaatcgc cctgttgaat t cgccaccac attcatattg tgacgcttga gcagtctga attgcctcag aatgtggaat catctgcccc ctgtgcga gc ccatgccagg catgtcgcgg gcgaggacac ccgccactg tacagcagac cattatgta cctcacaata gttcataaca gtgaccatat ttctgaagc tcccaacga gcacctccat gctctgagt gccaccccc ggccc tggtg ctgctggagg gcaggtcaac cggcatgggg ctaccgaaat ccccgaccgg atccaccac ccccgcg atg ggaagaatct ctccccggga tgtgggcca ccaccagcac aacctgctgg cccagcgag cgtcaaacc a taccacaaa atatecttg catcggcct gaattcttc tgccgctctg ctaccgggtg cttctgtccg aagca gggggt tctagggat cgtccgagt ccgcaaaccc ttgtcgtg gcggggcttg ttgagctg cagcatcgc t tgcacggggg ctatctcaa gtagtcggga agcatgggtc gcgacgtcga gacggcgag cgggcctatg g gtgagtcgg ccgtgcatgg ggagggctag cgccagggt cgggtcgccc cgtgcccgg ggtcttctg cgg ggtgcac gcggccgtgg gtgcgagaca ggttgcgtga cgaattctt ggaccccc caggatgcc tcacgc gggc gccccctcc ctgcagga ctccgcaag gtgtccggcc cgggtgtgtt ggccagcagc atgagcg ggt ctccatgta cgagctgtgc cgcgtggcg cggacaagct gattggcgag atcattgcc tcgagggaga cacggcgacc atccaggtga ggtggcgggg gaggtgctt ggtggaagt gggcgacaca acctccatt g gttggcggc caaggatgca ggggaagcga ttctgagcat tgcgcgtt ttgtgtgta catatgttg cttgca cccc atgtatttc caggtgtac aggatactc ggccctgac gtggcgaca cgtgtctcc ctccgcaa g gtgtggaggg tgtgagatt gatgtatg ctttggcgc cccctgtc ttgcacgtca gaggcgcacc tca gtgtc acaagtgtg ccgttagtga aactgaacaa gtgcttgact gttaggatt tccgtcgcc cttttttac c ctctccct ttctcttt ctcacccct caaaatctc agccctgtc tgtggagct ggccccggca tcatgggc ac catcttgac ggcatccage gccegtcaa gtgacgtcg gtggacagc attcgtct catccgcgc gg cgtggacg tgccggccct ggaccgaag gcctcgtggg agtttgacc ggtgtcaagc tcaaggtgg gcga ccgat cacggcggc gacatctac gcgtgtgca cgagaacacg ctgatggagc acaaggtgt gcttc gccg ggcgcgcgc gcaccatc ctacatcg ccggcgggga gctacagcat cacggacaag atcatg agg tctgatttg cggcgcgcgc aaggagtact ccatgctga gctctggccc gtgcgcgcgc cgcgccccg t ggcgcagaag ctgtggcca acacgccgt gctgacgggg cagcgcgtgc tggacgcgt cttccccggc gtgctgggc ggacgtgcgc catccgggc gcgtttgct gcgcaagac ggtcatctg caggcgtgt ca aagtacag taactggag ggcatcatct acgtcgctg cggcgagcgc ggcaacgaga tggccgaggt gctc atggac tccccgcgc tgacctgac catgccggac gggcgcgagg agagcatcat gcagcgacc acgtc gtgg ccaacacct caacatgccc gtgcgcgc gcgaggccag catctaca
18	Moltype = DNA length = 4448 Location/Qualifiers 1..4448 moltype = other DNA organism = synthetic construct	acgaattgt gtttgaagg gtgctgaatg gcaccagata ctgggtcgtt ggcatctgct caggcgtcat tttag gttc tgccaataat aattcatga ggacagcaag ccattgacaa gccactgac togaattta gagcaacct c gagaccaca ccaccgtgc acgcacctg aggtgtttt tggatagtg cagccgagcg catggccctg cgcc gtctc ggtagaggga gggcagcaca ggccctctc tttttttt actggccaag ctggacgtg ttcgctgc c gggatcaggt gagcttctg cggcgggggc gagcagggcc accaggtcga ggacctcct ctccaggcag

SEQ ID NO.	Source features	Sequence
		ccaatggtgt ggagcaggcc tctatgtgg tgatgagccg cttgtgctc ctggctgtcg gcaccggtcc tctg gatgg cggagggtggc gcaagaacac ccgtgtgtcg cacgttcac aacacgtctt gcgcagtgtc gcacaccg ca aggacgcagg tactttaaa agacacaagc gtgacatgtg cgtgatcatt gcaacggctg ccgcagagga cttaccaccag ctcagcaagc agtggtgaga cggccttgge caccaccatg agtcagcga gcagcttttc gtgct ggaag gccaggacct cgtttcttg ctgggtctcc acggctcgcc gtcgggattc ctggaggagg gagggagc cc aggggagagt ggtgttga tgacgacgcg tgtgcgcaga ggcaagtaga tgctcatgag tgcagcctct t tgcagccag gtttgcctgt cgcacagact tgagaggcac caccaaaagt ctactatcca aataatgtc cagagc atca gcaatgttg ccacctcag taccaacgt cccacctcg tctgcgcacc tgcgtgccc gcccccgcgc ggcctcgga gcgacggacg cccccgcgc gccctggcg gcctccgct gcgctcgag cgcctcacgt t cccttcca gcgcccag caggcctgg acgtccacca ggtgcttga ccgagcaaag ttgcgtctc gcag ccccgc caggttgcc accagctct ccatgtgga gtagcctgt gccgtggaga tctctgga agcgtcgg tg atggcctgct cgcgtgcgc gtgcgacga aagtggacgc ctggtgaggc cgcgcctgg ccgtcagg g aggaccccc gcccgcccgc gctcggcg gcgcgcgc gccgcccag gccctgccc ccgccgcg gc gccgcctg gcgtccagcc gcacgggcgc tgggtagtc tgcgagaaga gctgggcat ctctgctcg cgcgcggcac gctcgcgc gcgacggcg tcccgggtc cgcgtcctg ctggatgag gccgccagcc gccgccact ctgctcagc gcgccgact cgcgtccgc ctgcgccgc gcgtgcgcc ccagcgcct gg cctttcg ccgctcgt gcgctgct gatgtccac accaggtcca tgaggtctg ctgcgcgg ctgagcc act gcttcgtcg ggccgccaag aggagcatga gggaggactc ctggtccagg gtcctgact ggtcgcggc t ctgggagcg gccagcatca tctggtctg ccgaccgag gccgctcca actggtctc cagcagccgc a gtcgcgcc gacctggca gaggaagaca ggtgaggggg gtatgaattg tacagaaca ccacagcct tgt ctaggca gaatccctac cagtcatggc ttacctgga tgacggcctg cgaacagctg tccagcgacc ctgctg ccg ccgtctcc cgcacgttc ttccagcac cgtgatggc cgagccagc ccgacagctg gcgtgcgt tcgccgatct gaggacagt ggggaactct gatcgtcta aacccccctg cgcgttagtg ttgccatct ttgag accg gtgagagccg actgtgtg gcgccccc cacaccact cctccagac caattctgc accttttg c gaaggcatc ggcctcgcc tcagagagg acagcagtc ccagccgctg ggggttgcg gatgcagct ca gctgttt tccagaagga gttgtcct gagccttca tctcagct cgataacct caaagcgcct ctaattgtg agggggttc aatttaaa ctggaatg tggctcgtc gtctggaaca agccagact tgtgtctac tgggaa aagg accatcagct ccaaaaaact tgccgtcaa accggtacc tctgcttc cgcaatctc cctgttga t cgccaccac attcatattg tgacgttga gcagtctga attgcctcag aatgtggaat catctgccc ctgtgcga gc ccatgccag catgtcgcg gcgaggacac ccgacctg tacagcagac cattatgcta cctcacaata gttcataaca gtgacatat ttctgaagc tcccaacga gcacctcat gctctgagt gccaccccc ggccc tggtg ctgctggagg gcaggtcaac cggcatggg ctaccgaaat ccccgaccg atccaccac ccccgcg atg ggaagaatct cccccgga tggggccca ccaccagcac aacctgctg ccaggcgag cgtcaaac a taccacaaa atatcttg catcgccct gaattcttc tgcgctctg ctaccggtg cttctgctg aagca ggggt tctaggat cgtccgagt ccgcaaac ttgtcgtg gcggggttg ttgagctc cagcatcgc t tgcacgggg ctatctcaa gtatcgga agcatgggtc gcgacgtga gacggcgag cgggcctatg g gtgagtcgg ccgtgcatgg ggagggtag cgccagggt cgggtcgc cgtggccgg ggtcttctc cgg

SEQ ID NO.	Source features	Sequence
		ggtgcac gcggccgtgg gtgcgagaca ggtgctga cgaatttett ggcaccccc caccgatgcc tcacgc gggc gcccccctcc ctgcaggtga cgtccgcaag gtgtccggcc cgggtggtgt ggccagcagc atgagcg ggt ctgccatgta cgagctgttc cgcgtgggcg cggacaagct gattggcgag atcattgcc tcgagggaga cacggcgacc atccaggtga ggtggcgggg gaggtgctt gattggaagt gggcgacaca accctccatt g gttggcggc caaggatgca gggaagcgaa tgcgagcat tgcgcgttt tgcctgctga catatcgttg ctgca cccc atgtatttc caggtgtacg aggatactc gggcctgacg gtggcgacaca cggctgccg ctccggcaa g gtgtggagggt tgtgagatt gatgtatatg ctttggcgc cccctgtc tgccacgtca gaggcgacc ttca gtgctc acaagtgtg ccgttagtga aactgaacaa gtgcttgact gtgtaggatt tccgtcgcc cttttttac c ctctccct ttctcttt ctccacct caaaatctc agccctgtc tgtggagctg gggcccgga tcatgggc ac catcttgac ggcaccagc gcccgctcaa gtgcagcgc gtggacagc attcgtgct catcccgcg gg cgtggacg tgcggccct ggaccgcaag gcctcgtgg agtttgacc ggtgtcaagc ttcaagggtg gcga ccgcat cacggcgcg gacatctacg gcgtggtgca cgagaacacg ctgattggagc acaagggtgct gcttcc gccg ggcgcgcgc gcaccatct ctacatcgc cggcgggga gctacgcat caccgacaag atcatcg agg tcgagtttg cggcgcgcg aaggagtact ccatgctgca gctctggccc gtgcgcgcgc cgcgccccg t ggcgcagaag ctgctggcca acacgcgct gctgacgggg cagcgcgtg tggacgcgt cttccccggc gtgctggcg ggacgtgac catccggcg gcgtttgct gcggcaagac ggtcatctc caggcgctgt ca aagtacag taactggag ggcacatct acgtcggctg cggcgagcgc ggcaacaga tggcgaggt gctc atggac ttccccgc tgacatgac catccggac gggcgcgagg agagcatcat gcagcgacc acgtc gtgg ccaacacct caacatgccc gtcgcgcgc gcgaggccag catctaca
19	Moltype = DNA length = 4318 Location/Qualifiers 1..4318 moltype = other DNA organism = synthetic construct	gcaatctccc gcccaacct caccagcgg cgcaggacct ggagcgcgaa agaggatttc gggggtgtgg g ggagagtga gcgcgatgga agacactgat ttgcataaac aacctttggc cgtttcaaa gaccacctc accac cataa aacctccct ccatgcata cttgatgcc tggcgagatg gcttgggctt ggtggacct ggccatgtag tccggcgtgg tcgcgatgga cgcgcacgt tgtcgatga gcgccttgc cccgtctcg tggccgagcc cgc gtccggg cggctcggcc agccccgca gcagccccgg gtccagctc gcgttgaggg agcgtttgc cagc gcacgg atggactct acggcagcgt gatgaggtgg agcaggctgt cgagcaggcc cggctcgac cagag catgg tgaggatgc cgcgcgcgc ggcgtgccgg gctgcaccgc cagcagcaac cggcgagcc tctcg gccgc aaacaccgcc aggaagcagc cctccaggt cgggcccggc tccagcaggg gaccgcaggc gtgc actatg tccaggcgct ccgtgatgat gtcatggcc agcggcagca gcgagcgag cagctgctg gccagg tgcg tgagcgtgct cgtgccggcc gagtccgagg tctcgtccgt ggccaggtec atggcgcca gcgtgcgt c gctgcctgg tgcgccgtc agtgcgccg gtggcacagc agccgcagca ccgctggaa gagcggcag g cagcccgct ccaccgcgc ggcggcgccc gggcgccc cctctgcag gctggcgcg aggatggc cg ccgactccg cgcgcgcag agtgcctg tgtgtctaa aagtcgcac agcgtcgag agatgtcctg cggctccggc tccgggatgt gctggatgcc acgcacaccg gtcgaatca cctgttggat aatgggtggc agg atggttg gtaggttct gatcggctg gggaggcgga taagtcaca gccggtgat cttgctgacc aaacact gct ggccaccaa caaaagtcga ttgccaccga tgcacacggc atggcgcca gacagaagct cagcatacc c tcaggaggc caggccag ccaggccaca tgtcagctg gaagccggc tgggctgaag catgggcca g cagctcagg gcatggtttt gaaagcagg aacgtgctt gtcacgtca caaccaggga agaggcctgc g

SEQ ID NO.	Source features	Sequence
		acatcaagg ggcaaaagttg gaatgttcag catcatcagc ctgggaccaa tactctgctc acggttcagg gcaga aaggt ggcatagtc gggcctctgt cagaatttca agtcccagtt taaacttctg acctccgcga ggctgatecc ctgcccctga gcggactggg gccgggggagc cgggggtggga ccggtggtca cctgcgccat ggctagcccc c accctctgt gatctgggag ccctgcgcag ccccttaaat catctcagtc aggtttctgt gttcaactga gcctaaag gg ctctctcat gcgcacgagc acacgtatat cggccacgca gtttctcaaa agcggtagaa cagttcgcga gc cctcgtag gtcgaaaact tgcgccagta ctattaaatt aaattaattg atcgaacgag acgcgaaact ttgcagaa t gccaccgagt ttgccagag aatgggagtg gcgccattca ccatccgct gtgcccggct tgattcgccg ag acgatgga cggcgagacc agggagcggc ttgcagccc cgagccgga gcaggaacaa tgatcgacaa tct tctgtc caattactgg caaccattag aaagagccgg agcgcggtga aagtctgcaa tcgagtaatt ttctgatac g tcgggcctgc tgaacctaa ggctccggac ttgtttaag gcgatcaag atgcacgagg cccagggcac gt atctcaag cacaaccccc agccttagtt tcgagacttt gggagatagc gaccgatac tagtttgga tttgtatat taattacctc aagcaatgga gcgctctgat gcggtgcagc gtcggctgca gcacctggca gtggcgctag ggctc gcccta tcgctcgaa cctgttcagc tggctccgc ctctctgca gctcttcca cccgtagcgt ctgctgttg ggagctggag tcgtgggctt gacgacggcg ctgcagctgt tgcaggatgt gcctggcggt cgcgttcacg tcg tggctga gaaatatggc gacgaaactg tgacggctgg ggccggcggg ctgtgtagtc catacgatt gggtac gcgg ccattggatg ggattgatag gcttatggag ggataataga gtttttccc gatccaacgc atgtggatgc g gtatcccg tgggctgaaa gtgtggaagg atagtgcatt ggctattcac atgactgcc caccctttt ggcagg aaat gtgcccgcgt cgttggtgca ccgattggga aaatcgactg tcgaccata catgaagatt tatacgtctg a agatgcagc gactgcgggt gcgaaacgga tgacggtttg gtcgtgtatg tcacagcatg tgctggatct tgcgg gctaa ctccccctgc cagggccat tgcagggtgc atgttgactg gagggtacga ccttctgtec gcaaatcc cagaggagga cccgctctgg gccgacattg tggccacttt tcgcgcctg ggcaaggcag agctcctgac cta cgacccc agcggcaagt cgatgcaggg ctacggcttc accaccatca tcacgtacga tgctcaatga aagggg agtt cattgcctc actagagaaa catacatggg ttatgtataa ctgacgat ctccgaatt cctctgttg tgtctc tgggt tgtgcaggga aggtgcctc tactgcctt ggctgatgca gcagattcaa ggcttggggg ggaccttga acggcgccc atctccagcc tgcggagct gaaggagtat gatgccatg tcaattgcac aggtgggttg gagt tagtcc aggagatgc gatagaccaa tgcaaatcgt gtgcaacgca ttgacccaa tcatgctc cctcgccgc t ctacggcctg gagcgccaa agctggtgca ggacgagtc atgtaccgg tgcgcgggca cgttctgcgc g tacgggcgc cctgggtccg cactacatc aaccgcgacg ggggcaccta catcatccc aacacggaca cgg tgggtct gggcggcac acgcaaaagg gcaactggtc cctcgagccg accgaggagg atcgcgcgcg gatc ctggag cgctgtacg agatcctgcc cagcctgcgc aaggcgccga tctgcgcga gtgggtcggg ctccgg cccg gccgccaga catccgctg gagcgcaag atgcgcagct cgacggcaag tccgtaccg tcatccac aa ctacgggcac ggcggtccg ggctcacct gggctggggc tgcgccgag acgcccgcgc gctcgtgcg c agggcgctgg ccctctgacg gtgagagggc ctgttcagg ctgtcggtc cgtctatct caaaggcctt gg aattggct gggcgaatat tctgatcat tgatgtgat ctcttcgctt ttcttatgat gcaccactt ggagcgagta tt ctggtgcc aattgtctt tgaatcacg ctgcccacgg tggcaaatgc aatgagagcc gagtaagtgg actggtg gtt cctaaggcga cgggcagtc aaggaagcc agccccgat cccagctgc cacatcttg ttgcacgct gtatctggcg gtttgaatga gactcgcgc caccaggcc ggccctcgc ctgcggctcc gcccaact cga

SEQ ID NO.	Source features	Sequence
		gcagctt gtcctttgcc cctggcaaca tggctcgtat ggtgtcgtac acgggggcag gaaccagggtg caggcct gga gaagagggat ggaagtgcgc acagattcga gaagggtatg gccctcttgc cgaaatgcac accctcttcc cccagcctct gctccaccc accgtctgct cgggaaggga tgcggcagtc caggacaacg tccgggttgc gct ggtgtcg gagctccata tcgcggcccc gttctgaatg gatggggtgg gacagttggt gtgaccgg
20	Moltype = DNA length = 4317 Location/Qualifiers 1..4317 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	gcaatctccc gccaacctc caccagccgg cgcaggacct ggagcgcgaa agaggatttc ggggtgtggg g gagagttag cgcgatggaa gacactgatt tgcataaaca acctttggcc gctttcaaag caccacctca ccacc ataaa accctccctc ccatgcatac ttgatgcct gccgagatgg cctgggcttg gtggaccatg gccatgagt ccggcgtggt cgcgatggac gccgcacgtt gtgcgatgag cgccttgccc ccgtctctgt gcccagagccc gc gtccgggc ggctcggcca gccccgccag cagccccggg tccagctcgg cgttgaggga gcgctttgcc agc gcacgga tggactcga cggcagcgtg atgaggtgga gcaggctgtc gaggcagccc ggctgcgacc agag catggt gaggatgcgc gccgccggg gcgtgccggg ctgcaccgcc agcagcaacg cggcgaccgt ctgc gccgca aacaccgcca ggaagcagcc ctccgaggtc gggccgggct ccagcagggg accgcaggcg tgc actatgt ccaggcgctc cgtgatgatg tgcattggcca gggcagcag cgagcgcagc acgtctctgg ccagg tgcgt gaggctgctc gtccggccgg agtccgaggt ctgctccgtg gccagggtcca tgggcggcag cgtgcgct cg ctgcctgtgt gcgccgtcca gctgcggcgg tggcacagca gccgcagcac cgctggaag agcggcag gc agcccgctc caccgccgcg gcggcgggcg gggcgccgcc ctctgcagg ctggcgcgca ggatggc cgc cgactccgc gccgccgaca gctgctgcct gtgtctcaaa agctgcgaca gcgtcgaga gatgtcctgc ggctccggct ccgggatgtg ctggatgcca cgcacaccgg tcgcaatcac ctgttgata atgggtggca gga tggttgg ttagttctg atcggtcgg ggaggcggat aaggtcacag ccgctgtatc ttgtgacca aacactg ctg gccaccaaac aaaagtcgat tgccaccgat gcacacggca tggcggccag acagaagtc agcataccc t caggagggcc agggccaggc caggcccat gtgcagctgg aagccgggct gggctgaagc atgggccag c acgtcaggg catggtttt aaaggcagga acgtggcttg tcactcaac aaccaggga gaggcctgcg a catcaaggg gcaaaatttg aatgttcagc atcatcagcc tgggaccaat actctgtca cgttcaggg cagaa agggt gcatagtcag ggctcgttc agaattcaa gtccagttt aaactcga cctccgcag gctgatcccc t cgccctgag cggactgggg ccggggagcc ggggtgggac cgttggtcac ctgcgccatg gtagccgc a ccctgttg atctgggagc cctgcgcagc cccttaate atctcagta gttttctgt tcaactgag cctaaagg gc ttctgtatg cgcacagca cagtatatc ggccacgag ttctcaaaa gcgtagaac agttcgcgag cc ctctagg tcgaaaact gcgccagtac tattaatta aattaattga tcgaacgaga cgcgaaact ttgcagaat g ccaccagtt tggccagaga atgggagtg gccattcac catccgcctg tggccggctt gattcgcga ga cgatggac ggcgagacca gggagcggct tgcgagcccc gagccggtag caggaacaat gatcgacaat ctt cctgtcc aattactggc aaccattaga aagagccgga gcgcgttgaa agtctgcaat cgagtaattt ttcgatacg t cgggcctgct gaacctaaag gctccggact ttgttaagg cgatccaaga tgcacgggc cccaggcacg tat ctcaagc acaaaccca gccttagtt cgagacttt ggagatagcg accgatattc agtttggcat ttgtatatt a attacctca agcaatggag cgctctgatg cgtgcagcg tcggctgag cacttggcag tggcgtagg gtgc ccctat cgctcggaac ctgtcagct ggctccgcc tctgtctag cctcttcag ccgtagcgtc tgcgtgttg gagctggagt cgtgggctg acgacggcg tgcagctgt gcaggatgt cctggcgtgc gcgttcacgt cgt ggctgag aaatatggcg acgaaacgt gacggctggg gccggcgggc tgtggatgcc atacgcattg ggtac

SEQ ID NO.	Source features	Sequence
		gcggc cattggatgg gattgatagg cttatggagg gataatagag ttttgcggg atccaacgca tgggatgcg gtatcccggg gggctgaaag tgggaagga tagtgattg gctattcaca tgcactgccc acccctttg gcagg aaatg tgccggcatc gttggtgcac cgaaggga aatcgacgtt cgaccactac atgaagattt atacgtctga agatgcagcg actgcgggtg cgaacggat gacggtttg tctgtatgt cacagcatgt gctggatctt gcgg gctaac tccccctgcc acggccatt gcagggtca tttgactgg aggggtacgac ctttctccg tcaattccc agaggaggac ccgctctggg ccgacattgt gccactttt gcggcctgg gcaaggcaga gctcctggcc tac gaccca gcggcaagtc gatcgacgc tacgggtca ccaccatcat cactacgat gctcaatgaa agggga gttc attgcctca ctagagaaac atacatgggt ctatgtaaac ttgatgac tccgaatttc ctctgtgtg gtctct ggtt gtgcaggga ggtgcctct acctgccctg gctgatgcag cagattcaag gctgggggg gaccttgaa cggcgccgca tctccagct gtcggagctg aaggagtatg atgccatgt caattgcaca ggtgggttg agtt agtcca ggagatgtc atagaccaat gcaaatcgtg tgcaacgat tagaccaat cactgcctcc ctgcctgt c tcaggcctgg aggcgcaaa gctggtgcag gacgagtcca tttaccggg gcgggggac gttctgcgcg t acgggcgcc ctgggtccg cactacatca accgcgacgg gggcacctac atcatccca acacggacac ggt ggtgctg gcgggcatca cgaaaaggg caactggtcc ctgcagcca cggaggaga tggcgcggg atcc tggagc gctgctaca gatcctgcc agcctgcga aggcggcgt cctgcgcgag tgggtcgggc tccggc ccgg ccgcccagac atccgctgg agcgcaaga tgcgcagtc gacggcaagt ccgtaccctg catccaca ac tacgggcacg gcgggtccg gctaccctg ggtgggggt gcgcgcaga cggcgtcgcg ctctgcgc a gggcgctg cctctgacgg tgagaggcc tttgcaggc tttgcgtcc cgctcatc aaaggcctt ga attgctg ggcgaattt ctgatcatt gatgtgata tcttcgtt tcttatgat caccacttg gagcgagtat tc tgggtcca attgtctt gaaatcacg tcccacggg gcaaatgca atgagagccg agtaagtga ctggtg gttc ctaaggcgac gggcagctca aggaagcca gccccgat ccagctgcc acatcttgt tgcacgct g tatctgcggg ttgaatgag actcgcgtcc accagggcg gccctgcgc tgggctccg ccacaactc ga gcagctt tctttgcc ctgcaacat ggtctgatg gtgtctaca cgggggcagg aaccagtg aggcct ggag aagaggatg gaagtgcga cagattcag aaggtgatg cctcttgc gaaatgcaca cctcttc c ccagctctg cctcaccca ccgtctgctc gggaaggat gcggcagtc aggacaact ccgggttgcg c tgggtcgg agtccatat cgcggcccc ttctaatgg atgggggtgg acagttggtg tgaccgg
21	Moltype = DNA length = 4227 Location/Qualifiers 1..4227 moltype = other DNA organism = synthetic construct	cggctcgtg ctttgcgtc cgggtgcagc gatcagatcc aagtctgacg acttgcgtg atgtactgtg tctttg agt ccagagcgcc ggccgcacgc gtttctccc cttctcttc ctctgaaca tccagcgatg caagtgcagg g cgctggcg gctggcgtcc cgaaccggcc tggcgcacg cggctgaaat tgccaatgct ggcaatgtag tgc cgctcgg cccatccctc gatcaagttt ttacgcgtg gttggggat gatctgcgt catgggaaga taaaagg gt tctgaggtg ttgattgta cttaccgga agtactcata ttacatag actgatcca caaaaaaca aagct cactt caaagaaccg cgaatgcta cteccagca atcacttgc tcaccgtcg gttgtccc acgacaacgc cggtagaggg gtcggtggc tgcgacct gcgggcaca tcttcacg catgtctga taatctcag ctcata cgtc tggccgtcg acccaaaaat gacgggaccc tgcatgatat cgcccgagat ggggtccagg cattctctg gaggcgtcag cctgcggga gatgccgtc ccaccgatt ggaaaggcac aaagggggtg aatccccat tt catgaaat tttgggtcag cgaatgtgc cactcgtgc caatgaatat ggggtcacgc ggtggacgaa cgcgg agggg gcctggccga atctagctt gattctca gatcatttc tgcggcggt ccgggggttgc cgcgtcgcg

SEQ ID NO.	Source features	Sequence
		aacgctccgt ctccctagcc gctgcgcacc gcgcgtgcga cgcgaaggtc attttccaga acaacgacca tgg cttgtct tagcgatcgc tcgaatgact gctagttagt cgtacgctcg acccagtcgc tcgcaggaga acgcggc aac tggcgagctt cggttgcca gtcgtgactc gtatgtgac aggaatcatt ggcattgcta gcattataat tcgg cttccg cgctgtttat ggcatggca atgtctcatg cagtcgacct tagtcaacca attctgggtg gccagctccg ggcgaccggg ctccgtgctg cggggcacca cctcctgcca tgagtaacag ggccgccctc tctcccgcac gtt ggccca tgaataccgt gtcttggggc cctacatgat gggctgccta gtcgggcggg acgcgcaact gcccg gcaa tctgggacgt ggtctgaatc ctccaggcgg gttccccga gaaagaaagg gtgccgattt caaagcaga g ccatgtgccg ggccctgtgg cctgtgtgg cgcctatgta gtcaccccc ctcaccaat tctgccagt ttgc gcaatc cataaactca aaactgcagc ttctgagctg cgctgttcaa gaacacctct ggggtttgct caccgcga g gtcgacgcc agcatggcta tcaagacgaa caggcagcct gtggagaagc ctccgttacc gatcgggacg ctgcgaagg ccatccccg gcactgttc gagcgtcgg cgcttcgtag cagcatgtac ctgccccttg acatc gcggt catgtccctg ctctacgtg cgtcgacgta catcgacct gcgccggtgc ctacgtgggt caagtatggc gtcatgtggc cgctctactg gttcttcag gtgtgtgta gggttgtgt tgcctgatac gaggtcctgg tggcgc gcat gggggagaag gcgcctgtcc cgctgacccc ccggtctacc ctcccgac cttccagggc gccttcgg ca cgggtgtctg ggtgtgcgcg cacgagtgcg gccaccaggc ctttctcc agccaggcca tcaacgagg cgtgggcctg gtgttcaca gcctgtgtgt ggtgccctac tactctgga agcactcgca ccgccccac cact ccaaca cggggtgcct ggacaaggac gaggtgtttg tggcccgca ccgcgcagtg gcgcacagg gcctg gagtg ggaggagtgg ctgccatcc gcatgggcaa ggtgctggtc acctgaccc tgggctggcc gctgtac ctc atgttcaacg tcgcctcgcg gccgtaccg cgcttcgcca accatttga cccgtgtgc cccatttca gc aagcgcga gcgcatcgag gtggtcatct ccgacctggc gctggtggcg gtgctcagcg ggctcagcg gctg ggccgc accatgggct gggcctggt gtcaagacc tacgtgtgct cctacctgat cgtgaacatg ttgctcgt gc tcatcagct gctccagcac acgcacccgg cgctgccgca ctacttcgag aaggactggg actggtgctg cggcgccatg gccaccgtgg accgtccat gggcccgccc tcatggaca acatctgca ccacatctc gac accacg tctgcacca cctctcagc accatcccg actaccagc cgaggaggcc tccgccgca tcaggc ccat cctgggcaag tactaccagt ccgacagccg ctgggtcggc cgcgccctgt gggaggactg gcgcgact gc cgctacgtc tcccgagcg gcccgaggac gactccgcgc tctggttcca caagttagtg agtgatgc tactcagcg cgcctgcgc gggtatcgga acgcccgcc cgcctgtct tttcacgcg cgactccgtc gtt cgcggg tggcaccccc attgaaaaa acctcaattc ttttggga agacacgggt taccaccaac caccacc tg cacctctatt attgttatta ttgacgcggg agcgggcgtt gtactctaca acgtagcgc tctggtttc agctgg ctcc caccattgta aattctgtct aaaatagtc gtggttatgt gagaggatg gtgtaacagg gcgtcagta tgt tggtttt cgtgctgac tggggcaca ggctcgtc acgtgacgtg cccgtgatga gagcaatacc gcgtca aag ccgacgcatg gcctttact cgcactccaa acgactgtc ctcgtattt tcgatatct atttttaag agcga gcaca gcgccgggca tgggcctgaa aggcctcgc gccgtgctc tgggtggggc gcgcagcgcg tgggg catcg cggcagtgca ccaggcgag acggaggaa gcattgtgag tgcgcatcac aagatgcatg tctgttg tc tgtactataa tctagagca tcaccagggg cttagtcat gcacctgct tggcattac agaaattgca caag ggcgtc ctccgggatg aggagatgta ccagctcaag ctggagcggc ttcagccaa cgaggagcgc ggcgc atgac gacctacca catcgacaa aggggtctgg gtcgtacgac aaaccagtca ggaggcggcg gggtc

SEQ ID NO.	Source features	Sequence
		atga gctggccccg tcgcagcttc agcgcttcga gcatcgccgc gttgtccgcg atccaagcct cccgcagcg c ctcttgggg cccctgtgcg gagaggcggg tgccgagagg gctgtgaggg cggaggctgg aggacaggg t gcgaatcgtc gcgtgccta ggcgtctgcc gaggagaagc acaccacgcg ccgccgcttt catttcattg tgc ccgttgc tgtaatgcat gcgagcgatc ctgcactaaa gctgcgatcg ccacctcagc gttgttcgc tctagcacc g ctcacctgac cagcgcacg gaagagtggc ggtcggcgac ggcctgctgc acgtggcagt gcaggtccac gtggccgagg agcaggtctc gcagctgcag caccgcgcc tcgtgctcgc cgggggtccat ccagtcctat cc ccgcgcg ggcctcgag ccggacggcc agcagccagg gccgggggtc cgcgccgacg cccggcgggc ccgccgcga gctggggctg gccaggaaga gccactgggc ggccggcccg ggcgagcccc gcgccggcg c gggaccacc cgcagacca ccgctggtc caggtcgacg tcgtcgcg gcagggccgc gtgcggcg c tgcgccagcg tccaggcg gggggggcg gcgtgca
22	Moltype = DNA length = 4227 Location/Qualifiers 1..4227 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	cggtcgtg ctttgcgtc cgggtgcagc gatcagatcc aagtctgacg acttgtgctg atgtactgtg tctttg agt ccagagcgcc ggccgcacgc gcttctccc cttctcttc ctctgaaca tccagcgatg caagtgcagg g cgctggggc gctggcgcc cgaaccggcc tcggcgacg cggtgaaat tgccaatgc ggcaatgtag tgc cgctcgg cccatccctc gatcaatgtt ttacgcgcgt ggttggggat gatctgcgct catgggaaga taaaagg gt tctgaggtgg ttgattgta cttaccgga agtactcata ttcatacatg actgatccca caaaaaaca aagct cactt caaagaaccg cgcattgcta ctecccgca atcactcgc tcaccgtcgg gttgctccc acgacaacgc cggtagagg gtcggtggcc tcgcgacct cgcgggcaca tcttccagc catgtctgta taatctcagc ctata cgtc tggcccgctg accccaaat gacgggatcc tgcgatgat cgcgcgagat ggggtccagg cattctctg gagcgctcag cctcgcggga gatcccggtc ccaccgcat ggaaaggcac aaagggggtg aatcccccat tt catgaaat tgttggtcag cgatgtgcg cactcgtgcg caatgaatat ggggtcacgc ggtggacgaa cgcgg agggg gcctggccga atctaggctt gcattctca gatcacttgc tgcggcggt cgggggttg cgcgtcgcgc aacgtccgt ctcctagcc gctcgcacc gcgcgtgcga cgcgaaggtc atttccaga acaacgacca tgg cttgtct tagcgatcgc tcgaatgact gctagtgtg cgtacgctc acccagtcgc tcgcaggaga acgcggc aac tgcgagctt cggttgcca gtcgtgactc gtatgtgac aggaatcatt ggcattggtg gcattataat tcgg cttcg cgctgtttat ggcatggca atgtctcatg cagtcgacct tagtcaacca attctgggtg gccagctccg ggcgaccggg ctcgtgtc cggggacca cctcctgcca tgagtaacag ggccgcctc tctcccgac gtt ggccac tgaataccgt gtctggggc cctacatgat gggctgccta gtcggcggg acgcgcaact gcccg gcaa tctgggacgt ggtctgaat ctcaggcgg gttccccga gaaagaaagg gtgcccattt caaagcaga g ccatgtccg ggcctgtgg cctgtgttg cgcctatgta gtcaccccc ctcaccaat tgctgccagt ttgc gcaat cataactca aaactgcagc ttctgagctg cgctgttcaa gaacacctt ggggtttgct caccgcga g gtcgacgccc agcatggcta tcaagacgaa caggcagcct gtggagaagc ctcggttcac gatcgggacg ctgcgcaagg ccatccccg gcactgttc gagcgctcgg cgttcgtag cagcatgtac ctggccttg acatc gggt catgtccctg ctctacgtc gtcgacgta catcgacct gcgccggtgc ctacgtgggt caagtatggc gtcatgtggc cgtctactg gttcttcag gtgtgtgtga gggttgtgt tggcgatc gaggtcctgg tggcgc gcat gggggagaag gcgcctgtc cgctgacccc ccgggtacc ctcgggcac cttcagggc gccttcgg ca cgggtgtctg ggtgtgcgc cagagtgc gccaccagg ctttctcc agccaggcca tcaagacgg cgtgggcctg gtttcaca gcctgtgt ggtgcctac tactcctgga agcactgca ccgccccac cact

SEQ ID NO.	Source features	Sequence
		ccaaca cgggggtgcct ggacaaggac gaggtgtttg tgccgccgca ccgcgcagtg gcgcacgagg gcctg gagtg ggaggagtgg ctgccatcc gcatgggcaa ggtgctggc accctgacc tgggtggcc gctgtac ctc atgttcaac tcgctcgc gccgtaccg cgcttcgcca accacttga cccgtggtcg cccatttca gc aagcgcga gcgcacgag gtggtcatct cgcactggc gctggtggcg gtgctcagcg ggctcagcgt gctg ggccgc accatgggt gggcctggct ggtcaagacc tacgtggtgc cctacctgat cgtgaacatg tggctcgt gc tcacacgt gctccagcac acgcaccgg cgctgccgca ctactcgag aaggactggg actggtgctg cggcgccatg gccaccgtgg accgtccat gggcccgccc tcatggaca acatcctgca ccacatccc gac accacg tctgcacca cctctcagc accatcccgc actaccacgc cgaggaggcc tcgccgcca tcaggc ccat cctgggcaag tactaccagt ccgacagccg ctgggtcggc cgcgccctgt gggaggactg gcgcgact gc cgctacgtc tcccgagcg gcccgaggac gactccgcgc tctggtcca caagtgagtg agtgagtcgc tcactcagcg cgcctgctg gggatgcgga acgcccgcgc cgcttctt ttgcacgcg cgactccgtc gctt cgcggg tggcaccccc attgaaaaa acctcaattc tgttggga agacacggtg taccccaac caccacc tg cactctatt attggtatta tgacgcggg agcgggctt gtactctaca acgtagcgtc tctggttc agctgg ctcc caccattgta aattctgtc aaaatagtc gtggtatgt gagaggtatg gtgaacagg gcgtcagta tgt tggttt cgtgctgac tcgggcacaa ggcgtcgtc acgtgacgtg cccgtgatga gagcaatacc gcgtca aag ccgacgcatg gccttactc cgcactcaa acgactgtc ctcgtattt tcggatatc atttttaag agcga gcaca gcgccgggca tgggctgaa aggcctcgc gccgtgctg tgggtggggc gcgagcgcg tgggg catcg cggcagtga ccaggcgag acggaggaac gcatggtgag tgcgcatcac aagatgcatg tctgttg tc tctactataa tctagagca tcaccaggg ctagtcac gcacctgtt tggctattc agaaattgca caag ggcgtc ctccgggatg aggagatgta ccagctcaag ctggagcggc tcgagccaa gcaggagcgc ggcgc atgac gacctacca catgcgaca aggggtctg gtctagac aaaccagta ggaggcgcg gggtec atga gctggccgc tcgagcttc agccttcga gcatcgccg gttgtccgc atcaaacct cccgcagcg c ctcctgggg cccctgtgc gagaggcgg tccgagagg gctgtgagg cgaggctgg aggacaggc t gcgaatcgt gcgtgccta ggcgtctgc gaggagaagc acaccacgc ccgcgctt catttcattg tgc ccgttc tgtaatgcat gcgagcgtc ctgactaaa gctgcgac ccacctcagc gttgtcgc tctagacc g ctcactgac cagcgcacg gaagagtgg ggtggcgac ggctgtgc acgtggcagt gcaggtccac gtggccgagg agcaggtc gcagctgcag caccgcgcc tctgtcgc cggggtccat ccagtcctg cc ccgcgcg ggcctcgag ccggacggc agcagccagg gccgggggtc cgcgccgacg cccggcgcc ccggccgca gctggggtc gccaggaaga gccactggc gccggggccg gcgagcccc gcgccggc c gggaccacc cgcagacca ccgctggtc caggtcgac tgcctggcg gcaggcccc gtcggcgc c tgcgccagc tcaggggcag gcggggcg gcgtgca
23	Moltype = DNA length = 4067 Location/Qualifiers 1..4067 moltype = other DNA organism = synthetic construct	gcaatggcg tcggtacagg gtctgctcc gtgctgggt cctctcta cgtgcacaa gggagcgcc cgg ccagtc agcgcgtcca caacctccc tctcacaca cacactgcg gaaccaggcc gccatttgc tcttga gca tgccttgc atgtccggg ttcccatca tctgttgc gttctgggc tccagctt getccagcac acca tcctgt cgtcgaaga gaaggagaca tctgtacatt attggtgta gggcgctgaa tcggccattt tttaatgat cacgtcatg ccaatagac cggcacataa cgacttcaa acccccgca aagcccgga caaccacc cc tccacacc cccacacaa agaaccgcc accgtacc ttgccacga gtaggcctt tctgtcgca aaaccg

SEQ ID NO.	Source features	Sequence
		gcct cggatgatgaa tgcattgccc ttcttgacga gcgctgccc ggccaacacg ctcttttgcg gcgtctctc a ggcttgggg gcctccttg gcttgggtgc cgccatgac tgcgcgcac agagaaacgt tgcgtgtaaa aggag cgccc ggctgcgcaa tatatatata ggcattgccc cacagcccaa cctcactcgg gagcccgcc caccaccc cc aagtcgcgtg ccttgacggc atactgctgc agaagctta tgagaatgat gccgaacaag aggggcacga ggaccaatc ccggacatcc ttgtcgataa tgatctgtg agtccccac gtccgccga cgtccgggg agcc cgccga tgcataagc gagaggccc tcgaccagga ggggctggcc cgggcgggca ctggcgtcga aggt gcgccc gtcgttcgcc tgcagtcta tgcacaaaa caagtctct gacggggtgc gtttctccc gtgcgggca g gcaacagagg tattcacct ggctatgggg agatcgcgca tcgagctggg ataagata cggctcccg c aaggatgc tcactctgt ctgagccgga cagtattct ggcaagcaat gacaactgt caggaccgga ccgtg ccata tattctcac ctgagccgcg aaaacctaac aatttgggag tcactgtgcc actgagttc actgttagct g aatggagtc gctgtccac taaacgaatt gtcagcaccg ccagccggcc gaggaccgga gtcatacga ggg tagtagc gcgccatggc accgaccagc ctgcttgcca gtactggcg ctcttcgct tctctgtgt cctctgcgc g ctccagcgcg tgcgttttc cgttgatca tgcgtccgt ggcgcaccgc agcgccgct gccatgcag c gccgtgct tccgaacagt ggcggtcagg ggcgcaccg cggtagccgt ccgtccgaa cccgccaa ag ttttgga gcagcttgag ctctgcaaga tggcggagga caagcgcac ttctggagg agcaccggtg cgtgg aggtc cggggtgac cggccgtgc attcaacga atcaatcga tgatgatcag aggacacga gtctgtgtg g cgggtggccag aaacactgtc cattgcaagg gcataggat gcgttctc acctctatt tctcttct gaatcc ctcc ctgtcactc ttctctcc tcttcccgt tcacgcagca ttcggggcaa cgaggtgggc cctctgcagc g gtgacgat cagcgcggtg gccaacatcc tgcaggaggc ggcgggcaac cagcggtgg ccatgtggg c cggagctcg gaggtttcg cgacggacc ggagctgcag gagcggggc tcattttgt gatgacgcgc atgc agatcc agatgtaccg ctacccgcgc tggggcgacc tgatgcaggt ggagacctg ttccagacgg cgggca agct gggcgcgag cgcgagtggg tgcgcgcga caagctgacc ggcgagcggc tggcgcggc caccctc gagc tgggtcatga tcaacatcc cagcgccgg cgtgccga tgcggagct cgtccgcgc aagtcggc ct tcttcgcgc cgagccgccc gcctgggcgc tgcgcccgc ggtcacgct gccaaactgc ccaacatgc gacgcccgc cgtctgcgc ggcaccgcca ggtcgcgcgc gcacccgaca tggacatga cgggcacgtg aacaactgg cctactggc ctggtgctg gagccgtgc ccgagcagct ctcagcagc taccactct acca gatgga gatcgactc aaggccgagt gccacgccc cgacgtcac tcctccagg ccgagcagat cccgcc ccag gagcgctca cgcacaacgg cgcggccgc aaccctctc gttctgcca tagcattctg cgcgccga ga ccgagctcgt ccgcgcgcga accacatgt cggccccat cgacgcgccc gccgccaagc cgccaag gc gagccactga ggacagggtg gttgctgga tggggaaac ctggtcgcgg gattcgatcc tgcgttat a tctccctg gaagcacacc cagactctg aagaagaaaa cgtgcacaca cacaaccaa ccggccgaat attt gcttc ttatccggg tcaagagag actgcgatgc cccctcaat cagcatctc ctccctgcc cttcaattt ccctcttgc ctgcgcccgc ggtgcgcccgt ctgcccgcgc agtcagtcac tcctgcacag gcccttctg cgca gtgtc ctgtaccct tacgctctc tccattctg gagggccct attgaatga ttctgtcct gtgtggccaa gc gggctgct gggcgccgc cgtcgggca gtgtcggcg actttggcg aagccgattg ttctctgta agccac gcgc ttgtgcttt gggaagagaa ggggggggta ctgaatgat gagggagga aggggggta ttgtatta t ctgagttggg gaggcaggga gagttgaaa atgtaagtgg cagcagggc aaggagaatg gtgagcatgt

SEQ ID NO.	Source features	Sequence
		gcatggtgat gtcgttggtc gaggacgac ctgcacgcgt gtatctgatg tagaatacgg caatcacctt agtcta catc tataacctct ccgataacg ccccttccaa atgcccctcc gtttctctc tattcttgat ccacatgatg acctg gcac tattcaagg gctggacatt tcaagaagg ttgcgtatct gaagaaggat tggttggag aggtggccga t gaaagtggg gtcaagctgc gtggagccct ctgcacggat ttctatgga atctgcgtcc acgtcatcag tagccg tacg cctctgcggc gtcgccgtgc tctcagccg ccgttcacc accagccagg agccccatac gcacgcgt c ccagacggcg cgcgcgacgg ggacgatcat gagcagcgc atctggggcg tctcgtcga gctgagccg c acaaccacgt tgcctggcc gcgcgtggtg tactggcagt tggcagac catggtgtgc gacttgaagt ag gagccgtc aaagcgcggg aggcgcgtct cggccgagac gttgccgcg atgaccaggc tctggcgctt ctgg cgcggg tcgcggccct gccactgcgc ggagccgccc aggagcaggc gcgactgtc acggccgcgc agct tgactg cggccgccc cgccgtgccc ggctcgtagg ccgacccat cttgacgaag agggcgccga tcgtg cgcg gacctgacc gcgtcggaca ggcgagggcg ctgtccagc ttgacccga gcgccgctt gccgggct tg aacggcagcg agaactctc gccagcttg gagccgagca ggccagcgc cagcttctgc cgttgcgc gcgcgcgag gcgcgagtcg acgcgcagcg ttagag
24	Moltype = DNA length = 4069 Location/Qualifiers 1..4069 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	gcaatggcg tcggtacagg gtctgcgtcc gtgctgggct cctctctca cgatgcacaa gggagcggcc cgg ccagtc agcgcgtcca caactcccc tcgtcacaca cacactcgc gaaccaggcc gccatttgc tgettga gca tgcctgcat catgtccggg ttcccatca tctcgtttag gttcttggc tccagcttct gctccagcac acca tctgt cgatgaaga gaaggagaca tgtgtacatt attggtgta gggcgctgaa tcggccattt ttaaaatga tcacgtcat gccaatagac gcggcacata acgacgtca aaccccccc aaagcccgcg acaacccat cc ctccacac cccccacaca aagaaccgc caccgttac cttgccacg aggtaggcct ttgttgcgc aaaacc ggcc tcggtgatga atgcatccc gtctcagc agcgtgccc ggccaacac gctctttgc tgcgtctct c aggcttggg ggctccttg ggttgggtg ccgcatgat ctgcgcgat cagagaaacg ttgctgtaa aaagg agcgc ccggtgcgc aatatata tagcatgcc aacacagccc aacctactc gggagcccg cccaccac cc ccaagtcgc tgcctgacg gcatactgct gcagaagct catgagaatg atgccgaaca agaggggc gaggaccaa tcccgacat cctgtgat aatgatcgc tgcgtccca tcgtccccc gacgtcccg ggag cccgc gatgctaa acgagagggc cctcagcag gaggggctg cccggcggg cactggcgc gaag gtgcgc ccgtgttcg cctcagtc tatgccaca aacaagtct ctgacgggtg cgtttgtc ccgtgcggg c aggcacaga ggtattacc ctggtcatgg ggagatcgc gatcagctg ggataagaga tacgtcccg c gcaaggatc gctcactg gtctgagccg gacagtatt ctggcaagca atgacaact gtcaggaccg gaccg tgcca tatattctc acctagccc gcaaaccta acaatttgg agtactgt cactgagt cactgtag c tgaatggag tcgtgtcc actaacgaa ttgtcagc cgcagccgg ccgaggacc gatcatagc gag ggtagta gcgcgcatg gcaccacca gctgttgc cagtactggc gtctctccg cttctgtg gtctctgc g cgtccagcg cgtgcgtt tccggtggt catgggtcc gtggcgacc gcagcgccg ctgccatgc a gcgccgtg tccgaaca gtggcggtca gggccgacc cgcggtagc gtccgtccg aaccgcca ag agtttgg gacagctg agccctgcaa gatggcgag gacaagcga tctcttga ggagaccgg tgcgt ggagg tccggggctg accggcgct gcattcaacg taatcaatc catgatgac agaggacacg aagtctg gt ggcggtggcc agaaactg tccattgcaa ggcataggg atcggttct teaccteca ttctcatt ctga atccct cctgctcac tcttctct cctctccc gttcacgcg cattcggggc aacgaggtgg gccctcga

SEQ ID NO.	Source features	Sequence
		gcggctgacg atcacggcgg tggccaacat cctgcaggag gcggcgggca accacgcggt ggccatgtgg ggccggagct cggagggttt cgcgacggac ccggagctgc aggagggcggg tctcatcttt gtgatgacgc gc atgcagat ccagatgtac cgctaccgc gctggggcga cctgatgcag gtggagacct gggtccagac ggcg ggcaag ctgggcgcgc agcgcgagtg ggtgctgcgc gacaagctga ccggcgaggc gctgggcgcgc gcc acctga gctgggtcat gatcaacatc cgcacgcgcc ggccgtgccg catgccggag ctctgccgcg tcaagt cggc ctcttcgcg cgcgaggcgc cgccctggc gctcccgcg cggtcacgc gtgccaagct gcccaaca tc gcgacggcgg cgccgctgcg cgggcaccgc caggtcgcgc gccgcaccga catggacatg aacgggca cg tgaacaacgt ggcctacctg gcctggtgcc tggaggccgt gcccgagcac gtcttcagcg actaccacct c taccagatg gagatcact tcaaggccga gtgccacgcg ggcgacgtca tctctccca ggccgagcag atcc cgcccc aggagggcgt caccacaaac ggccgcggcc gcaaccctc ctgcttcgc catagcatc tgcgcg ccga gaccgagctc gtcgcgcgc gaaccacatg gtcggccccc atcgacgcgc ccgccccaa gccgcc caag gcgagccact gaggacaggg tgggtggctg gatggggaaa cgctggtcgc gggattcgat cctgctgt t atactctcc tgggaagcaca cccacgactc tgaagaagaa aacgtgcaca cacacaacc aaccggccga a tatttgctt cttatcccg ggtccaagag agactcgat gccccctca atcagcatc tctctctgc cgttcaa tc ttcctgctt gcctgcgcgc cgggtgcgcc gtctgccgc ccagtcagtc actctgcac aggcccttg tgc gcagtc tctgtaccc ttaccgctc ctccattct gcgaggcccc ctattgaatg tattcgtgc ctgtgtgcc a agcgggctg ctgggcgcgc cgccgtcggg cagtgtcgg cgactttgg ggaagccgat tgttctctg taag ccacgc gcttgctgt tgggaagag aagggggggg tactgaatgg atgaggagga gaaggagggg tattgg tatt atctgagttg gggaggcagg gagagttgga aaatgtaagt ggcacgacgg gcaaggagaa tggtagca t gtgcatggg atgtcgttg tcaggacga tctgcacgc gtgtatctga ttagaatac ggcaatcacc ctatg ctaca tctatactt ctccgtataa cgcccttcc aaatgcctc ccgtttctt ctattcttg atccacatg tgacc ctggc actattcaa gggtgggaca ttcaagaag gttgcgtat ctgaagaagg attggtttg agaggtggcc gatgaaagt gggtaagct gcgtggagcc ctctgcacgg atttctatg taatctgcgt ccacgtcact agtag ccgta cgcctctgcg gcgtgccgt gctctcagc cgccgttca ccaccagcca ggagcccat acgcacgc gc tcccagacgg cgcgcgcgc ggggacgac atgagcagcg cgatctgggg cgtgtcgtgc gagctgagc c gcacaaccac gttgccctg ccgcgcgtggt gtactggca gttgtcag accatggtgt cggacttgaa gt aggagccg tcaaagcgc ggaggcgcgt ctgcccgag acgttgcgc cgatgaccag gctctggcgc ttct ggcgcg ggtcgcggcc ctgccactgc cgggagccgc ccaggagcag gcgcgactcgc tcaggccgc gca gcttgac gtcggccgcc agcgcctgc ccgctcgta ggccgaccc atcttgacga agaggcgcc gatcg ctgc cggacctga ccgcgtcga caggcgcagg cgtcgtcca gcttgaccc gagcgccgc ttgccgg gct tgaacggcag cgagaactc tcgccagct tggagccgag caggccagc gccagcttct gccgcttgc c gcgcgcgc aggcgcgagt cgacgcgcag cgtgtagag
25	Moltype = DNA length = 4492 Location/Qualifiers 1..4492 moltype = other DNA organism = synthetic construct	ttgcgattcc tcttgacca gccacacga taaatttaa cgcaagctgg cccaatatgg tatccgctgt gtcaga aca gcacctcgt gcacctctg gctctgaatc tcacgtcac ggtcactc ctcctagcct gccagcgtt gg gcgcagcc atcgggcctg catagagaaa ggtacaaa caacatatat aaatacact ggtattcata atactcg cgc atgtttgtct tgcgcacga tggccagct cgtgcacgc ctcacctctg gccaggcagc ccccttgcgg g gtgccccgc agcgggcacc gcgaggcctg aggcgcacc agcagcatgt tgaattatt tatacgagcgt ttttgg

SEQ ID NO.	Source features	Sequence
		aggt acagagaaac cccttgcat tctctcttg tgtgggtgt cagtgccat tgcagtggc tggccaccg ca tatcggaat aagattgctt gcggatgtcg ttcgtgccag cgtcatcggt gcgagagttt ataacgcgtg cccacgc cag aagccctcac gctttgact tggcgaggta caaaaaatac caaacgcgt cctgaagcat tatggcggcg aggacgagcg tgttcaaggt gcggatcgcg acctgagtg agagacgctc cgtggtcgc ggcagaacc cc agcgcccgc taaaagccg gtcctcacgc cccagggtg agtgcgcgcc cgcaaacccc tttgcagga actc aaggag gtggctcgt ccaaggacaa caacggcatc gagctcatt ccaatgact caacatctc ctctggcg gg cgggtattga ggcaaggggg tctgctcat tgggtggagg gagatgagtt gtgtgcagg cgctcaaagg aggttttgg cgcctattgc ccaaggtcat gcatccgtgt gtatcccggt cccctccct tcaacaggcg cccgcg gact cgcctatca ggggggtcac ttgaactca gcatccaagt gcctgagcag taccattgg tcccggcga agtcaagttc aagaccaagc tcttcaccc gaacgtgcac ttaagggtga gctgtttgga gacaaagggt gcag ggagag tgctagcaag caggggttga cgggcgttgc tgtgagttt tgatgtatag actcgccgg gaacttga gt ggaagccgac agaccttca tctgctctc tgcccatgca gacgggagag atctgcctcg acattctca ga acgctgg agcccagctt ggacgtgca ggtacgattg tttggttag gagaggcgcc ccttggtcg tccctg gca ctacgactc ttgggaaat atccgatct ccgacttct gtctccagac gcctgctcg tgcgctgtg ctc gatgaag cattcttct ccccgaatg ccgcagtcg gtgtgccagg ccacgtggc gctcttgaca gattctgcg c cagactccc cctcaactgc gacgtggga acctctccg ggcaggcgat gtgcgtgggt ttgctcctt ggc acggctc tacaccgtcg agtacgcat gaggcggtga tggctgtgc gtttgccat tctccagag acggcaa gtc gtccatctc tgcgtgtgt gcgcgacgt gcagcagtc ctctgcagca gatgagcgtg accttgcca tt tcacgcac tcaggtgac acaatccatt tttctaaag caaatgactg ctgattgacc agatactga acgtgattt cgctccagat cgcacagata gcgacatgt tgctgcgtct gaaaatctgg attccaatt cgaccctggc gtcc atcca tgcaacagat ggcgacactt gttacaattc ctgtaccca tggcatgga gcaggtccac ttgattccc gatcacccac gcacatctc ctaatagta ttcgtctgt tcttcgata atctcaagt agtggtcatg gatcttgg tt gacgatggg tatgggttg cgcgctggc tgcagggtct geccaaggca agtaacca gctctctcc cc gacaatac tctgcagc aaagccggc acttgcttc cagattgcca ataaactca ttatggctc tgcatgc ca tccatgggtc tgatgaatgg tcacgtcgt gtctgaccg tccccagcc tctggctcc cctgcccgc cca ccagccc acggcgcg gcagtcgtg ccaaggctgt ctcggatgcc cagcgccatg ccacgacctt tgatg gttc aagtacgatt acggtgttg attgtgtt tttgctgtag tgtgcatggt ttagaataat acattgatt tcttg ctcac ggcaatctc gttgtccgc aggttcaacc ccatttcgga gtctcaggtc agccgcgcaa tgaccagcc g ctactcaag gacttcacg acaacgccga ggtgagctat gtttaggact tgattggaaa ttgtcgtcga cgca tattcg cgtccgcga cagcaccaa gcaaatgtc aagtgcgtc cgatttgcgt ccgcaggtc atgttgtat cgtcggcgcc ggatccgccc gtctgtctg cgcttacgag ctgaccaagc acctgacgt ccgggtacgc ga gctgagat tcgattagac ataaattgaa gattaaacc gtagaaaaat ttgatggtc cgaaactgtg ctcgattgc a agaaattgat cgtctccac tccgaggtc gccatcatc agcaggcgct gtctccggcg gggcgctg gc tgggggga cagctgtct cggccatgtg tgtagtaga aggatgaatt tcagctggtt tctgtgcac agctgtttgt gcatgattg ttacagacta ttgtgaatg ttttagatt tcttaggatg catgattgt ctgcatgcga ctttagcta g cacccttac acaccatcca acatttctt gcttgctc cctgcgcag atccgcaagc cggcgcaaa gctgat ggac gagtcaaca tccctacga cgacgaggtg ggtattggcg agctagaacg catgctgtct gtgcgactg

SEQ ID NO.	Source features	Sequence
		g gtcgatccat tgtgcgaaac gtgtggcgga atacgtgcgc gtccccggcc atgcaccgac cccccctccc ac cacccacac aaatataaca gggtagcatg gtgtgggtca agcacgcgcg cctggtagc tccacgtgc tgtec aaggt gctggcgccc cccaacatca agcttttcaa cgccaccgcg gcggaggacc tgatcgtcaa gtccaagc cg gccggcgggg cgcccgctcc gcacgtggcc ggggcccgtga ccaactggac cctgggtgcc ctcaaccac g acaccagat gtcatggac cccaacacca tctgagcaa ggtcatggtc tctccaccg gccacgacgg c cccatgggc gcctccgggg tcaagcgcct ggccaagctg ggctgatcg agcgcgcgcc gggcatgggc g cgctggaca tgaacagcgc cagggacgcg gtgtgcgacc ggacgcggga gatcgtccct ggcattgtca tct gcggcat ggaggtgcc gagctggacg ggtgcccgcg catgggcccc acctttggcg ccatgtttgt gtccgg ggtc aaggccgcg acgtggccct ggcctcgtc cgccgccagc aggaggagga ggggctccgc gccaaag accg cagacacct gcgcagcgcg gcgccgtcca tggccatggc gtgacctgt gatcccaag agggcgg aca tgcgtcgcgt ctccctccg tgcacagaga tctgtccac ggtgtgtgc gatgtttca aaactgttat tta ctaaaa tgactcgatt ctctctaaa aaaaacatga tgtgatagt atctccaca aagcatggcg actcagatgc ctctactga ctgtactcaa gaatacaact ctggagacc ggggccgtt cccatttggc gccgaatgt gatcg ccac tttgtggc catgcctgag gtattctga caaagaagg agcggccaat gcgggtttc aagcagacgg cgctggtggg cctggaatc aaaaggcgt agtacacggg agagagtcgg attgatgat agggatttaa aaa aggccag gcagtgaagt gagatttga cgattgtgc ccaggattg ccttttagg gggggacggc cagggg cgca gttattggg caccgcacat atagatccgc gcaagcaat ggcctcaga gcaggcgca tggtaaga a aggaaggacg agtagcgat gacgtgatac tgcatacga gaacgtacta gtgaggagag tgcttgaac a acatatgat gctgccccct ctctgtcca ac
26	Moltype = DNA length = 4494 Location/Qualifiers 1..4494 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	ttgcatcc tcttgacca gccacacgca taaatttaa cgcaagctgg cccaatatgg tatcgtgtgt gtcaga aca gcacctcgt gcacctctg gctctgaac tcaccgtcac ggtcatctc ctctagcct gccagcgtt gg gcgcagcc atcgggcctg catagagaaa gggtaaaaag caacatatat aaatacagt ggtattcata atactc cgc atgtttgtct tgcgcacgca tgcacacgt cgtgcacgcg ctacacctg gccaggcagc cccttgcgg g gtgcccgc agcgggcacc gcgaggcctg aggcgccacc agcagcatgt tgaattatt tatacgcatg ttttg aggt acagagaaac ccctgtcat tctctctt gttgggtgt cagtgccat tgcagtgtc tggccaccgg ca tatcggtat aagattgcct gcggtgtgc ttcgtccag cgtcatcggg gcgagagttt ataacgcgtg cccacgc cag aagccctcac gcttgcact tggcgaggta caaaaaatac caaacgcgt cctgaagcat tatggcggcg aggacgaggc tgttaaggt gcggtatcgc acctgagtg agagacgctc cgtggtgcg ggcagaacc cc agcgcgcc taaaagcgc gtcctcacgc cccaggcgtg agtgcgcgcc cgcaaacccc tttgcagga actc aaggag gtggtcgtat ccaaggacaa caacggcatc gagctcatt ccaatgactc caacatctc ctctggcg gg cggtgattga ggcaaggggg tctgtctcat tgggtggagg gagatgagtt gtgtgcaggc cgctcaaagg aggttttgt cgcctattgc ccaaggtcat gcacccgtgt gtatccctg cccctccct tcaacaggc cccgcg gact cgcctatca ggggggtcac ttgaactca gcacccaagt gcctgagcag taccatttg tcccgccga agtcaagttc aagaccaagc tcttcaccc gaacgtgcac ttcaagggtga gctgtttgga gacaagggtg gcag ggagag tgctagcaag caggggttga cgggcgttgc tgtgagttt tgatgtatag actcgccgg gaacttga gt ggaagccgac agaccttca tctgcctc tgcctatgca gacgggagag atctgcctg acatttcaa ga acgcctgg agccagcgt ggacgtgca ggtacgattg ttggttag gagaggcgcc ccccttgcg gtcctc

SEQ ID NO.	Source features	Sequence
		ggc actcacgact ctgggaaaa tatccgac tccgcattcc tgtctccaga cgcctgctct gtgcgcgtg tct cgatgaa geattcttt cctcccacat gccgcagtc ggtgtgccag gccatcgagg cgctcttgac agattctgc g ccagactccc cctcaactg cgacgtggg aacctctcc gggcaggcga tgtgcgtggg ttgcctct tgg cacggct ctacaccgtc gactacgcca tgaggcgggtg atggctgtgt cggttgccac ttgtccaga gacggca agt cgtccatcct ctgcgtgtgt ggcgcgacgc tgcagcagtc cctctgcagc agatgagcgt gactttggcc a tttcacgca ctcgagtga cacaatccat tttctaaa gcaaatgact gctgattgac cagatactgt aacgctgatt tcgtccaga tcgcacagat agcgaccatg ttgctgcgtc tgaatatctg gattccgaat tcgacctgg cgtc catcc atgcaacaga tggcgacact tgttacaatt cctgtcacc atcgcatgg agcaggcca ctgattcc cgatcaccca cgcacatctc gctaatagtc attcgttcgt gtcttcgac aatcacaagt gactgtgcat ggacttg gt tgacgatgcg gtatgggttt gcgcgcgtgg ctgcagggtc tgccaaggc aagctaacc agctcctc cc cgacaata ctctgcagg caaagccgt cactgcctt ccagattgcc aataactca attatggcct ctgcatg cc atccatgggt ctgatgaatg gtcacgctc tgtctgacc gtccccagc ctctggcgtc cctgcccc cc accagcc cagccgcgc ggcagtcgt gccaggctg tctcgatgc ccagcccat gccacgcc ttgatg gctt caagtacgat tacggtgttg gattgtgtgt ttgttcgta gtgtcatgg tttagaataa tacattgat ttctg ctca cggcaatctc ggctgtccg caggttcaac cccatttcgg agtctcaggt cagccgcga atgaccgac gctactcaa ggacttcac gacaacgcc aggtgagcta tgttaggac ttgattgaa attgctgc acgcat atc ggcctccgc acagaccca agcaaatgt caagtgcgtt ccgatttggc tccgcagtc gatgttga t cgtcggcgc cggatccgc ggtctgtct gcgcttaca gctgaccaag caccctgac tccgggtacg cgag ctgaga ttcgattaga cataaattga agattaacc cgtagaaaaa tttagtgtc gcgaaactgt gctcattgc aagaaattga tctctcca ctccgaggt cgcctcctc gagcaggcgc ttgtcccg cggcggcgcc tgg ctggggg gacagctgt ctgcgccatg tgtgtacgta gaaggatgaa ttacagtcg tttctgtg acagctgtt gtcatgatt tttttcagac tattttgaa ttttttaga ttcttagga tcatgattt gtctcatgc gacttttag ta gcacctt acacaccac caacatttg ctgccttgc tccctgcgc agatccgaa gccgcgcac aagctg atgg acgagctcaa catccttac gacgacgagg tgggtattgg gcagctagaa cgcatgcgtg ctgtgcgact gggtcgtac attgtcgaa acgtgtggc gaatacgtc gcgtcccg ccctgcacc acccccccc cc accacccc acaatataa cagggtgaca tgggtgtgt caagcagcc gccctggtga cgtccacgt gctgt ccaag gtgctggcgc ccccaacat caagctttc aacgccacc cggcggagga cctgacgtc aagtcaa gc cggccggcgc ggcggccgtc ccgcacgtg ccggggccgt gaccaactgg accctggtgt cctcaacc a cgacaccag atgtcatg accccaac catcctgac aaggtcatg tctctccac cggccacgac g gccccatg gcgcctccg ggtcaagcgc ctggccaagc tgggcctgat cagcgcgc cgggcatg g cgcgctga catgaacgc gccgaggac cgggtgtcga ccgacgcgc gagatcgtc ctggcatg ca tctgcgc atggagtcg ccgagctga cgggtgccc ccatgggccc ccactttg cgcctgtt gtgtcc gggg tcaaggccgc gcacgtggc ctggcctgc tgcgcgcca gcaggaggag gagggtcc gcgcca agac cgcagacac ctgcgcgc cggcgccgt catggccatg gctgacctg gtgaccca agagggc gga catgcgtc gtcctctc ggtgccaga gatctgtcc acggtgtct gcgatgctt caaaactgt attt cactaa atgactcga ttctctcta aaaaaacat gatgttgata gtatccac aaaagcatg cgactcagat gccttactg cactgtact aagaataca ctctggagac ccgggtccc tccccattg gcgccaatg ttgatg

SEQ ID NO.	Source features	Sequence
		cgcc acttttctg gccatgcctg aggtattctg tacaaaagaa ggagcggcca atcggggtt tcaagcagac ggcgcgctgg ggccctggaaa tcaaaaggcg gtagtacacg ggagagagtc ggattgatga tgagggattt aa aaaaggcc aggcagtgag gtgagagttg tacgattgtt gcccaggat tgcctttta gggggggacg gccag ggcg cagttattgg ggcaccgcac atatatgcc gcgcaagcaa atggcctcga gagcaggcgc gatggtt aag aaaggaagga cgagtagcgg atgacgtgat actgcatacg aagaacgtac tagtgaggag agtgcttga a acaacatatg atgctgcccc cctctctgc caac
30	Moltype = DNA length = 573 Location/Qualifiers 1..573 moltype = unassigned DNA note = <i>Prototheca wickerhamii</i> organism = unidentified	tgttgaagaa tgagccggcg acttaaaata aatggcaggc taagagaatt aataactcga aacctaaagc aaag caagtc ttaatagggc gctaatttaa caaaacatta aataaaatct aaagtcattt atttagacc cgaacctgag tg atctaacc atggtcagga tgaacttgg gtgacacaa gtggaagtcc gaaccgaccg atgttgaata atcggc ggat gaactgtggt tagtggtgaa ataccagtcg aactcagagc tagctgttc tccccgaaat gcgttgaggc gcagcaatat atctgtcta tctaggggta aagcactgtt tgggtcggg ctatgaaat ggtaccaat cgtggc aaac tctgaatact agaaatgacg atatattagt gagactatgg gggataagct ccatagtcga gagggaaaca gcccagacca ccagttaagg ccccaaatg ataataaggt ggtaaaggag gtgaaatgc aaatacaacc agg aggttg cttagaagca gccatcctt aaagagtgcg taatagctca ctg
31	Moltype = DNA length = 531 Location/Qualifiers 1..531 moltype = unassigned DNA note = <i>Prototheca moriformis</i> organism = unidentified	ttaaaataaa tggcaggcta agagaattaa taactcga cctaagcgaa agcaagtctt aataggcgcg taattt aaca aaacattaaa taaaatctaa agtcattat ttagaccg aacctgagtg atctaaccat gtcaggatg aaa cttgggt gacaccaagt ggaagtccga accgaccgat gttgaaaaat cggcgatga actgtggtta gtgtgta aat accagtcgaa ctcagagcta gctgttctc cccgaaatgc gttgagcgcg agcaatatat ctcgtctatc tag gggtaaa gcactgttc ggtgcgggct atgaaatgg taccaaatcg tggcaaac tgaatactag aaatgacg at atattagtga gactatggg gataagctcc atagtcgaga gggaaacagc ccagaccacc agttaaggcc c caaaatgat aatgaagtgg taaaggaggt gaaaatgcaa atacaaccag gaggttgct tagaagcagc catc ctttaa a

[0176] The various embodiments disclosed herein have been described and illustrated with reference to a variety of examples, but it should be understood that the features of the embodiments are susceptible to modification, alteration, changes or substitution without departing significantly from the spirit of the invention. One skilled in the art will recognize that numerous variations and modifications may be made to the examples described above without departing from the scope of the present disclosure.

EMBODIMENTS

[0177] The following provides various embodiments exemplifying the compositions, methods of making, and uses of the methods and compositions described herein:

[0178] Embodiment [1]. An oil derived from a microalgal cell comprising a triacylglyceride (TAG) component, wherein the TAG component has a fatty acid content of at least 30% C16:0 fatty acids and at least 40% C18:1 fatty acids, and wherein the oil is produced through classical strain improvement of the microalgal cell.

- [0179] Embodiment [2]. The oil of Embodiment [1], wherein the fatty acid content of the TAG component comprises at least 35% C16:0 fatty acids.
- [0180] Embodiment [3]. The oil of Embodiment [1], wherein the fatty acid content of the TAG component comprises at least 40% C16:0 fatty acids.
- [0181] Embodiment [4]. The oil of any one of Embodiments [1] to [3], wherein the fatty acid content of the TAG component further comprises less than 8% C18:2 fatty acids.
- [0182] Embodiment [5]. The oil of any one of Embodiments [1] to [4], wherein the microalgal cell is of the genus *Prototheca*.
- [0183] Embodiment [6]. The oil of any one of Embodiments [1] to [5], wherein the fatty acid content of the TAG component comprises at least 45% C18:1 fatty acids.
- [0184] Embodiment [7]. The oil of any one of Embodiments [1] to [6], further comprising less than 1% monoacylglycerols (MAG).
- [0185] Embodiment [8]. The oil of any one of Embodiments [1] to [7], further comprising less than 1% of diacylglycerols (DAG).
- [0186] Embodiment [9]. The oil of any one of Embodiments [1] to [8], further comprising less than 1% tripalmitin.
- [0187] Embodiment [10]. The oil of any one of Embodiments [1] to [9], wherein the oil comprises ergosterol.
- [0188] Embodiment [11]. An unpurified oil derived from a *Prototheca* microalgal cell strain obtained by classical strain improvement, wherein the unpurified oil comprises a triacylglyceride (TAG) component, wherein the TAG component has a fatty acid content comprising 30% or more C16:0 fatty acid and less than 1% tripalmitin.
- [0189] Embodiment [12]. The oil of Embodiment [11], wherein the oil is produced from classical strain improvement of the microalgal cell.
- [0190] Embodiment [13]. The oil of any one of Embodiments [10] to [12], wherein the fatty acid content of the TAG component further comprises at least 40% C18:1 fatty acids.
- [0191] Embodiment [14]. The oil of any one of Embodiments [10] to [13], wherein the oil further comprises ergosterol.
- [0192] Embodiment [15]. The oil of any one of Embodiments [10] to [14], wherein the microalgal cell does not comprise an exogenous gene.
- [0193] Embodiment [16]. The oil of any one of Embodiments [10] to [15], further comprising less than 1% of monoacylglycerols (MAG).

[0194] Embodiment [17]. The oil of any one of Embodiments [10] to [16], further comprising less than 1% of diacylglycerols (DAG).

[0195] Embodiment [18]. The oil of any one of Embodiments [10] to [14], wherein the fatty acid content of the TAG component further comprises less than 8% C18:2 fatty acids.

[0196] Embodiment [19]. An unpurified microalgal oil comprising: a triacylglyceride (TAG) component having a fatty acid content comprising at least 30% C16:0 fatty acids, less than about 1% diacylglycerols, and less than about 1% monoacylglycerols.

[0197] Embodiment [20]. The oil of any one of Embodiment [19], wherein the oil is derived from a *Prototheca* microalgal cell.

[0198] Embodiment [21]. The oil of any one of Embodiments [19] or [20], wherein the oil is derived from an oleaginous microorganism.

[0199] Embodiment [22]. The oil of any one of Embodiments [19] or [21], wherein the microalgal cell does not comprise an exogenous gene.

[0200] Embodiment [1A]. An isolated microalgal cell, wherein the microalgal cell is a *Prototheca* *sp.* cell obtained through classical microalgal cell improvement, wherein the microalgal cell produces an oil characterized as comprising:

- (a) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 17% of a triacylglyceride (TAG component), or of at least 23%, or of at least 29%, or of at least 32%, or of at least 35%, or at least 45%, and
- (b) a tripalmitin (PPP) content less than 0.3% of the TAG component, or less than about 0.6%, or less than about 1.0%, or less than at least 1.2%, or less than about 1.5%, or less than about 2.7% of the TAG component.

[0201] Embodiment [2A]. The isolated microalgal cell of Embodiment [1A], wherein the POP content is at least 45% and the PPP content is less than about 2.7% of the TAG component.

[0202] Embodiment [3A]. The isolated microalgal cell of Embodiment [1A] or Embodiment [2A], wherein the microalgal cell oil comprises:

- (a) the 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least about 50%, and
- (b) the tripalmitin (PPP) content of less than about 2.7% of the TAG component.

[0203] Embodiment [4A]. A microalgal oil comprising a triacylglyceride (TAG) component, wherein the TAG component comprises: (a) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 17% of the TAG component, or at least 23%, or at least 29%, or at

least 32%, or at least 35%, or at least 45% of the TAG component, (b) a tripalmitin (PPP) content less than 0.3% of the TAG component, or less than about 0.6%, or less than about 1.0%, or less than at least 1.2%, or less than about 1.5%, or less than about 2.7% of the TAG component, and wherein the microalgal oil is produced by a microalgal *Prototheca* cell obtained through classical strain improvement.

[0204] Embodiment [5A]. The microalgal oil of any one of Embodiments [4A] to [5A] or the microalgal oil produced by the isolated microalgal cell of any one of Embodiments [1A] to [3A], wherein the *Prototheca sp.* cell has an exogenous nucleic acid introduced recombinantly.

[0205] Embodiment [6A]. The microalgal oil of any one of Embodiments [4A] and [5A], wherein the TAG component comprises: (a) the 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 45%, or at least 50% or at least 55%, or at least 60%, and (b) the tripalmitin (PPP) content of less than about 2.7% of the TAG.

[0206] Embodiment [7A]. The microalgal oil of Embodiment [5A], wherein the microalgal oil has (a) the POP content of at least 35% and the PPP content of less than about 1.5%, or (b) the POP content of at least 45% and the PPP content of less than about 2.7%.

[0207] Embodiment [8A]. The microalgal oil of any one Embodiments [4A] to [7A], wherein the oil further comprises at least about 30% C16:0 fatty acids, or at least about 35% C16:0 fatty acids, or at least about 40% C16:0 fatty acids, or at least about 45% C16:0 fatty acids, or at least about 50% C16:0 fatty acids.

[0208] Embodiment [9A]. A method of obtaining a microalgal cell that produces the oil of any of Embodiments [4A] to [8A], the method comprising the steps of:

- (a) mutagenizing microalgal cells with at least one or more agents, which promote errors in DNA excision repair known (e.g., ultraviolet light (e.g., 254 nm), 4-nitroquinoline 1-oxide (4-NQO), methyl methanesulfonate (MMS), Ni, Ar, or polycyclic aromatic hydrocarbons while culturing the microalgal cells at about 28 °C to about 38 °C in a vegetative or a lipid production media;
- (b) enriching the mutagenized microalgal cells by exposing the mutagenized cells to physical conditions (e.g., an elevated temperature of about 50 to about 68°C or by using density fractionation on sucrose or percoll gradients to isolate specific subpopulations based on density) or exposure to a chemical compound (e.g., herbicides, fungicides, or inhibitors of lipid biosynthesis, such as tebutam or cofenstrole);

- (c) growing the mutagenized, enriched microalgal cells in a lipid production medium;
- (d) harvesting the microalgal oil from the microalgal cells and interrogating the microalgal oil using a FAME analysis and a triacylglyceride (TAG) analysis to determine the component of the screened microalgal cells; and

wherein the TAG component of the screened microalgal cells comprises:

- (i) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 17% of the TAG component, or of at least 23%, or of at least 29%, or of at least 32%, or of at least 35%, or at least 45%, and
- (ii) a tripalmitin (PPP) content less than 0.3%, or less than about 0.6%, or less than about 1.0%, or less than at least 1.2%, or less than about 1.5%, or less than about 2.7% of the TAG component.

[0209] Embodiment [10A]. The method of Embodiment [9A], wherein the microalgal cell is a *Prototheca moriformis*.

[0210] Embodiment [11A]. An isolated microalgal cell obtained by the method of Embodiments [9A] to [10A].

[0211] Embodiment [12A]. The microalgal cell of Embodiments [10A] to [11A], wherein the oil of the microalgal cell further comprises at least about 30% C16:0 fatty acids, or at least about 35% C16:0 fatty acids, or at least about 40% C16:0 fatty acids, or at least about 45% C16:0 fatty acids, or at least about 50% C16:0 fatty acids.

[0212] Embodiment [13A]. A method of obtaining a microalgal oil of any of Embodiments [1A] to [3A] comprising the steps of:

- (a) mutagenizing cells with agents that introduce mutations through promoting errors in DNA excision repair known to include, but not limited to exposure to UV light (254 nm), 4-nitroquinoline 1-oxide (4-NQO), methyl methanesulfonate (MMS), Ni, Ar, or polyaromatic hydrocarbons, culturing the exposed microalgal cells at about 28 °C to about 38 °C in vegetative or lipid production media for a sufficient time;
- (b) enriching said cell populations using physical conditions (e.g., elevated temperatures of about 50, about 60, or up to about 68 °C, or fractionation on sucrose or percoll gradients to isolate specific subpopulations based on density) or chemical compound exposure (e.g., herbicides, fungicides or inhibitors of lipid biosynthesis, such as tebutam or cofenstrole);

- (c) plating the mutagenized and enriched cell populations on a solid vegetative medium;
- (d) selecting and growing selected cells in a lipid production medium;
- (e) harvesting the microalgal cells from said lipid production medium; and
- (f) obtaining and interrogating a microalgal oil from the selected microalgal cells using FAME analysis and determining if the selected cells have a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 17% of the TAG component, or of at least 23%, or of at least 29%, or of at least 32%, or of at least 35%, or at least 45%; and a tripalmitin (PPP) content less than 0.3%, or less than about 0.6%, or less than about 1.0%, or less than at least 1.2%, or less than about 1.5%, or less than about 2.7% of the TAG component.

[0213] Embodiment [14A]. An isolated microalgal cell, wherein the microalgal cell is a *Prototheca sp.* that produces a microalgal oil, said oil is characterized as comprising:

- (a) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 60%, and
- (b) a tripalmitin (PPP) content of less than about 2% of the triacylglyceride (TAG) component, and the *Prototheca sp.* producing said microalgal oil was obtained through classical microalgal cell improvement of an isolated *Prototheca sp.*

[0214] Embodiment [15A]. A microalgal oil comprising a triacylglyceride (TAG) component, wherein the TAG component comprises:

- (a) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 60% of the TAG component, and
- (b) a tripalmitin (PPP) content of less than about 2% of the TAG component, and wherein the microalgal oil is produced by a microalgal *Prototheca sp.* cell obtained through a classical microalgal strain improvement method of the isolated *Prototheca sp.*

CLAIMS

WHAT IS CLAIMED IS:

1. An isolated microalgal cell, wherein the microalgal cell is a *Prototheca sp.* cell obtained through classical microalgal cell improvement, wherein the microalgal cell produces an oil characterized as comprising:
 - (a) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 17% of a TAG (triacylglyceride) component, or of at least 23%, or of at least 29%, or of at least 32%, or of at least 35%, or at least 45% of the TAG component, and
 - (b) a tripalmitin (PPP) content less than 0.3%, or less than about 0.6%, or less than about 1.0%, or less than at least 1.2%, or less than about 1.5%, or less than about 2.7% of the TAG component.
2. The isolated microalgal cell of claim 1, wherein the POP content is at least 45% and the PPP content is less than about 2.7% of the TAG component.
3. The isolated microalgal cell of claim 1 or claim 2, wherein the microalgal cell oil comprises:
 - (a) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 50%, and
 - (b) a tripalmitin (PPP) content of less than about 2.7% of the TAG component.
4. A microalgal oil comprising a triacylglyceride (TAG) component, wherein the triacylglyceride (TAG) component comprises:
 - (a) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 17% of the TAG component, or at least 23%, or at least 29%, or at least 32%, or at least 35%, or at least 45% of the TAG component,
 - (b) a tripalmitin (PPP) content of less than 0.3% , or less than about 0.6%, or less than about 1.0%, or less than at least 1.2%, or less than about 1.5%, or less than about 2.7% of the TAG component, andwherein the microalgal oil is produced by a microalgal *Prototheca sp.* cell obtained through classical strain improvement.

5. The microalgal oil of claim 4 or produced by the isolated microalgal cell of any one of claims 1 to 3, wherein the *Prototheca* cell has an exogenous nucleic acid introduced recombinantly.
6. The microalgal oil of claim 4 or claim 5 or of the isolated microalgal cell of any of claims 1 to 3, wherein the TAG component comprises:
 - (a) the 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 45% of the TAG component, and
 - (b) the tripalmitin (PPP) content of less than about 2.7% of the TAG component.
7. The microalgal oil of claim 5, wherein the microalgal oil has (a) the POP content of at least 35% and the PPP content of less than about 1.5%, or (b) the POP content of at least 45% and the PPP content of less than about 2.7%.
8. The microalgal oil of any one of claims 4 to 7, wherein the oil comprises at least about 30% C16:0 fatty acids, or at least about 35% C16:0 fatty acids, or at least about 40% C16:0 fatty acids, or at least about 45% C16:0 fatty acids, or at least about 50% C16:0 fatty acids.
9. A method of obtaining a microalgal cell that produces a microalgal oil as defined by any of claims 4 to 8 comprising the steps of:
 - (a) mutagenizing microalgal cells with at least one or more agents, which promote errors in DNA excision repair known (e.g., ultraviolet light (e.g., 254 nm), 4-nitroquinoline 1-oxide (4-NQO), methyl methanesulfonate (MMS), Ni, Ar, or polycyclic aromatic hydrocarbons while culturing the microalgal cells at about 28 °C to about 38 °C in a vegetative or a lipid production media;
 - (b) enriching the mutagenized microalgal cells by exposing the mutagenized cells to physical conditions (e.g., an elevated temperature of about 50 to about 68°C, or by using density fractionation on sucrose or percoll gradients to isolate specific subpopulations based on density) or a chemical compound exposure (e.g., herbicides, fungicides, or inhibitors of lipid biosynthesis, such as tebutam or cofenstrole);
 - (c) growing the mutagenized, enriched microalgal cells in a lipid production medium;

(c) harvesting the microalgal oil from the microalgal cells and interrogating the microalgal oil using a FAME analysis and a triacylglyceride (TAG) analysis to determine the component of the screened microalgal cells;

wherein the TAG component of the screened microalgal cells comprises:

- (i) a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 17% of the TAG component, or of at least 23%, or of at least 29%, or of at least 32%, or of at least 35%, or at least 45%, and
- (ii) a tripalmitin (PPP) content less than 0.3%, or less than about 0.6%, or less than about 1.0%, or less than at least 1.2%, or less than about 1.5%, or less than about 2.7% of the TAG component.

10. The method of claim 9, wherein the microalgal cell is a *Prototheca* sp.

11. An isolated microalgal cell obtained by the method of claim 9 or 10.

12. A microalgal oil obtained from the microalgal cell of any of claims 1 to 3 and 10, wherein the oil of the microalgal cell further comprises at least about 30% C16:0 fatty acids, or at least about 35% C16:0 fatty acids, or at least about 40% C16:0 fatty acids, or at least about 45% C16:0 fatty acids, or at least about 50% C16:0 fatty acids.

13. A method of obtaining a microalgal cell that produces a microalgal oil of any of claims 1 to 3 comprising the steps of:

- (a) mutagenizing cells with agents that introduce mutations through promoting errors in DNA excision repair known to include, but not limited to, UV light (254nm), 4-nitroquinoline 1-oxide (4-NQO), methyl methanesulfonate (MMS), Ni, Ar, or polycyclic aromatic hydrocarbons culturing the microalgal cells at about 28 °C to about 38 °C in vegetative or lipid production media and
- (b) Enriching said cell populations using a physical change (e.g., elevated temperatures of about 50, about 60, or up to about 68 °C, or fractionation on sucrose or percoll gradients to isolate specific subpopulations based on density) or chemical compound (e.g., herbicides, fungicides, or inhibitors of lipid biosynthesis such as tebutam or cofenstrole),

- (c) plating the mutagenized and enriched cell populations to a solid vegetative medium,
- (d) growing selected clones from the plated cells in a lipid production medium,
- (e) harvesting the selected clones from said lipid production medium, and
- (f) obtaining and characterizing a microalgal oil from the selected cells using FAME analysis and TAG analysis and determining if the selected cells have a 1, 3-dipalmitoyl-2-oleoyl-glycerol (POP) content of at least 17% of the TAG component, or of at least 23%, or of at least 29%, or of at least 32%, or of at least 35%, or at least 45%; and a tripalmitin (PPP) content less than 0.3%, or less than about 0.6%, or less than about 1.0%, or less than at least 1.2%, or less than about 1.5%, or less than about 2.7% of the TAG component.