



(51) International Patent Classification:

C12N 1/16 (2006.01) C12P 7/64 (2006.01)
C12N 15/87 (2006.01)

(21) International Application Number:

PCT/US2015/033211

(22) International Filing Date:

29 May 2015 (29.05.2015)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/004,506 29 May 2014 (29.05.2014) US
62/090,169 10 December 2014 (10.12.2014) US

(71) Applicant: **NOVOGY, INC.** [US/US]; 85 Bolton Street,
Cambridge, MA 02140 (US).

(72) Inventors: **BREVENOVA, Elena, E.**; 33 Hamilton Road,
Belmont, MA 02478 (US). **SHAW, Arthur, J., IV**; 36 Pal-
trey Road, Belmont, MA 02478 (US). **VAN DIJKEN, Jo-
hannes, P.**; Spotvogellaan 38, Leidschendam, 2261 CP
(NL).

(74) Agents: **GORDON, Dana, M.** et al.; Foley Hoag LLP,
Seaport West, 155 Seaport Blvd., Boston, MA 02210-2600
(US).

(81) Designated States (*unless otherwise indicated, for every
kind of national protection available*): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,
MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,
PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC,
SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: INCREASING LIPID PRODUCTION AND OPTIMIZING LIPID COMPOSITION

(57) Abstract: Disclosed are nucleotide sequences and corresponding amino acid sequences of *Arxula adenivorans* genes that can be utilized to manipulate the lipid content and/or composition of a cell. Methods and compositions for utilizing this information are disclosed to increase the lipid content or modify the lipid composition of a cell by either increasing or decreasing the activity of certain genetic targets.



***Increasing lipid production and optimizing lipid
composition***

PRIORITY CLAIM

This application claims the benefit of priority to U.S. Provisional Patent Application No. 62/004,506, filed May 29, 2014; and U.S. Provisional Patent Application No. 62/090,169, filed December 10, 2014.

5

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on May 18, 2015, is named NGX-03325_SL.txt and is 657,774 bytes in size.

10

BACKGROUND

Lipids are indispensable ingredients in the food and cosmetics industries, and they are important precursors in the biodiesel and biochemical industries. Many oleaginous microorganisms produce lipids, including the well-characterized yeast *Yarrowia lipolytica*.

15

Arxula adeninivorans (*Blastobotrys adeninivorans*) is an oleaginous yeast that is less well characterized than *Y. lipolytica*. In fact, currently, there are no industrial *A. adeninivorans*-based processes, perhaps because *Arxula adeninivorans* is a yeast with a number of unusual characteristics. It is dimorphic and can utilize a wide range of substrates as sources of carbon, nitrogen, phosphorous, and energy, including starch, xylose, acetic acid, ethanol, arabinose, cellobiose, phytic acid, adenine, xanthine, uric acid, putrescine, and n-alkylamines. *A. adeninivorans* is also nitrate-assimilating and thermo- and osmo-tolerant. A distinctive feature is its ability to grow efficiently at 37°C, and a temperature-dependent dimorphism with mycelial structures formed at temperatures above 42°C.

20

25

SUMMARY

Disclosed are nucleotide sequences and corresponding amino acid sequences of *Arxula adeninivorans* genes that can be utilized to manipulate the lipid content and composition of a cell. Methods and compositions for utilizing this information are disclosed to increase the lipid content or modify the lipid composition of a cell by either increasing or decreasing the activity of certain genetic targets.

30

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 depicts a map of the pNC363 construct used to overexpress the NG167 gene (AaDgal) in *A. adeninivorans* strain NS252 (ATCC 76597). Vector pNC363 was linearized by a PmeI/AscI restriction digest before transformation. “2u ori” denotes the *S. cerevisiae* origin of replication from the 2 μ m circle plasmid; “pMB1 ori” denotes the *E. coli* pMB1 origin of replication from the pBR322 plasmid; “AmpR” denotes the *bla* gene used as a marker for selection with ampicillin; “Sc URA3” denotes the *S. cerevisiae* URA3 auxotrophic marker for selection in yeast; PR26 PGK1p denotes the *A. adeninivorans* PGK1 promoter - 524 to -1; “NG3 NatR” denotes the *Streptomyces noursei* NatI gene used as a marker for selection with nourseothricin; “ScFBA1t” denotes the *S. cerevisiae* FBA1 terminator 205 bp after stop; “PR25 AaADH1p” denotes the *A. adeninivorans* ADH1 promoter - 877 to -1; “NG167 AaDGA1” denotes the *A. adeninivorans* DGA1 gene ORF (NG167); “TER16 CYC1t” denotes the *A. adeninivorans* CYC1 terminator 301 bp after stop codon.

Figure 2 depicts a map of the deletion cassette used to delete the NG166 gene in *A. adeninivorans* strain NS252 (ATCC 76597). “NG166 5' Flank” denotes the 500 base pairs upstream of the NG166 open reading frame used as a recombination flank for NG166 deletion; “PR26 PGK1p” denotes the *A. adeninivorans* PGK1 promoter - 524 to -1; “NG3 NatR” denotes the *Streptomyces noursei* NatI gene used as marker for selection with nourseothricin; “ScFBA1t” denotes the *S. cerevisiae* FBA1 terminator 205 bp after stop; “NG166 3' Flank” denotes the 500 base pairs downstream of the NG166 open reading frame used as recombination flank for NG166 deletion.

Figure 3 depicts a map of the pNC243 construct used to overexpress the diacylglycerol acyltransferase DGA1 gene NG66. Vector pNC243 is linearized by a PacI/NotI restriction digest before transformation. “2u ori” denotes the *S. cerevisiae* origin of replication from the 2 μ m circle plasmid; “pMB1 ori” denotes the *E. coli* pMB1 origin of replication from the pBR322 plasmid; “AmpR” denotes the *bla* gene used as a marker for selection with ampicillin; “PR2” denotes the *Y. lipolytica* GPD1 promoter -931 to -1; “NG66” denotes the native *Rhodospiridium toruloides* DGA1 cDNA synthesized by GenScript; “TER1” denotes the *Y. lipolytica* CYC1 terminator 300 base pairs after stop; “PR22” denotes the *S. cerevisiae* TEF1 promoter -412 to -1; “NG3” denotes the *Streptomyces noursei* NatI gene used as a marker for selection with nourseothricin;

“TER2” denotes the *S. cerevisiae* CYC1 terminator 275 base pairs after stop; and “Sc URA3” denotes the *S. cerevisiae* URA3 auxotrophic marker for selection in yeast.

Figure 4 depicts the percentage of C16 and C18 fatty acids that are palmitate, palmitoleate, stearate, oleate, and linoleate for *A. adenivorans* strain NS252 after

5 transforming the strain with nucleic acids comprising a *DGAT2* gene from various species. The *DGAT2* gene encodes the DGA1 protein.

Figure 5 depicts a map of the pNC104 construct used to overexpress the NG15 gene (YIDga1). Vector pNC243 is linearized by a *PacI*/*NotI* restriction digest before

transformation. “2u ori” denotes the *S. cerevisiae* origin of replication from the 2 μ m circle plasmid; “pMB1 ori” denotes the *E. coli* pMB1 origin of replication from the pBR322 plasmid; “AmpR” denotes the *bla* gene used as a marker for selection with ampicillin; “ScGPM1p” denotes the *Y. lipolytica* GPD1 promoter -764 to -1; “hygR” denotes the *Escherichia coli* hph gene expression cassette used as marker for selection with hygromycin B; “ScGPM1t” denotes the *S. cerevisiae* GPD1 terminator 406 bp after stop codon; “ARS68” and “CEN1-1” denote *Y. lipolytica* chromosomal origins of replication; “YITEF1p” denotes the *Y. lipolytica* TEF promoter -406 to +125; “YIDGA1” denotes the *Y. lipolytica* DGA1 gene ORF (NG15); “YICYC1t” denotes the *Y. lipolytica* CYC1 terminator 300 base pairs after stop; “ScTEF1p” denotes the *S. cerevisiae* TEF1 promoter -412 to -1; “NAT” denotes the *Streptomyces noursei* Nat1 gene used as a marker for selection with nourseothricin; “ScCYC1t” denotes the *S. cerevisiae* CYC1 terminator 275 base pairs after stop; and “URA3p-ScURA3-URA3t” denotes the *S. cerevisiae* URA3 auxotrophic marker for selection in yeast.

Figure 6 depicts a map of the pNC281 construct used to overexpress the NG16 gene. Vector pNC281 is linearized by a *PacI*/*AscI* restriction digest before transformation.

25 “2u ori” denotes the *S. cerevisiae* origin of replication from the 2 μ m circle plasmid; “pMB1 ori” denotes the *E. coli* pMB1 origin of replication from the pBR322 plasmid; “AmpR” denotes the *bla* gene used as a marker for selection with ampicillin; “PR3” denotes the *Y. lipolytica* TEF1 promoter -406 to +125; “NG16” denotes the native *Y. lipolytica* DGA2 gene amplified from gDNA of *Y. lipolytica* strain NRRL YB-437; “TER1” denotes the *Y. lipolytica* CYC1 terminator 300 bp after stop; “PR1” denotes the *Y. lipolytica* TEF1 promoter -406 to -1; “NG76” denotes the *Streptoalloteichus hindustanus* BLE gene used as marker for selection with Zeocin; “TER7” denotes the *Y. lipolytica* TEF1 terminator 400 bp after stop; and “Sc URA3” denotes the *S. cerevisiae* URA3 auxotrophic

marker for selection in yeast.

Figure 7 depicts the relative lipid contents of wild type *Arxula adeninivorans* (strain NS252), *Arxula adeninivorans* strain NS458, which comprises a *TGL3* knockout, *Arxula adeninivorans* strain NS481, which comprises a *TGL4* knockout, and *Arxula adeninivorans* strain NS478, which comprises a *Δ12* desaturase knockout, as assessed by a fluorescence-based lipid assay. The *TGL3* and *TGL4* knockouts each comprise a higher lipid content than the wild type strain.

Figure 8 depicts the percentage of C16 and C18 fatty acids that are palmitate, palmitoleate, stearate, oleate, and linoleate for *Y. lipolytica* strains expressing either the *A. adeninivorans* or *Y. lipolytica* *Δ9* desaturase gene.

Figure 9 depicts lipid accumulation as measured using a fluorescence-based assay for *Y. lipolytica* strains NS18 (wild type), NS563 (deletion of *Y. lipolytica* *SCT1*) and NS804 (*A. adeninivorans* *SCT1* overexpression in the NS563 background). Strains were grown in duplicate in 96-well plates.

Figure 10 depicts the percentage of C16 and C18 fatty acids that are palmitate, palmitoleate, stearate, oleate, and linoleate for *A. adeninivorans* strains carrying deletions of the native *MFE1*, *TGL3*, *Δ12*, *PEX10*, *POX*, and *TGL4* genes.

Figure 11 consists of four panels, labeled panels (A), (B), (C), and (D). The panels depict results from a 96-well plate lipid assay for *Y. lipolytica* (panels (A) and (B)) and *A. adeninivorans* (panels (C) and (D)) transformants with randomly integrated *Arxula adeninivorans* genes DGA1 (NG167), DGA2 (NG168), GPD1 (NG170), SLC1 (NG177), SCT1 (NG183), ZWF1 (NG211), GND1 (NG212), and PAH1 (NG218).

Figure 12 depicts the relative proportion of C16 and C18 fatty acids for *Arxula adeninivorans* strain NS554 (*Δ12* knockout in wild type *Arxula adeninivorans* strain NS252) and NS631 (*Arxula adeninivorans* *ELO1* overexpression in a NS554 background). The expression of *Arxula adeninivorans* *ELO1* in a *Δ12* knockout increased the fraction of C18 fatty acids.

DETAILED DESCRIPTION

Overview

Many strategies exist to increase or modify the lipid production or content of oleaginous organisms by genetic engineering. However, most oleaginous organisms grow relatively slowly at mesophilic temperatures, and many require additional engineering to

utilize pentose sugars. The oleaginous yeast *Arxula adeninivorans* has the potential to generate lipids from various substrates with fast growth at high temperature, which may result in reduced biorefinery operating and capital expenses as well as a reduced risk of microbial contamination.

5 The lipid yield and lipid composition of an oleaginous organism can be modified by up-regulating and/or down-regulating or deleting genes that are involved in or regulate various lipid pathways. In addition to the manipulation of native genes, exogenous genes can be introduced into the host genome to modulate lipid production and composition.

 This invention describes compositions and methods to produce lipids by engineering
10 *A. adeninivorans*, and these compositions and methods are generally applicable to other organisms to increase lipid production levels and to change an organism's lipid-composition profile. The method of increasing lipid production levels involves up-regulating genetic targets that have a positive effect on lipid content and down-regulating targets that have a negative effect on lipid content. Similarly, the method of changing a
15 lipid-composition profile involves up-regulating genetic targets that have a positive effect on a particular fatty acid of interest and down-regulating targets that have a negative effect on the particular fatty acid. For example, genetic targets can be modified to increase the palmitic acid, palmitoleic acid, stearic acid, oleic acid, or linoleic acid content of a cell. The genetic targets can be modified to increase the concentration of lipids,
20 triacylglycerides, fatty alcohols, fatty acids, alkanes, alkenes, isoprenoids, isoprene, squalene, farnasene, alcohols, isopropanol, n-propanol, n-butanol, isobutanol, 2-butanol, butadiene, diols, 1,3 propanediol, 1,4 propanediol, succinic acid, adipic acid, nylon precursors, citric acid, malic acid, polyols, or erythritol in the cell.

Definitions

25 The articles "a" and "an" are used herein to refer to one or to more than one (*i.e.*, to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

 The term "activity" refers to the total capacity of a cell to perform a function. For example, a genetic modification that decreases the activity of a triacylglycerol lipase in a
30 cell may reduce the amount of triacylglycerol lipase in a cell or reduce the efficiency of triacylglycerol lipase. A triacylglycerol lipase knockout reduces the amount of triacylglycerol lipase in the cell. Alternatively, a mutation to a triacylglycerol lipase gene may reduce the efficiency of its triacylglycerol lipase protein product with little effect on

the amount of cellular triacylglycerol lipase. Mutations that reduce the efficiency of triacylglycerol lipase may affect the active site, for example, by changing one or more active site residues; they may impair the enzyme's kinetics, for example, by sterically blocking substrates or products; they may affect protein folding or dynamics, for example, by reducing the proportion of properly-folded enzymes; they may affect protein localization, for example, by preventing the lipase from localizing to lipid particles; or they may affect protein degradation, for example, by adding one or more protein cleavage sites or by adding one or more residues or amino acid sequences that target the protein for proteolysis. These mutations affect coding regions. Mutations that decrease triacylglycerol lipase activity may instead affect the transcription or translation of the gene. For example, mutation to a triacylglycerol lipase enhancer or promoter can reduce triacylglycerol lipase activity by reducing its expression. Mutating or deleting the non-coding portions of a triacylglycerol lipase gene, such as its introns, may also reduce transcription or translation. Additionally, mutations to the upstream regulators of a triacylglycerol lipase may affect triacylglycerol lipase activity; for example, the over-expression of one or more repressors may decrease triacylglycerol lipase activity, and a knockout or mutation of one or more activators may similarly decrease triacylglycerol lipase activity. A genetic modification that increases the activity of a diacylglycerol acyltransferase in a cell may increase the amount of triacylglycerol acyltransferase in a cell or increase the efficiency of diacylglycerol acyltransferase. For example, the genetic modification may simply insert an additional copy of diacylglycerol acyltransferase into the cell such that the additional copy is transcribed and translated into additional functional diacylglycerol acyltransferase. The added diacylglycerol acyltransferase gene can be native to the host organism or from a different organism. Alternatively, mutating or deleting the non-coding portions of a native diacylglycerol acyltransferase gene, such as its introns, may also increase translation. A native diacylglycerol acyltransferase gene can be altered by adding a new promoter that causes more transcription. Similarly, enhancers may be added to the diacylglycerol acyltransferase gene that increase transcription, or silencers may be mutated or deleted from the diacylglycerol acyltransferase gene to increase transcription. Mutations to a native gene's coding region might also increase diacylglycerol acyltransferase activity, for example, by preventing interactions with inhibitory proteins or molecules. The over-expression of one or more activators may increase diacylglycerol acyltransferase activity by increasing its expression, and a knockout or mutation of one or more repressors may

similarly increase diacylglycerol acyltransferase activity.

The term “*biologically-active portion*” refers to an amino acid sequence that is less than a full-length amino acid sequence, but exhibits at least one activity of the full length sequence. For example, a biologically-active portion of a diacylglycerol acyltransferase
5 may refer to one or more domains of DGA1 or DGA2 having biological activity for converting acyl-CoA and diacylglycerol to triacylglycerol. Biologically-active portions of a DGA1 protein include peptides or polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the DGA1 protein, *e.g.*, the amino acid sequence as set forth in SEQ ID NOs: 15, 117, 119, 121, 123, 125, 127, 129,
10 131 or 133, which include fewer amino acids than the full length DGA1, and exhibit at least one activity of a DGA1 protein. Typically, biologically active portions comprise a domain or motif having the catalytic activity of converting acyl-CoA and diacylglycerol to triacylglycerol. A biologically active portion of a DGA1 protein can be a polypeptide which is, for example, 262 amino acids in length.

15 The term “*DGAT1*” refers to a gene that encodes a type 1 diacylglycerol acyltransferase protein, such as a gene that encodes a DGA2 protein.

The term “*DGAT2*” refers to a gene that encodes a type 2 diacylglycerol acyltransferase protein, such as a gene that encodes a DGA1 protein.

20 “*Diacylglyceride*,” “*diacylglycerol*,” and “*diglyceride*,” are esters comprised of glycerol and two fatty acids.

The terms “*diacylglycerol acyltransferase*” and “*DGA*” refer to any protein that catalyzes the formation of triacylglycerides from diacylglycerol. Diacylglycerol acyltransferases include type 1 diacylglycerol acyltransferases (DGA2), type 2 diacylglycerol acyltransferases (DGA1), and all homologs that catalyze the above-
25 mentioned reaction.

The terms “*diacylglycerol acyltransferase, type 1*” and “*type 1 diacylglycerol acyltransferases*” refer to DGA2 and DGA2 orthologs.

The terms “*diacylglycerol acyltransferase, type 2*” and “*type 2 diacylglycerol acyltransferases*” refer to DGA1 and DGA1 orthologs.

30 The term “*domain*” refers to a part of the amino acid sequence of a protein that is able to fold into a stable three-dimensional structure independent of the rest of the protein.

The term “*drug*” refers to any molecule that inhibits cell growth or proliferation, thereby providing a selective advantage to cells that contain a gene that confers resistance

to the drug. Drugs include antibiotics, antimicrobials, toxins, and pesticides.

"Dry weight" and "dry cell weight" mean weight determined in the relative absence of water. For example, reference to oleaginous cells as comprising a specified percentage of a particular component by dry weight means that the percentage is calculated based on the weight of the cell after substantially all water has been removed.

The term "*encode*" refers to nucleic acids that comprise a coding region, portion of a coding region, or compliments thereof. Both DNA and RNA may encode a gene. Both DNA and RNA may encode a protein.

The term "*exogenous*" refers to anything that is introduced into a cell. An "*exogenous nucleic acid*" is a nucleic acid that entered a cell through the cell membrane. An exogenous nucleic acid may contain a nucleotide sequence that exists in the native genome of a cell and/or nucleotide sequences that did not previously exist in its genome. Exogenous nucleic acids include exogenous genes. An "*exogenous gene*" is a nucleic acid that codes for the expression of an RNA and/or protein that has been introduced into a cell (*e.g.*, by transformation/transfection), and is also referred to as a "transgene." A cell comprising an exogenous gene may be referred to as a recombinant cell, into which additional exogenous gene(s) may be introduced. The exogenous gene may be from the same or different species relative to the cell being transformed. Thus, an exogenous gene can include a native gene that occupies a different location in the genome of the cell or is under different control, relative to the endogenous copy of the gene. An exogenous gene may be present in more than one copy in the cell. An exogenous gene may be maintained in a cell as an insertion into the genome (nuclear or plastid) or as an episomal molecule.

The term "*expression*" refers to the amount of a nucleic acid or amino acid sequence (*e.g.*, peptide, polypeptide, or protein) in a cell. The increased expression of a gene refers to the increased transcription of that gene. The increased expression of an amino acid sequence, peptide, polypeptide, or protein refers to the increased translation of a nucleic acid encoding the amino acid sequence, peptide, polypeptide, or protein.

The term "*gene*," as used herein, may encompass genomic sequences that contain introns, particularly polynucleotide sequences encoding polypeptide sequences involved in a specific activity. The term further encompasses synthetic nucleic acids that did not derive from genomic sequence. In certain embodiments, the genes lack introns, as they are synthesized based on the known DNA sequence of cDNA and protein sequence. In other embodiments, the genes are synthesized, non-native cDNA wherein the codons have been

optimized for expression in *Y. lipolytica* based on codon usage. The term can further include nucleic acid molecules comprising upstream, downstream, and/or intron nucleotide sequences.

The term "*genetic modification*" refers to the result of a transformation. Every
5 transformation causes a genetic modification by definition.

The term "*homolog*", as used herein, refers to (a) peptides, oligopeptides, polypeptides, proteins, and enzymes having amino acid substitutions, deletions and/or insertions relative to the unmodified protein in question and having similar biological and functional activity as the unmodified protein from which they are derived, and (b) nucleic
10 acids which encode peptides, oligopeptides, polypeptides, proteins, and enzymes with the same characteristics described in (a).

"*Inducible promoter*" is a promoter that mediates transcription of an operably linked gene in response to a particular stimulus.

The term "*integrated*" refers to a nucleic acid that is maintained in a cell as an
15 insertion into the genome of the cell, such as insertion into a chromosome, including insertions into a plastid genome.

"*In operable linkage*" is a functional linkage between two nucleic acid sequences, such a control sequence (typically a promoter) and the linked sequence (typically a sequence that encodes a protein, also called a coding sequence). A promoter is in operable
20 linkage with a gene if it can mediate transcription of the gene.

The term "*knockout mutation*" or "*knockout*" refers to a genetic modification that prevents a native gene from being transcribed and translated into a functional protein.

The term "*native*" refers to the composition of a cell or parent cell prior to a transformation event. A "*native gene*" refers to a nucleotide sequence that encodes a
25 protein that has not been introduced into a cell by a transformation event. A "*native protein*" refers to an amino acid sequence that is encoded by a native gene.

The terms "*nucleic acid*" refers to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. Polynucleotides may have any three-dimensional structure, and may perform any function. The following are
30 non-limiting examples of polynucleotides: coding or non-coding regions of a gene or gene fragment, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA

of any sequence, nucleic acid probes, and primers. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs. If present, modifications to the nucleotide structure may be imparted before or after assembly of the polymer. A polynucleotide may be further modified, such as by conjugation with a labeling
5 component. In all nucleic acid sequences provided herein, U nucleotides are interchangeable with T nucleotides.

The acronym "*ORF*" stands for open reading frame.

The term "*parent cell*" refers to every cell from which a cell descended. The genome of a cell is comprised of the parent cell's genome and any subsequent genetic
10 modifications to its genome.

As used herein, the term "*plasmid*" refers to a circular DNA molecule that is physically separate from an organism's genomic DNA. Plasmids may be linearized before being introduced into a host cell (referred to herein as a linearized plasmid). Linearized plasmids may not be self-replicating, but may integrate into and be replicated with the
15 genomic DNA of an organism.

The term "*portion*" refers to peptides, oligopeptides, polypeptides, protein domains, and proteins. A nucleotide sequence encoding a "portion of a protein" includes both nucleotide sequences that can be transcribed and/or translated and nucleotide sequences that must undergo one or more recombination events to be transcribed and/or translated. For
20 example, a nucleic acid may comprise a nucleotide sequence encoding one or more amino acids of a selectable marker protein. This nucleic acid can be engineered to recombine with one or more different nucleotide sequences that encode the remaining portion of the protein. Such nucleic acids are useful for generating knockout mutations because only recombination with the target sequence is likely to reconstitute the full-length selectable
25 marker gene whereas random-integration events are unlikely to result in a nucleotide sequence that can produce a functional marker protein. A "biologically-active portion" of a polypeptide is any amino acid sequence found in the polypeptide's amino acid sequence that is less than the full amino acid sequence but can perform the same function as the full-length polypeptide. A biologically-active portion of a diacylglycerol acyltransferase
30 includes any amino acid sequence found in a full-length diacylglycerol acyltransferase that can catalyze the formation of triacylglycerol from diacylglycerol and acyl-CoA. A biologically-active portion of a polypeptide includes portions of the polypeptide that have the same activity as the full-length peptide and every portion that has more activity than

background. For example, a biologically-active portion of a diacylglycerol acyltransferase may have 0.1, 0.5, 1, 2, 3, 4, 5, 10, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.5, 99.6, 99.7, 99.8, 99.9, 100, 100.1, 100.2, 100.3, 100.4, 100.5, 100.6, 100.7, 100.8, 100.9, 101, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 160, 170, 180, 190, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400 percent activity relative to the full-length polypeptide or higher. A biologically-active portion of a polypeptide may include portions of a peptide that lack a domain that targets the polypeptide to a cellular compartment.

A "*promoter*" is a nucleic acid control sequence that directs transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription. A promoter also optionally includes distal enhancer or repressor elements, which can be located as much as several thousand base pairs from the start site of transcription.

"*Recombinant*" refers to a cell, nucleic acid, protein, or vector, which has been modified due to the introduction of an exogenous nucleic acid or the alteration of a native nucleic acid. Thus, *e.g.*, recombinant cells can express genes that are not found within the native (non-recombinant) form of the cell or express native genes differently than those genes are expressed by a non-recombinant cell. Recombinant cells can, without limitation, include recombinant nucleic acids that encode for a gene product or for suppression elements such as mutations, knockouts, antisense, interfering RNA (RNAi), or dsRNA that reduce the levels of active gene product in a cell. A "recombinant nucleic acid" is a nucleic acid originally formed *in vitro*, in general, by the manipulation of nucleic acid, *e.g.*, using polymerases, ligases, exonucleases, and endonucleases, or otherwise is in a form not normally found in nature. Recombinant nucleic acids may be produced, for example, to place two or more nucleic acids in operable linkage. Thus, an isolated nucleic acid or an expression vector formed *in vitro* by ligating DNA molecules that are not normally joined in nature, are both considered recombinant for the purposes of this invention. Once a recombinant nucleic acid is made and introduced into a host cell or organism, it may replicate using the *in vivo* cellular machinery of the host cell; however, such nucleic acids, once produced recombinantly, although subsequently replicated intracellularly, are still considered recombinant for purposes of this invention. Similarly, a "recombinant protein" is a protein made using recombinant techniques, *i.e.*, through the expression of a recombinant nucleic acid.

The term "*regulatory region*" refers to nucleotide sequences that affect the transcription or translation of a gene but do not encode an amino acid sequence. Regulatory regions include promoters, operators, enhancers, and silencers.

"*Transformation*" refers to the transfer of a nucleic acid into a host organism or the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "recombinant", "transgenic" or "transformed" organisms. Thus, isolated polynucleotides of the present invention can be incorporated into recombinant constructs, typically DNA constructs, capable of introduction into and replication in a host cell. Such a construct can be a vector that includes a replication system and sequences that are capable of transcription and translation of a polypeptide-encoding sequence in a given host cell. Typically, expression vectors include, for example, one or more cloned genes under the transcriptional control of 5' and 3' regulatory sequences and a selectable marker. Such vectors also can contain a promoter regulatory region (*e.g.*, a regulatory region controlling inducible or constitutive, environmentally- or developmentally-regulated, or location-specific expression), a transcription initiation start site, a ribosome binding site, a transcription termination site, and/or a polyadenylation signal.

The term "*transformed cell*" refers to a cell that has undergone a transformation. Thus, a transformed cell comprises the parent's genome and an inheritable genetic modification.

The terms "*triacylglyceride*," "*triacylglycerol*," "*triglyceride*," and "*TAG*" are esters comprised of glycerol and three fatty acids.

The term "*triacylglycerol lipase*" refers to any protein that can catalyze the removal of a fatty acid chain from a triacylglycerol. Triacylglycerol lipases include TGL3.

The term "*vector*" refers to the means by which a nucleic acid can be propagated and/or transferred between organisms, cells, or cellular components. Vectors include plasmids, linear DNA fragments, viruses, bacteriophage, pro-viruses, phagemids, transposons, and artificial chromosomes, and the like, that may or may not be able to replicate autonomously or integrate into a chromosome of a host cell.

Microbe Engineering

A. Overview

In certain embodiments of the invention, a microorganism is genetically modified to increase its triacylglycerol content or modify its lipid profile.

Genes and gene products may be introduced into microbial host cells. Suitable host cells for expression of the genes and nucleic acid molecules are microbial hosts that can be found broadly within the fungal or bacterial families. Examples of suitable host strains include but are not limited to fungal or yeast species, such as *Arxula*, *Aspergillus*,
5 *Aurantiochytrium*, *Candida*, *Claviceps*, *Cryptococcus*, *Cunninghamella*, *Geotrichum*, *Hansenula*, *Kluyveromyces*, *Kodamaea*, *Leucosporidiella*, *Lipomyces*, *Mortierella*, *Ogataea*, *Pichia*, *Prototheca*, *Rhizopus*, *Rhodospiridium*, *Rhodotorula*, *Saccharomyces*, *Schizosaccharomyces*, *Tremella*, *Trichosporon*, *Wickerhamomyces*, and *Yarrowia*, or bacterial species, such as members of proteobacteria and actinomycetes, as well as the
10 genera *Acinetobacter*, *Arthrobacter*, *Brevibacterium*, *Acidovorax*, *Bacillus*, *Clostridia*, *Streptomyces*, *Escherichia*, *Salmonella*, *Pseudomonas*, and *Cornyebacterium*. *Yarrowia lipolytica* and *Arxula adenivorans* are well-suited for use as the host microorganism because they can accumulate a large percentage of their weight as triacylglycerols.

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

20 For example, a gene encoding an enzyme can be cloned in a suitable plasmid, and the aforementioned starting parent strain as a host can be transformed with the resulting plasmid. This approach can increase the copy number of each of the genes encoding the enzymes and, as a result, the activities of the enzymes can be increased. The plasmid is not particularly limited so long as it renders a desired gene inheritable to the microorganism's
25 progeny.

Vectors or cassettes useful for the transformation of suitable host cells are well known in the art. Typically the vector or cassette contains sequences directing transcription and translation of the relevant gene, a selectable marker, and sequences allowing autonomous replication or chromosomal integration. Suitable vectors comprise a region 5'
30 of the gene harboring transcriptional initiation controls and a region 3' of the DNA fragment which controls transcriptional termination. Both control regions may be derived from genes homologous to the transformed host cell, although it is to be understood that such control regions need not be derived from the genes native to the specific species chosen as a

production host.

Promoters, cDNAs, and 3'UTRs, as well as other elements of the vectors, can be generated through cloning techniques using fragments isolated from native sources (Green & Sambrook, Molecular Cloning: A Laboratory Manual, (4th ed., 2012); U.S. Patent No. 4,683,202; incorporated by reference). Alternatively, elements can be generated synthetically using known methods (Gene 164:49-53 (1995)).

B. Homologous Recombination

Homologous recombination is the ability of complementary DNA sequences to align and exchange regions of homology. Transgenic DNA ("donor") containing sequences homologous to the genomic sequences being targeted ("template") are introduced into the organism and then undergo recombination into the genome at the site of the corresponding genomic homologous sequences.

The ability to carry out homologous recombination in a host organism has many practical implications for what can be carried out at the molecular genetic level and is useful in the generation of a microbe that can produce a desired product. By its very nature homologous recombination is a precise gene targeting event, hence, most transgenic lines generated with the same targeting sequence will be essentially identical in terms of phenotype, necessitating the screening of far fewer transformation events. Homologous recombination also targets gene insertion events into the host chromosome, potentially resulting in excellent genetic stability, even in the absence of genetic selection. Because different chromosomal loci will likely impact gene expression, even from heterologous promoters/UTRs, homologous recombination can be a method of querying loci in an unfamiliar genome environment and to assess the impact of these environments on gene expression.

A particularly useful genetic engineering approach using homologous recombination is to co-opt specific host regulatory elements such as promoters/UTRs to drive heterologous gene expression in a highly specific fashion.

Because homologous recombination is a precise gene targeting event, it can be used to precisely modify any nucleotide(s) within a gene or region of interest, so long as sufficient flanking regions have been identified. Therefore, homologous recombination can be used as a means to modify regulatory sequences impacting gene expression of RNA and/or proteins. It can also be used to modify protein coding regions in an effort to modify enzyme activities such as substrate specificity, affinities and K_m , and thus affecting the

desired change in metabolism of the host cell. Homologous recombination provides a powerful means to manipulate the host genome resulting in gene targeting, gene conversion, gene deletion, gene duplication, gene inversion and exchanging gene expression regulatory elements such as promoters, enhancers and 3'UTRs.

5 Homologous recombination can be achieved by using targeting constructs containing pieces of endogenous sequences to "target" the gene or region of interest within the endogenous host cell genome. Such targeting sequences can either be located 5' of the gene or region of interest, 3' of the gene/region of interest or even flank the gene/region of interest. Such targeting constructs can be transformed into the host cell either as a
10 supercoiled plasmid DNA with additional vector backbone, a PCR product with no vector backbone, or as a linearized molecule. In some cases, it may be advantageous to first expose the homologous sequences within the transgenic DNA (donor DNA) by cutting the transgenic DNA with a restriction enzyme. This step can increase the recombination efficiency and decrease the occurrence of undesired events. Other methods of increasing
15 recombination efficiency include using PCR to generate transforming transgenic DNA containing linear ends homologous to the genomic sequences being targeted.

C. Vectors and Vector Components

Vectors for transformation of microorganisms in accordance with the present invention can be prepared by known techniques familiar to those skilled in the art in view
20 of the disclosure herein. A vector typically contains one or more genes, in which each gene codes for the expression of a desired product (the gene product) and is operably linked to one or more control sequences that regulate gene expression or target the gene product to a particular location in the recombinant cell.

1. Control Sequences

25 Control sequences are nucleic acids that regulate the expression of a coding sequence or direct a gene product to a particular location inside or outside a cell. Control sequences that regulate expression include, for example, promoters that regulate transcription of a coding sequence and terminators that terminate transcription of a coding sequence. Another control sequence is a 3' untranslated sequence located at the end of a
30 coding sequence that encodes a polyadenylation signal. Control sequences that direct gene products to particular locations include those that encode signal peptides, which direct the protein to which they are attached to a particular location inside or outside the cell.

Thus, an exemplary vector design for expression of a gene in a microbe contains a coding sequence for a desired gene product (for example, a selectable marker, or an enzyme) in operable linkage with a promoter active in yeast. Alternatively, if the vector does not contain a promoter in operable linkage with the coding sequence of interest, the coding sequence can be transformed into the cells such that it becomes operably linked to an endogenous promoter at the point of vector integration.

The promoter used to express a gene can be the promoter naturally linked to that gene or a different promoter.

A promoter can generally be characterized as constitutive or inducible. Constitutive promoters are generally active or function to drive expression at all times (or at certain times in the cell life cycle) at the same level. Inducible promoters, conversely, are active (or rendered inactive) or are significantly up- or down-regulated only in response to a stimulus. Both types of promoters find application in the methods of the invention. Inducible promoters useful in the invention include those that mediate transcription of an operably linked gene in response to a stimulus, such as an exogenously provided small molecule, temperature (heat or cold), lack of nitrogen in culture media, etc. Suitable promoters can activate transcription of an essentially silent gene or upregulate, preferably substantially, transcription of an operably linked gene that is transcribed at a low level.

Inclusion of termination region control sequence is optional, and if employed, then the choice is primarily one of convenience, as the termination region is relatively interchangeable. The termination region may be native to the transcriptional initiation region (the promoter), may be native to the DNA sequence of interest, or may be obtainable from another source (*See, e.g., Chen & Orozco, Nucleic Acids Research 16:8411 (1988)*).

2. *Genes and Codon Optimization*

Typically, a gene includes a promoter, coding sequence, and termination control sequences. When assembled by recombinant DNA technology, a gene may be termed an expression cassette and may be flanked by restriction sites for convenient insertion into a vector that is used to introduce the recombinant gene into a host cell. The expression cassette can be flanked by DNA sequences from the genome or other nucleic acid target to facilitate stable integration of the expression cassette into the genome by homologous recombination. Alternatively, the vector and its expression cassette may remain unintegrated (*e.g., an episome*), in which case, the vector typically includes an origin of replication, which is capable of providing for replication of the vector DNA.

A common gene present on a vector is a gene that codes for a protein, the expression of which allows the recombinant cell containing the protein to be differentiated from cells that do not express the protein. Such a gene, and its corresponding gene product, is called a selectable marker or selection marker. Any of a wide variety of selectable markers can be employed in a transgene construct useful for transforming the organisms of the invention.

For optimal expression of a recombinant protein, it is beneficial to employ coding sequences that produce mRNA with codons optimally used by the host cell to be transformed. Thus, proper expression of transgenes can require that the codon usage of the transgene matches the specific codon bias of the organism in which the transgene is being expressed. The precise mechanisms underlying this effect are many, but include the proper balancing of available aminoacylated tRNA pools with proteins being synthesized in the cell, coupled with more efficient translation of the transgenic messenger RNA (mRNA) when this need is met. When codon usage in the transgene is not optimized, available tRNA pools are not sufficient to allow for efficient translation of the transgenic mRNA resulting in ribosomal stalling and termination and possible instability of the transgenic mRNA.

D. Transformation

Cells can be transformed by any suitable technique including, e.g., biolistics, electroporation, glass bead transformation, and silicon carbide whisker transformation. Any convenient technique for introducing a transgene into a microorganism can be employed in the present invention. Transformation can be achieved by, for example, the method of D. M. Morrison (Methods in Enzymology 68:326 (1979)), the method by increasing permeability of recipient cells for DNA with calcium chloride (Mandel & Higa, J. Molecular Biology, 53:159 (1970)), or the like.

Examples of expression of transgenes in oleaginous yeast (e.g., *Yarrowia lipolytica*) can be found in the literature (Bordes et al., J. Microbiological Methods, 70:493 (2007); Chen et al., Applied Microbiology & Biotechnology 48:232 (1997)). Examples of expression of exogenous genes in bacteria, such as *E. coli*, are well known (Green & Sambrook, Molecular Cloning: A Laboratory Manual, (4th ed., 2012)).

Vectors for transformation of microorganisms in accordance with the present invention can be prepared by known techniques familiar to those skilled in the art. In one embodiment, an exemplary vector design for expression of a gene in a microorganism contains a gene encoding an enzyme in operable linkage with a promoter active in the

microorganism. Alternatively, if the vector does not contain a promoter in operable linkage with the gene of interest, the gene can be transformed into the cells such that it becomes operably linked to a native promoter at the point of vector integration. The vector can also contain a second gene that encodes a protein. Optionally, one or both gene(s) is/are
 5 followed by a 3' untranslated sequence containing a polyadenylation signal. Expression cassettes encoding the two genes can be physically linked in the vector or on separate vectors. Co-transformation of microbes can also be used, in which distinct vector molecules are simultaneously used to transform cells (Protist 155:381-93 (2004)). The transformed cells can be optionally selected based upon the ability to grow in the presence
 10 of the antibiotic or other selectable marker under conditions in which cells lacking the resistance cassette would not grow.

Exemplary Nucleic Acids, Cells, and Methods

A. Diacylglycerol acyltransferase nucleic acid molecules and vectors

Diacylglycerol acyltransferase is used as a model for increasing the activity of a
 15 protein, and the compositions and methods described below are generally applicable to other proteins in various organisms.

The diacylglycerol acyltransferase may be a type 1 diacylglycerol acyltransferase, type 2 diacylglycerol acyltransferase, or any other protein that catalyzes the conversion of diacylglycerol into a triacylglycerol. In some embodiments, the diacylglycerol
 20 acyltransferase is DGA1. For example, the diacylglycerol acyltransferase may be a DGA1 protein encoded by a *DGAT2* gene selected from the group consisting of *Yarrowia lipolytica*, *Arxula adeninivorans*, *Rhodospiridium toruloides*, *Lipomyces starkeyi*, *Aspergillus terreus*, *Claviceps purpurea*, and *Aurantiochytrium limacinum*.

The *DGAT2* gene may have a nucleotide sequence set forth in SEQ ID NOs: 16,
 25 118, 120, 122, 124, 126, 128, 130, 132, or 134. In other embodiments, the *DGAT2* gene is substantially identical to SEQ ID NOs: 16, 118, 120, 122, 124, 126, 128, 130, 132, or 134, and the nucleotide sequence encodes a protein that retains the functional activity of a protein encoded by SEQ ID NOs: 16, 118, 120, 122, 124, 126, 128, 130, 132, or 134, yet differs in nucleotide sequence due to natural allelic variation or mutagenesis. In another
 30 embodiment, the *DGAT2* gene comprises a nucleotide sequence at least about 70%, 72%, 64%, 76%, 78%, 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9% or more identical to SEQ ID NOs: 16, 118, 120, 122, 124, 126, 128, 130, 132, or 134.

The DGA1 protein may have an amino acid sequence set forth in SEQ ID NOs: 15, 117, 119, 121, 123, 125, 127, 129, 131 or 133. In other embodiments, the DGA1 protein is substantially identical to SEQ ID NOs: 15, 117, 119, 121, 123, 125, 127, 129, 131 or 133, and retains the functional activity of the protein of SEQ ID NOs: 15, 117, 119, 121, 123, 125, 127, 129, 131 or 133, yet differs in amino acid sequence due to natural allelic variation or mutagenesis. In another embodiment, the DGA1 protein comprises an amino acid sequence at least about 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9% or more identical to SEQ ID NO: SEQ ID NOs: 15, 117, 119, 121, 123, 125, 127, 129, 131 or 133.

In some embodiments, the diacylglycerol acyltransferase is DGA2. For example, the diacylglycerol acyltransferase may be a DGA2 protein encoded by a *DGAT1* gene found in an organism selected from the group consisting of *Yarrowia lipolytica*, *Rhodospiridium toruloides*, *Lipomyces starkeyi*, *Aspergillus terreus*, *Arxula adeninivorans*, *Claviceps purpurea*, and *Chaetomium globosum*.

The *DGAT1* gene may have a nucleotide sequence set forth in SEQ ID NO: 18, 136, 138, 140, 142, 144, or 146. In other embodiments, the *DGAT1* gene is substantially identical to SEQ ID NO: 18, 136, 138, 140, 142, 144, or 146, and the nucleotide sequence encodes a protein that retains the functional activity of a protein encoded by SEQ ID NO: 18, 136, 138, 140, 142, 144, or 146, yet differs in nucleotide sequence due to natural allelic variation or mutagenesis. In another embodiment, the *DGAT1* gene comprises a nucleotide sequence at least about 70%, 72%, 74%, 76%, 78%, 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9% or more identical to SEQ ID NO: 18, 136, 138, 140, 142, 144, or 146.

The DGA2 protein may have an amino acid sequence set forth in SEQ ID NO: 17, 135, 137, 139, 141, 143, or 145. In other embodiments, the DGA2 protein is substantially identical to SEQ ID NO: 17, 135, 137, 139, 141, 143, or 145, and retains the functional activity of the protein of SEQ ID NO: 17, 135, 137, 139, 141, 143, or 145, yet differs in amino acid sequence due to natural allelic variation or mutagenesis. In another embodiment, the DGA2 protein comprises an amino acid sequence at least about 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9% or more identical to SEQ ID NO: 17, 135, 137, 139, 141, 143, or 145.

The *DGAT2* and *DGAT1* genes may comprise conservative substitutions, deletions, and/or insertions while still encoding a protein that has functional diacylglycerol acyltransferase activity. For example, the *DGAT2* and *DGAT1* codons may be optimized for a particular host cell, different codons may be substituted for convenience, such as to
5 introduce a restriction site or create optimal PCR primers, or codons may be substituted for another purpose. Similarly, the nucleotide sequence may be altered to create conservative amino acid substitutions, deletions, and/or insertions.

The DGA1 and DGA2 polypeptides may comprise conservative substitutions, deletions, and/or insertions while still maintaining functional diacylglycerol acyltransferase
10 activity. Conservative substitution tables are well known in the art (Creighton, Proteins (2d. ed., 1992)).

Amino acid substitutions, deletions and/or insertions may readily be made using peptide synthetic techniques well known in the art, such as solid phase peptide synthesis and/or any other synthetic techniques, or by recombinant DNA manipulation. Methods for
15 the manipulation of DNA sequences to produce substitution, insertion or deletion variants of a protein are well known in the art. For example, techniques for making substitution mutations at predetermined sites in DNA are well known to those skilled in the art and include M13 mutagenesis, T7-Gen in vitro mutagenesis (USB, Cleveland, OH), Quick Change Site Directed mutagenesis (Stratagene, San Diego, CA), PCR-mediated site-
20 directed mutagenesis or other site-directed mutagenesis protocols.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences can be aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-identical sequences can be disregarded for comparison
25 purposes). The length of a reference sequence aligned for comparison purposes can be at least 95% of the length of the reference sequence. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions can then be compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that
30 position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number

of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In one embodiment, the percent identity between two amino acid sequences can be determined using the Needleman and Wunsch (*J. Molecular Biology* 48:444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another embodiment, the percent identity between two nucleotide sequences can be determined using the GAP program in the GCG software package (available at <http://www.gcg.com>), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. In another embodiment, the percent identity between two amino acid or nucleotide sequences can be determined using the algorithm of E. Meyers and W. Miller (*Computer Applications in the Biosciences* 4:11-17 (1988)) which has been incorporated into the ALIGN program (version 2.0 or 2.0U), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

Exemplary computer programs which can be used to determine identity between two sequences include, but are not limited to, the suite of BLAST programs, *e.g.*, BLASTN, MEGABLAST, BLASTX, TBLASTN, TBLASTX, and BLASTP, and Clustal programs, *e.g.*, ClustalW, ClustalX, and Clustal Omega.

Sequence searches are typically carried out using the BLASTN program, when evaluating a given nucleic acid sequence relative to nucleic acid sequences in the GenBank DNA Sequences and other public databases. The BLASTX program is effective for searching nucleic acid sequences that have been translated in all reading frames against amino acid sequences in the GenBank Protein Sequences and other public databases.

An alignment of selected sequences in order to determine "% identity" between two or more sequences is performed using for example, the CLUSTAL-W program.

A "coding sequence" or "coding region" refers to a nucleic acid molecule having sequence information necessary to produce a protein product, such as an amino acid or polypeptide, when the sequence is expressed. The coding sequence may comprise and/or consist of untranslated sequences (including introns or 5' or 3' untranslated regions) within translated regions, or may lack such intervening untranslated sequences (*e.g.*, as in cDNA).

The abbreviation used throughout the specification to refer to nucleic acids comprising and/or consisting of nucleotide sequences are the conventional one-letter abbreviations. Thus when included in a nucleic acid, the naturally occurring encoding nucleotides are abbreviated as follows: adenine (A), guanine (G), cytosine (C), thymine (T) and uracil (U). Also, unless otherwise specified, the nucleic acid sequences presented herein is the 5'→3' direction.

As used herein, the term "complementary" and derivatives thereof are used in reference to pairing of nucleic acids by the well-known rules that A pairs with T or U and C pairs with G. Complement can be "partial" or "complete". In partial complement, only some of the nucleic acid bases are matched according to the base pairing rules; while in complete or total complement, all the bases are matched according to the pairing rule. The degree of complement between the nucleic acid strands may have significant effects on the efficiency and strength of hybridization between nucleic acid strands as well known in the art. The efficiency and strength of said hybridization depends upon the detection method.

As used herein, "DGA1" means a diacylglycerol acyltransferase type 2 (DGAT2). DGA1 is an integral membrane protein that catalyzes the final enzymatic step in oil biosynthesis and the production of triacylglycerols in plants, fungi, and mammals. The DGA1 may play a key role in altering the quantity of long-chain polyunsaturated fatty acids produced in oils of oleaginous organisms. DGA1 is related to the acyl-coenzyme A:cholesterol acyltransferase ("ACAT"). This enzyme is responsible for transferring an acyl group from acyl-coenzyme-A to the sn-3 position of 1,2-diacylglycerol ("DAG") to form triacylglycerol ("TAG") (thereby involved in the terminal step of TAG biosynthesis). DGA1 is associated with membrane and lipid body fractions in plants and fungi, particularly, in oilseeds where it contributes to the storage of carbon used as energy reserves. TAG is believed to be an important chemical for storage of energy in cells. DGA1 is known to regulate TAG structure and direct TAG synthesis.

The DGA1 polynucleotide and polypeptide sequences may be derived from highly oleaginous organisms having very high, native levels of lipid accumulation. (Bioresource Technology 144:360-69 (2013); Progress Lipid Research 52:395-408 (2013); Applied Microbiology & Biotechnology 90:1219-27 (2011); European Journal Lipid Science & Technology 113:1031-51 (2011); Food Technology & Biotechnology 47:215-20 (2009); Advances Applied Microbiology 51:1-51 (2002); Lipids 11:837-44 (1976)). The list of organisms with reported lipid content about 50% and above are shown in Table 1. R.

toruloides and *L. starkeyi* have the highest lipid content. Among the organisms in Table 1, only five had publicly accessible sequence for DGA1, *R. toruloides*, *L. starkeyi*, *A. limacinum*, *A. terreus*, and *C. purpurea* (bolded in Table 1).

- 5 **Table 1.** List of oleaginous fungi with reported lipid content about 50% and above. Organisms with publicly accessible sequence for DGA1 gene are in bold.

Fungi with reported high lipid content
<i>Aspergillus terreus</i>
<i>Aurantiochytrium limacinum</i>
<i>Claviceps purpurea</i>
<i>Cryptococcus albidus</i>
<i>Cryptococcus curvatus</i>
<i>Cryptococcus ramirezgomezianus</i>
<i>Cryptococcus terreus</i>
<i>Cryptococcus wieringae</i>
<i>Cunninghamella echinulata</i>
<i>Cunninghamella japonica</i>
<i>Leucosporidiella creatinivora</i>
<i>Lipomyces lipofer</i>
<i>Lipomyces starkeyi</i>
<i>Lipomyces tetrasporus</i>
<i>Mortierella isabellina</i>
<i>Prototheca zopfii</i>
<i>Rhizopus arrhizus</i>
<i>Rhodospiridium babjevae</i>
<i>Rhodospiridium paludigenum</i>
<i>Rhodospiridium toruloides</i>
<i>Rhodotorula glutinis</i>
<i>Rhodotorula mucilaginosa</i>
<i>Tremella encephala</i>
<i>Trichosporon cutaneum</i>
<i>Trichosporon fermentans</i>

- 10 Nucleic acid constructs for increasing the activity of DGA1 were described in USSN 61/943,664 and PCT Patent Application No. US15/017227 (both hereby incorporated by reference). Figure 3 shows expression construct pNC243 used for overexpression of the *R.*

toruloides DGA1 gene NG66 (SEQ ID NO: 6) in *Y. lipolytica*. DGA1 expression constructs were linearized before transformation by PacI/NotI restriction digest. The linear expression constructs each included the expression cassette for DGA1 gene and for the NatI gene, used as marker for selection with nourseothricin (NAT).

- 5 Nucleic acid constructs for increasing the activity of DGA2 and/or other diacylglycerol acyltransferases may be created using the methods described above and/or other methods known in the art. Similarly, nucleic acid constructs that increase the activity of other proteins may be engineered as described for DGA1 above and/or by other methods known in the art.

10 B. Triacylglycerol lipase nucleic acid molecules and vectors

Triacylglycerol lipase is used as a model for decreasing the activity of a protein, and the compositions and methods described below are generally applicable to other proteins in various organisms.

- 15 Triacylglycerol lipase depletes a cell's triacylglycerol by removing one or more fatty acid chains. Thus, decreasing the net triacylglycerol lipase activity of a cell may increase the cell's triacylglycerol. This decrease may be accomplished by reducing the efficiency of the enzyme, *e.g.*, by mutating amino acids in its active site, or by reducing the expression of the enzyme. For example, a TGL3 knockout mutation will decrease the activity of a triacylglycerol lipase because it prevents the cell from transcribing *TGL3*.

- 20 In some embodiments, the triacylglycerol lipase is TGL3.

- The *TGL3* gene in *Y. lipolytica* encodes the triacylglycerol lipase protein TGL3 (SEQ ID NO: 147). SEQ ID NO: 148 contains the *TGL3* nucleotide sequence, 100 upstream nucleotides, and 100 downstream nucleotides. Thus, the SEQ ID NO: 148 nucleotide sequence may be used to design a nucleic acid capable of recombining with a
25 nucleic acid sequence in the native *Y. lipolytica* triacylglycerol lipase gene.

- Knockout cassettes SEQ ID NOS: 160 and 161 are capable of recombining with the native *TGL3* gene in *Y. lipolytica*. Thus, in some embodiments, the nucleic acids encoded by SEQ ID NOS: 160 and 161 may be used to generate a triacylglycerol lipase knockout mutation in *Y. lipolytica*. SEQ ID NOS: 160 and 161 each contain portions of a
30 hygromycin resistance gene *hph*. Neither isolated sequence encodes a functional protein, but the two sequences are capable of encoding a functional kinase that confers hygromycin resistance upon successful recombination. Further, neither SEQ ID NO: 160 nor SEQ ID NO: 161 contains a promoter or terminator, and thus, they rely on homologous

recombination with the *Y. lipolytica* *TGL3* gene in order for the *hph* gene to be transcribed and translated. In this way, successfully transformed oleaginous cells may be selected by growing the cells on medium containing hygromycin.

Knockout cassette SEQ ID NO: 160 may be prepared by amplifying a hygromycin resistance gene *hph* (SEQ ID NO: 152) with primer NP1798 (SEQ ID NO: 157) and primer NP656 (SEQ ID NO: 154). Knockout cassette SEQ ID NO: 161 may be prepared by amplifying a hygromycin resistance gene *hph* (SEQ ID NO: 152) with primer NP655 (SEQ ID NO: 153) and primer NP1799 (SEQ ID NO: 158).

Different approaches may be used to design nucleic acids that reduce the activity of the *TGL3* gene in *Y. lipolytica* (Biochimica Biophysica Acta 1831:1486-95 (2013)). The methods disclosed herein and other methods known in the art may be used to reduce the activity of triacylglycerol lipase and other genes in various species, including *Arxula adeninivorans*.

C. Transformed Cell

In some embodiments, the transformed cell is a prokaryotic cell, such as a bacterial cell. In some embodiments, the cell is a eukaryotic cell, such as a yeast cell or a filamentous fungi cell.

The cell may be selected from the group consisting of *Arxula*, *Aspergillus*, *Aurantiochytrium*, *Candida*, *Claviceps*, *Cryptococcus*, *Cunninghamella*, *Geotrichum*, *Hansenula*, *Kluyveromyces*, *Kodamaea*, *Leucosporidiella*, *Lipomyces*, *Mortierella*, *Ogataea*, *Pichia*, *Prototheca*, *Rhizopus*, *Rhodospiridium*, *Rhodotorula*, *Saccharomyces*, *Schizosaccharomyces*, *Tremella*, *Trichosporon*, *Wickerhamomyces*, and *Yarrowia*.

In some embodiments, the cell is selected from the group consisting of *Arxula adeninivorans*, *Aspergillus niger*, *Aspergillus oryzae*, *Aspergillus terreus*, *Aurantiochytrium limacinum*, *Candida utilis*, *Claviceps purpurea*, *Cryptococcus albidus*, *Cryptococcus curvatus*, *Cryptococcus ramirezgomezianus*, *Cryptococcus terreus*, *Cryptococcus wieringae*, *Cunninghamella echinulata*, *Cunninghamella japonica*, *Geotrichum fermentans*, *Hansenula polymorpha*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Kodamaea ohmeri*, *Leucosporidiella creatinivora*, *Lipomyces lipofer*, *Lipomyces starkeyi*, *Lipomyces tetrasporus*, *Mortierella isabellina*, *Mortierella alpina*, *Ogataea polymorpha*, *Pichia cifferrii*, *Pichia guilliermondii*, *Pichia pastoris*, *Pichia stipites*, *Prototheca zopfii*, *Rhizopus arrhizus*, *Rhodospiridium bahjevae*, *Rhodospiridium toruloides*, *Rhodospiridium paludigenum*, *Rhodotorula glutinis*, *Rhodotorula mucilaginis*, *Saccharomyces cerevisiae*,

Schizosaccharomyces pombe, *Tremella encephala*, *Trichosporon cutaneum*, *Trichosporon fermentans*, *Wickerhamomyces ciferrii*, and *Yarrowia lipolytica*.

In certain embodiments, the transformed cell is a high-temperature tolerant yeast cell. In some embodiments the transformed cell is *Kluyveromyces marxianus*.

5 In certain embodiments, the cell is *Yarrowia lipolytica* or *Arxula adeninivorans*.

D. Increasing the activity of a protein in a cell

A protein's activity may be increased by overexpressing the protein. Proteins may be overexpressed in a cell using a variety of genetic modifications. In some embodiments, the genetic modification increases the expression of a native protein. A native protein may
10 be overexpressed by modifying the upstream transcription regulators of a native gene, for example, by increasing the expression of a transcription activator or decreasing the expression of a transcription repressor. Alternatively, the promoter of a native gene may be substituted with a constitutively active or inducible promoter by recombination with an exogenous nucleic acid.

15 In some embodiments, the genetic modification encodes at least one copy of a gene. The gene may be a gene native to the cell or from a different species. In certain embodiments, the gene is inheritable to the progeny of a transformed cell. In some embodiments, the gene is inheritable because it resides on a plasmid. In certain embodiments, the gene is inheritable because it is integrated into the genome of the
20 transformed cell.

In certain embodiments, the gene is a type 2 diacylglycerol acyltransferase gene ("DGAT2") from *Aspergillus terreus*, *Arxula adeninivorans*, *Aurantiochytrium limacinum*, *Claviceps purpurea*, *Lipomyces starkeyi*, *Rhodospiridium toruloides*, or *Yarrowia lipolytica*. In certain embodiments, diacylglycerol acyltransferase is overexpressed by
25 transforming a cell with a gene encoding a diacylglycerol acyltransferase gene. The genetic modification may encode one or more than one copy of a diacylglycerol acyltransferase gene. In certain embodiments, the genetic modification encodes at least one copy of the DGA1 protein from *R. toruloides*. In some embodiments, the genetic modification encodes at least one copy of the DGA1 protein from *R. toruloides* and the transformed cell is *Y. lipolytica*. In some embodiments, the genetic modification encodes at least one copy of the
30 DGA1 protein from *R. toruloides*, and the transformed cell is *Arxula adeninivorans*.

In certain embodiments, the gene is a type 1 diacylglycerol acyltransferase gene ("DGAT1") from *Aspergillus terreus*, *Arxula adeninivorans*, *Chaetomium globosum*,

Claviceps purpurea, *Lipomyces starkeyi*, *Rhodospiridium toruloides*, or *Yarrowia lipolytica*. In certain embodiments, diacylglycerol acyltransferase is overexpressed by transforming a cell with a gene encoding a diacylglycerol acyltransferase gene. The genetic modification may encode one or more than one copy of a diacylglycerol acyltransferase gene. In certain embodiments, the genetic modification encodes at least one copy of the DGA2 protein from *Claviceps purpurea*. In some embodiments, the genetic modification encodes at least one copy of the DGA2 protein from *C. purpurea* and the transformed cell is *Y. lipolytica*. In some embodiments, the genetic modification encodes at least one copy of the DGA2 protein from *C. purpurea*, and the transformed cell is *Arxula adeninivorans*.

In some embodiments, the DGA1 protein is from *R. toruloides* and the DGA2 protein is from *C. purpurea*. In some embodiments, the DGA1 protein is from *R. toruloides*, the DGA2 protein is from *C. purpurea*, and the transformed cell is *Y. lipolytica*. In some embodiments, the DGA1 protein is from *R. toruloides*, the DGA2 protein is from *C. purpurea*, and the transformed cell is *Arxula adeninivorans*.

E. Decreasing the activity of a protein in a cell

A protein's activity may be decreased using a variety of different genetic modifications. In some embodiments, the transformed cell comprises a genetic modification that decreases the activity of a native protein. Such genetic modifications may affect a protein that regulates the transcription of a gene, including modifications that decrease the expression of a transcription activator and/or increase the expression of a transcription repressor. Modifications that affect a regulator protein may both decrease the expression of target protein and alter other gene expression profiles that shift the cellular equilibrium toward increased lipid accumulation or modified lipid composition. Alternatively, the genetic modification may be the introduction of a small interfering RNA, or a nucleic acid that encodes a small interfering RNA. In other embodiments, the genetic modification consists of the homologous recombination of a nucleic acid and the regulatory region of a gene, including an operator, promoter, sequences upstream from the promoter, enhancers, and sequences downstream of the gene.

In some embodiments the transformed oleaginous cell comprises a genetic modification consisting of a homologous recombination event. In certain embodiments, the transformed cell comprises a genetic modification consisting of a homologous recombination event between a gene and a nucleic acid. Thus, the genetic modification deletes the gene, prevents its transcription, or prevents the transcription of a gene that can

be transcribed into a fully-active protein. A homologous recombination event may mutate or delete a portion of a gene. For example, the homologous recombination event may mutate one or more residues in the active site of an enzyme, thereby reducing the efficiency of the enzyme or rendering it inactive. Alternatively, the homologous recombination event may affect post-translational modification, folding, stability, or localization within the cell. In some embodiments, the homologous recombination event replaces the promoter with a promoter that drives less transcription. In other embodiments, the homologous recombination event mutates the promoter to impair its ability to drive transcription. In certain embodiments, the genetic modification is a knockout mutation.

A knockout mutation may delete a gene. Additionally, the knockout mutation may substitute the native gene with an exogenous gene that encodes a different protein. The exogenous gene may be operably linked to an exogenous promoter. In certain embodiments, the exogenous gene is not linked to an exogenous promoter, and instead, the exogenous gene is configured to recombine with the native gene such that the native gene's promoter drives transcription of the exogenous gene. Thus, the exogenous gene is less likely to be expressed if it randomly integrates into the cell's genome. Methods for creating knockouts are well-known in the art (*See, e.g., Fickers et al., J. Microbiological Methods 55:727 (2003)*).

In certain embodiments, the genetic modification comprises two homologous recombination events. In the first event, a nucleic acid encoding a portion of an exogenous gene recombines with the native gene, and in the second event, a nucleic acid encoding the remaining portion of the exogenous gene recombines with the native gene. The two portions of the exogenous gene are designed such that neither portion is functional unless they recombine with each other. These two events further reduce the likelihood that the exogenous gene can be expressed following random integration events.

In certain embodiments, the exogenous gene encodes a dominant selectable marker. Thus, knockout cells may be selected by screening for the marker. In some embodiments, the dominant selectable marker is a drug resistance marker. A drug resistance marker is a dominant selectable marker that, when expressed by a cell, allows the cell to grow and/or survive in the presence of a drug that would normally inhibit cellular growth and/or survival. Cells expressing a drug resistance marker can be selected by growing the cells in the presence of the drug. In some embodiments, the drug resistance marker is an antibiotic resistance marker. In some embodiments, the drug resistance marker confers resistance to a

- drug selected from the group consisting of Amphotericin B, Candicidin, Filipin, Hamycin, Natamycin, Nystatin, Rimocidin, Bifonazole, Butoconazole, Clotrimazole, Econazole, Fenticonazole, Isoconazole, Ketoconazole, Luliconazole, Miconazole, Omoconazole, Oxiconazole, Sertaconazole, Sulconazole, Tioconazole, Albaconazole, Fluconazole,
- 5 Isavuconazole, Itraconazole, Posaconazole, Ravuconazole, Terconazole, Voriconazole, Abafungin, Amorolfine, Butenafine, Naftifine, Terbinafine, Amidulafungin, Caspofungin, Micafungin, Benzoic acid, Ciclopirox, Flucytosine, 5-fluorocytosine, Griseofulvin, Haloprogin, Polygodial, Tolnaftate, Crystal violet, Amikacin, Gentamicin, Kanamycin, Neomycin, Netilmicin, Tobramycin, Paromomycin, Spectinomycin, Geldanamycin,
- 10 Herbimycin, Rifaximin, Streptomycin, Loracarbef, Ertapenem, Doripenem, Imipenem, Meropenem, Cefadroxil, Cefazolin, Cefalotin, Cefalexin, Cefaclor, Cefamandole, Cefoxitin, Cefprozil, Cefuroxime, Cefixime, Cefdinir, Cefditoren, Cefoperazone, Cefotaxime, Cefpodoxime, Ceftazidime, Ceftibuten, Ceftizoxime, Ceftriaxone, Cefepime, Ceftaroline fosamil, Ceftobiprole, Teicoplanin, Vancomycin, Telavancin, Clindamycin,
- 15 Lincomycin, Daptomycin, Azithromycin, Clarithromycin, Dirithromycin, Erythromycin, Roxithromycin, Troleandomycin, Telithromycin, Spiramycin, Aztreonam, Furazolidone, Nitrofurantoin, Linezolid, Posizolid, Radezolid, Torezolid, Amoxicillin, Ampicillin, Azlocillin, Carbenicillin, Cloxacillin, Dicloxacillin, Flucloxacillin, Mezlocillin, Methicillin, Nafcillin, Oxacillin, Penicillin G, Penicillin V, Piperacillin, Penicillin G, Temocillin,
- 20 Ticarcillin, clavulanate, sulbactam, tazobactam, clavulanate, Bacitracin, Colistin, Polymyxin B, Ciprofloxacin, Enoxacin, Gatifloxacin, Levofloxacin, Lomefloxacin, Moxifloxacin, Nalidixic acid, Norfloxacin, Ofloxacin, Trovafloxacin, Grepafloxacin, Sparfloxacin, Temafloxacin, Mafenide, Sulfacetamide, Sulfadiazine, Silver sulfadiazine, Sulfadimethoxine, Sulfamethizole, Sulfamethoxazole, Sulfanilimide, Sulfasalazine,
- 25 Sulfisoxazole, Trimethoprim-Sulfamethoxazole, Co-trimoxazole, Sulfonamidochrysoidine, Demeclocycline, Doxycycline, Minocycline, Oxytetracycline, Tetracycline, Clofazimine, Dapsone, Capreomycin, Cycloserine, Ethambutol, Ethionamide, Isoniazid, Pyrazinamide, Rifampicin, Rifabutin, Rifapentine, Streptomycin, Arsphenamine, Chloramphenicol, Fosfomycin, Fusidic acid, Metronidazole, Mupirocin, Platensimycin, Quinupristin,
- 30 Dalbapristin, Thiamphenicol, Tigecycline, Tinidazole, Trimethoprim, Geneticin, Nourseothricin, Hygromycin, Bleomycin, and Puromycin.

In some embodiments, the dominant selectable marker is a nutritional marker. A nutritional marker is a dominant selectable marker that, when expressed by the cell, enables

the cell to grow or survive using one or more particular nutrient sources. Cells expressing a nutritional marker can be selected by growing the cells under limiting nutrient conditions in which cells expressing the nutritional marker can survive and/or grow, but cells lacking the nutrient marker cannot. In some embodiments, the nutritional marker is selected from the group consisting of Orotidine 5-phosphate decarboxylase, Phosphite specific oxidoreductase, Alpha-ketoglutarate-dependent hypophosphite dioxygenase, Alkaline phosphatase, Cyanamide hydratase, Melamine deaminase, Cyanurate amidohydrolase, Biuret hydrolyase, Urea amidolyase, Ammelide aminohydrolase, Guanine deaminase, Phosphodiesterase, Phosphotriesterase, Phosphite hydrogenase, Glycerophosphodiesterase, Parathion hydrolyase, Phosphite dehydrogenase, Dibenzothiophene desulfurization enzyme, Aromatic desulfinase, NADH-dependent FMN reductase, Aminopurine transporter, Hydroxylamine oxidoreductase, Invertase, Beta-glucosidase, Alpha-glucosidase, Beta-galactosidase, Alpha-galactosidase, Amylase, Cellulase, and Pullulanase.

In some embodiments, a genetic modification decreases the expression of a native gene by 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9, or 100 percent.

In some embodiments, a genetic modification decreases the efficiency of a native gene by 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9, or 100 percent.

In some embodiments, a genetic modification decreases the activity of a native gene by 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9, or 100 percent.

F. Decreasing acyl-CoA oxidase activity in a cell with concomitant overexpression of diacylglycerol acyltransferase

In some embodiments, the transformed oleaginous cell comprises a acyl-CoA oxidase knockout mutation and a genetic modification that increase the expression of a native diacylglycerol acyltransferase. In certain embodiments, the transformed oleaginous cell comprises a acyl-CoA oxidase knockout mutation and a genetic modification that encodes at least one copy of a diacylglycerol acyltransferase gene that is either native to the cell or from a different species of cell. In some embodiments, a *POX* gene is disrupted and

a DGA1 protein is overexpressed. In other embodiments, a *POX* gene is disrupted and a DGA2 protein is overexpressed.

In some embodiments, one nucleic acid increases the expression of a native diacylglycerol acyltransferase or encodes at least one copy of a diacylglycerol acyltransferase gene and a second nucleic acid decreases the activity of a acyl-CoA oxidase in the cell. In some embodiments the same nucleic acid encodes at least one copy of a diacylglycerol acyltransferase gene and decreases the activity of a acyl-CoA oxidase in the cell. For example, the nucleic acid designed to knock out a acyl-CoA oxidase gene may also contain a copy of a diacylglycerol acyltransferase gene.

10 G. Decreasing both triacylglycerol lipase and acyl-CoA oxidase activity in a cell

In some embodiments, the transformed oleaginous cell comprises a acyl-CoA oxidase knockout mutation and a triacylglycerol lipase knockout mutation. In certain embodiments, a *TGL3*, *TGL3/4*, and/or *TGL4* gene is disrupted. In certain embodiments, a *POX* gene is disrupted. In some embodiments, a first nucleic acid decreases the activity of a triacylglycerol acyltransferase in the cell and a second nucleic acid decreases the activity of a acyl-CoA oxidase in the cell.

H. Triacylglycerol production

In certain embodiments, the transformed cells are grown in the presence of exogenous fatty acids, glucose, ethanol, xylose, sucrose, starch, starch dextrin, glycerol, cellulose, and/or acetic acid. These substrates may be added during cultivation to increase lipid production. The exogenous fatty acids may include stearate, oleic acid, linoleic acid, γ -linolenic acid, dihomo- γ -linolenic acid, arachidonic acid, α -linolenic acid, stearidonic acid, eicosatetraenoic acid, eicosapentaenoic acid, docosapentaenoic acid, eicosadienoic acid, and/or eicosatrienoic acid.

In certain embodiments, the present invention relates to a product produced by a modified host cell described herein. In certain embodiments, the product is an oil, lipid, or triacylglycerol. In some embodiments, the product is palmitic acid, palmitoleic acid, stearic acid, oleic acid, or linoleic acid. In certain embodiments, the product is a saturated fatty acid. Thus, the product may be caprylic acid, capric acid, lauric acid, myristic acid, palmitic acid, stearic acid, arachidic acid, behenic acid, lignoceric acid, or cerotic acid. In some embodiments, the product is an unsaturated fatty acid. Thus, the product may be myristoleic acid, palmitoleic acid, sapienic acid, oleic acid, elaidic acid, vaccenic acid,

linoleic acid, linoelaidic acid, α -linolenic acid, arachidonic acid, eicosapentaenoic acid, erucic acid, or docosahexaenoic acid.

The product may be selected from the group consisting of lipids, triacylglycerides, fatty alcohols, fatty acids, alkanes, alkenes, isoprenoids, isoprene, squalene, farnasene, alcohols, isopropanol, n-propanol, n-butanol, isobutanol, 2-butanol, butadiene, diols, 1,3 propanediol, 1,4 propanediol, succinic acid, adipic acid, nylon precursors, citric acid, malic acid, polyols, and erythritol.

I. Increasing the activity of a protein to increase or modify the lipid content of a cell

In some embodiments, the invention relates to a nucleic acid comprising a nucleotide sequence set forth in SEQ ID NO: 12, 22, 86, 94, 100, 104, 110, 173, or 177. In some embodiments, the invention relates to a nucleic acid comprising a nucleotide sequence having at least 70%, 72%, 64%, 76%, 78%, 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with a nucleotide sequence set forth in SEQ ID NO: 12, 22, 86, 94, 100, 104, 110, 173, or 177. In some embodiments, the invention relates to a nucleic acid that encodes an amino acid sequence having at least 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with an amino acid sequence set forth in SEQ ID NO: 11, 21, 85, 93, 99, 103, 109, 172, or 176. In some embodiments, the nucleic acid further comprises a promoter. In certain embodiments, the nucleotide sequence and promoter are operably linked. In some embodiments, the promoter is not an *Arxula adeninivorans* promoter.

In some embodiments, the invention relates to a plasmid vector comprising a gene encoding a protein selected from Acetyl-CoA carboxylase ("ACC," Gene NG160, SEQ ID:2), ATP-citrate lyase, subunit 1 ("ACL1," Gene NG161, SEQ ID: 4), ATP-citrate lyase, subunit 2 ("ACL2," Gene NG162, SEQ ID NO:6), Acetyl-coA synthetase isoform 1 ("ACS1," Gene NG163, SEQ ID NO:8), Acetyl-coA synthetase isoform 2, ("ACS2," Genes NG164, NG165, SEQ ID NOs:10,12), Diacylglycerol acyltransferase type 2 ("DGA1," Gene NG167, SEQ ID NO:16), Diacylglycerol acyltransferase type 1 ("DGA2," Gene NG168, SEQ ID NO:18), Glycerol-3-phosphate dehydrogenase ("GPD1," Genes NG169, NG170, SEQ ID NOs:20, 22), Malic enzyme ("MAE1," Gene NG171, SEQ ID NO:24), 1-acyl-sn-glycerol-3-phosphate acyltransferase ("SLC1," Gene NG177, SEQ ID NO:36), Phospholipid:diacylglycerol acyltransferase ("LRO1," Gene NG178, SEQ ID NO:38),

Glycerol kinase ("GUT1," Gene NG182, SEQ ID NO:46), Glycerol-3-phosphate acyl transferase ("SCT1," Gene NG183, SEQ ID NO:48), Phosphatidate phosphatase ("PAH1," Gene NG218, SEQ ID NO:50), Fatty acid synthase, subunit alpha ("FAS2," Gene NG186, SEQ ID NO:54), Fatty acid synthase, subunit beta ("FAS1," Gene NG187, SEQ ID NO:56),

5 Type 2A-related protein phosphatase ("SIT4," Gene NG188, SEQ ID NO:58), Regulatory subunit of type 1 protein phosphatase Glc7p ("REG1," Gene NG189, SEQ ID NO:60), Regulatory subunit of Sit4 phosphatase ("SAP190," Gene NG190, SEQ ID NO:62), Dephospho-CoA kinase ("DPCK," Gene NG191, SEQ ID NO:64), HAC1 basic leucine zipper transcription factor ("HAC1," Gene NG192, SEQ ID NO:66), CAT8 zinc finger

10 transcription factor ("CAT8," Gene NG193, SEQ ID NO:68), Ubiquitin ("UBI4," Gene NG194, SEQ ID NO:70), STB5 transcription factor ("STB5," Gene NG199, SEQ ID NO:80), Pyruvate carboxylase ("PYC1," Genes NG201, NG202 SEQ ID NOs:84, 86), Pyruvate decarboxylase ("PDC1," Gene NG203, SEQ ID NO:88), Malate dehydrogenase 1 ("MDH1," Gene NG204, SEQ ID NO:90), Malate dehydrogenase 2 ("MDH2," Genes

15 NG205, NG206 SEQ ID NOs:92, 94), Pyruvate dehydrogenase alpha ("PDA1," Gene NG207, SEQ ID NO:96), Pyruvate dehydrogenase beta ("PDB1," Genes NG208, NG209 SEQ ID NOs:98, 100), Glucose-6-phosphate dehydrogenase ("ZWF1," Genes NG210, NG211 SEQ ID NOs:102, 104), 6-phosphogluconate dehydrogenase ("GND1," Gene NG212, SEQ ID NO:106), Citrate synthase 1 ("CIT1," Genes NG213, NG214, SEQ ID

20 NOs:108, 110), Citrate synthase 2 ("CIT2," Genes NG215, NG216 SEQ ID NOs:112, 114), Aldehyde dehydrogenase ("ALD6," Gene NG217, SEQ ID NO:116), Glycerol-3-phosphate/dihydroxyacetone phosphate sn-1 acyltransferase ("GPT2," Gene NG219, SEQ ID NO:167), Lysophospholipid acyltransferase ("SLC4," Gene NG220, SEQ ID NO:169), and Lysophosphatidic acid acyltransferase ("LOA1," Genes NG221, NG222 SEQ ID

25 NOs:171, 173).

In some embodiments, the invention relates to a vector comprising a nucleotide sequence having at least 70%, 72%, 64%, 76%, 78%, 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with a nucleotide sequence set forth in SEQ

30 ID NO: 2, 4, 6, 8, 10, 12, 16, 18, 20, 22, 24, 36, 38, 46, 48, 50, 54, 56, 58, 60, 62, 64, 66, 68, 70, 80, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 167, 169, 171, or 173. In some embodiments, the invention relates to a vector that encodes an amino acid sequence having at least 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%,

96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with an amino acid sequence set forth in SEQ ID NO: 1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 35, 37, 45, 47, 49, 53, 55, 57, 59, 61, 63, 65, 67, 69, 79, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 166, 168, 170, or 172. In

5 some embodiments, the vector is a plasmid. In some embodiments, vector further comprises a promoter. In certain embodiments, the nucleotide sequence and promoter are operably linked. In some embodiments, the promoter is not an *Arxula adeninivorans* promoter.

In some embodiments, the invention relates to a transformed cell comprising a
 10 genetic modification that increases the activity of a protein selected from an Acetyl-CoA carboxylase; ATP-citrate lyase, subunit 1; ATP-citrate lyase, subunit 2; Acetyl-coA synthetase isoform 1; Acetyl-CoA carboxylase; ATP-citrate lyase, subunit 1; ATP-citrate lyase, subunit 2; Acetyl-coA synthetase isoform 1; Acetyl-coA synthetase isoform 2; Diacylglycerol acyltransferase type 2; Diacylglycerol acyltransferase type 1; Glycerol-3-
 15 phosphate dehydrogenase; Malic enzyme; 1-acyl-sn-glycerol-3-phosphate acyltransferase; Phospholipid:diacylglycerol acyltransferase; Glycerol kinase; Glycerol-3-phosphate acyl transferase; Phosphatidate phosphatase; Fatty acid synthase, subunit alpha; Fatty acid synthase, subunit beta; Type 2A-related protein phosphatase; Regulatory subunit of type 1 protein phosphatase Glc7p; Regulatory subunit of Sit4 phosphatase; Dephospho-CoA
 20 kinase; HAC1 basic leucine zipper transcription factor; CAT8 zinc finger transcription factor; Ubiquitin; STB5 transcription factor; Pyruvate carboxylase; Pyruvate decarboxylase; Malate dehydrogenase; Pyruvate dehydrogenase; Glucose-6-phosphate dehydrogenase; 6-phosphogluconate dehydrogenase; Citrate synthase; Aldehyde dehydrogenase; Phosphatidate phosphatase; Glycerol-3-phosphate/dihydroxyacetone phosphate sn-1
 25 acyltransferase; Lysophospholipid; and Lysophosphatidic acid acyltransferase.

In some embodiments, the invention relates to a transformed cell comprising a genetic modification, wherein the genetic modification is transformation with a nucleic acid comprising a nucleotide sequence that encodes Acetyl-CoA carboxylase ("ACC," Gene NG160, SEQ ID:2), ATP-citrate lyase, subunit 1("ACL1," Gene NG161, SEQ ID: 4), ATP-
 30 citrate lyase, subunit 2 ("ACL2," Gene NG162, SEQ ID NO:6), Acetyl-coA synthetase isoform 1 ("ACSI," Gene NG163, SEQ ID NO:8), Acetyl-coA synthetase isoform 2, ("ACS2," Genes NG164, NG165, SEQ ID NOs:10,12), Diacylglycerol acyltransferase type 2 ("DGA1," Gene NG167, SEQ ID NO:16), Diacylglycerol acyltransferase type 1

("DGA2," Gene NG168, SEQ ID NO:18), Glycerol-3-phosphate dehydrogenase ("GPD1,"
 Genes NG169, NG170, SEQ ID NOs:20, 22), Malic enzyme ("MAE1," Gene NG171, SEQ
 ID NO:24), 1-acyl-sn-glycerol-3-phosphate acyltransferase ("SLC1," Gene NG177, SEQ
 ID NO:36), Phospholipid:diacylglycerol acyltransferase ("LRO1," Gene NG178, SEQ ID
 5 NO:38), Glycerol kinase ("GUT1," Gene NG182, SEQ ID NO:46), Glycerol-3-phosphate
 acyl transferase ("SCT1," Gene NG183, SEQ ID NO:48), Phosphatidate phosphatase
 ("PAH1," Gene NG218, SEQ ID NO:50), Fatty acid synthase, subunit alpha ("FAS2,"
 Gene NG186, SEQ ID NO:54), Fatty acid synthase, subunit beta ("FAS1," Gene NG187,
 SEQ ID NO:56), Type 2A-related protein phosphatase ("SIT4," Gene NG188, SEQ ID
 10 NO:58), Regulatory subunit of type 1 protein phosphatase Glc7p ("REG1," Gene NG189,
 SEQ ID NO:60), Regulatory subunit of Sit4 phosphatase ("SAP190," Gene NG190, SEQ
 ID NO:62), Dephospho-CoA kinase ("DPCK," Gene NG191, SEQ ID NO:64), HAC1 basic
 leucine zipper transcription factor ("HAC1," Gene NG192, SEQ ID NO:66), CAT8 zinc
 finger transcription factor ("CAT8," Gene NG193, SEQ ID NO:68), Ubiquitin ("UBI4,"
 15 Gene NG194, SEQ ID NO:70), STB5 transcription factor ("STB5," Gene NG199, SEQ ID
 NO:80), Pyruvate carboxylase ("PYC1," Genes NG201, NG202 SEQ ID NOs:84, 86),
 Pyruvate decarboxylase ("PDC1," Gene NG203, SEQ ID NO:88), Malate dehydrogenase 1
 ("MDH1," Gene NG204, SEQ ID NO:90), Malate dehydrogenase 2 ("MDH2," Genes
 NG205, NG206 SEQ ID NOs:92, 94), Pyruvate dehydrogenase alpha ("PDA1," Gene
 20 NG207, SEQ ID NO:96), Pyruvate dehydrogenase beta ("PDB1," Genes NG208, NG209
 SEQ ID NOs:98, 100), Glucose-6-phosphate dehydrogenase ("ZWF1," Genes NG210,
 NG211 SEQ ID NOs:102, 104), 6-phosphogluconate dehydrogenase ("GND1," Gene
 NG212, SEQ ID NO:106), Citrate synthase 1 ("CIT1," Genes NG213, NG214, SEQ ID
 NOs:108, 110), Citrate synthase 2 ("CIT2," Genes NG215, NG216 SEQ ID NOs:112, 114),
 25 Aldehyde dehydrogenase ("ALD6," Gene NG217, SEQ ID NO:116), Glycerol-3-
 phosphate/dihydroxyacetone phosphate sn-1 acyltransferase ("GPT2," Gene NG219, SEQ
 ID NO:167), Lysophospholipid acyltransferase ("SLC4," Gene NG220, SEQ ID NO:169),
 or Lysophosphatidic acid acyltransferase ("LOA1," Genes NG221, NG222 SEQ ID
 NOs:171, 173).

30 In some embodiments, the invention relates to a transformed cell comprising a
 genetic modification, wherein the genetic modification is transformation with a nucleic acid
 comprising a nucleotide sequence having at least 70%, 72%, 64%, 76%, 78%, 80%, 82%,
 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%,

99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with a nucleotide sequence set forth in SEQ ID NO: 2, 4, 6, 8, 10, 12, 16, 18, 20, 22, 24, 36, 38, 46, 48, 50, 54, 56, 58, 60, 62, 64, 66, 68, 70, 80, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 167, 169, 171, or 173. In some embodiments, the invention relates to a transformed cell comprising a genetic modification, wherein the genetic modification is transformation with a nucleic acid that encodes an amino acid sequence having at least 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with an amino acid sequence set forth in SEQ ID NO: 1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 35, 37, 45, 47, 49, 53, 55, 57, 59, 61, 63, 65, 67, 69, 79, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 166, 168, 170, or 172.

In some embodiments, the invention relates to a method of modifying the lipid content and/or composition of a cell, comprising transforming a parent cell with a nucleic acid comprising a nucleotide sequence having at least 70%, 72%, 74%, 76%, 78%, 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with a nucleotide sequence set forth in SEQ ID NO: 2, 4, 6, 8, 10, 12, 16, 18, 20, 22, 24, 36, 38, 46, 48, 50, 54, 56, 58, 60, 62, 64, 66, 68, 70, 80, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 167, 169, 171, or 173. In some embodiments, the invention relates to a method of modifying the lipid content and/or composition of a cell, comprising transforming a parent cell with a nucleic acid encoding an amino acid sequence having at least 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with an amino acid sequence set forth in SEQ ID NO: 1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 35, 37, 45, 47, 49, 53, 55, 57, 59, 61, 63, 65, 67, 69, 79, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 166, 168, 170, or 172.

I. Decreasing the activity of a protein to increase or modify the lipid content of a cell

In some embodiments, the invention relates to a transformed *Axula adenivorans* cell comprising a genetic modification that decreases the activity of a protein selected from Multi-functional enzyme involved in B-oxidation ("MFE1," Gene NG172, SEQ ID NO:26), Triacylglycerol lipase ("TGL3," Gene NG174, SEQ ID NO:30), Triacylglycerol lipase ("TGL3/4," Gene NG175, SEQ ID NO:32), Triacylglycerol lipase ("TGL4," Gene NG176, SEQ ID NO:34), Glycerol-3-phosphate dehydrogenase ("GUT2," Gene NG179, SEQ ID

NO:40), AMP-activated kinase ("SNF1," Gene NG180, SEQ ID NO:42), Acyl-CoA oxidase ("POX," Gene NG181, SEQ ID NO:44), Peroxisomal membrane E3 ubiquitin ligase ("PEX10," Gene NG196, SEQ ID NO:74), Dihydrolipoamide acetyltransferase component (E2) of pyruvate dehydrogenase complex ("LAT1," Gene NG197, SEQ ID NO:76), Transcription factor involved in glucose repression ("MIG1," Gene NG198, SEQ ID NO:78), 2-methylcitrate dehydratase ("PHD1," Gene NG200, SEQ ID NO:82), Phosphatidyl-ethanolamine methyltransferase ("CHO2," Gene NG226, SEQ ID NO:165), Methylene-fatty-acyl-phospholipid synthase ("OPI3," Gene NG223, SEQ ID NO:175), and INO4 transcription factor ("INO4," Gene NG225, SEQ ID NO: 179). In some
 10 embodiments, the genetic modification is a knockout mutation.

K. Increasing or decreasing the activity of a protein to modify lipid content and/or composition

In some embodiments, the invention relates to a plasmid vector comprising a gene encoding a protein selected from Delta-12 fatty acid desaturase ("D12," Gene NG166, SEQ ID NO:14), Delta-9 fatty acid desaturase ("OLE1," Gene NG173, SEQ ID NO:28), Fatty acid elongase 2 ("ELO2," Gene NG185, SEQ ID NO:52), and Fatty acid elongase 1 ("ELO1," Gene NG195, SEQ ID NO:72).

In some embodiments, the invention relates to a vector comprising a nucleotide sequence having at least 70%, 72%, 64%, 76%, 78%, 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with a nucleotide sequence set forth in SEQ ID NO: 14, 28, 52, or 72. In some embodiments, the invention relates to a vector that encodes an amino acid sequence having at least 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with an amino acid sequence set forth in SEQ ID NO: 13, 27, 51, or 71. In some embodiments, the vector is a plasmid.

In some embodiments, the invention relates to a transformed cell comprising a genetic modification that increases the activity of a protein selected from a Delta-12 fatty acid desaturase, Delta-9 fatty acid desaturase, and Fatty acid elongase.

In some embodiments, the invention relates to a transformed cell comprising a genetic modification, wherein the genetic modification is transformation with a nucleic acid comprising a nucleotide sequence that encodes Delta-12 fatty acid desaturase ("D12," Gene NG166, SEQ ID NO:14), Delta-9 fatty acid desaturase ("OLE1," Gene NG173, SEQ ID

NO:28), Fatty acid elongase 2 ("ELO2," Gene NG185, SEQ ID NO:52), or Fatty acid elongase 1 ("ELO1," Gene NG195, SEQ ID NO:72).

In some embodiments, the invention relates to a transformed cell comprising a genetic modification, wherein the genetic modification is transformation with a nucleic acid comprising a nucleotide sequence having at least 70%, 72%, 64%, 76%, 78%, 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with a nucleotide sequence set forth in SEQ ID NO: 14, 28, 52, or 72. In some embodiments, the invention relates to a transformed cell comprising a genetic modification, wherein the genetic modification is transformation with a nucleic acid encoding an amino acid sequence having at least 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with an amino acid sequence set forth in SEQ ID NO: 13, 27, 51, or 71.

In some embodiments, the invention relates to a method of modifying the lipid content and/or composition of a cell, comprising transforming a parent cell with a nucleic acid comprising a nucleotide sequence having at least 70%, 72%, 64%, 76%, 78%, 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with a nucleotide sequence set forth in SEQ ID NO: 14, 28, 52, or 72. In some embodiments, the invention relates to a method of modifying the lipid content and/or composition of a cell, comprising transforming a parent cell with a nucleic acid encoding an amino acid sequence having at least 80%, 82%, 84%, 85%, 87%, 88%, 90%, 92%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% sequence homology with an amino acid sequence set forth in SEQ ID NO: 13, 27, 51, or 71.

In some embodiments, the invention relates to a transformed *Arxula adenitivorans* cell comprising a genetic modification that decreases the activity of a protein selected from Delta-12 fatty acid desaturase ("D12," Gene NG166, SEQ ID NO:14), Delta-9 fatty acid desaturase ("OLE1," Gene NG173, SEQ ID NO:28), Fatty acid elongase 2 ("ELO2," Gene NG185, SEQ ID NO:52), and Fatty acid elongase 1 ("ELO1," Gene NG195, SEQ ID NO:72). In some embodiments, the genetic modification is a knockout mutation.

The present description is further illustrated by the following examples, which should not be construed as limiting in any way. The contents of all cited references (including literature references, issued patents, published patent applications and GenBank

Accession numbers as cited throughout this application) are hereby expressly incorporated by reference. When definitions of terms in documents that are incorporated by reference herein conflict with those used herein, the definitions used herein govern.

5 EXEMPLIFICATION

Example 1: Sequencing the *Arxula adeninivorans* genome and selecting genetic targets

In order to engineer an *A. adeninivorans* strain with optimal lipid content and composition, native *A. adeninivorans* genes (genetic targets) were screened for a positive or negative effect on lipid production in *A. adeninivorans*.

10 First, in order to access DNA and protein sequences of selected targets, the genome of *A. adeninivorans* strain NS252 (ATCC 76597) was sequenced and annotated by Synthetic Genomics Inc. (CA, USA). Strain NS252 was also selected as a host for screening the genetic targets.

Genetic targets that affect lipid production efficiency and composition were
15 enumerated based on published data assessing lipid pathways in oleaginous yeast and other organisms (Ratledge & Wynn, *Advances Applied Microbiology* 51:1-51 (2002); Tai & Stephanopoulos, *Metabolic Engineering* 15:1-9 (2013), Beopoulos et al., *Applied & Environmental Microbiology*, 74:7779-89 (2008); Beopoulos et al., *Progress Lipid Research*, 48:375-87 (2009); Courchesne et al., *J. Biotechnology* 141:31-41 (2009); Morin
20 et al., *PLoS One*, 6(11):e27966 (2011); Bozaquel-Morais et al., *PLoS One*, 5(10):e13692 (2010)). The sequences of the genetic targets selected for screening in *A. adeninivorans* are shown in SEQ ID NO:1-116 and 164-179, and listed in Table 2 below.

Table 2.

25 *Arxula adeninivorans* genes that may be upregulated or downregulated in *Arxula adeninivorans* and other organisms to increase the organism's lipid content, modify its lipid composition, or both.

Gene	ID	Function	DNA	
			SEQ ID	Modification
			NO	
ACC	NG160	Acetyl-CoA carboxylase	2	Increase activity to increase lipid content
ACLI	NG161	ATP-citrate lyase, subunit 1	4	Increase activity to increase lipid content

ACL2	NG162	ATP-citrate lyase, subunit 2	6	<i>Increase activity to increase lipid content</i>
ACS1	NG163	Acetyl-coA synthetase isoform 1	8	<i>Increase activity to increase lipid content</i>
ACS2	NG164, NG165*	Acetyl-coA synthetase isoform 2	10, 12*	<i>Increase activity to increase lipid content</i>
D12	NG166	Delta-12 fatty acid desaturase	14	<i>Increase or decrease activity to modify lipid composition</i>
DGA1	NG167	Diacylglycerol acyltransferase type 2	16	<i>Increase activity to increase lipid content</i>
DGA2	NG168	Diacylglycerol acyltransferase type 1	18	<i>Increase activity to increase lipid content</i>
GPD1	NG169, NG170*	Glycerol-3-phosphate dehydrogenase	20, 22*	<i>Increase activity to increase lipid content</i>
MAE1	NG171	Malic enzyme	24	<i>Increase activity to increase lipid content</i>
MFE1	NG172	Multi-functional enzyme involved in B-oxidation	26	<i>Decrease activity to increase lipid content</i>
OLE1	NG173	Delta-9 fatty acid desaturase	28	<i>Increase or decrease activity to modify lipid composition</i>
TGL3	NG174	Triacylglycerol lipase	30	<i>Decrease activity to increase lipid content</i>
TGL3/4	NG175	Triacylglycerol lipase	32	<i>Decrease activity to increase lipid content</i>
TGL4	NG176	Triacylglycerol lipase	34	<i>Decrease activity to increase lipid content</i>
SLC1	NG177	1-acyl-sn-glycerol-3-phosphate acyltransferase	36	<i>Increase activity to increase lipid content</i>
LRO1	NG178	Phospholipid:diacylglycerol acyltransferase	38	<i>Increase activity to increase lipid content</i>
GUT2	NG179	Glycerol-3-phosphate dehydrogenase	40	<i>Decrease activity to increase lipid content</i>

SNF1	NG180	AMP-activated kinase	42	<i>Decrease activity to increase lipid content</i>
POX	NG181	Acyl-CoA oxidase	44	<i>Decrease activity to increase lipid content</i>
GUT1	NG182	Glycerol kinase	46	<i>Increase activity to increase lipid content</i>
SCT1	NG183	Glycerol-3-phosphate acyl transferase	48	<i>Increase activity to increase lipid content</i>
PAH1	NG218	Phosphatidate phosphatase	50	<i>Increase activity to increase lipid content</i>
ELO2	NG185	Fatty acid elongase	52	<i>Increase or decrease activity to modify lipid composition</i>
FAS2	NG186	Fatty acid synthase, subunit alpha	54	<i>Increase activity to increase lipid content</i>
FAS1	NG187	Fatty acid synthase, subunit beta	56	<i>Increase activity to increase lipid content</i>
SIT4	NG188	Type 2A-related protein phosphatase	58	<i>Increase activity to increase lipid content</i>
REG1	NG189	Regulatory subunit of type 1 protein phosphatase Glc7p	60	<i>Increase activity to increase lipid content</i>
SAP190	NG190	Regulatory subunit of Sit4 phosphatase	62	<i>Increase activity to increase lipid content</i>
DPCK	NG191	Dephospho-CoA kinase	64	<i>Increase activity to increase lipid content</i>
HAC1	NG192	Basic leucine zipper transcription factor	66	<i>Increase activity to increase lipid content</i>
CAT8	NG193	Zinc finger transcription factor	68	<i>Increase activity to increase lipid content</i>
UBI4	NG194	Ubiquitin	70	<i>Increase activity to increase lipid content</i>
ELO1	NG195	Fatty acid elongase	72	<i>Increase or decrease activity to modify lipid composition</i>

		Peroxisomal		
PEX10	NG196	membrane E3 ubiquitin ligase	74	<i>Decrease activity to increase lipid content</i>
		Dihydrolipoamide acetyltransferase		
LAT1	NG197	component (E2) of pyruvate dehydrogenase complex	76	<i>Decrease activity to increase lipid content</i>
MIG1	NG198	Transcription factor involved in glucose repression	78	<i>Decrease activity to increase lipid content</i>
STB5	NG199	Transcription factor	80	<i>Increase activity to increase lipid content</i>
PHD1	NG200	2-methylcitrate dehydratase	82	<i>Decrease activity to increase lipid content</i>
PYC1	NG201, NG202*	Pyruvate carboxylase	84, 86*	<i>Increase activity to increase lipid content</i>
PDC1	NG203	Pyruvate decarboxylase	88	<i>Increase activity to increase lipid content</i>
MDH1	NG204	Malate dehydrogenase	90	<i>Increase activity to increase lipid content</i>
MDH2	NG205, NG206*	Malate dehydrogenase	92, 94*	<i>Increase activity to increase lipid content</i>
PDA1	NG207	Pyruvate dehydrogenase alpha	96	<i>Increase activity to increase lipid content</i>
PDB1	NG208, NG209*	Pyruvate dehydrogenase beta	98, 100*	<i>Increase activity to increase lipid content</i>
ZWF1	NG210, NG211*	Glucose-6-phosphate dehydrogenase	102, 104*	<i>Increase activity to increase lipid content</i>
GND1	NG212	6-phosphogluconate dehydrogenase	106	<i>Increase activity to increase lipid content</i>
CIT1	NG213, NG214*	Citrate synthase	108, 110*	<i>Increase activity to increase lipid content</i>

CIT2	NG215,	Citrate synthase	112,	<i>Increase activity to increase lipid content</i>
	NG216*		114*	
ALD6	NG217	Aldehyde dehydrogenase	116	<i>Increase activity to increase lipid content</i>
		Phosphatidyl-ethanolamine methyltransferase		
CHO2	NG226	Glycerol-3-phosphate/	165	<i>Decrease activity to increase lipid content</i>
		dihydroxyacetone phosphate sn-1 acyltransferase		
GPT2	NG219	Lysophospholipid acyltransferase	167	<i>Increase activity to increase lipid content</i>
		Lysophosphatidic acid acyltransferase		
SLC4	NG220	Methylene-fatty-acyl-phospholipid synthase	169	<i>Increase activity to increase lipid content</i>
LOA1	NG221,	Transcription factor involved in phospholipid synthesis	171, 173*	<i>Increase activity to increase lipid content</i>
	NG222*			
OPI3	NG223,	involved in phospholipid synthesis	175, 177*	<i>Decrease activity to increase lipid content</i>
	NG224*			
INO4	NG225	involved in phospholipid synthesis	179	<i>Decrease activity to increase lipid content</i>

*denotes nucleotide sequences without introns, corresponding to *Arxula adeninivorans* genes that contain introns.

The genetic targets were either up-regulated or down-regulated to increase lipid production and/or modify lipid composition depending on the target's role in the lipid pathway. For example, the targets involved in the final step of triacylglycerol synthesis, such as NG167(DGA1) and NG168(DGA2) were overexpressed, while targets involved in peroxisomal lipid degradation, such as NG173(TGL3) and NG181(POX), were down-regulated by deletion. The targets were applied individually or in combination.

For changing lipid composition, the same target can be either up-regulated or down-regulated depending on the desired composition. For example, the combination of NG173 (delta-9 fatty acid desaturase "OLE1") overexpression and NG166 (delta-12 fatty acid

desaturase “D12”) down-regulation or deletion could result in an increased oleic acid content in *A. adeninivorans*. Alternatively, overexpression of both NG173 and NG166 could result in an increase of linoleic acid content.

Example 2: Increasing the activity of a target protein

5 The map of the expression construct used to overexpress the up-regulation targets is shown in Figure 1 with the NG167 target (type 2 diacylglycerol acyltransferase “DGA1”) as an example. The constructs for all other up-regulated targets were the same except for the target ORFs. The overexpression constructs were assembled by the yeast mediated ligation method. The constructs were linearized by PmeI/AscI restriction digest before
10 transformation into NS252 by the transformation procedure described by Chen (Applied Microbiology & Biotechnology 48:232-35 (1997)). The linear expression construct included an expression cassette for a genetic target and for the *Streptomyces noursei* Nat1 gene used as marker for selection with nourseothricin (NAT).

 The up-regulated targets that contain introns were expressed in two forms: native
15 genes amplified from the genomic DNA of NS252 strain and intron-free synthetic cDNA synthesized by Genscript (NJ, USA).

Example 3: Decreasing the activity of a target protein

 The map of the deletion construct used to delete the down-regulation targets is shown on the Figure 2 with the NG166 target (delta-12 fatty acid desaturase “D12”) as an
20 example. The deletion constructs for all other down-regulation targets were the same except for the target’s recombination flanks that were specific to the targets. The deletion constructs were assembled by the yeast mediated ligation method, and they included an expression cassette for the *Streptomyces noursei* Nat1 gene used as marker for selection with NAT flanked by upstream and downstream DNA regions of the genetic targets. The
25 recombination flanks were either amplified from the genomic DNA of NS252 strain or synthesized by Integrated DNA Technologies, Inc. (Iowa, USA) as part of the oligonucleotides used for the yeast mediated ligation plasmid assembly.

Example 4: Screening strategies

 Both overexpression and deletion transformants were selected on YPD plates with
30 50 µg/ml NAT. Additionally, the transformants were analyzed by a fluorescent lipid assay described below to screen for increased lipid content. To screen for an optimized lipid composition, the transformants were analyzed by gas chromatography using standard methods for the lipid composition analysis.

Example 5: Modifying multiple targets

After a first round of screening targets, the best transformants are transformed again with the overexpression or deletion construct that contains either the same target (for overexpression) or different targets for further increasing lipid production or optimizing lipid composition. In the second round of screening, a similar approach is used as in the first round, except for the selection marker. The *Escherichia coli hph* gene expression cassette can be used as a marker for selection with hygromycin B for the second round of screening. For combinations of more than two targets, additional dominant markers can be used, or alternatively, a marker recycling procedure can be used similar to the approaches used with other yeast species (Akada et al., Yeast, 23:399-05 (2006); Sauer, Biotechniques 16:1086-88 (1994); Hartzog et al., Yeast, 22:789-98 (2005)).

Example 6: Optimizing target activity

After the optimal combination of native *A. adeninivorans* genetic targets is determined, each target in the selected set can be further optimized. For example, in the case of gene deletion, when complete gene removal has a negative effect on an organisms' growth rate or another culturing parameter, a target can be integrated back into the NS252 genome under a weak or inducible promoter to express a minimal amount of the target protein. In such cases, decreasing the target's expression or timing its expression at particular bioprocess stages can be more beneficial for the overall product yield than complete removal of the target.

In the case of an overexpression target, each selected target can be further optimized by screening heterologous targets with the same function from different organisms for better expression or specific activity. The targets can also be optimized by protein engineering, by introducing mutations that improve protein expression or activity. Another optimization approach for the overexpression targets involves screening regulatory elements such as promoters that are more optimal for a particular target.

Example 7: Lipid assay

Each well of an autoclaved, multi-well plate was filled with filter-sterilized media containing 0.5 g/L urea, 1.5 g/L yeast extract, 0.85 g/L casamino acids, 1.7 g/L YNB (without amino acids and ammonium sulfate), 100 g/L glucose, and 5.11 g/L potassium hydrogen phthalate (25 mM). 1.5 mL of media was used per well for 24-well plates and 300 µl of media was used per well for 96-well plates. Alternatively, the yeast cultures were used to inoculate 50 mL of sterilized media in an autoclaved 250 mL flask. Yeast strains

that had been incubated for 1-2 days on YPD-agar plates at 30°C were used to inoculate each well of the multiwell plate.

Multi-well plates were covered with a porous cover and incubated at 30°C, 70-90% humidity, and 900 rpm in an Infors Multitron ATR shaker. Alternatively, flasks were covered with aluminum foil and incubated at 30°C, 70-90% humidity, and 900 rpm in a New Brunswick Scientific shaker. After 96 hours, 20 µL of 100% ethanol was added to 20 µL of cells in an analytical microplate and incubated at 4°C for 30 minutes. 20 µL of cell/ethanol mix was then added to 80 µL of a pre-mixed solution containing 50 µL 1 M potassium iodide, 1 mM µL Bodipy 493/503, 0.5 µL 100% DMSO, 1.5 µL 60% PEG 4000, and 27 µL water in a Costar 96-well, black, clear-bottom plate and covered with a transparent seal. Bodipy fluorescence was monitored with a SpectraMax M2 spectrophotometer (Molecular Devices) kinetic assay at 30°C, and normalized by dividing fluorescence by absorbance at 600 nm.

Example 8: Method to overexpress the DGA1 protein (DGAT2 gene) in yeast

Nucleic acid constructs for overexpressing the DGA1 gene were described in USSN 61/943,664 and PCT Patent Application No. US15/017227 (both hereby incorporated by reference). Figure 3 shows expression construct pNC243 used for overexpression of the *R. toruloides* DGA1 gene NG66 (SEQ ID NO: 122). DGA1 expression constructs were linearized before transformation by a PacI/NotI restriction digest. The linear expression constructs each included an expression cassette for the *DGAT2* gene and for the *NatI* gene, used as a marker for selection with nourseothricin (NAT).

DGA1 expression constructs were randomly integrated into the genome of *Y. lipolytica* strain NS18 (obtained from ARS Culture Collection, NRRL# YB 392) using a transformation protocol as described in Chen (Applied Microbiology & Biotechnology 48:232-35 (1997)). Transformants were selected on YPD plates with 500 µg/mL NAT and screened for the ability to accumulate lipids by a fluorescent staining lipid assay as described in Example 7 above. For each expression construct, eight transformants were analyzed.

For most constructs, there was significant colony variation between the transformants, likely due to the lack of a functional DGA1 expression cassette in cells that only obtained a functional *NatI* cassette, or due to a negative effect of the site of DGA1 integration on DGA1 expression. Nevertheless, all transformants had a significant increase in lipid content.

Overexpression of native *Y. lipolytica* DGA1 (NG15) under a strong promoter increased the transformant lipid content by about 2-fold compared to the parental strain NS18 as measured by the fluorescence assay described in Example 7. Transformants that demonstrated the highest fluorescence (about 3-fold higher compared to NS18) were generated by the overexpression of *R. toruloides* DGA1 (NG66, NG67) and *L. starkeyi* DGA1 (NG68).

In certain experiments, the effect of native *R. toruloides* DGA1 (NG49) overexpression on lipid production was not as high as the effect of synthetic versions of *R. toruloides* DGAT2 genes that did not contain introns. This result may indicate that the gene splicing of the *R. toruloides* DGAT2 gene in *Y. lipolytica* was not very efficient. In certain experiments, codon optimization of the *R. toruloides* DGA1 gene for expression in *Y. lipolytica* did not have a positive effect on lipid production.

Similar strategies may be used to express the *A. adeninivorans* DGAT2 gene or other *A. adeninivorans* gene in *Yarrowia lipolytica* or another species. Type 2 diacylglycerol acyltransferases from *Yarrowia lipolytica* (SEQ ID NO:118), *Rhodospiridium toruloides* (SEQ ID NO:120), and *A. adeninivorans* (SEQ ID NO:52) were expressed in *A. adeninivorans* strain NS252, and the effect of each gene on various C16/C18 lipid fractions is shown in Figure 4.

Example 9: Analysis and screening of yeast that overexpress DGA1

In order to select strains with the highest lipid production level, *Y. lipolytica* strain NS18 transformants expressing NG15 (*Y. lipolytica* DGA1) or NG66 (*R. toruloides* DGA1) were screened. For NG15, about 50 colonies were screened by a lipid assay for the highest lipid accumulation, and the best transformant was named NS249. For NG66, 80 colonies were screened, and the 8 best colonies were selected for further analysis.

Strain NS249 and the 8 selected NG66 transformants were grown in shake flasks and analyzed by the lipid assay for lipid content and by HPLC for glucose consumption. *Y. lipolytica* strains overexpressing *R. toruloides* DGA1 had significantly higher lipid contents than *Y. lipolytica* strains with a native *Y. lipolytica* DGAT2 gene expressed under the same promoter as *R. toruloides* DGAT2. At the same time, NG66 transformants used significantly more glucose than NS249, demonstrating that NG66 was more efficient in converting glucose to lipids. The difference in efficiency between the two DGAT2 genes may be attributed to either a higher level of expression of *R. toruloides* DGA1 in *Y. lipolytica* or a higher level of *R. toruloides* DGA1 specific activity, or both.

Strain NS125 is a derivative of *Y. lipolytica* strain NS18 (obtained from ARS Culture Collection, NRRL# YB 392) that was transformed with a *Y. lipolytica* DGA1 expression cassette from the pNC104 vector (Figure 5). The pNC104 construct was linearized by a PacI/NotI restriction digest prior to transformation. The linear expression
 5 construct included the expression cassette for the *Y. lipolytica* DGAT2 gene and for the *Nat1* gene used as a marker for selection with nourseothricin (NAT). The expression construct was randomly integrated into the genome of *Y. lipolytica* strain NS18 using the transformation protocol as described in Chen (Applied Microbiology & Biotechnology 48:232-35 (1997)). Transformants were selected on YPD plates with 500 µg/mL NAT and
 10 screened for the ability to accumulate lipids by a fluorescent staining lipid assay as described in Example 7. The best transformant out of about 100 transformants screened was named NS125.

The NS281 strain was obtained using a similar process as strain NS125, except that the pNC243 construct was used for the transformation of the NS18 strain (Figure 3). The
 15 pNC243 construct contained the *Rhodospiridium toruloides* DGAT2 gene (NG66) instead of *Y. lipolytica* DGAT2 gene used to make NS125. The NS281 strain contains a *Rhodospiridium toruloides* DGAT2 gene that is integrated into the *Y. lipolytica* genome.

Example 10: Method to overexpress the DGA2 protein (DGAT1 gene)

Several DGAT1 gene sequences that encode DGA2 proteins were selected from
 20 public databases and their native intron-free cDNAs were synthesized (GenScript, USA). DGAT1 genes from six selected donors (*Yarrowia lipolytica*, *Rhodospiridium toruloides*, *Lipomyces starkeyi*, *Aspergillus terreus*, *Claviceps purpurea*, *Chaetomium globosum*) were expressed in *Y. lipolytica* strains NS125 and NS281, which each overexpress a different DGA1 protein. Thus, the method provides an example of increasing the activity of two
 25 proteins in the same cell.

NS125 and NS281 strains each were transformed with a set of DGA2 expression constructs. Figure 6 shows the map of the expression construct used to overexpress a *Y. lipolytica* DGAT1 gene (NG16) in the NS125 and NS281 strains. The constructs for all
 30 other DGAT1 genes used were the same except for the DGAT1 ORF. DGAT1 constructs were linearized before transformation by a PacI/AscI restriction digest. The linear expression construct included an expression cassette for the DGAT1 gene and for a BLE gene used as a marker for selection with Zeocin (NG76 ZEO).

For each expression construct, 15 transformants were analyzed by the fluorescent lipid assay described below. In this experiment, the presence of heterologous *DGAT1* in *Y. lipolytica* was not confirmed by colony PCR. For most constructs, there was a significant colony variation between transformants, probably due to the lack of a functional *DGAT1* expression cassette in some transformants that only obtained a functional ZEO cassette or due to negative effect of the *DGAT1* expression cassette site of integration on DGA2 expression. Nevertheless, most of the tested *DGAT1* genes increased the lipid content in *Y. lipolytica* beyond the effect displayed by a *DGAT2* gene alone. The positive effect of DGA2 overexpression was more noticeable in the NS125 strain background, probably because the NS125 strain, which overexpresses its native DGA1, was less oleaginous than the NS281 strain, which overexpresses a more-highly active DGA1. In both the NS125 and NS281 strains, the NG112 *DGAT1* gene (*Claviceps purpurea* DGA2) was the most efficient in increasing the transformants' lipid content, possibly due to a higher expression level and/or higher specific activity compared to other *DGAT1* genes tested.

Similar strategies may be used to express the *A. adenivorans* *DGAT1* gene or other *A. adenivorans* gene in *Yarrowia lipolytica* or another species.

Example 11: Method to reduce triacylglycerol lipase activity in yeast

The *TGL3* gene was deleted in *Y. lipolytica* wild-type strain NS18 (obtained from NRLL# YB-392) and its DGA1 overexpressing derivative NS281. NS281 overexpresses the DGA1 gene from *Rhodospiridium toruloides* as described above. The *Y. lipolytica* *TGL3* gene (YALI0D17534g, SEQ ID NO: 148) was deleted as follows: A two-fragment deletion cassette was amplified by PCR from a plasmid containing the hygromycin resistance gene ("*hph*," SEQ ID NO: 152) using primer pairs NP1798-NP656 and NP655-NP1799 (SEQ ID NOs: 157, 154, 153, 158). The resulting PCR fragments (SEQ ID NOs: 160 & 161) were co-transformed into NS18 and NS281 according to the protocol developed in USSN 61/819,746 and PCT Patent Application Publication No. WO14/182657 (both hereby incorporated by reference). The omission of a promoter and terminator in the *hph* cassette and the splitting of the *hph* coding sequence into two PCR fragments reduce the probability that random integration of these pieces will confer hygromycin resistance. The *hph* gene should only be expressed if it integrates at the *TGL3* locus by homologous recombination so that the *TGL3* promoter and terminator can direct its transcription. Hygromycin resistant colonies were screened by PCR to confirm the absence of *TGL3* and the presence of a *tgl3::hyg* specific product.

Similar strategies may be used to reduce triacylglycerol lipase activity, or the activity of another protein, in *A. adeninivorans* (see Example 13).

Example 12: Method to reduce glycerol-3-phosphate dehydrogenase activity in yeast

The *GUT2* gene, which encodes glycerol-3-phosphate dehydrogenase, was deleted
 5 from *Y. lipolytica* wild-type strain NS18 (obtained from NRLL# YB-392). The *Y. lipolytica* *GUT2* gene (YALI0B13970g, SEQ ID NO: 150) was deleted as follows: A two-fragment deletion cassette was amplified by PCR from a plasmid containing the hygromycin resistance gene ("*hph*," SEQ ID NO: 152) using primer pairs NP1563-NP656 and NP655-
 NP1800 (SEQ ID NOs: 156, 154, 153, and 159, respectively). The resulting PCR fragments
 10 (SEQ ID NOs: 162 & 163) were co-transformed into NS18. The omission of a promoter and terminator in the *hph* cassette and the splitting of the *hph* coding sequence into two PCR fragments reduce the probability that random integration of these pieces will confer hygromycin resistance.

Yarrowia lipolytica strain NS18 was grown overnight on a YPD plate at 30 °C. 1
 15 mL of water was applied to the plate to collect the cells using a L-shaped spreader and pipet. Cells were diluted to OD₆₀₀ = 0.5 in two flasks each containing 25 mL YPD and shaken at 30 °C for 30 min to acclimate the cells to liquid culture. 95 mg hydroxyurea was added to a final concentration of 50 mM. Shaking was continued for two hours to allow for cell cycle arrest as determined by microscopy (all cells were arrested at the budded stage).
 20 Cells were collected by centrifugation, washed with water, and resuspended in a pellet volume of water. 50 µL was aliquoted per transformation, cells were collected by centrifugation, and the supernatant was discarded. The cells were resuspended in 100 µL transformation mix (80 µL of filter sterilized 60% PEG 4000; 5 µL of filter sterilized 2 M lithium acetate, pH adjusted to 6.0; 5 µL of filter sterilized 2 M dithiothreitol (DTT); 10 µL
 25 of 2 mg/mL salmon sperm DNA, boiled 10 min prior to use). 9 µL of unpurified PCR product for each of SEQ ID NOs: 162 & 163 was added. The cells were subjected to heat shock at 39 °C for 1 h, collected by centrifugation, resuspended in 1 mL YPD, and cultured overnight at 30 °C to allow *hph* gene expression. 100 µL of cells was spread onto dry selective plates (YPD/agar containing 300 µg/mL hygromycin) the next day, and
 30 transformants were screened for growth on glycerol a day later.

Hygromycin resistant colonies were patched onto minimal media containing glucose or glycerol to screen for isolates that had lost the ability to grow on glycerol due to loss of

GUT2 function. Seven correct integrations were obtained when the transformed cells were first arrested in S phase with hydroxyurea (48 colonies screened).

Similar strategies may be used to reduce glycerol-3-phosphate dehydrogenase activity, or the activity of another protein, in *A. adenivorans* and other species.

5 Example 13: Method to reduce triacylglycerol acyltransferase activity in yeast

In order to test the idea that lipase deletion leads to higher lipid accumulation in *A. adenivorans*, *TGL3* and *TGL4* were deleted in *A. adenivorans* wild-type strain NS252 (obtained from ATCC, #76597). The *A. adenivorans* *TGL3* (SEQ ID NO: 30), *TGL4* (SEQ ID NO: 34), and *$\Delta 12$ desaturase* (SEQ ID NO:14) genes were deleted as described in
 10 PCT Patent Application Publication No. WO14/182657 (hereby incorporated by reference). Deletion of *TGL3* in NS252 resulted in strain NS458. Deletion of *TGL4* in NS252 resulted in strain NS481. Deletion of *$\Delta 12$* in NS252 resulted in strain NS478. These strains were grown using a batch glucose process in 50-mL flasks and lipid content was analyzed by a fluorescence-based assay after 96 hours. Specifically, each flask was filled with 50 mL of
 15 filter-sterilized media containing 0.5 g/L urea, 1.5 g/L yeast extract, 0.85 g/L casamino acids, 1.7 g/L YNB (without amino acids and ammonium sulfate), 100 g/L glucose, and 5.11 g/L potassium hydrogen phthalate (25 mM). Yeast strains that had been incubated for 1-2 days on YPD-agar plates at 30°C were used to inoculate each flask. After 96 hours of incubation with agitation at 30°C, 20 μ L of 100% ethanol was added to 20 μ L of cells in an
 20 analytical microplate and incubated at 4°C for 30 minutes. 20 μ L of cell/ethanol mix was then added to 80 μ L of a pre-mixed solution containing 50 μ L 1 M potassium iodide, 1 μ L 1 mM Bodipy 493/503, 0.5 μ L 100% DMSO, 1.5 μ L 60% PEG 4000, and 27 μ L water in a Costar 96-well, black, clear-bottom plate and covered with a transparent seal. Bodipy fluorescence was monitored with a SpectraMax M2 spectrophotometer (Molecular Devices)
 25 kinetic assay at 30°C, and normalized by dividing fluorescence by absorbance at 600 nm. Data was averaged in duplicate growth experiments. Both deletion strains showed increased lipid content compared to the parent strain (Figure 7).

Example 14: Expressing *A. adenivorans* $\Delta 9$ desaturase in *Y. lipolytica*.

The native $\Delta 9$ desaturase of *Y. lipolytica* uses both C16 and C18 saturated fatty
 30 acids as substrates. *Arxula adenivorans* $\Delta 9$ enzyme was tested for higher C18 specificity by introducing the gene (SEQ ID NO:28) as the sole $\Delta 9$ activity in *Y. lipolytica*. This was achieved by first deleting *Y. lipolytica* $\Delta 9$ in NS18 to produce strain NS418. NS418 required supplementation with unsaturated fatty acids for growth due to the absence of $\Delta 9$

activity. The exogenous *A9* gene was then inserted into the native locus through targeted integration and selected for the ability to grow without supplementation. Expression of *A. adeninivorans A9* in the absence of the native enzyme resulted in a switch in substrate specificity to overwhelmingly C18:0 substrate, thus reducing C16:1 content to minimal levels (Figure 8).

Example 15: Expressing *A. adeninivorans* SCT1 in *Y. lipolytica*

The glycerol-3-phosphate acyltransferase SCT1 from *A. adeninivorans* was expressed in a *Y. lipolytica* strain containing a knockout of the native *Y. lipolytica* SCT1 gene. Specifically, the glycerol acyltransferase SCT1 was deleted in *Y. lipolytica* strain NS18 to produce strain NS563, resulting in a lower accumulation of lipids (Figure 9). The SCT1 gene from *A. adeninivorans* (SEQ ID NO:48) was then expressed in a strain NS563 to produce strain NS804, which produced a *Y. lipolytica* strain that could accumulate more lipids than either strain NS18 or NS563 (Figure 9).

Example 16: Decreasing the activity of various proteins in *A. adeninivorans*

The native *MFE1*, *TGL3*, *A12*, *PEX10*, *POX* and *TGL4* genes were individually knocked out of wild type *A. adeninivorans* strain 252 using methods similar to those described in PCT Patent Application Publication No. WO14/182657 (hereby incorporated by reference), and the percentage of various C16 and C18 fatty acids, relative to total C16 and C18 fatty acids, for each knockout is shown in Figure 10.

Example 17: Increasing the activity of various proteins in *A. adeninivorans* and *Y. lipolytica*

Eight *Arxula adeninivorans* genes involved in lipid synthesis (Table 3) were selected for overexpression in *Arxula adeninivorans* strain NS252 and *Yarrowia lipolytica* strain NS18. A map of the expression construct used to overexpress the up-regulation targets is shown in Figure 1, with *Arxula adeninivorans* DGA1 (NG167) as example. The constructs for all other targets were the same, except for the open reading frame. The overexpression constructs were assembled by yeast mediated ligation. The constructs were linearized with a PmeI/AscI restriction digest before transformation. Linear expression constructs included each expression cassette for a genetic target and the *Streptomyces noursei* *Nat1* gene, used as marker for selection with nourseothricin (NAT).

Table 3. Eight *Arxula adeninivorans* genes involved in lipid synthesis that were expressed in *Arxula adeninivorans* and *Yarrowia lipolytica*.

Gene ID	SEQ ID NO	Gene Name	Function
NG167	16	DGA1	diacylglycerol acyltransferase type 2
NG168	18	DGA2	diacylglycerol acyltransferase type 1
NG170	22	GPD1	glycerol-3-phosphate dehydrogenase
NG177	36	SLC1	1-acyl-sn-glycerol-3-phosphate acyltransferase
NG183	48	SCT1	glycerol-3-phosphate acyltransferase
NG211	104	ZWF1	glucose-6-phosphate dehydrogenase
NG212	106	GND1	6-phosphogluconate dehydrogenase
NG218	50	PAH1	phosphatidate phosphatase

The expression constructs were randomly integrated into the genome of *Y. lipolytica* strain NS18 (obtained from ARS Culture Collection, NRRL# YB 392) by the transformation protocol described in Chen et al. (Applied Microbiology & Biotechnology 48:232-35 (1997)). The expression constructs were integrated into *A. adeninivorans* (ATCC #76597) using a protocol specifically adapted to *A. adeninivorans*. Briefly, 2 mL of YPD was inoculated with the parent *A. adeninivorans* culture and grown overnight at 37°C in a rotary shaker. 0.5 mL of the overnight culture was then used to inoculate 25 mL of fresh YPD in a 250 flask, which was then grown at 37°C for 3.5 to 4 hours. The cells were pelleted at 3000 rpm for 2 minutes, and the supernatant was discarded. The cells were washed in sterile water, and pelleted again at 3000 rpm for 2 minutes. The pellet was suspended in 2 mL of 100 mM lithium acetate comprising 40 µM dithiothreitol and transferred into a microcentrifuge tube. The suspension was incubated for one hour at 37°C on a rotary shaker. The cells were pelleted at 10,000 rpm for 10 seconds and the supernatant was discarded. The cells were resuspended in 1 mL of water with gentle pipetting, centrifuged again at 10,000 rpm for 10 seconds, and the water was discarded. The cells were washed by pipetting with 1 M cold sorbitol, then centrifuged again at 10,000 rpm for 10 seconds, and the supernatant discarded. 2 mL of cold 1 M sorbitol was added to the pellet, and the tube was placed on ice. 40 µL of the cells were then added to pre-chilled 0.2 cm electroporation cuvettes along with 5 µL of DNA. The cells were electroporated at 25 µF, 200 ohms, 1.5 kV, with a ~4.9-5.0 ms time constant. The cells were added to 1 mL

of YPD, incubated at 37°C overnight, and 100 µL to 500 µL of cells were plated onto YPD agar.

Y. lipolytica transformants were selected on YPD plates with 500 µg/mL NAT. *A. adeninivorans* transformants were selected on YPD plates with 50 µg/mL NAT. The transformants for both yeast were screened for the ability to accumulate lipids by the fluorescence lipid assay described above. For each expression construct, eight transformants were analyzed. Parental strains NS18 and NS252 were used as negative controls and analyzed in 4 replicates. The results of the lipid assays are shown in Figure 11. Figure 11 demonstrates that *A. adeninivorans* DGA1 and DGA2 have a noticeable effect on lipids at 72 hrs of sampling. At 96 hrs, all eight selected genes increased the lipid contents of *A. adeninivorans*. Interestingly, 4 targets (DGA1, DGA2, SLC1, PAH1), encoding for enzymes involved into the last 3 steps of TAG synthesis, had the largest effect on lipid content in *A. adeninivorans*. In *Y. lipolytica*, only DGA1 and DGA2 had a noticeable positive effect on lipids at 96 hrs.

Example 18: Increasing the activity of ELO1 in *A. adeninivorans*

The *A. adeninivorans* *Δ12* desaturase gene (SEQ ID NO:14) was knocked out of *A. adeninivorans*, resulting in strain NS554, and the *A. adeninivorans* *ELO1* gene (SEQ ID NO:72) was overexpressed in the *Δ12* knockout background, resulting in strain NS631. The total C16 and C18 lipids were quantified for each strain, and Figure 12 shows that the *A. adeninivorans* *ELO1* gene increased the fraction of C18 lipids in the cell relative to the parent strain NS554.

INCORPORATION BY REFERENCE

All of the U.S. patents and U.S. published patent applications cited herein are hereby incorporated by reference.

EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

What is claimed is:

1. A nucleic acid, comprising a nucleotide sequence encoding a protein from *Arxula adeninivorans*, wherein said protein comprises an amino acid sequence having at least 95% sequence homology with the amino acid sequence set forth in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 164, 166, 168, 170, 172, 174, 176, or 178, or a biologically-active portion of any one of them.
2. The nucleic acid of claim 1, wherein said protein comprises the amino acid sequence set forth in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 164, 166, 168, 170, 172, 174, 176, or 178, or a biologically-active portion of any one of them.
3. The nucleic acid of claim 1 or 2, wherein said nucleotide sequence has at least 70% sequence homology with the nucleotide sequence set forth in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 165, 167, 169, 171, 173, 175, 177, or 179.
4. The nucleic acid of claim 3, wherein said nucleotide sequence comprises the nucleotide sequence set forth in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 165, 167, 169, 171, 173, 175, 177, or 179.
5. The nucleic acid of any one of claims 1-4, further comprising a promoter, wherein the nucleotide sequence and promoter are operably linked.
6. The nucleic acid of claim 5, wherein said promoter is not an *Arxula adeninivorans* promoter.
7. A vector, comprising the nucleic acid of any one of claims 1-6, wherein said vector is a plasmid.
8. A method for modifying the lipid content and/or composition of a cell, comprising transforming the cell with the nucleic acid of any one of claims 1-6 or the vector of claim 7.

9. The method of claim 8, wherein the cell is selected from the group consisting of algae, bacteria, molds, fungi, plants, and yeasts.
10. The method of claim 9, wherein said cell is selected from the group consisting of *Arxula*, *Aspergillus*, *Aurantiochytrium*, *Candida*, *Claviceps*, *Cryptococcus*, *Cunninghamella*, *Geotrichum*, *Hansenula*, *Kluyveromyces*, *Kodamaea*, *Leucosporidiella*, *Lipomyces*, *Mortierella*, *Ogataea*, *Pichia*, *Prototheca*, *Rhizopus*, *Rhodospiridium*, *Rhodotorula*, *Saccharomyces*, *Schizosaccharomyces*, *Tremella*, *Trichosporon*, *Wickerhamomyces*, and *Yarrowia*.
11. The method of claim 10, wherein said cell is selected from the group consisting of *Arxula adeninivorans*, *Aspergillus niger*, *Aspergillus oryzae*, *Aspergillus terreus*, *Aurantiochytrium limacinum*, *Candida utilis*, *Claviceps purpurea*, *Cryptococcus albidus*, *Cryptococcus curvatus*, *Cryptococcus ramirezgomezianus*, *Cryptococcus terreus*, *Cryptococcus wieringae*, *Cunninghamella echinulata*, *Cunninghamella japonica*, *Geotrichum fermentans*, *Hansenula polymorpha*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Kodamaea ohmeri*, *Leucosporidiella creatinivora*, *Lipomyces lipofer*, *Lipomyces starkeyi*, *Lipomyces tetrasporus*, *Mortierella isabellina*, *Mortierella alpina*, *Ogataea polymorpha*, *Pichia cifferri*, *Pichia guilliermondii*, *Pichia pastoris*, *Pichia stipites*, *Prototheca zopfii*, *Rhizopus arrhizus*, *Rhodospiridium babjevae*, *Rhodospiridium toruloides*, *Rhodospiridium paludigenum*, *Rhodotorula glutinis*, *Rhodotorula mucilaginosa*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Tremella encephala*, *Trichosporon cutaneum*, *Trichosporon fermentans*, *Wickerhamomyces cifferri*, and *Yarrowia lipolytica*.
12. The method of claim 11, wherein said cell is *Yarrowia lipolytica* or *Arxula adeninivorans*.
13. A transformed cell, comprising a genetic modification, wherein said genetic modification is transformation with the nucleic acid of any one of claims 1-6 or the vector of claim 7.
14. A transformed cell, comprising a genetic modification, wherein said genetic modification increases the activity of a protein selected from an Acetyl-CoA carboxylase, ATP-citrate lyase, subunit 1, ATP-citrate lyase, subunit 2, Acetyl-coA synthetase isoform 1, Acetyl-CoA carboxylase, ATP-citrate lyase, subunit 1, ATP-citrate lyase, subunit 2, Acetyl-coA synthetase isoform 1, Acetyl-coA synthetase isoform 2, Diacylglycerol acyltransferase

type 2, Diacylglycerol acyltransferase type 1, Glycerol-3-phosphate dehydrogenase, Malic enzyme, 1-acyl-sn-glycerol-3-phosphate acyltransferase, Phospholipid:diacylglycerol acyltransferase, Glycerol kinase, Glycerol-3-phosphate acyl transferase, Phosphatidate phosphatase, Fatty acid synthase, subunit alpha, Fatty acid synthase, subunit beta, Type 2A-related protein phosphatase, Regulatory subunit of type 1 protein phosphatase Glc7p, Regulatory subunit of Sit4 phosphatase, Dephospho-CoA kinase, HAC1 basic leucine zipper transcription factor, CAT8 zinc finger transcription factor, Ubiquitin, STB5 transcription factor, Pyruvate carboxylase, Pyruvate decarboxylase, Malate dehydrogenase, Pyruvate dehydrogenase, Glucose-6-phosphate dehydrogenase, 6-phosphogluconate dehydrogenase, Citrate synthase, Aldehyde dehydrogenase, Phosphatidate phosphatase, Glycerol-3-phosphate/dihydroxyacetone phosphate sn-1 acyltransferase, Lysophospholipid, and Lysophosphatidic acid acyltransferase.

15. The transformed cell of claim 13 or 14, wherein said cell is selected from the group consisting of algae, bacteria, molds, fungi, plants, and yeasts.

16. The transformed cell of claim 15, wherein said cell is selected from the group consisting of *Arxula*, *Aspergillus*, *Aurantiochytrium*, *Candida*, *Claviceps*, *Cryptococcus*, *Cunninghamella*, *Geotrichum*, *Hansenula*, *Kluyveromyces*, *Kodamaea*, *Leucosporidiella*, *Lipomyces*, *Mortierella*, *Ogataea*, *Pichia*, *Prototheca*, *Rhizopus*, *Rhodospiridium*, *Rhodotorula*, *Saccharomyces*, *Schizosaccharomyces*, *Tremella*, *Trichosporon*, *Wickerhamomyces*, and *Yarrowia*.

17. The transformed cell of claim 16, wherein said cell is selected from the group consisting of *Arxula adeninivorans*, *Aspergillus niger*, *Aspergillus oryzae*, *Aspergillus terreus*, *Aurantiochytrium limacinum*, *Candida utilis*, *Claviceps purpurea*, *Cryptococcus albidus*, *Cryptococcus curvatus*, *Cryptococcus ramirezgomezianus*, *Cryptococcus terreus*, *Cryptococcus wieringae*, *Cunninghamella echinulata*, *Cunninghamella japonica*, *Geotrichum fermentans*, *Hansenula polymorpha*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Kodamaea ohmeri*, *Leucosporidiella creatinivora*, *Lipomyces lipofer*, *Lipomyces starkeyi*, *Lipomyces tetrasporus*, *Mortierella isabellina*, *Mortierella alpina*, *Ogataea polymorpha*, *Pichia ciferrii*, *Pichia guilliermondii*, *Pichia pastoris*, *Pichia stipites*, *Prototheca zopfii*, *Rhizopus arrhizus*, *Rhodospiridium babjevae*, *Rhodospiridium toruloides*, *Rhodospiridium paludigenum*, *Rhodotorula glutinis*, *Rhodotorula mucilaginosa*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Tremella*

enchepala, *Trichosporon cutaneum*, *Trichosporon fermentans*, *Wickerhamomyces ciferrii*, and *Yarrowia lipolytica*.

18. The transformed cell of claim 17, wherein said cell is *Yarrowia lipolytica* or *Arxula adeninivorans*.

19. The transformed cell of claim 18, wherein said cell is *Arxula adeninivorans*, further comprising a second genetic modification, wherein:

said second genetic modification decreases the activity of a protein that comprises the amino acid sequence set forth in SEQ ID NO: 13, 25, 27, 29, 31, 33, 39, 41, 43, 51, 71, 73, 75, 77, 81, 164, 174, 176, or 178; or

said second genetic modification decreases the activity of a protein that is encoded by the nucleic acid sequence set forth in SEQ ID NO: 14, 26, 28, 30, 32, 34, 40, 42, 44, 52, 72, 74, 76, 78, 82, 165, 175, 177, or 179.

20. The transformed cell of claim 19, wherein said second genetic modification is a knockout mutation.

21. An *Arxula adeninivorans* cell, comprising a genetic modification, wherein said genetic modification decreases the activity of a protein selected from delta-12 fatty acid desaturase, multi-functional β -oxidation enzyme, delta-9 fatty acid desaturase, triacylglycerol lipase, glycerol-3-phosphate dehydrogenase, AMP-activated kinase, acyl-CoA oxidase, fatty acid elongase, peroxisomal membrane E3 ubiquitin ligase, dihydrolipoamide acetyltransferase component (E2) of pyruvate dehydrogenase complex, MIG1 transcription factor involved in glucose repression, 2-methylcitrate dehydratase, phosphatidyl-ethanolamine methyltransferase, methylene-fatty-acyl-phospholipid synthase, and INO4 transcription factor.

22. The cell of claim 21, wherein said genetic modification is a knockout mutation.

23. The cell of claim 21 or 22, wherein said protein comprises the amino acid sequence set forth in SEQ ID NO: 13, 25, 27, 29, 31, 33, 39, 41, 43, 51, 71, 73, 75, 77, 81, 164, 174, 176, or 178.

24. The cell of any one of claims 21-23, wherein said protein is encoded by the nucleic acid sequence set forth in SEQ ID NO: 14, 26, 28, 30, 32, 34, 40, 42, 44, 52, 72, 74, 76, 78, 82, 165, 175, 177, or 179.

25. The cell of any one of claims 21-24, further comprising a second genetic modification, wherein said second genetic modification is transformation with the nucleic acid of any one of claims 1-6 or the vector of claim 7.

Figure 1

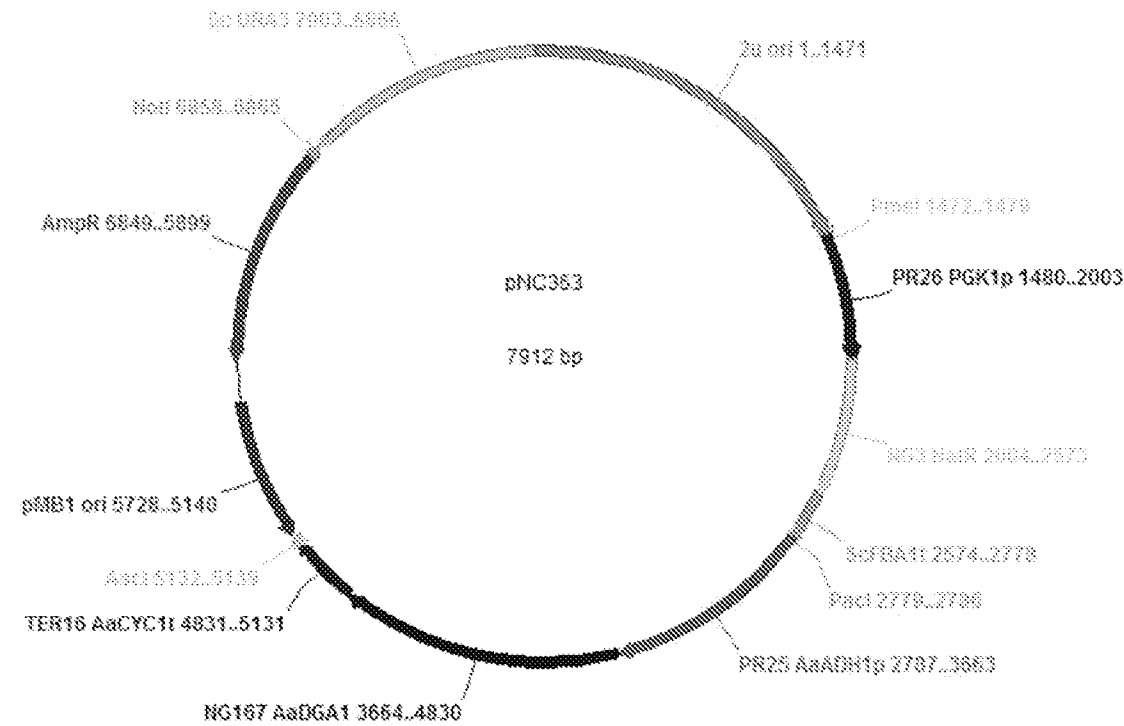


Figure 2

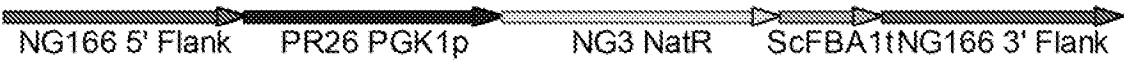


Figure 3

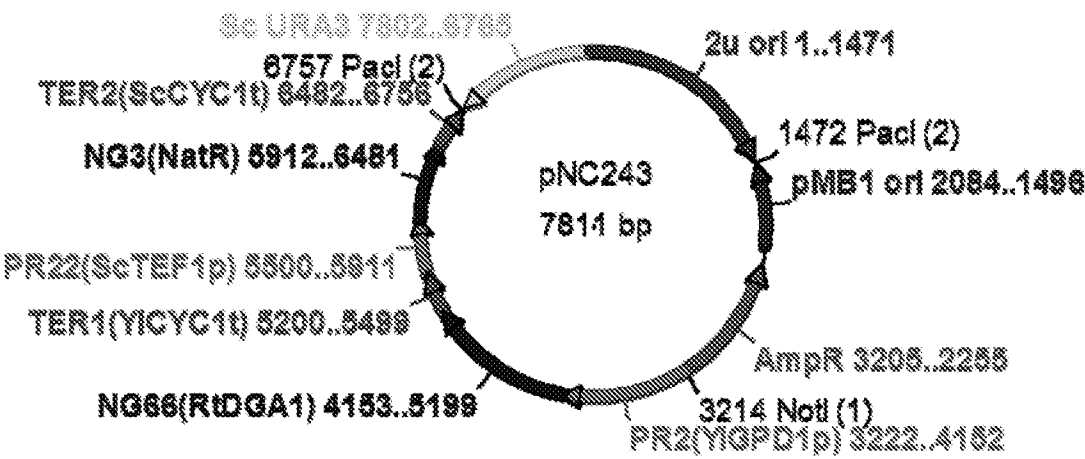


Figure 4

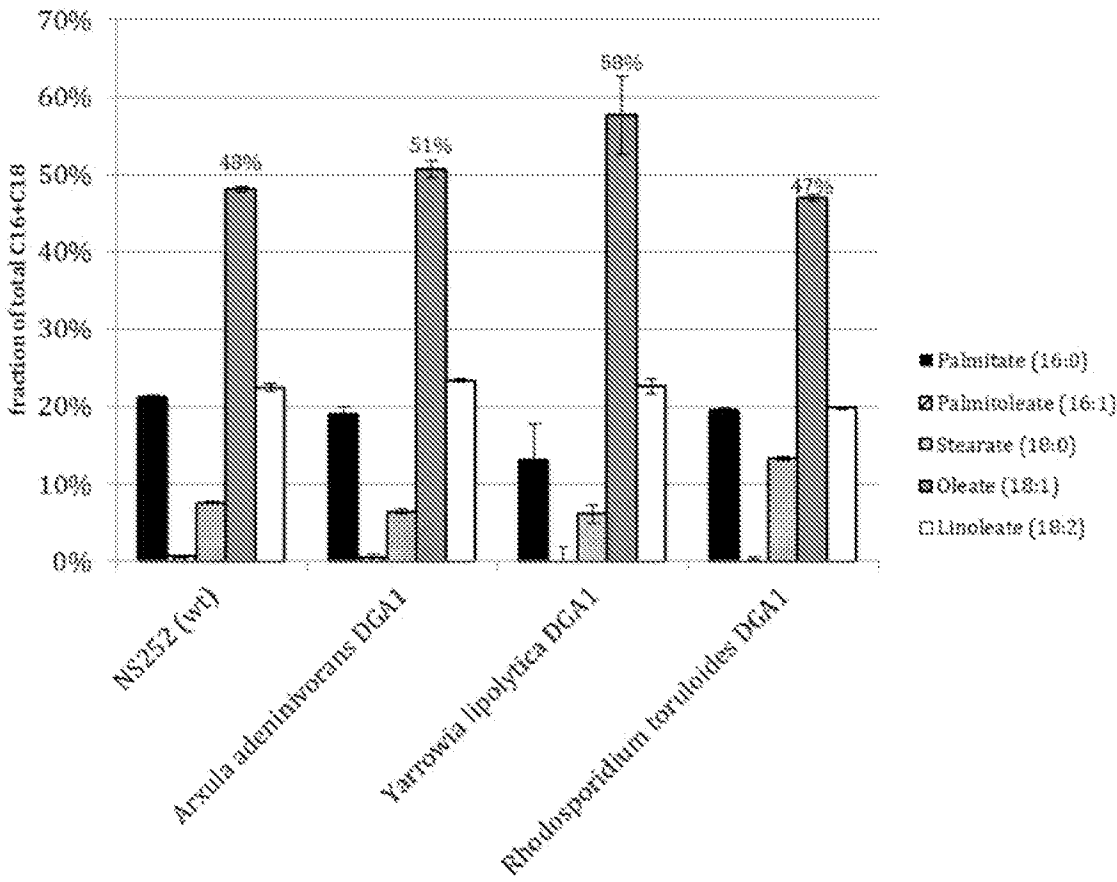


Figure 5

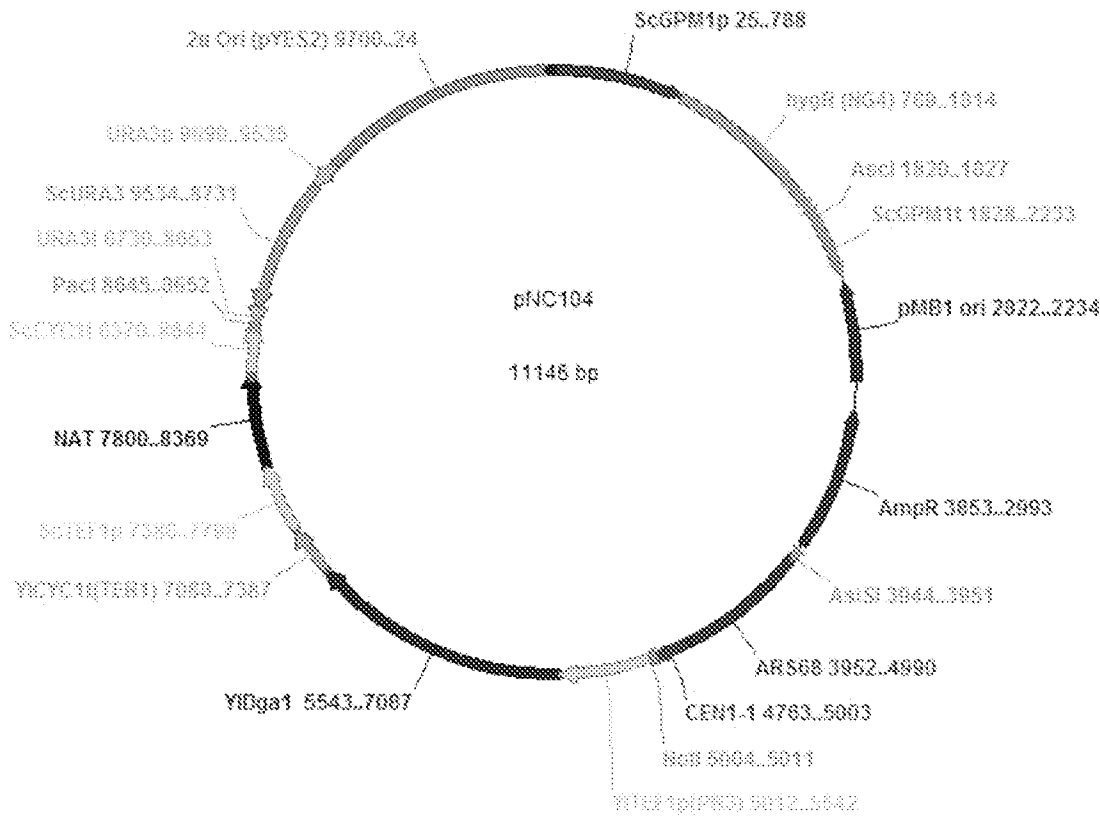


Figure 6

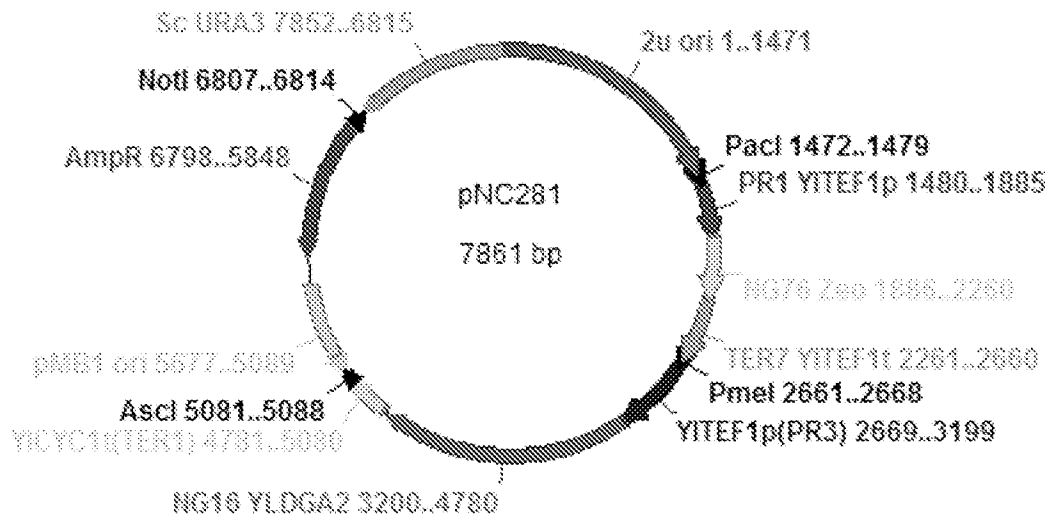


Figure 7

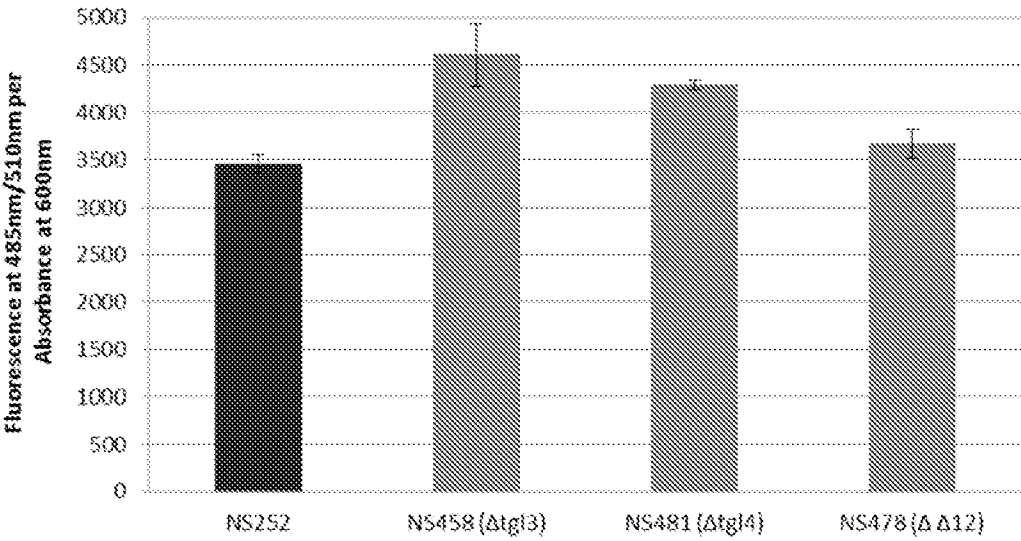


Figure 8

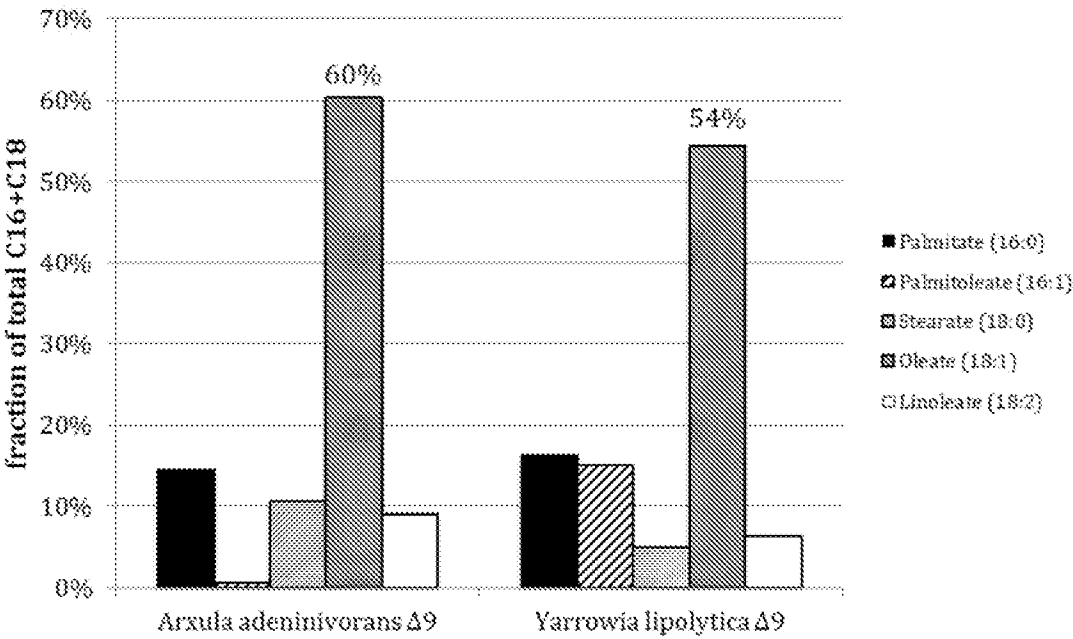


Figure 9

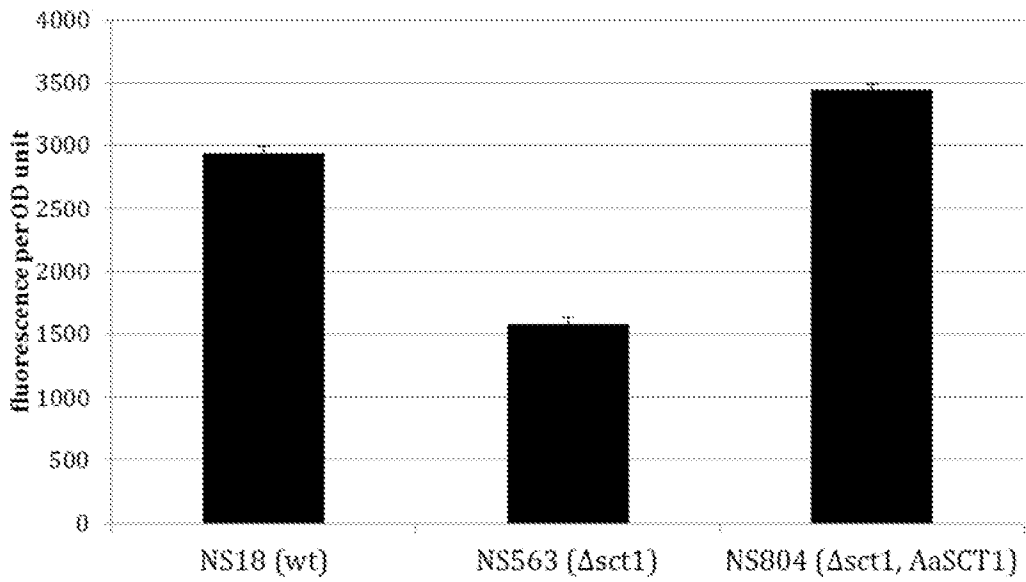


Figure 10

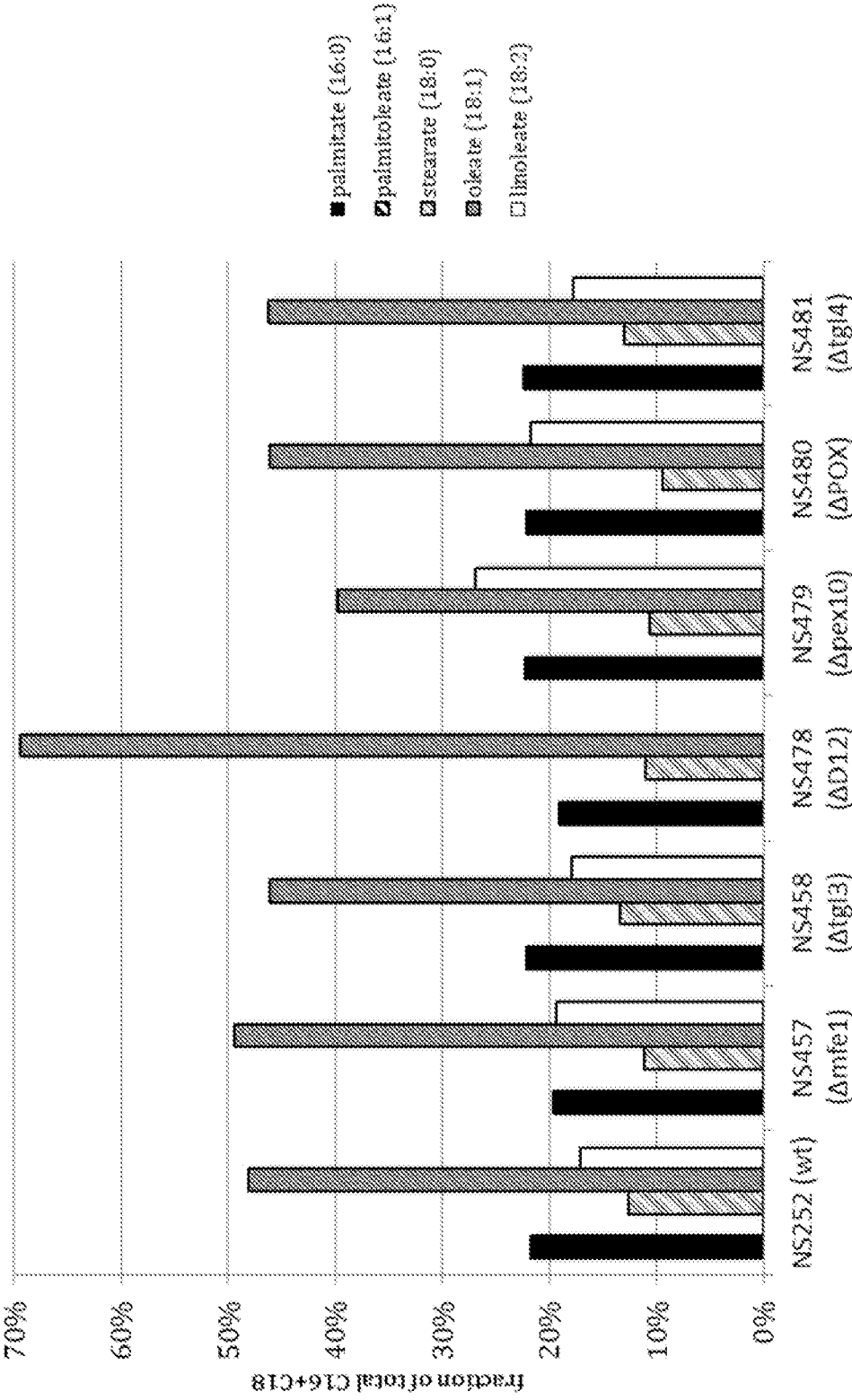
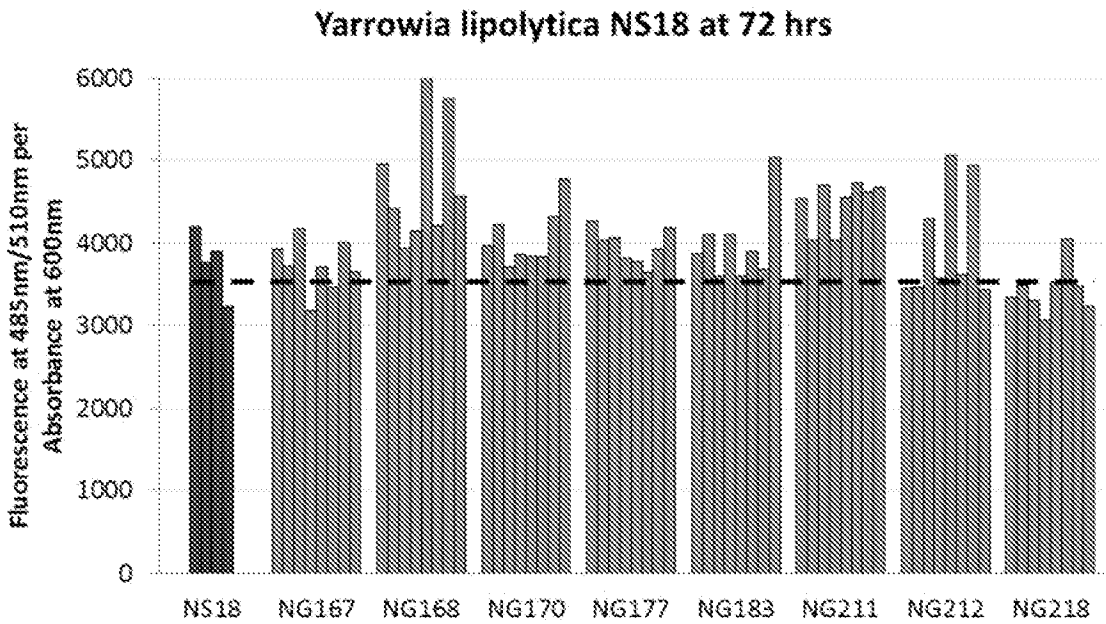


Figure 11

A



B

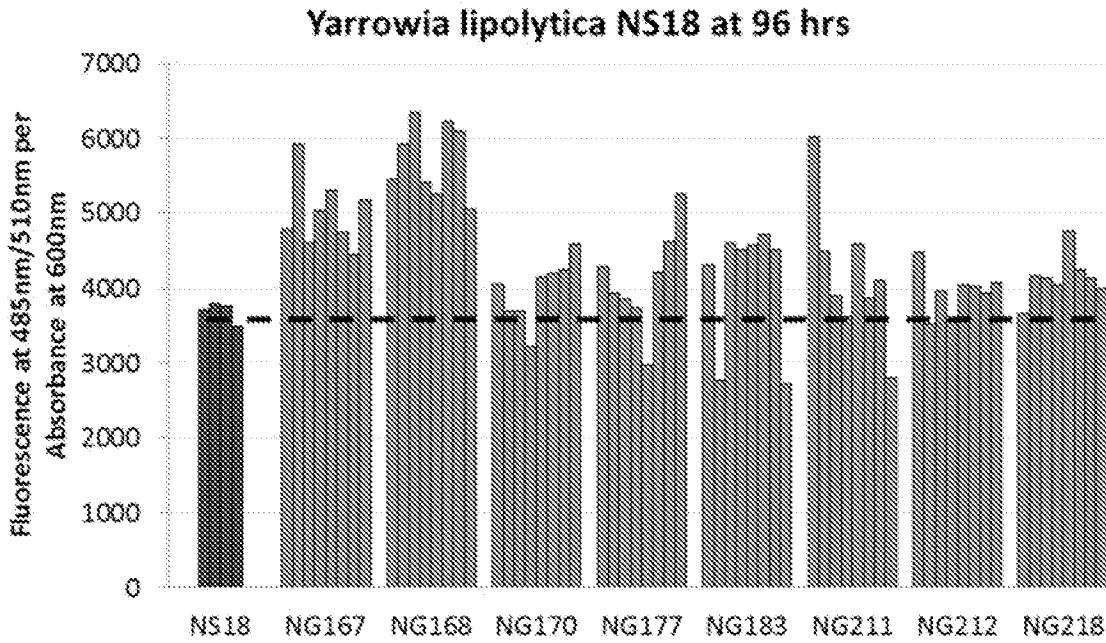
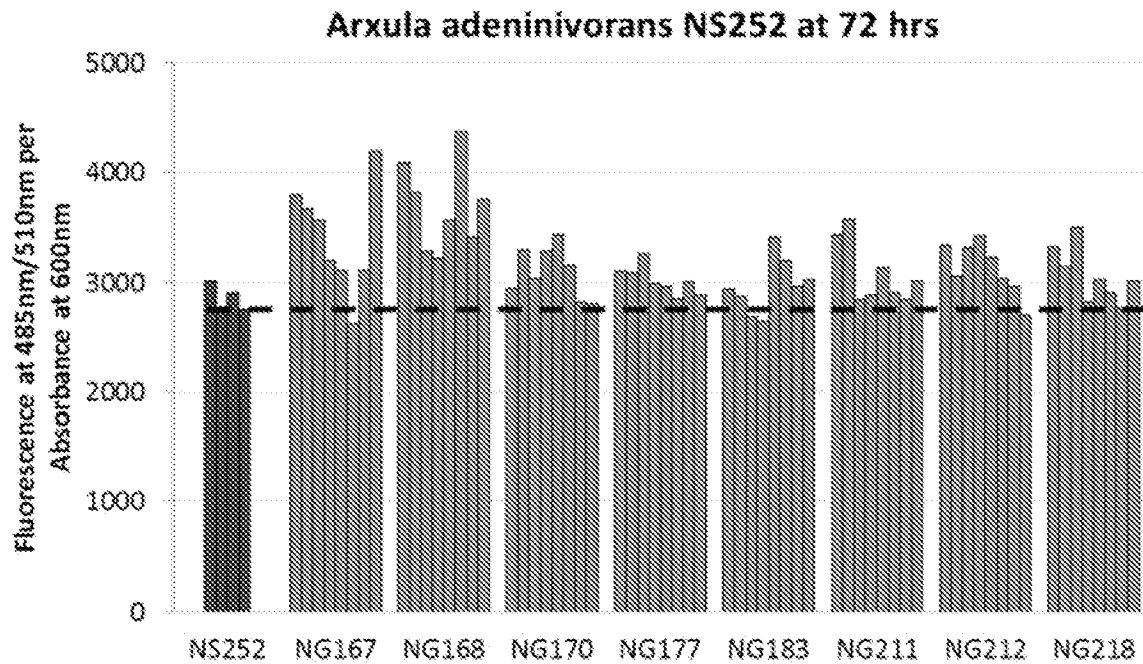


Figure 11 (continued)

C



D

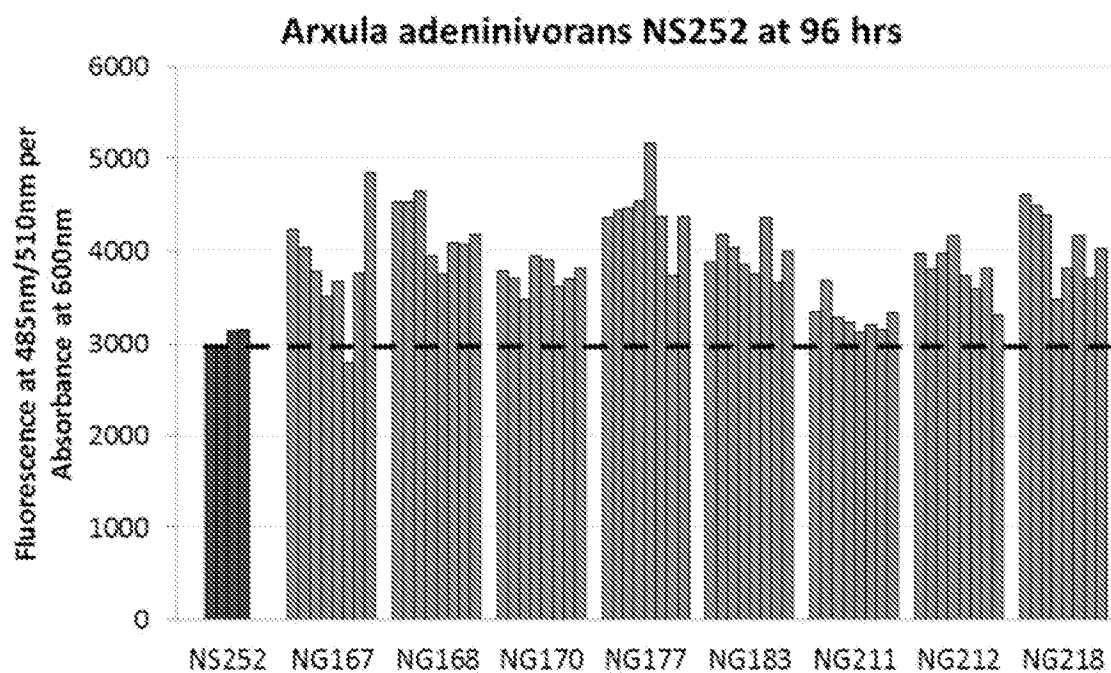
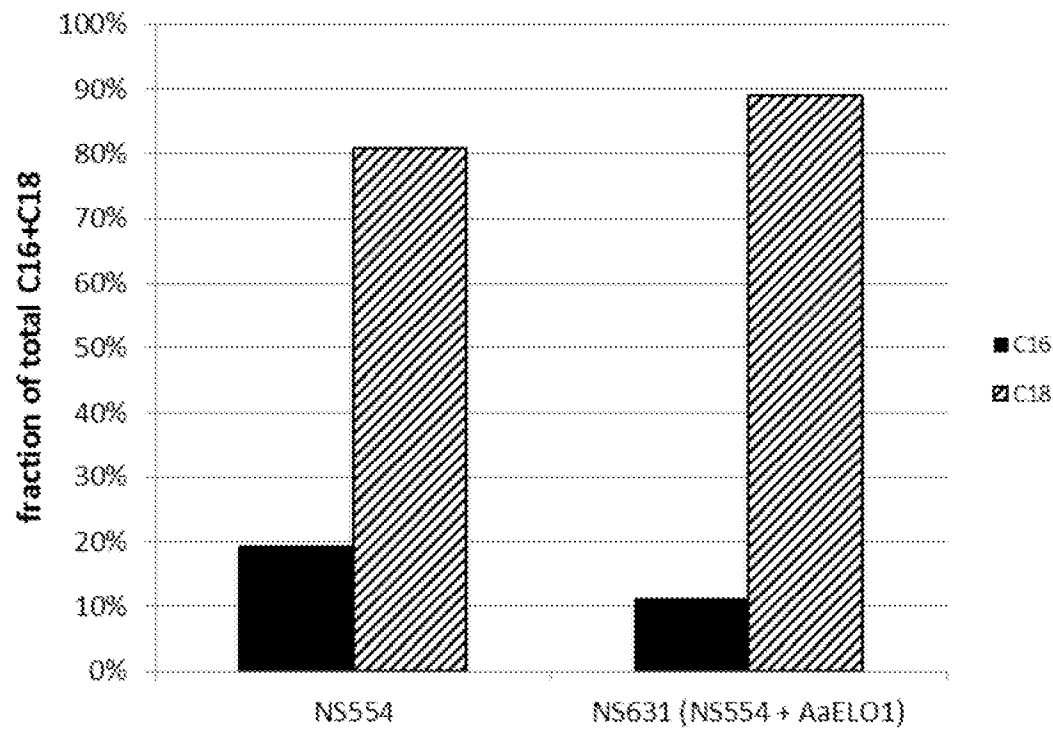


Figure 12



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US15/33211

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12N 1/16, 15/87, C12P 7/64 (2015.01)

CPC - C12N 1/16, 15/87, C12P 7/64

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC(8): C12P 7/64 C12N 1/19 C07C 57/00, C12N 15/87 (2015.01)

CPC: C12N1/16, C12P 7/64 C07C 57/00, C12N 15/87

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PatSeer (US, EP, WO, JP, DE, GB, CN, FR, KR, ES, AU, IN, CA, INPADOC Data); PubMed; NCBI Blast; Google; Google Scholar;
Search Terms Used: 'Arxula adenivoran,' 'Arxul,' 'Sacchromyces,' lipid, 'fatty acid,' 'acetyl CoA carboxylase,' 'yeast lipase'

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BOER, E., et al. An Extracellular Lipase From The Dimorphic Yeast <i>Arxula Adenivorans</i> : Molecular Cloning Of The ALIPI Gene And Characterization Of The Purified Recombinant Enzyme. Yeast. 2005, Vol. 22, pages 523-535; published online in Wiley InterScience; www.interscience.wiley.com. DOI: 10.1002/yea.1230.	1-4
A	WO 2010/025374A2 (E.I. DUPONT DE NEMOURS AND COMPANY) March 4, 2010; page 187, Example 5	1-4
A	US 2003/0233675A1 (CAO, Y, et al) December 18, 2003; paragraph [0099]; Table 13A	1-4

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

14 October 2015 (14.10.2015)

Date of mailing of the international search report

06 NOV 2015

Name and mailing address of the ISA/

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Shane Thomas

PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US15/33211

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 5-13, 15-20, 24, 25
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

-Please See Supplemental Page-

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

-Please See Supplemental Page-

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/US15/33211

---Continued from Box No. III: Observations Where Unity Of Invention Is Lacking:

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Groups I+: Claims 1-4 are directed toward a nucleic acid, comprising a nucleotide sequence encoding a protein from *Arxula adeninivorans*, wherein said protein comprises an amino acid sequence having at least 95% sequence homology with the amino acid sequence set forth in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 164, 166, 168, 170, 172, 174, 176, or 178, or a biologically-active portion of any one of the sequences.

The nucleic acid will be searched to the extent it encodes a sequence encompassing SEQ ID NO: 1 (*Arxula adeninivorans* amino acid sequence) and encompasses SEQ ID NO: 2 (*Arxula adeninivorans* nucleic acid sequence). It is believed that Claims 1 (in-part), 2 (in-part), 3 (in-part) and 4 (in-part) encompass this first named invention and thus these claims will be searched without fee to the extent that they encompass SEQ ID NOs: 1 (*Arxula adeninivorans* amino acid sequence) and 2 (*Arxula adeninivorans* nucleic acid sequence). Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined. An Exemplary Election would be: SEQ ID NOs: 3 (*Arxula adeninivorans* amino acid sequence), 4 (*Arxula adeninivorans* nucleic acid sequence).

Group II: Claim 14 is directed toward a transformed cell, comprising a genetic modification.

Groups III+: Claims 21-23 are directed toward an *Arxula adeninivorans* cell, comprising a genetic modification, wherein said genetic modification decreases the activity of a protein selected from delta-12 fatty acid desaturase, multi-functional B-oxidation enzyme, delta-9 fatty acid desaturase, triacylglycerol lipase, glycerol-3-phosphate dehydrogenase, AMP-activated kinase, acyl-CoA oxidase, fatty acid elongase, peroxisomal membrane E3 ubiquitin ligase, dihydrolipoamide acetyltransferase component (E2) of pyruvate dehydrogenase complex, MIG1 transcription factor involved in glucose repression, 2-methylcitrate dehydratase, phosphatidyl-ethanolamine methyltransferase, methylene-fatty-acyl-phospholipid synthase and INO4 transcription factor.

The *Arxula adeninivorans* cell can be searched to the extent that it encompasses a protein of SEQ ID NO: 13 (*Arxula adeninivorans* amino acid sequence). It is believed that Claims 21, 22 and 23 (in-part) encompass this first named invention and thus these claims will be searched without fee to the extent that they encompass SEQ ID NO: 13 (*Arxula adeninivorans*). Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined. An Exemplary Election would be: SEQ ID NO: 25 (*Arxula adeninivorans* amino acid sequence).

The inventions listed as Groups I+-III+ do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: the special technical features of Groups I+ include SEQ ID NO: 1 (*Arxula adeninivorans* amino acid sequence), which is not present in either of Groups II or III+, the special technical features of Group II including acetyl-CoA carboxylase, which is not present in either of Group.

No technical features are shared between the amino acid or nucleic acid sequences of Groups I+ and III+ and, accordingly, these sequences and groups lack unity a priori.

Groups I+-III+ share the technical features including proteins. Groups II and III share the technical features including a cell comprising a genetic modification. Groups I+ share the technical features including a nucleic acid, comprising a nucleotide sequence encoding a protein from *Arxula adeninivorans*, wherein said protein comprises an amino acid sequence. Groups III+ share the technical features including: an *Arxula adeninivorans* cell, comprising a genetic modification, wherein said genetic modification decreases the activity of a protein selected from delta-12 fatty acid desaturase, multi-functional B-oxidation enzyme, delta-9 fatty acid desaturase, triacylglycerol lipase, glycerol-3-phosphate dehydrogenase, AMP-activated kinase, acyl-CoA oxidase, fatty acid elongase, peroxisomal membrane E3 ubiquitin ligase, dihydrolipoamide acetyltransferase component (E2) of pyruvate dehydrogenase complex, MIG1 transcription factor involved in glucose repression, 2-methylcitrate dehydratase, phosphatidyl-ethanolamine methyltransferase, methylene-fatty-acyl-phospholipid synthase and INO4 transcription factor.

However, these shared technical features are previously disclosed by the publication entitled 'An Extracellular Lipase From The Dimorphic Yeast *Arxula Adeninivorans*: Molecular Cloning Of The ALIPI Gene And Characterization Of The Purified Recombinant Enzyme' by Boer, et al. (hereinafter 'Boer') in view of the publication entitled 'Arxula Adeninivorans Xanthine Oxidoreductase And Its Application In The Production Of Food With Low Purine Content' by Jankowska, et al. (hereinafter 'Jankowska').

---Continued Within the Next Supplemental Box---

-----Continued from Previous Supplemental Page:

Boer discloses proteins (lipase enzyme (proteins); abstract); a nucleic acid (an ALIP1 gene (a nucleic acid); abstract), comprising a nucleotide sequence (comprising a sequence encoding a lipase; abstract, figure 2b) encoding a protein (encoding a lipase (a protein); abstract) from *Arxula adeninivorans* (abstract), wherein said protein comprises an amino acid sequence (abstract, figure 2b); an *Arxula adeninivorans* cell (*A. adeninivorans* utilized as a host for recombinant expression (an *Arxula adeninivorans* cell); page 525, column 1, paragraph 3 to column 2, paragraph 1); and triacylglycerol lipase (triacylglycerol acylhydrolase (lipase); page 523, column 1). Boer does not disclose a cell comprising a genetic modification and an *Arxula adeninivorans* cell, comprising a genetic modification, wherein said genetic modification decreases the activity of a triacylglycerol lipase. Jankowska discloses an *Arxula adeninivorans* cell comprising a genetic modification (an *A. adeninivorans* strain, including a delta-axor knockout mutant (an *Arxula adeninivorans* cell comprising a genetic modification); page 805, column 1, paragraph 2), wherein said genetic modification decreases the activity of a protein (wherein said genetic modification is a gene knockout (wherein said genetic modification decreases the activity of a protein); page 805, column 1, paragraph 2). It would have been obvious to a person of ordinary skill in the art, at the time of the invention, to have modified the previous disclosure of Boer, for inclusion of knocking-out the target gene in the host, as previously disclosed by Jankowska, for ensuring: a) that the target gene was actually responsible for the targeted activity; and b) for ensuring the capacity to produce recombinant enzyme, as previously disclosed by Boer, in a null host background, thus producing an *A. adeninivorans* cell having a genetic modification which decreases the activity of the triacylglycerol lipase previously disclosed by Boer, particularly in the event that mutations are introduced into the recombinant enzyme, for further enabling characterization of the enzyme, where a null host background would have been essential.

Since none of the special technical features of the Groups I-III+ inventions is found in more than one of the inventions, and since all of the shared technical features are previously disclosed by a combination of the Boer and Jankowska references, unity of invention is lacking