(54) Title: INCREASED OIL CONTENT BY INCREASING YAP1 TRANSCRIPTION FACTOR ACTIVITY IN OLEAGINOUS YEASTS

(57) Abstract:
Transgenic oleaginous yeast having increased oil content comprising increased Yap1 transcription factor activity, wherein the increased oil content is compared to the oil content of a non-transgenic oleaginous yeast, are described herein. The increased Yap1 transcription factor activity results from overexpressing a Yap1 transcription factor, by increasing the interaction between the transcription factor and a protein that is capable of activating the transcription factor, or by a combination thereof. Methods of using these yeast strains are also described.
**Title:** INCREASED OIL CONTENT BY INCREASING YAP1 TRANSCRIPTION FACTOR ACTIVITY IN OLEAGINOUS YEASTS

**Abstract:** Transgenic oleaginous yeast having increased oil content comprising increased Yap1 transcription factor activity, wherein the increased oil content is compared to the oil content of a non-transgenic oleaginous yeast, are described herein. The increased Yap1 transcription factor activity results from overexpressing a Yap1 transcription factor, by increasing the interaction between the transcription factor and a protein that is capable of activating the transcription factor, or by a combination thereof. Methods of using these yeast strains are also described.
(15) Information about Correction:
see Notice of 27 December 2012
TITLE
INCREASED OIL CONTENT BY INCREASING YAP1 TRANSCRIPTION FACTOR ACTIVITY IN OLEAGINOUS YEASTS

This application claims the benefit of U.S. Provisional Application No. 61/428,655, filed December 30, 2010, which is hereby incorporated by reference in its entirety.

FIELD OF THE INVENTION
This invention is in the field of biotechnology. More specifically, this invention pertains to oleaginous yeast strains comprising increased Yap1 transcription factor activity, resulting in increased oil content.

BACKGROUND OF THE INVENTION
Reactive oxygen species ["ROS"] are chemically reactive molecules containing oxygen and comprising unpaired valence shell electrons. ROS, such as hydroxyl radicals, superoxide anions, and hydrogen peroxide ["H_2O_2"], are generated continually as by-products of aerobic metabolism in cells, e.g., via incomplete reduction of oxygen to water during respiration. ROS are also produced during beta-oxidation of fatty acids by exposure to radiation, light, metals, and redox active drugs. Since ROS may perturbate the cellular redox status and ultimately cause toxic damage to cellular components, including lipids, proteins, and DNA, cells must possess a variety of means to sense levels of ROS and transduce the signal such that the cell is protected against the effects of oxidative stress and cellular integrity is maintained.

Typically, levels of ROS are controlled by use of the glutathione reduction-oxidation (re-doX) cycle and thioredoxin system, such that electrons are accepted from NADPH and utilized to reduce H_2O_2 to water. More specifically, the electrons are transferred from NADPH to thioredoxin reductase to thioredoxin to peroxiredoxins to H_2O_2, yielding water.

Regulation of the multiple genes in this pathway is complex. The adaptive response to H_2O_2 in the yeast Saccharomyces cerevisiae has been found to involve a change in the expression of at least 167 proteins (Godon, C. et al., J. Biol. Chem., 273:22480-22489 (1998)).
One means to sense levels of $\text{H}_2\text{O}_2$ in the yeast *S. cerevisiae* relies on a signaling pathway based on the master transcription factor for the oxidative stress response, i.e., the transcription factor protein Yap1. In response to $\text{H}_2\text{O}_2$ stress, a multi-step conformational change in Yap1 occurs based on the formation of at least one intra-molecular disulfide bond, a reaction catalyzed by peroxiredoxins such as Tsa1 and Gpx3 and facilitated by other proteins such as Ybp1. In this active oxidized form, Yap1 controls the expression of a large regulon of at least 32 different proteins, including those involved in cellular antioxidant defenses and glutathione/ NADPH regeneration (Lee, J. et al., *J. Biol. Chem.*, 274:16040-16046 (1999)). Deactivation of Yap1 occurs by enzymatic reduction with Yap1-controlled thioredoxins, thus providing a mechanism for autoregulation. Mutant strains of *S. cerevisiae* lacking a functional Yap1 protein are hypersensitive to killing by $\text{H}_2\text{O}_2$.

It is known that fatty acids having more double bonds are more susceptible to lipid peroxidation. Thus, polyunsaturated fatty acids ["PUFAs"] are more susceptible to oxidative degradation by ROS because they contain multiple double bonds in between which lie methylene-CH$_2$-groups that possess especially reactive hydrogens. Avery, A.M. and S.V. Avery (J. Biol. Chem., 276:33730-33735 (2001)) reported that a *S. cerevisiae gpx1Δ gpx2Δ gpx3Δ* mutant was defective for growth in medium supplemented with the PUFA alpha-linolenic acid ["ALA"; 18:3], wherein ALA can comprise up to 60% of the total membrane fatty acids; *gpx1Δ, gpx2Δ* and *gpx3Δ* mutants also demonstrated toxicity to the 18:3, although the effect was delayed based on the slower incorporation rate of exogenous 18:3 into membrane lipids.

Since ROS are continually produced in cells performing aerobic metabolism and since ROS can lead to cell damage and death, one of skill in the art will appreciate methods that increase the capacity of recombinantly engineered organisms to defend against ROS. This is especially true in those organisms that produce microbial oils, since the generation of ROS in certain microbial strains during production of these
oils can lead to lower yields and/or reduced efficiency in microbial oil production.

It has been found that engineering oleaginous yeast to have increased Yap1 transcription factor activity and to produce PUFAs results in both increased lipid content ["TFAs % DCW"] and increased average PUFA titer ["PUFA % DCW"].

**SUMMARY OF THE INVENTION**

In one embodiment, the invention concerns a transgenic oleaginous yeast having increased oil content comprising increased Yap1 transcription factor activity wherein the increased oil content is compared to the oil content of a non-transgenic oleaginous yeast cell.

In a second embodiment, the increased Yap1 transcription factor activity results from overexpressing the Yap1 transcription factor, by increasing the interaction between the transcription factor and a protein that is capable of activating the transcription factor, or by a combination thereof.

In a third embodiment, the protein that is capable of activating the transcription factor is selected from the group consisting of: Gpx3, Ybp1 and Tsa1.

In a fourth embodiment, the Yap1 transcription factor comprises a nucleotide sequence encoding a polypeptide having transcription factor activity and comprising: (a) a bZIP leucine zipper motif; (b) an N-terminal Cys-rich domain comprising a sequence of at least two cysteine residues that are separated by at least 6 amino acids; and, (c) a C-terminal Cys-rich domain comprising a sequence of at least two cysteine residues that are separated by at least 8 amino acids.

In a fifth embodiment, the Gpx3 protein comprises: (a) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the polypeptide has at least 70% amino acid identity, based on the BLASTP method of alignment, when compared to a sequence selected from the group consisting of SEQ ID NO:26 [ScGpx3] or SEQ ID NO:28
[YIGpx3]; (b) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the nucleotide sequence has at least 70% sequence identity, based on the BLASTN method of alignment, when compared to a sequence selected from the group consisting of SEQ ID NO:25 [ScGpx3] or SEQ ID NO:27 [YIGpx3]; and, (c) a complement of the nucleotide sequence of (a) or (b), wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

In a sixth embodiment, the Tsa1 protein comprises: (a) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the polypeptide has at least 70% amino acid identity, based on the BLASTP method of alignment, when compared to a sequence selected from the group consisting of SEQ ID NO:34 [ScTsa1] or SEQ ID NO:36 [YITsa1]; (b) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the nucleotide sequence has at least 70% sequence identity, based on the BLASTN method of alignment, when compared to a sequence selected from the group consisting of SEQ ID NO:33 [ScTsa1] or SEQ ID NO:35 [YITsa1]; and, (c) a complement of the nucleotide sequence of (a) or (b), wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

In a seventh embodiment, the Ybp1 protein comprises: (a) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the polypeptide is selected from the group consisting of SEQ ID NO:38 [ScYbp1] or SEQ ID NO:40 [YIYbp1]; (b) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the polypeptide sequence is classified within a kinetochor_Ybp2 super family, based on the conserved domain method of analysis; or, (c) a complement
of the nucleotide sequence of (a) or (b), wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

In an eighth embodiment, the transgenic oleaginous yeast cell is from a genus selected from the group consisting of: *Yarrowia, Candida, Rhodotorula, Rhodosporidium, Cryptococcus, Trichosporon,* and *Lipomyces.* Preferably, the transgenic oleaginous yeast cell is *Yarrowia lipolytica."

In a ninth embodiment, the transgenic oleaginous yeast cell produces at least one polyunsaturated fatty acid.

In a tenth embodiment, the invention concerns a method of increasing oil content in an oleaginous yeast comprising:

a) engineering the oleaginous yeast to overexpress a protein selected from the group consisting of:

(i) a Yap1 transcription factor;

(ii) a protein that is capable of activating the transcription factor;

(iii) a combination of (a) and (b); and,

b) growing the oleaginous yeast under suitable conditions to result in increased oil content when compared to the oil content of a non-transgenic oleaginous yeast.

**BRIEF DESCRIPTION OF THE DRAWINGS AND SEQUENCE DESCRIPTIONS**

**FIG. 1** is a diagram of the mechanism by which *Saccharomyces cerevisiae* GPX3 [*ScGPX3*] activates *S. cerevisiae* YAP1 [*"ScYap1".*] ScGPX3 comprises Cys36 and Cys82, which either form an intermolecular disulfide bond (-S-S-) or are reduced to contain thiol groups (-SH). ScYap1 comprises a N-terminal and a C-terminal Cys-rich domain (each shown in black); Cys 303, Cys310 and Cys315 and Cys598, Cys620 and Cys629 within these Cys-rich domains are shown as 6 vertical black lines. The thiol group (-SH) of Cys36 ScGpx3 reacts with *H₂O₂*, resulting in the release of water and formation of a sulphenic acid (-SOH). The -
SOH then condenses with the -SH of Cys598 of ScYap1 (reduced form), forming an inter-molecular disulfide bond (-S-S-), which is then converted into an intra-molecular disulfide bond between Cys303 and Cys598 of ScYap1, thereby producing a conformational change in the oxidized ScYap1 protein.

FIG. 2 is a sequence comparison between ScYAP1 (SEQ ID NO:2) and YIYAP1 (SEQ ID NO:4). Underlined, bolded basic amino acids at positions 69-115 of ScYAP1 (corresponding to positions 115-166 of YIYAP1) represent the basic region of the bZIP domain for DNA binding. Bold leucine residues, shown with a star over the alignment, at positions 87, 94, 108, and 115 of ScYAP1 (corresponding to positions 138, 145, 159, and 166 of YIYAP1) are the leucine zipper motif of the bZIP domain. Boxed cysteine residues at positions 303, 310, 315, 598, 620 and 629 of ScYAP1 (corresponding to positions 309, 316, 483, 505 and 514 of YIYap1) are important (or likely important) for inter- and intra-molecular interactions.

FIG. 3 provides plasmid maps for the following: (A) pYRH60; and, (B) pYRH61.

FIG. 4 shows $H_2O_2$ sensitivity assay results on YPD plates under increasing $H_2O_2$ concentrations, i.e., from 0 mM to 50 mM $H_2O_2$. (A) compares growth of *Y. lipolytica* strains Y4184 (control) and Y4184U (yap1Δ) cells. (B) compares growth of *S. cerevisiae* strains BY4743 (control) and BY4743 (yap1Δ) cells, transformed with either plasmid pRS316 (control) or pYRH61.

FIG. 5 provides plasmid maps for the following: (A) pYRH43; and, (B) pYRH65.

FIG. 6 is a sequence comparison between ScGPX3 (SEQ ID NO:26) and YIGPX3 (SEQ ID NO:28). Boxed cysteine residues at positions 36, 64 and 82 of ScGpx3 (corresponding to positions 42, 70 and 88 of YIGpx3) are important (or likely important) for inter- and intra-molecular interactions.

FIG. 7 is a sequence comparison between *ScTsa1* (SEQ ID NO:34) and *YITsa1* (SEQ ID NO:36). Boxed cysteine residues at positions 48 and
171 of ScTsa1 (corresponding to positions 48 and 169 of YITsa1) are important (or likely important) for inter- and intra-molecular interactions.

FIG. 8 is a sequence comparison between ScYbp1 (SEQ ID NO:38) and YIYbp1 (SEQ ID NO:40).

FIG. 9 is a sequence comparison between ScYbp1 (SEQ ID NO:38), YIYbp1 (SEQ ID NO:40), the Candida glabrata Ybp1 ["CgYbp1"] (SEQ ID NO:43), the Kluyveromyces lactis NRRL Y-1140 Ybp1 ["KlYbp1"] (SEQ ID NO:44), the Scheffersomyces stipitis CBS 6054 Ybp1 ["SsYbp1"] (SEQ ID NO:45), the Zygosaccharomyces rouxii CBS 732 Ybp1 ["ZrYbp1"] (SEQ ID NO:46), and the Candida albicans SC5314 Ybp1 ["CaYbp1"] (SEQ ID NO:47).

The following sequences comply with 37 C.F.R. §§ 1.821-1.825 ("Requirements for Patent Applications Containing Nucleotide Sequences and/or Amino Acid Sequence Disclosures - the Sequence Rules") and are consistent with World Intellectual Property Organization (WIPO) Standard ST.25 (1998) and the sequence listing requirements of the EPO and PCT (Rules 5.2 and 49.5(a-bis), and Section 208 and Annex C of the Administrative Instructions). The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. § 1.822.

SEQ ID NOs:1-47 are ORFs encoding genes, proteins (or portions thereof), primers or plasmids, as identified in Table 1.

### Table 1. Summary Of Nucleic Acid And Protein SEQ ID Numbers

<table>
<thead>
<tr>
<th>Description</th>
<th>Nucleic acid SEQ ID NO.</th>
<th>Protein SEQ ID NO.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saccharomyces cerevisiae Yap1 (GenBank Accession No. NM_001182362)</td>
<td>1 (1953 bp)</td>
<td>2 (650 AA)</td>
</tr>
<tr>
<td>Yarrowia lipolytica Yap1 (GenBank Accession No. XM_504945)</td>
<td>3 (1605 bp)</td>
<td>4 (534 AA)</td>
</tr>
<tr>
<td>Plasmid pYRH60</td>
<td>5 (7412 bp)</td>
<td>--</td>
</tr>
<tr>
<td>Plasmid pYPS161</td>
<td>6 (7966 bp)</td>
<td>--</td>
</tr>
<tr>
<td>Yarrowia lipolytica Yap1 promoter region</td>
<td>7 (940 bp)</td>
<td>--</td>
</tr>
<tr>
<td>Description</td>
<td>Accession Numbers</td>
<td>Lengths</td>
</tr>
<tr>
<td>----------------------------------------------------------------------------</td>
<td>--------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Primers YI-EF-1214F, YI-EF-1270R, YAP1-346F and YAP1-409R</td>
<td></td>
<td>8-11</td>
</tr>
<tr>
<td>Primer YL-EF-MGB-1235T containing reporter dye 5′-6-FAM and quencher 3′-TAMRA</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>Primer YAP1-366T containing reporter dye 5′-6-FAM and quencher 3′-TAMRA</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>Plasmid pYRH61</td>
<td>14</td>
<td>(8043 bp)</td>
</tr>
<tr>
<td>Primers YI.Yap1-F-Spel and Yap1-R</td>
<td>15-16</td>
<td></td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em> FBA1 promoter region</td>
<td>17</td>
<td>(601 bp)</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em> FBA1 terminator region</td>
<td>18</td>
<td>(1022 bp)</td>
</tr>
<tr>
<td>Plasmid pRS316</td>
<td>19</td>
<td>(4887 bp)</td>
</tr>
<tr>
<td>Plasmid pYRH43</td>
<td>20</td>
<td>(8597 bp)</td>
</tr>
<tr>
<td>Primer Yap1-F</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Primers ef-324F and ef-392R</td>
<td>22, 23</td>
<td></td>
</tr>
<tr>
<td>Primer ef-345T containing reporter dye 5′-6-FAM and quencher 3′-TAMRA</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em> GPX3 (GenBank Accession No. NM_001179559)</td>
<td>25</td>
<td>(492 bp)</td>
</tr>
<tr>
<td><em>Yarrowia lipolytica</em> GPX3 (GenBank Accession No. XM_503454)</td>
<td>27</td>
<td>(507 bp)</td>
</tr>
<tr>
<td>Plasmid pYRH65</td>
<td>29</td>
<td>(7651 bp)</td>
</tr>
<tr>
<td>Primers GPX3-F and GPX3-R</td>
<td>30-31</td>
<td></td>
</tr>
<tr>
<td><em>Yarrowia lipolytica</em> Yap1 terminator region</td>
<td>32</td>
<td>(1164 bp)</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em> Tsa1 (GenBank Accession No. NP_013684)</td>
<td>33</td>
<td>(591 bp)</td>
</tr>
<tr>
<td><em>Yarrowia lipolytica</em> Tsa1 (GenBank Accession No. XM_500915)</td>
<td>35</td>
<td>(591 bp)</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em> Ybp1 (GenBank Accession No. NP_009775)</td>
<td>37</td>
<td>(2025 bp)</td>
</tr>
<tr>
<td><em>Yarrowia lipolytica</em> Ybp1 (GenBank Accession No. XM_500469)</td>
<td>39</td>
<td>(2025 bp)</td>
</tr>
<tr>
<td>Mutant delta-5 desaturase motif: HPGs</td>
<td>--</td>
<td>41</td>
</tr>
<tr>
<td>Mutant delta-5 desaturase motif: HaGG</td>
<td>--</td>
<td>42</td>
</tr>
<tr>
<td><em>Candida glabrata</em> Ybp1 (GenBank Accession No. CAG61477.1)</td>
<td>--</td>
<td>43</td>
</tr>
<tr>
<td><em>Kluyveromyces lactis</em> NRRL Y-1140 Ybp1 (GenBank Accession No. XP_452453.1)</td>
<td>--</td>
<td>44</td>
</tr>
<tr>
<td><em>Scheffersomyces stipitis</em> CBS 6054 Ybp1 (GenBank Accession No. XP_001386941.2)</td>
<td>--</td>
<td>45</td>
</tr>
<tr>
<td><em>Zygosaccharomyces rouxii</em> CBS 732 Ybp1 (GenBank Accession No. XP_002495870.1)</td>
<td>--</td>
<td>46</td>
</tr>
</tbody>
</table>

(HPGs: 196 AA, HaGG: 655 AA, Ybp1: 702 AA, Ybp1: 673 AA, Ybp1: 664 AA)
**DETAILED DESCRIPTION OF THE INVENTION**

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

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

“Open reading frame” is abbreviated as “ORF”.

“Polymerase chain reaction” is abbreviated as “PCR”.

“American Type Culture Collection” is abbreviated as “ATCC”.

“Polyunsaturated fatty acid(s)” is abbreviated as “PUFA(s)”.

“Triacylglycerols” are abbreviated as “TAGs”.

“Total fatty acids” are abbreviated as "TFAs".

“Fatty acid methyl esters” are abbreviated as “FAMEs”.

“Dry cell weight” is abbreviated as “DCW”.

“Weight percent” is abbreviated as “wt %”.

“Reactive oxygen species” is abbreviated as “ROS”.

“Hydrogen peroxide” is abbreviated as “$H_2O_2$”.

As used herein and in the appended claims, the singular forms “a”, “an”, and “the” include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to “a cell” includes one or more cells and equivalents thereof known to those skilled in the art, and so forth.

A “transcription factor” refers to a protein (or the DNA encoding that protein) which interacts with a DNA regulatory element to affect expression of a structural gene or expression of a second regulatory gene.

More specifically, the transcription factor (either alone or in a complex with other proteins) affects transcription of DNA to mRNA, by e.g., activation or repression of transcription initiation. A transcription factor may comprise one or more DNA-binding domains which attach to specific sequences of DNA adjacent to the genes that they regulate.

“Yap1 transcription factor activity” refers to activity that occurs as a result of a Yap1 transcription factor, a transcriptional regulator of the AP-1 family involved in a cellular pathway that controls the oxidative stress...
response. Increased Yap1 transcription factor activity results in regulation of a family of proteins, which typically enables increased tolerance to ROS (e.g., when H₂O₂ stress is encountered). According to the present invention described herein, increased Yap1 transcription factor activity results in increased oil content.

A “Yap1 transcription factor” refers to a transcription factor having Yap1 transcription factor activity. In general, such a protein should have: (a) a bZIP leucine zipper motif; (b) an N-terminal Cys-rich domain comprising a sequence of at least two cysteine residues that are separated by at least six (6) amino acids; and, (c) a C-terminal Cys-rich domain comprising a sequence of at least two cysteine residues that are separated by at least eight (8) amino acids.

A “bZIP leucine zipper motif” is characterized as comprising: (i) a basic DNA binding region spanning approximately fourteen to sixteen amino acids (i.e., comprising arginine and lysine residues); and, (ii) an adjacent leucine-rich zipper region (i.e., comprising evenly spaced leucine residues allowing dimerization) (Hurst, H.C. Transcription factors 1: bZIP proteins. *Protein Profile*, 2:101–168 (1995)). Typically, the leucine residues are spaced at seven amino acid intervals; although other hydrophobic amino acids such as methionine, isoleucine, valine and phenylalanine have been reported to form zippers in combination with leucine.

The term “ScYAP1” (SEQ ID NO:1; GenBank Accession No. NM_001182362.1) refers to a Yap1 transcription factor of the AP-1 family isolated from *Saccharomyces cerevisiae* S288c, encoded by SEQ ID NO:2. As annotated in GenBank, ScYAP1 is required for oxidative stress tolerance, and is activated by H₂O₂ through the multistep formation of disulfide bonds and transit from the cytoplasm to the nucleus. ScYAP1 also mediates resistance to cadmium.

The term “YlYAP1” (SEQ ID NO:4; YALI0F03388p; GenBank Accession No. XP_504945) refers to a Yap1 transcription factor isolated from *Yarrowia lipolytica*, encoded by SEQ ID NO:3 herein.
The term "a protein that is capable of activating the Yap1 transcription factor" refers to a protein that interacts with the Yap1 transcription factor in a manner that facilitates oxidation of the Yap1 transcription factor, such that the transcription factor comprises at least one intra-molecular disulfide bond and is thus in an "activated state". Preferred proteins that are capable of activating the Yap1 transcription factor include Yap1 binding protein (Ybp1) and the peroxiredoxin proteins, Gpx3 and Tsa1, although this should not be construed as limiting to the invention herein.

"Yap1 binding protein" or "Ybp1" refers to a binding protein that binds to a Yap1 transcription factor. As described in Gulshan, K. et al. (J. Biol. Chem., 286(39):34071-34081 (2011)), Yap1 and Ybp1 are likely to directly interact in the cell, but further localization of the sites or domains of interaction has not been achieved.

The term “ScYbp1” (SEQ ID NO:38; GenBank Accession No. NP_009775.1) refers to a Yap1 binding protein isolated from Saccharomyces cerevisiae S288c, encoded by SEQ ID NO:37 herein (Veal, E.A. et al., J Biol Chem., 278(33):30896-30904 (2003); Gulshan, K. et al., supra). As annotated in GenBank, ScYbp1 functions as a "protein required for oxidation of specific cysteine residues of the transcription factor Yap1p, resulting in the nuclear localization of Yap1p in response to stress".

The term “YlYbp1” (SEQ ID NO:40; YALI0B03762g; GenBank Accession No. XP_500469.1) refers to a Yap1 binding protein isolated from Yarrowia lipolytica, encoded by SEQ ID NO:39 herein.

A "peroxiredoxin protein" or "Prx protein" comprises redox-active cysteine residues. During catalysis, the peroxidatic cysteine is oxidized (e.g., by H₂O₂) to a sulfinic acid, which condenses with a resolving cysteine residue to form a disulfide (wherein the resolving cysteine residue is either within the same Prx molecule or within another Prx molecule, resulting in dimer formation). This disulfide bond is reduced by thioredoxin to regenerate the active Prx. Thus, Prx proteins are active in a redox cycle, accepting electrons from NADPH via thioredoxin and thioredoxin
reductase. As defined by T. Tachibana et al. (J. Biol. Chem., 284 (7):4464-4472 (2009)), proteins that show thioredoxin-dependent peroxidase activity in budding yeast include five Prx family proteins [i.e., Tsa1, Tsa2, Prx1, Ahp1, Dot5] and two glutathione peroxidase (Gpx)-like proteins [i.e., Gpx2, Gpx3], although “Prx” will be used herein to refer to both the Prx proteins and the Gpx-like proteins. Preferred Prx proteins that are capable of activating the Yap1 transcription factor include Gpx3 and Tsa1.

The term “ScGpx3” (SEQ ID NO:26; GenBank Accession No. NM_001179559.1; E.C. 1.11.1.15) refers to a thiol peroxidase isolated from Saccharomyces cerevisiae S288c, encoded by SEQ ID NO:25. As annotated in GenBank, ScGpx3 functions as a hydroperoxide receptor to sense intracellular H₂O₂ levels and transduce a redox signal to the Yap1p transcription factor.

The term “YIGpx3” (SEQ ID NO:28; YALI0E02310p; GenBank Accession No. XP_503454) refers to a thiol peroxidase isolated from Yarrowia lipolytica, encoded by SEQ ID NO:27 herein.

The term “ScTsa1” (SEQ ID NO:34; E.C. 1.11.1.15; GenBank Accession No. NP_013684) refers to a thioredoxin peroxidase isolated from Saccharomyces cerevisiae S288c, encoded by SEQ ID NO:33 herein (Trotter, E.W. et al., Biochem J., 412(1):73-80 (2008)). As annotated in GenBank, ScTsa1 is of the peroxiredoxin (PRX) 2-Cys subfamily, wherein peroxiredoxins function as “thiol-specific antioxidant (TSA) proteins, which confer a protective role in cells through its peroxidase activity by reducing H₂O₂, peroxynitrite, and organic hydroperoxides”.

The term “YITsa1” (SEQ ID NO:36; YALI0B15125g; GenBank Accession No. XP_500915.1) refers to a thioredoxin peroxidase isolated from Yarrowia lipolytica, encoded by SEQ ID NO:35 herein.

Generally, the term “oleaginous” refers to those organisms that tend to store their energy source in the form of oil (Weete, In: Fungal Lipid Biochemistry, 2nd Ed., Plenum, 1980). During this process, the cellular oil content of oleaginous microorganisms generally follows a sigmoid curve, wherein the concentration of lipid increases until it reaches a maximum at
the late logarithmic or early stationary growth phase and then gradually decreases during the late stationary and death phases (Yongmanitchai and Ward, *Appl. Environ. Microbiol.*, 57:419-25 (1991)). For the purposes of the present application, the term “oleaginous” refers to those microorganisms that can accumulate at least about 25% of their dry cell weight ["DCW"] as oil.

The term “oleaginous yeast” refers to those oleaginous microorganisms classified as yeasts that can make oil, i.e., wherein oil can accumulate in excess of about 25% of their DCW. Examples of oleaginous yeast include, but are no means limited to, the following genera: *Yarrowia, Candida, Rhodotorula, Rhodosporidium, Cryptococcus, Trichosporon* and *Lipomyces*. The ability to accumulate oil in excess of about 25% of the DCW of the yeast may be through efforts of recombinant engineering or through the natural abilities of the organism.

The term “transgenic oleaginous yeast” generically refers to an oleaginous yeast that contains a foreign or heterologous nucleic acid fragment as a result of a transformation procedure. However, for the purposes herein, the term “transgenic oleaginous yeast” will specifically refer to an oleaginous yeast that contains a foreign or heterologous nucleic acid fragment(s) as a result of a transformation procedure, wherein expression of the foreign or heterologous nucleic acid(s) results in increased Yap1 transcription factor activity in the oleaginous yeast. Thus, for example, a transgenic oleaginous yeast of the invention herein may be genetically engineered to overexpress a chimeric gene encoding either a Yap1 transcription factor or a protein that is capable of activating the Yap1 transcription factor, wherein the Yap1 transcription factor or the protein that is capable of activating the Yap1 transcription is either a native gene or a foreign gene.

In contrast, a non-transgenic oleaginous yeast herein will refer to an oleaginous yeast having a genotype identical to the transgenic oleaginous yeast to which it is compared, with the exception that the non-transgenic oleaginous yeast has not been transformed with the foreign or heterologous nucleic acid(s) that results in increased Yap1 transcription
factor activity in the transgenic oleaginous yeast and thus lacks this particular foreign or heterologous nucleic acid(s). To be clear, the non-transgenic oleaginous yeast of the invention herein may express at least one foreign gene or heterologous nucleic acid(s), but this does not result in increased Yap1 transcription factor activity.

"Transformation" refers to the transfer of a nucleic acid molecule into a host organism. The nucleic acid molecule may be a plasmid that replicates autonomously; or, it may integrate into the genome of the host organism. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" or "recombiant" or "transformed" organisms or "transformants".

The term "lipids" refer to any fat-soluble (i.e., lipophilic), naturally-occurring molecule. A general overview of lipids is provided in U.S. Pat. Appl. Pub. No. 2009-0093543-A1 (see Table 2 therein).

The term "oil" refers to a lipid substance that is liquid at 25 °C; the oil and is hydrophobic but is soluble in organic solvents. In oleaginous organisms, oil constitutes a major part of the total lipid. "Oil" is composed primarily of triacylglycerols ["TAGs"] but may also contain other neutral lipids, phospholipids and free fatty acids. The fatty acid composition in the oil and the fatty acid composition of the total lipid are generally similar; thus, an increase or decrease in the concentration of fatty acids in the total lipid will correspond with an increase or decrease in the concentration of fatty acids in the oil, and vice versa.

"Neutral lipids" refer to those lipids commonly found in cells in lipid bodies as storage fats and are so called because at cellular pH, the lipids bear no charged groups. Generally, they are completely non-polar with no affinity for water. Neutral lipids generally refer to mono-, di-, and/or triesters of glycerol with fatty acids, also called monoacylglycerol, diacylglycerol or triacylglycerol, respectively, or collectively, acylglycerols. A hydrolysis reaction must occur to release free fatty acids from acylglycerols.

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

The term "total fatty acids" ["TFAs"] herein refer to the sum of all cellular fatty acids that can be derivitized to fatty acid methyl esters ["FAMEs"] by the base transesterification method (as known in the art) in a given sample, which may be the biomass or oil, for example. Thus, total fatty acids include fatty acids from neutral lipid fractions (including diacylglycerols, monoacylglycerols and TAGs) and from polar lipid fractions (including, e.g., the phosphatidylcholine and the phosphatidylethanolamine fractions) but not free fatty acids.

The terms "total lipid content" and "oil content" are used interchangeably herein, to refer to the lipid/oil content of cells as a measure of TFAs as a percent of the dry cell weight ["DCW"], although total lipid content can be approximated as a measure of FAMEs as a percent of the DCW ["FAMEs % DCW"]. Thus, total lipid content ["TFAs % DCW"] is equivalent to, e.g., milligrams of total fatty acids per 100 milligrams of DCW.

Oil content of the transgenic oleaginous yeast of the invention must be compared to the oil content of the non-transgenic oleaginous yeast of the invention under comparable conditions of growth (e.g., type/amount of carbon source, type/amount of nitrogen source, carbon-to-nitrogen ratio, amount of mineral ions, oxygen level, growth temperature, pH, length of the biomass production phase, length of the oil accumulation phase and the time/method of cell harvest).

As used herein, an "isolated nucleic acid fragment" is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. An isolated nucleic acid fragment in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

A "substantial portion" of an amino acid or nucleotide sequence is that portion comprising enough of the amino acid sequence of a polypeptide or the nucleotide sequence of a gene to putatively identify that polypeptide or gene, either by manual evaluation of the sequence by one
skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul, S. F., et al., J. Mol. Biol. 215:403-410 (1993)). In general, a sequence of ten or more contiguous amino acids or thirty or more nucleotides is necessary in order to identify putatively a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene-specific oligonucleotide probes comprising 20-30 contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., in situ hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12-15 bases may be used as amplification primers in polymerase chain reaction ("PCR") in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises enough of the sequence to specifically identify and/or isolate a nucleic acid fragment comprising the sequence.

The term "complementary" describes the relationship between two sequences of nucleotide bases that are capable of Watson-Crick base-pairing when aligned in an anti-parallel orientation. For example, with respect to DNA, adenosine is capable of base-pairing with thymine and cytosine is capable of base-pairing with guanine.

"Codon degeneracy" refers to the nature in the genetic code permitting variation of the nucleotide sequence without affecting the amino acid sequence of an encoded polypeptide. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a gene for improved expression in a host cell, it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

"Synthetic genes" can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form gene segments that are then enzymatically assembled to construct the
entire gene. Accordingly, the genes can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell, where sequence information is available. For example, the codon usage profile for *Y. lipolytica* is provided in U.S. Patent 7,125,672, incorporated herein by reference.

“Gene” refers to a nucleic acid fragment that expresses a specific protein, and which may refer to the coding region alone or may include regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. “Native gene” refers to a gene as found in nature with its own regulatory sequences. “Chimeric gene” refers to any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. “Endogenous gene” refers to a native gene in its natural location in the genome of an organism. A “foreign” gene (or “exogenous” gene) refers to a gene that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, native genes introduced into a new location within the native host, or chimeric genes. A “transgene” is a gene that has been introduced into the genome by a transformation procedure. A “codon-optimized gene” is a gene having its frequency of codon usage designed to mimic the frequency of preferred codon usage of the host cell.

“Coding sequence” refers to a DNA sequence that codes for a specific amino acid sequence.

“Suitable regulatory sequences” refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3'
non-coding sequences) of a coding sequence, and which influence the
transcription, RNA processing or stability, or translation of the associated
coding sequence. Regulatory sequences may include promoters,
enhancers, silencers, 5′ untranslated leader sequence (e.g., between the
transcription start site and the translation initiation codon), introns,
polyadenylation recognition sequences, RNA processing sites, effector
binding sites and stem-loop structures.

“Promoter” refers to a DNA sequence capable of controlling the
expression of a coding sequence or functional RNA. In general, a coding
sequence is located 3′ to a promoter sequence. Promoters may be
derived in their entirety from a native gene, or be composed of different
elements derived from different promoters found in nature, or even
comprise synthetic DNA segments. It is understood by those skilled in the
art that different promoters may direct the expression of a gene in different
tissues or cell types, or at different stages of development, or in response
to different environmental or physiological conditions. Promoters that
cause a gene to be expressed in most cell types at most times are
commonly referred to as “constitutive promoters”. It is further recognized
that since in most cases the exact boundaries of regulatory sequences
have not been completely defined, DNA fragments of different lengths may
have identical promoter activity.

The terms “3′ non-coding sequences”, “transcription terminator” and
“terminator” are used interchangeably herein and refer to DNA sequences
located 3′ downstream of a coding sequence. This includes
polyadenylation recognition sequences and other sequences encoding
regulatory signals capable of affecting mRNA processing or gene
expression. The polyadenylation signal is usually characterized by
affecting the addition of polyadenylic acid tracts to the 3′ end of the mRNA
precursor. The 3′ region can influence the transcription, RNA processing
or stability, or translation of the associated coding sequence.

The term “operably linked” refers to the association of nucleic acid
sequences on a single nucleic acid fragment so that the function of one is
affected by the other. For example, a promoter is operably linked with a
coding sequence when it is capable of effecting the expression of that coding sequence. That is, the coding sequence is under the transcriptional control of the promoter. Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

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

"Stable transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, including both nuclear and organellar genomes, resulting in genetically stable inheritance (i.e., the nucleic acid fragment is "stably integrated"). In contrast, "transient transformation" refers to the transfer of a nucleic acid fragment into the nucleus, or DNA-containing organelle, of a host organism resulting in gene expression without integration or stable inheritance.

The terms "plasmid" and "vector" refer to an extra chromosomal element often carrying genes that are not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA fragments. Such elements may be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, and may be linear or circular, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction that is capable of introducing an expression cassette(s) into a cell.

The term "expression cassette" refers to a fragment of DNA comprising the coding sequence of a selected gene and regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence that are required for expression of the selected gene product. Thus, an expression cassette is typically composed of: 1) a promoter sequence; 2) a coding sequence (i.e., an open reading frame ("ORF")); and, 3) a 3' untranslated region (i.e., a terminator) that, in eukaryotes, usually contains a polyadenylation site. The expression cassette(s) is usually included within a vector, to facilitate cloning and transformation. Different expression cassettes can be
transformed into different organisms including bacteria, yeast, plants and mammalian cells, as long as the correct regulatory sequences are used for each host.

The term "sequence analysis software" refers to any computer algorithm or software program that is useful for the analysis of nucleotide or amino acid sequences. "Sequence analysis software" may be commercially available or independently developed. Typical sequence analysis software will include, but is not limited to: 1) the GCG suite of programs (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI); 2) BLASTP, BLASTN, BLASTX (Altschul et al., J. Mol. Biol. 215:403-410 (1990)); 3) DNASTAR (DNASTAR, Inc. Madison, WI); 4) Sequencher (Gene Codes Corporation, Ann Arbor, MI); and 5) the FASTA program incorporating the Smith-Waterman algorithm (W. R. Pearson, Comput. Methods Genome Res., [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor(s): Suhai, Sandor. Plenum: New York, NY). Within the context of this application it will be understood that where sequence analysis software is used for analysis, that the results of the analysis will be based on the "default values" of the program referenced, unless otherwise specified. As used herein "default values" will mean any set of values or parameters that originally load with the software when first initialized.

"Sequence identity" or "identity" in the context of nucleic acid or polypeptide sequences refers to the nucleic acid bases or amino acid residues in two sequences that are the same when aligned for maximum correspondence over a specified comparison window. Thus, "percentage of sequence identity" or "percent identity" refers to the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield
the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the results by 100 to yield the percentage of sequence identity.


Sequence alignments and percent identity or similarity calculations may be determined using a variety of comparison methods designed to detect homologous sequences.

Multiple alignment of sequences can be performed using the “Clustal method of alignment” which encompasses several varieties of the algorithm including the “ClustalV method of alignment” and the “ClustalW method of alignment” (described by Higgins and Sharp, CABIOS, 5:151-153 (1989); Higgins, D.G. et al., Comput. Appl. Biosci., 8:189-191(1992)) and found in the MegAlign™ (version 8.0.2) program, above. After alignment of the sequences using either Clustal program, it is possible to obtain a “percent identity” by viewing the “sequence distances” table in the program.

The term “conserved domain” or “motif” means a set of amino acids conserved at specific positions along an aligned sequence of evolutionarily related proteins. While amino acids at other positions can vary between homologous proteins, amino acids that are highly conserved at specific positions indicate amino acids that are essential in the structure, the stability, or the activity of a protein. Because they are identified by their high degree of conservation in aligned sequences of a
family of protein homologues, they can be used as identifiers, or "signatures", to determine if a protein with a newly determined sequence belongs to a previously identified protein family.


The term "kinetochor_Ybp2 super family" refers to the Pfam08568 family of proteins described in the Pfam protein database (Finn, R.D. et al., *Nucleic Acids Res.*, 36(Database issue):D281–D288 (2008)).


The activator protein 1 ["AP-1"] is a transcription factor which is a heterodimeric protein composed of subunits that are the products of at least three different proto-oncogene families: the Jun (c-Jun, v-Jun, JunB, JunD), Fos (c-Fos, v-Fos, FosB, FosB2, Fra-1, Fra-2) and activating transcription factor (B-ATF, ATF2, ATF3/LRF1) families. AP-1 regulates gene expression in response to a variety of stimuli, including cytokines, growth factors, stress, and bacterial and viral infections; thus, AP-1 controls a number of cellular processes by upregulating transcription of genes containing the AP-1 recognition element, having the sequence set forth as TGA(C/G)TCA. AP-1 binds to this DNA
sequence via a basic amino acid region, while the dimeric structure is formed by a leucine zipper.

YAP1 is the *Saccharomyces cerevisiae* equivalent of AP-1, and has been concluded to function as the master transcription factor for the oxidative stress response (Moye-Rowley et al., *Genes Dev.*, 3:283-292 (1989)) (SEQ ID NO:2). Structurally, this protein possesses a bZip structural motif consisting of a leucine-rich zipper region and an adjacent basic region (i.e., comprising arginine and lysine residues), as well as an N-terminal Cys-rich domain (i.e., Cys303, Cys310 and Cys315) and a C-terminal Cys-rich domain (i.e., Cys598, Cys620 and Cys629).

Functionally, at least one intra-molecular disulfide bond forms between Cys303 and Cys598 in response to H$_2$O$_2$ stress, thereby causing a multi-step conformational change in Yap1 and nuclear accumulation of Yap1 (due to modifications to the nuclear export signal). In this active oxidized form, Yap1 controls the expression of a large regulon of at least 32 different proteins, including cellular antioxidants and enzymes of the glutathione and pentose phosphate pathways (Lee, J. et al., *J. Biol. Chem.*, 274:16040-16046 (1999)). Deactivation of Yap1 occurs by enzymatic reduction with Yap1-controlled thioredoxins, thus providing a mechanism for autoregulation.

The oxidation of the *S. cerevisiae* Yap1 protein does not occur directly in response to H$_2$O$_2$; instead, a constitutively expressed thiol peroxidase protein (e.g., Gpx3; SEQ ID NO:26) transduces the H$_2$O$_2$ signal and is responsible for catalyzing the formation of the intra-molecular disulfide bond(s) within Yap1 (Inoue et al., *J. Biol. Chem.*, 274:27002-27009 (1999); Delaunay, A., et al., *Cell*, 111:471-481 (2002)). Cys36 of this glutathione peroxidase (Gpx)-like protein initially bridges Cys598 of Yap1 by a disulfide bond, which is converted into the Yap1 intra-molecular disulfide bond (FIG. 1, recreated from Tachibana, T. et al., *J. Biol. Chem.*, 284:4464-4472 (2009); Okazaki et al., *Mol. Cell*, 27:675-688 (2007)). Gpx3 proteins not reacting with Yap1 are able to reduce H$_2$O$_2$ directly to water, resulting in formation of an intra-molecular disulfide bond between Cys36 and Cys82 of the Gpx3 protein. A Gpx3-independent pathway for

In addition to Gpx3, a suite of other peroxiredoxin (Prx) proteins comprising at least one redox-active cysteine residue may be capable of activating the Yap1 transcription factor (via direct or indirect means). Specifically, T. Tachibana et al. (*J. Biol. Chem.*, 284 (7):4464-4472 (2009)) identifies five Prx family proteins [i.e., Tsa1, Tsa2, Prx1, Ahp1, Dot5] and two Gpx-like proteins [i.e., Gpx2, Gpx3] as having thioredoxin-dependent peroxidase activity in budding yeast (all of which will be referred to generically herein as Prx proteins). The exact nature by which these proteins interact with Yap1 continues to be investigated. Tachibana, T. et al. (*J. Biol. Chem.*, 284:4464-4472 (2009)) report that *S. cerevisiae* Tsa1 (SEQ ID NO:34) interacts with Yap1 in a manner similar to that of Gpx3, based on Cys-48 and Cys-171.

Although the exact mechanism by which Yap1 and Ybp1 interact is unknown, Ybp1 has been demonstrated to also affect activation of the Yap1 transcription factor. Gulshan, K. et al. (*J. Biol. Chem.*, 286(39):34071-34081 (2011)) studied the interaction between Ybp1 and Yap1 in both *S. cerevisiae* and *Candida glabrata*; they report that “Yap1 and Ybp1 are likely to directly interact in the cell... efforts to further localize the interaction motifs of these two proteins were unsuccessful”. It is hypothesized therein that the interaction of Yap1 and Ybp1 likely involves multiple, low-affinity interactions while oxidation of Yap1 likely triggers release of the folded protein from its Ybp1 partner. Ybp1 overproduction in *S. cerevisiae* was also reported to lead to increased H$_2$O$_2$ tolerance.

Mechanisms of oxidative stress response are relatively well conserved (although there are some differences among species) and Yap1p homologs, such as Cap1p in *Candida albicans* and Pap1p in *Schizosaccharomyces pombe*, are also known to transcriptionally regulate some anti-oxidant genes in response to oxidative stress (lkner, A. and K. Shiozaki, *Mutat. Res.*, 569:13–27 (2005)). However, the means by which the oxidative stress response functions in oleaginous yeast is much less
well characterized. Oleaginous yeast are naturally capable of oil synthesis and accumulation, wherein the total oil content can comprise greater than about 25% of the dry cell weight ["DCW"], more preferably greater than about 30% of the DCW, more preferably greater than about 40% of the DCW, more preferably greater than about 50% of the DCW, and most preferably greater than about 60% of the DCW (wherein this rate of oil accumulation is prior to any efforts to increase the native Yap1 transcription factor activity in the yeast, according to the invention herein). Various yeast are naturally classified as oleaginous; however, in alternate embodiments, a non-oleaginous organism can be genetically modified to become oleaginous, e.g., yeast such as *Saccharomyces cerevisiae* (see, Intl'. App. Pub. No. WO 2006/102342).

Genera typically identified as oleaginous yeast include, but are not limited to: *Yarrowia*, *Candida*, *Rhodotorula*, *Rhodospiridium*, *Cryptococcus*, *Trichosporon* and *Lipomyces*. More specifically, illustrative oil-synthesizing yeasts include: *Rhodospiridium toruloides*, *Lipomyces starkeyii*, *L. lipoferus*, *Candida revkaufi*, *C. pulcherrima*, *C. tropicalis*, *C. utilis*, *Trichosporon pullans*, *T. cutaneum*, *Rhodotorula glutinis*, *R. graminis*, and *Yarrowia lipolytica* (such as, for example, but not limited to the *Y. lipolytica* strains designated as ATCC #20362, ATCC #8862, ATCC #18944, ATCC #76982 and/or LGAM S(7)1 (Papanikolaou S., and Aggelis G., *Bioresour. Technol.*, 82(1):43-9 (2002)).

Like other organisms that conduct aerobic metabolism and thus rely on defenses against ROS (created via incomplete reduction of oxygen to water during respiration), obligatory aerobic oleaginous yeast also require various means to sense and respond to oxidative stress. In the present invention, homologs of *S. cerevisiae* Yap1, Gpx3, Tsa1 and Ybp1 genes have been identified in the oleaginous yeast, *Yarrowia lipolytica*, as summarized in the Table below. Alignments of each pair of proteins were created with CLUSTAL W (1.81) multiple sequence alignment (Thompson J.D., et al., *Nucleic Acids Res.* 22:4673-4680 (1994)).

**Table 2. Yap1, Gpx3, Tsa1 And Ybp1 Homologs**
Surprisingly, when the *Y. lipolytica* Yap1 ["YfYap1"] and Gpx3 ["YfGpx3"] proteins were over-expressed to result in increased Yap1 transcription factor activity, the transgenic *Y. lipolytica* was found to have increased oil content (measured as TFAs % DCW), as compared to the oil content in a non-transgenic *Y. lipolytica*.

Thus, the instant invention concerns a transgenic oleaginous yeast having increased oil content comprising increased Yap1 transcription factor activity wherein the increased oil content is compared to the oil content of a non-transgenic oleaginous yeast.

It is hypothesized herein that increased oil content is observed in the transgenic oleaginous yeast since increased Yap1 transcription factor activity provides increased resistance to oxidative stresses. One beneficial outcome of this increased resistance to oxidative stresses is increased protection against lipid peroxidation, which thereby results in increased oil/lipid content in the transgenic oleaginous yeast. Among lipid molecules, PUFAs are particularly sensitive to ROS, and it was shown that the susceptibility of fatty acids to lipid peroxidation increased with the degree of fatty acyl chain unsaturation (Porter, N.A. et al., *Lipids*, 30:277–290 (1995)). The lipid peroxidation was shown to affect cell viability via generation of polar hydroperoxides which affect membrane integrity (Howlett, N.G. and S.V. Avery, *Appl. Microbiol. Biotechnol.*, 48(4):539-545 (1997); Howlett, N.G. and S.V. Avery, *Appl. Environ. Microbiol.*,)
63(8):2971-2976 (1997)). However, no study has previously shown the effect of the Yap1 transcription factor overexpression in oil content.

Preferably, the transgenic oleaginous yeast of the present invention will be capable of producing at least 10-25% greater oil content than the oil content of a non-transgenic oleaginous yeast. More preferably, the increase in oil content is at least 25-45% greater, and most preferably the increase in oil content is at least 45-65% greater than the oil content of a non-transgenic oleaginous yeast. Thus, those skilled in the art will appreciate that the increase in oil content can be any integer percentage (or fraction thereof) from 10% up to and including 100% or greater, i.e., specifically, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% increase in oil content when compared to the oil content of a non-transgenic oleaginous yeast.

As described above, the microbial oil will comprise triacylglycerols (comprising long chain polyunsaturated and/or saturated fatty acids, as well as shorter chain saturated and/or unsaturated fatty acids), as well as other neutral lipids, phospholipids and free fatty acids.

In one embodiment, the increased Yap1 transcription factor activity in the transgenic oleaginous yeast having increased oil content results from overexpressing a Yap1 transcription factor. A suitable Yap1 transcription factor will preferably comprise a nucleotide sequence encoding a polypeptide having transcription factor activity and comprising:

a) a bZIP leucine zipper motif;

b) an N-terminal Cys-rich domain comprising a sequence of at least two cysteine residues that are separated by at least 6 amino acids; and,
c) a C-terminal Cys-rich domain comprising a sequence of at least two cysteine residues that are separated by at least 8 amino acids.

A bZIP leucine zipper motif comprises a basic DNA binding region spanning approximately fourteen to sixteen amino acids (i.e., comprising arginine and lysine residues) and an adjacent leucine-rich zipper region (i.e., comprising evenly spaced leucine residues allowing dimerization) (Hurst, H.C., Protein Profile, 2:101-168 (1995)).

One preferred Yap1 transcription factor is the Yarrowia lipolytica Yap1 ["YIYap1"] polypeptide sequence, as set forth in SEQ ID NO:4. In alternate embodiments, the ScYap1 (SEQ ID NO:2) or any of the sequences set forth in Table 4 (Example 1), or homologs or codon-optimized derivatives thereof, may be used in the present invention.

In another embodiment, the increased Yap1 transcription factor activity in the transgenic oleaginous yeast having increased oil content results by increasing the interaction between the transcription factor and a protein that is capable of activating the transcription factor (i.e., by overexpressing the protein capable of activating the transcription factor itself). Preferably, the protein that is capable of activating the transcription factor is selected from the group consisting of: Gpx3, Ybp1 and Tsa1.

For example, a suitable Gpx3 protein will comprise:

a) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the polypeptide has at least 70% amino acid identity, based on the BLASTP method of alignment, when compared to a sequence selected from the group consisting of SEQ ID NO:26 [ScGpx3] or SEQ ID NO:28 [YIGpx3];

b) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the nucleotide sequence has at least 70% sequence identity, based on the BLASTN method of alignment, when compared to a sequence selected
from the group consisting of SEQ ID NO:25 [ScGpx3] or SEQ ID NO:27 [YIGpx3]; or

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

A suitable Gpx3 protein will comprise at least one redox-active cysteine residue, such as Cys36 and Cys82 of ScGpx3 (SEQ ID NO:26).

Preferably, the polypeptide sequence encoding Gpx3 is set forth in SEQ ID NO:28 (“YIGpx3”). In alternate embodiments, the polypeptide sequence encoding Gpx3 has at least 70% sequence identity based on the CLUSTALW method of alignment, when compared to SEQ ID NO:26 or SEQ ID NO:28, i.e., the polypeptide may have at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity when compared thereto. In alternate embodiments, the sequences set forth in Table 9 (Example 5), or homologs or codon-optimized derivatives thereof, may be used in the present invention.

Similarly, a suitable Tsa1 protein will comprise:

a) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the polypeptide has at least 70% amino acid identity, based on the BLASTP method of alignment, when compared to a sequence selected from the group consisting of SEQ ID NO:34 [ScTsa1] or SEQ ID NO:36 [YITsa1];

b) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the nucleotide sequence has at least 70% sequence identity, based on the BLASTN method of alignment, when compared to a sequence selected from the group consisting of SEQ ID NO:33 [ScTsa1] or SEQ ID NO:35 [YITsa1]; or,
c) a complement of the nucleotide sequence of (a) or (b), wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

A suitable Tsa1 protein will comprise at least one redox-active cysteine residue, such as Cys48 and Cys171 of ScTsa1 (SEQ ID NO:34). Preferably, the polypeptide sequence encoding Tsa1 is set forth in SEQ ID NO:36 ("YITsa1"). In alternate embodiments, the polypeptide sequence encoding Tsa1 has at least 70% sequence identity based on the CLUSTALW method of alignment, when compared to SEQ ID NO:34 or SEQ ID NO:36, i.e., the polypeptide may have at least 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity when compared thereto. In alternate embodiments, the sequences set forth in Table 12 (Example 9), or homologs or codon-optimized derivatives thereof, may be used in the present invention.

A suitable Ybp1 protein will comprise a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the polypeptide sequence is classified within a kinetochor_Ybp2 super family, based on the conserved domain method of analysis.

Preferably, the polypeptide sequence encoding Ybp1 is set forth in SEQ ID NO:40 ("YIYbp1"). In alternate embodiments, the ScYbp1 (SEQ ID NO:38) or any of the sequences set forth in Table 13 or Table 14 (Example 10, i.e., including SEQ ID NOs:43-48), or homologs or codon-optimized derivatives thereof, may be used in the present invention.

For clarity, the increased Yap1 transcription factor activity in the transgenic oleaginous yeast of the present invention can be achieved by overexpression of a native Yap1 transcription factor, a foreign Yap1 transcription factor, a native protein that is capable of activating the transcription factor, a foreign protein that is capable of activating the transcription factor, or any combination thereof. Overexpression may occur, for example, by introducing additional copies of appropriate genes into the host cell on multicopy plasmids. Such genes may also be
integrated into the chromosome with appropriate regulatory sequences that result in increased activity of their encoded functions. The target genes may be modified so as to be under the control of non-native promoters or altered native promoters. Endogenous promoters can be altered in vivo by mutation, deletion, and/or substitution.

As noted above, it may be desirable to codon-optimize any one of the Yap1, Gpx3, Tsa1 or Ybp1 proteins described above for expression in the oleaginous yeast of interest. For example, one could codon-optimize any of the sequences set forth in Tables 4, 9, 12, 13 or 14 for expression in Y. lipolytica. This is possible based on previous determination of the Y. lipolytica codon usage profile, identification of those codons that are preferred, and determination of the consensus sequence around the ‘ATG’ initiation codon (see U.S. Pat. 7,238,482).

In another embodiment, the sequences set forth in Tables 4, 9, 12, 13 or 14, or portions of thereof, may be used to search for Yap1, Gpx3, Tsa1 or Ybp1 homologs in the same or other species using sequence analysis software. In general, such computer software matches similar sequences by assigning degrees of homology to various substitutions, deletions, and other modifications. Use of software algorithms, such as the BLASTP method of alignment with a low complexity filter and the following parameters: Expect value = 10, matrix = Blosum 62 (Altschul, et al., Nucleic Acids Res., 25:3389-3402 (1997)), is well-known as a means for comparing any Yap1, Gpx3, Tsa1 or Ybp1 protein in Tables 4, 9, 12, 13 or 14 against a database of nucleic or protein sequences and thereby identifying similar known sequences within a preferred organism.

Use of a software algorithm to comb through databases of known sequences is particularly suitable for the isolation of homologs having a relatively low percent identity to publicly available Yap1, Gpx3, Tsa1 or Ybp1 sequences, such as those described in Tables 4, 9, 12, 13 or 14. It is predictable that isolation would be relatively easier for Yap1, Gpx3, Tsa1 or Ybp1 homologs of at least about 70%-75% identity and more preferably at least about 80%-85% identity to publicly available Yap1, Gpx3, Tsa1 or Ybp1 sequences. Further, those sequences that are at
least about 85%-90% identical would be particularly suitable for isolation and those sequences that are at least about 90%-95% identical would be the most easily isolated.

Some Yap1, Gpx3, Tsa1 or Ybp1 homologs have also been isolated by the use of motifs unique to these enzymes. As one will appreciate, this is particularly useful with transcription factors, which share relatively low sequence homology with one another, despite sharing several conserved sequence motifs. Motifs (e.g., the basic DNA binding region and adjacent leucine-rich zipper region of the bZIP leucine zipper motif, N-terminal Cys-rich domain and C-terminal Cys-rich domain of a Yap1 transcription factor) are identified by their high degree of conservation in aligned sequences of a family of protein homologues. As unique "signatures", they can determine if a protein with a newly determined sequence belongs to a previously identified protein family.

Similarly, Gpx3 and Tsa1 homologs are expected to comprise at least one redox-active cysteine residue, whose relative position within the protein sequence will be conserved. These motifs are useful as diagnostic tools for the rapid identification of novel homologous genes.

Any of the Yap1, Gpx3, Tsa1 or Ybp1 nucleic acid fragments described herein or in public literature, or any identified homologs, may be used to isolate genes encoding homologous proteins from the same or other species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to: 1) methods of nucleic acid hybridization; 2) methods of DNA and RNA amplification, as exemplified by various uses of nucleic acid amplification technologies, such as polymerase chain reaction ["PCR"] (U.S. Pat. No.4,683,202); ligase chain reaction ["LCR"] (Tabor, S. et al., Proc. Natl. Acad. Sci. U.S.A., 82:1074 (1985)); or strand displacement amplification ["SDA"] (Walker, et al., Proc. Natl. Acad. Sci. U.S.A., 89:392 (1992)); and, 3) methods of library construction and screening by complementation.

The present invention is also drawn to methods of increasing oil content in an oleaginous yeast, wherein said method comprises:
a) engineering the oleaginous yeast to overexpress a protein selected from the group consisting of:
   (i) a Yap1 transcription factor;
   (ii) a protein that is capable of activating the transcription factor;
   (iii) a combination of (a) and (b); and,

b) growing the oleaginous yeast under suitable conditions to result in increased oil content when compared to the oil content of a non-transgenic oleaginous yeast.

One of ordinary skill in the art is aware of standard resource materials that describe: 1) specific conditions and procedures for construction, manipulation and isolation of macromolecules, such as DNA molecules, plasmids, etc.; 2) generation of recombinant DNA fragments and recombinant expression constructs; and, 3) screening and isolating of clones. See, Maniatis, Sihavy, and Ausubel, as cited above.

In general, the choice of sequences included in a recombinant expression construct depends on the desired expression products, the nature of the host cell and the proposed means of separating transformed cells versus non-transformed cells. Typically, a vector contains at least one expression cassette, a selectable marker and sequences allowing autonomous replication or chromosomal integration. Suitable expression cassettes typically comprise a promoter, the coding sequence of a selected gene (e.g., encoding a polypeptide whose expression results in increased Yap1 transcription factor activity), and a terminator (i.e., a chimeric gene). Preferably, both control regions are derived from genes from the transformed host cell.

Virtually any promoter (i.e., native, synthetic, or chimeric) capable of directing expression of an ORF encoding a polypeptide of the invention herein will be suitable, although transcriptional and translational regions from Y. lipolytica are particularly useful. Expression can be accomplished in an induced or constitutive fashion. Induced expression can be accomplished by inducing the activity of a regulatable promoter operably linked to the gene of interest (e.g., Yap1, Gpx3, Tsa1, Ybp1), while
constitutive expression can be achieved by the use of a constitutive promoter operably linked to the gene of interest.

A terminator can be derived from the 3' region of a gene from which the promoter was obtained or from a different gene. A large number of terminators are known and function satisfactorily in a variety of hosts, when utilized both in the same and different genera and species from which they were derived. The terminator usually is selected more as a matter of convenience rather than because of any particular property. Preferably, the terminator is derived from a yeast gene. The terminator can also be synthetic, as one of skill in the art can utilize available information to design and synthesize a terminator. A terminator may be unnecessary, but it is highly preferred.


Many specialized expression vectors have been created to obtain a high expression rate. Such vectors are made by adjusting certain properties that govern transcription, RNA stability, translation, protein stability and location, and secretion from the host cell. These properties include: the nature of the relevant transcriptional promoter and terminator sequences; the number of copies of the cloned gene (wherein additional copies may be cloned within a single expression construct and/or additional copies may be introduced into the host cell by increasing the
plasmid copy number or by multiple integration of the cloned gene into the genome; whether the gene is plasmid-borne or integrated into the host cell genome; the efficiency of translation and correct folding of the protein in the host organism; the intrinsic stability of the mRNA and protein of the cloned gene within the host cell; and, the codon usage within the cloned gene, such that its frequency approaches the frequency of preferred codon usage of the host cell.

Once a DNA cassette (e.g., comprising a chimeric gene comprising a promoter, an ORF encoding a polypeptide whose expression results in increased Yap1 transcription factor activity [e.g., Yap1, Gpx3, Tsa1, Ybp1], and a terminator) suitable for expression in an oleaginous yeast has been obtained, it is placed in a plasmid vector capable of autonomous replication in the host cell, or DNA fragment containing the chimeric gene is directly integrated into the genome. Integration of expression cassettes can occur randomly within the host genome or can be targeted through the use of constructs containing regions of homology with the genome sufficient to target recombination to a particular locus. Where constructs are targeted to an endogenous locus, all or some of the transcriptional and translational regulatory regions can be provided by the endogenous locus.

Constructs comprising a chimeric gene(s) of interest may be introduced into oleaginous yeast by any standard technique. These techniques include transformation (e.g., lithium acetate transformation [*Methods in Enzymology*, 194:186-187 (1991)], bolistic impact, electroporation, microinjection, or any other method that introduces the gene(s) of interest into the host cell. More specific teachings applicable for *Y. lipolytica* include U.S. Pat. No. 4,880,741 and U.S. Pat. No. 5,071,764 and Chen, D. C. et al. (*Appl. Microbiol. Biotechnol.*, 48(2):232-235 (1997)). Preferably, integration of a linear DNA fragment into the genome of the host is favored in transformation of *Y. lipolytica* host cells. Integration into multiple locations within the genome can be particularly useful when high level expression of genes are desired. Preferred loci include those taught in U.S. Pat. Pub. No. 2009-0093543-A1.
The terms "transformed", "transformant" or "recombinant" are used interchangeably herein. A transformed host will have at least one copy of an expression construct and may have two or more, depending upon whether the expression cassette is integrated into the genome, amplified or is present on an extrachromosomal element having multiple copy numbers. The transformed host cell can be identified by selection for a marker contained on the introduced construct. Alternatively, a separate marker construct may be co-transformed with the desired construct, as many transformation techniques introduce many DNA molecules into host cells. Typically, transformed hosts are selected for their ability to grow on selective media, which may incorporate an antibiotic or lack a factor necessary for growth of the untransformed host, such as a nutrient or growth factor. An introduced marker gene may confer antibiotic resistance, or encode an essential growth factor or enzyme, thereby permitting growth on selective media when expressed in the transformed host. Selection of a transformed host can also occur when the expressed marker protein can be detected, either directly or indirectly. Additional selection techniques are described in U.S. Pat. 7,238,482, U.S. Pat. 7,259,255 andWO 2006/052870.

Stability of an integrated DNA fragment in oleaginous yeast is dependent on the individual transformants, the recipient strain and the targeting platform used. Thus, multiple transformants of a particular recombinant microbial host should be screened in order to obtain a strain displaying the desired expression level and pattern. Southern analysis of DNA blots (Southern, *J. Mol. Biol.*, 98:503 (1975)), Northern analysis of mRNA expression (Kroczek, *J. Chromatogr. Biomed. Appl.*, 618 (1-2):133-145 (1993)), Western analysis of protein expression, phenotypic analysis or GC analysis are suitable screening methods.

Suitable host cells for use in the invention herein are oleaginous yeast, capable of accumulating oil in excess of about 25% of their DCW, as defined above. In some embodiments herein, the oleaginous yeast host is a wildtype strain; in alternate embodiments, the oleaginous yeast host is a transformed or recombinant strain that was previously subjected
to transformation with an expression construct that does not affect the native level of Yap1 transcription factor activity. For example, in some embodiments, the oleaginous yeast has been previously modified such that it is capable of producing at least one non-native product of interest, wherein examples of suitable non-native products of interest include, e.g., polyunsaturated fatty acids, carotenoids, amino acids, vitamins, sterols, flavonoids, organic acids, polyols and hydroxyesters, quinone-derived compounds and resveratrol, although this is not intended to be limiting herein.

It is noted that an oleaginous yeast host may produce "polyunsaturated fatty acids" (or "PUFAs") within its microbial oils (either through natural abilities or genetic modifications). Although the health benefits associated with PUFAs, especially omega-3 and omega-6 PUFAs, have been well documented, these molecules are particularly susceptible to lipid peroxidation within the cell since they contain multiple double bonds in between which lie methylene-CH₂- groups that possess especially reactive hydrogens. More specifically, PUFAs refer herein to fatty acids having at least 18 carbon atoms and 2 or more double bounds. The term "fatty acids" refers to long chain aliphatic acids (alkanoic acids) of varying chain lengths, from about C₁₂ to C₂₂, although both longer and shorter chain-length acids are known. The predominant chain lengths are between C₁₆ and C₂₂. The structure of a fatty acid is represented by a simple notation system of "X:Y", where X is the total number of carbon ["C"] atoms in the particular fatty acid and Y is the number of double bonds. Additional details concerning the differentiation between "saturated fatty acids" versus "unsaturated fatty acids", "monounsaturated fatty acids" versus "polyunsaturated fatty acids" ["PUFAs"], and "omega-6 fatty acids" ["ω-6"] versus "omega-3 fatty acids" ["ω-3"] are provided in U.S. Patent 7,238,482, which is hereby incorporated herein by reference. U.S. Pat. App. Pub. No. 2009-0093543-A1, Table 3, provides a detailed summary of the chemical and common names of omega-3 and omega-6 PUFAs and their precursors, and well as commonly used abbreviations.
Some examples of PUFAs, however, include, but are not limited to, linoleic acid ["LA", 18:2 omega-6], gamma-linolenic acid ["GLA", 18:3 omega-6], eicosadienoic acid ["EDA", 20:2 omega-6], dihomo-gamma-linolenic acid ["GLA", 20:3 omega-6], arachidonic acid ["ARA", 20:4 omega-6], docosatetraenoic acid ["DTA", 22:4 omega-6], docosapentaenoic acid ["DPAn-6", 22:5 omega-6], alpha-linolenic acid ["ALA", 18:3 omega-3], stearidonic acid ["STA", 18:4 omega-3], eicosatrienoic acid ["ETA", 20:3 omega-3], eicosatetraenoic acid ["ETrA", 20:4 omega-3], eicosapentaenoic acid ["EPA", 20:5 omega-3], docosapentaenoic acid ["DPAn-3", 22:5 omega-3] and docosahexaenoic acid ["DHA", 22:6 omega-3].

Much effort has been invested towards engineering strains of Y. lipolytica for PUFA production. For example, U.S. Patent No. 7,238,482 demonstrated the feasibility of producing omega-6 and omega-3 fatty acids in the yeast. U.S. Patent No. 7,932,077 demonstrated recombinant production of 28.1% EPA of total fatty acids; U.S. Patent No. 7,588,931 demonstrated recombinant production of 14% ARA of total fatty acids; U.S. Patent No. 7,550,286 demonstrated recombinant production of 5% DHA of total fatty acids; and, U.S. Pat. Appl. Pub. No. 2009-0093543-A1 describes optimized recombinant strains for EPA production and demonstrated production of up to 55.6% EPA of total fatty acids. U.S. Pat. Appl. Pub. No. 2010-0317072-A1 describes further optimized recombinant Y. lipolytica strains producing microbial oils comprising up to 50% EPA of TFAs and having a ratio of at least 3.1 of EPA, measured as a weight percent of TFAs, to linoleic acid, measured as a weight percent of TFAs. The transformant Y. lipolytica express various combinations of desaturase (i.e., delta-12 desaturase, delta-6 desaturase, delta-8 desaturase, delta-5 desaturase, delta-17 desaturase, delta-15 desaturase, delta-9 desaturase, delta-4 desaturase) and elongase (i.e., C_{14/16} elongase, C_{16/18} elongase, C_{18/20} elongase, C_{20/22} elongase and delta-9 elongase) genes for PUFA production.

Table 3 provides information about some of the specific Y. lipolytica strains described in the above cited references, wherein said strains
possess various combinations of desaturases and elongases. It is to be recognized that these are exemplary strain which could be used as suitable host cells in the invention herein, although the specific strain and the specific strains and the specific PUFA produced (or quantities thereof) are by no means limiting to the invention herein.
### Table 3. Lipid Profile of Representative Y. lipolytica Strains Engineered to Produce Omega-3/Omega-6 PUFAs

<table>
<thead>
<tr>
<th>Strain</th>
<th>Reference</th>
<th>ATCC Deposit No.</th>
<th>Fatty Acid Content (As A Percent [%] of Total Fatty Acids)</th>
<th>TFAs % DCW</th>
</tr>
</thead>
<tbody>
<tr>
<td>pDMW208</td>
<td>U.S. Pat. No. 7,465,564</td>
<td></td>
<td>11.9 8.6 1.5 24.4 17.8 0 25.9 -- -- -- -- -- -- -- --</td>
<td></td>
</tr>
<tr>
<td>pDMW208D62</td>
<td></td>
<td></td>
<td>16.2 1.5 0.1 17.8 22.2 0 34 -- -- -- -- -- -- -- --</td>
<td></td>
</tr>
<tr>
<td>M4</td>
<td>U.S. Pat. No. 7,932,077</td>
<td></td>
<td>-- 15 4 2 5 27 0 35 -- 8 0 0 0 -- -- -- --</td>
<td></td>
</tr>
<tr>
<td>Y2034</td>
<td></td>
<td></td>
<td>-- 13.1 8.1 1.7 7.4 14.8 0 25.2 -- 8.3 11.2 -- -- -- --</td>
<td></td>
</tr>
<tr>
<td>Y2047</td>
<td>U.S. Pat. No. 7,586,931</td>
<td>PTA-7186</td>
<td>15.9 6.6 0.7 8.9 16.6 0 29.7 -- 0 10.9 -- -- -- --</td>
<td></td>
</tr>
<tr>
<td>Y2214</td>
<td></td>
<td></td>
<td>-- 7.9 15.3 0 13.7 37.5 0 0 -- 7.9 14 -- -- -- --</td>
<td></td>
</tr>
<tr>
<td>EU</td>
<td></td>
<td></td>
<td>-- 19 10.3 2.3 15.8 12 0 18.7 -- 5.7 0.2 3 10.3 -- -- 36</td>
<td></td>
</tr>
<tr>
<td>Y2072</td>
<td></td>
<td></td>
<td>-- 7.6 4.1 2.2 16.8 13.9 0 27.8 -- 3.7 1.7 2.2 15 -- --</td>
<td></td>
</tr>
<tr>
<td>Y2102</td>
<td></td>
<td></td>
<td>-- 9 3 3.5 5.6 18.6 0 29.6 -- 3.8 2.8 2.3 18.4 -- --</td>
<td></td>
</tr>
<tr>
<td>Y2095</td>
<td></td>
<td></td>
<td>-- 13 0 2.6 5.1 16 0 29.1 -- 3.1 1.9 2.7 19.3 -- --</td>
<td></td>
</tr>
<tr>
<td>Y2090</td>
<td></td>
<td></td>
<td>-- 6 1 6.1 7.7 12.6 0 26.4 -- 6.7 2.4 3.6 26.6 -- 22.9</td>
<td></td>
</tr>
<tr>
<td>Y2096</td>
<td></td>
<td>PTA-7184</td>
<td>8.1 1 6.3 8.5 11.5 0 25 -- 5.8 2.1 2.5 28.1 -- -- 20.8</td>
<td></td>
</tr>
<tr>
<td>Y2201</td>
<td></td>
<td>PTA-7185</td>
<td>11 16.1 0.7 18.4 27 0 -- 3.3 3.3 1 3.8 9 -- -- --</td>
<td></td>
</tr>
<tr>
<td>Y3000</td>
<td>U.S. Pat. No. 7,550,286</td>
<td>PTA-7187</td>
<td>5.9 12 5.5 7.7 11.7 0 30.1 -- 2.6 1.2 1.2 4.7 18.3 5.6 --</td>
<td></td>
</tr>
<tr>
<td>Y4001</td>
<td></td>
<td></td>
<td>-- 4.3 4.4 3.9 35.9 23 0 -- 23.8 0 0 0 -- -- -- --</td>
<td></td>
</tr>
<tr>
<td>Y4036</td>
<td></td>
<td></td>
<td>-- 7.7 3.6 1.1 14.2 32.6 0 -- 15.6 18.2 0 0 -- -- -- --</td>
<td></td>
</tr>
<tr>
<td>Y4070</td>
<td></td>
<td></td>
<td>-- 8 5.3 3.5 14.6 42.1 0 -- 6.7 2.4 11.9 -- -- -- --</td>
<td></td>
</tr>
<tr>
<td>Y4086</td>
<td></td>
<td></td>
<td>-- 3.3 2.2 4.6 26.3 27.9 6.9 -- 7.6 1 0 2 9.8 -- -- 28.6</td>
<td></td>
</tr>
<tr>
<td>Y4128</td>
<td></td>
<td>PTA-8514</td>
<td>6.6 4 2 8.8 19 2.1 -- 4.1 3.2 0 5.7 42.1 -- -- 18.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Y4158</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y4184</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y4259</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y4305</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y4127</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTA-8802</td>
<td>4.1</td>
<td>2.3</td>
<td>2.9</td>
<td>15.4</td>
</tr>
<tr>
<td>Y4184</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y8402</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y8404</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y8406</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTA-10025</td>
<td>2.6</td>
<td>0.5</td>
<td>2.9</td>
<td>5.7</td>
</tr>
<tr>
<td>Y8412</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y8413</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y8647</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y9028</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y9477</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y9479</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y9502</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y9508</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y8145</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTA-10026</td>
<td>2.5</td>
<td>0.4</td>
<td>2.6</td>
<td>4.3</td>
</tr>
<tr>
<td>Y8259</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y8370</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y8672</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**

The terms "lipid profile" and "lipid composition" are interchangeable and refer to the amount of individual fatty acids contained in a particular lipid fraction, such as in the total lipid or the oil, wherein the amount is expressed as a wt % of TFAs. The sum of each individual fatty acid present in the mixture should be 100.
The term "total fatty acids" ("TFAs") refer to the sum of all cellular fatty acids that can be derivitized to fatty acid methyl esters ("FAMEs") by the base transesterification method (as known in the art) in a given sample, which may be the biomass or oil, for example. Thus, total fatty acids include fatty acids from neutral lipid fractions (including diacylglycerols, monoacylglycerols and triacylglycerols) and from polar lipid fractions but not free fatty acids. The concentration of a fatty acid in the total lipid is expressed herein as a weight percent of TFAs ["% TFAs"], e.g., milligrams of the given fatty acid per 100 milligrams of TFAs. Unless otherwise specifically stated in the disclosure herein, reference to the percent of a given fatty acid with respect to total lipids is equivalent to concentration of the fatty acid as % TFAs (e.g., % EPA of total lipids is equivalent to EPA % TFAs).

Fatty acids are 16:0 (palmitate), 16:1 (palmitoleic acid), 18:0 (stearic acid), 18:1 (oleic acid), 18:2 (linoleic acid), 18:3 (ALA or alpha-linolenic acid), GLA (gamma-linolenic acid), 20:2 (EDA or eicosadienoic acid), DGLA (dihomo-gamma-linolenic acid), ARA (arachidonic acid), ETA (eicosatetraenoic acid), EPA (eicosapentaenoic acid), DPA (docosapentaenoic acid) and DHA (docosahexaenoic acid).
It will be obvious to one of ordinary skill in the art that means to reduce reactive oxygen species ["ROS"] in oleaginous yeast producing at least one PUFA will be particularly desirable. Thus, one embodiment of the present invention concerns a transgenic oleaginous yeast having increased oil content and producing at least one PUFA, wherein said transgenic oleaginous yeast comprises increased Yap1 transcription factor activity and wherein the increased oil content is compared to the oil content of a non-transgenic oleaginous yeast. Increased Yap1 transcription factor activity, via overexpression of the Yap1 transcription factor itself or by overexpression of a protein that is capable of activating the Yap1 transcription factor (e.g., Gpx3, Tsa1 Ybp1), may additionally result in increased content of a given PUFA(s) in a cell as its weight percent of the dry cell weight ["% DCW"].

For example, a measure of EPA productivity or EPA titer ["EPA % DCW"] is determined according to the following formula: (EPA % TFAs) * (TFAs % DCW)}/100. In any of the strains set forth above in Table 3, producing primarily EPA, it is expected that genetic manipulation that results in increased Yap1 transcription factor activity in the yeast will result in both increased oil content ["TFAs % DCW"] and increased EPA titer ["EPA % DCW"].

In preferred embodiments, a transgenic oleaginous yeast of the present invention that produces at least one PUFA will be capable of producing at least 10-25% greater content of a given PUFA(s) as its weight percent of the DCW than the content of the given PUFA(s) as its weight percent of the DCW in a non-transgenic oleaginous yeast (i.e., whose Yap1 transcription factor activity has not been increased). More preferably, the increase in the given PUFA(s) is at least 25-45%, and most preferably the increase in the given PUFA(s) is at least 45-65% greater. Thus, those skilled in the art will appreciate that the increase in the given PUFA(s) as its weight percent of the DCW can be any integer percentage (or fraction thereof) from 10% up to and including 100% or greater, i.e., specifically, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19%, 20%, 21%, 22%, 23%, 24%, 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%,
33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100%.

A transformed host cell can be grown under conditions that optimize expression of chimeric genes (e.g., encoding a polypeptide whose expression results in increased Yap1 transcription factor activity [e.g., Yap1, Gpx3, Tsa1, Ybp1], etc.) and produce the greatest and the most economical yield of the microbial oils. In general, media conditions that may be optimized include: the type and amount of carbon source, the type and amount of nitrogen source, the carbon-to-nitrogen ratio, the amount of different mineral ions, the oxygen level, growth temperature, pH, length of the biomass production phase, length of the oil accumulation phase and the time and method of cell harvest. Oleaginous yeast are grown in a complex medium (e.g., yeast extract-peptone-dextrose broth (YPD)) or a defined minimal medium that lacks a component necessary for growth and thereby forces selection of the desired expression cassettes (e.g., Yeast Nitrogen Base (DIFCO Laboratories, Detroit, MI)).

Fermentation media for the methods and host cells described herein must contain a suitable carbon source, such as are taught in U.S. Patent 7,238,482 and U.S. Pat. Pub. No. 2011-0059204-A1. Although it is contemplated that the source of carbon utilized in the present invention may encompass a wide variety of carbon-containing sources, preferred carbon sources are sugars, glycerol and/or fatty acids. Most preferred is glucose, sucrose, invert sucrose, fructose and/or fatty acids containing between 10-22 carbons.

Nitrogen may be supplied from an inorganic (e.g., (NH₄)₂SO₄) or organic (e.g., urea, glutamate, or yeast extract) source. In addition to sucrose and nitrogen sources, the fermentation medium also contains suitable minerals, salts, cofactors, buffers, vitamins and other components known to those skilled in the art suitable for the growth of the
microorganism and promotion of the enzymatic pathways necessary for microbial oil production. Particular attention is given to several metal ions (e.g., Fe$^{+2}$, Cu$^{+2}$, Mn$^{+2}$, Co$^{+2}$, Zn$^{+2}$, Mg$^{+2}$) that promote synthesis of lipids and PUFAs (Nakahara, T. et al., *Ind. Appl. Single Cell Oils*, D. J. Kyle and R. Colin, eds. pp 61-97 (1992)).

Preferred growth media in the present invention are common commercially prepared media, such as Yeast Nitrogen Base (DIFCO Laboratories, Detroit, MI). Other defined or synthetic growth media may also be used and the appropriate medium for growth of the particular microorganism will be known by one skilled in the art of microbiology or fermentation science. A suitable pH range for the fermentation is typically between about pH 4.0 to pH 8.0, wherein pH 5.5 to pH 7.5 is preferred as the range for the initial growth conditions. The fermentation may be conducted under aerobic or anaerobic conditions, wherein microaerobic conditions are preferred.

Typically, accumulation of high levels of PUFAs in oleaginous yeast cells requires a two-stage fermentation process, since the metabolic state must be “balanced” between growth and synthesis/storage of fats. Thus, most preferably, a two-stage fermentation process is employed for the production of PUFAs in oleaginous yeast. This process is described in U.S. Patent 7,238,482, as are various suitable fermentation process designs (i.e., batch, fed-batch and continuous) and considerations during growth.

Example 10 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1 also provides a detailed description of parameters required for a 2-L fermentation of the recombinant *Yarrowia lipolytica* strain Y4305 (whose maximum production was 12.1 EPA % DCW [i.e., 55.6 EPA % TFAs, with a ratio of EPA % TFAs to LA % TFAs of 3.03], over a period of 162 hours). This disclosure includes a description of means to prepare inocula from frozen cultures to generate a seed culture, initially culture the yeast under conditions that promoted rapid growth to a high cell density, and then culture the yeast to promote lipid and PUFA accumulation (via starving for nitrogen and continuously feeding glucose). Process variables including
temperature (controlled between 30-32 °C), pH (controlled between 5-7), dissolved oxygen concentration and glucose concentration were monitored and controlled per standard operating conditions to ensure consistent process performance and final PUFA oil quality. In particular, the data of Example 10 of U.S. Pat. Appl. Pub. No. 2009-0093543-A1 are useful to demonstrate that the oil profile of the recombinant microbial host cell will depend on the fermentation run itself, media conditions, process parameters, scale-up, etc., as well as the particular time-point in which the culture is sampled. Thus, the particular engineered strain therein was capable of producing microbial oil having a variety of different lipid contents and compositions (i.e., based on EPA % TFAs, LA % TFAs and EPA:LA ratio).

These factors should be considered when culturing the transgenic oleaginous yeast described herein, to realize the full potential of the yeast in any particular fermentation run. Transgenic oleaginous yeast and non-transgenic oleaginous yeast should be grown and sampled under similar conditions when oil content is to be compared.

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

However, certain end uses and/or product forms may require partial and/or complete isolation/purification of the microbial oil from the biomass, to result in partially purified biomass, purified oil, and/or purified lipid fractions thereof. For example, PUFAs may be found in the host microorganism as free fatty acids or in esterified forms such as acylglycerols, phospholipids, sulfolipids or glycolipids, and may be extracted from the host cell through a variety of means well-known in the art. One review of extraction techniques, quality analysis and acceptability standards for yeast lipids is that of Z. Jacobs (Critical Reviews in Biotechnology 12(5/6):463-491 (1992)). A brief review of downstream processing is also provided by A. Singh and O. Ward (Adv. Appl. Microbiol., 45:271-312 (1997)).
In general, methods for the recovery and purification of microbial lipids and/or PUFAs from microbial biomass may include extraction (e.g., U.S. Patent Nos. 6,797,303 and 5,648,564) with organic solvents, sonication, supercritical fluid extraction (e.g., using carbon dioxide), saponification and physical means such as presses, bead beaters, or combinations thereof. One is referred to the teachings of U.S. Patent 7,238,482 for additional details.

There are a plethora of food and feed products incorporating omega-3 and/or omega-6 fatty acids, particularly e.g., ALA, GLA, ARA, EPA, DPA and DHA. It is contemplated that the microbial biomass comprising long-chain PUFAs, partially purified microbial biomass comprising PUFAs, purified microbial oil comprising PUFAs, and/or purified PUFAs will function in food and feed products to impart the health benefits of current formulations. More specifically, oils containing omega-3 and/or omega-6 fatty acids will be suitable for use in a variety of food and feed products including, but not limited to: food analogs, meat products, cereal products, baked foods, snack foods and dairy products (see U.S. Pat. Appl. Pub. No. 2006-0094092). Feed products also include those for animal uses.

The present compositions may be used in formulations to impart health benefit in medical foods including medical nutritionals, dietary supplements, infant formula and pharmaceuticals. One of skill in the art of food processing and food formulation will understand how the amount and composition of the present oils may be added to the food or feed product. Such an amount will be referred to herein as an “effective” amount and will depend on the food or feed product, the diet that the product is intended to supplement or the medical condition that the medical food or medical nutritional is intended to correct or treat.

The present compositions may be used in formulations to impart animal health benefit in medical foods including medical nutritionals, dietary supplements, and pharmaceuticals.
EXAMPLES

The meaning of abbreviations is as follows: "sec" means second(s), "min" means minute(s), "h" means hour(s), "d" means day(s), "μL" means microliter(s), "mL" means milliliter(s), "L" means liter(s), "μM" means micromolar, "mM" means millimolar, "M" means molar, "mmol" means millimole(s), "μmole" mean micromole(s), "g" means gram(s), "μg" means microgram(s), "ng" means nanogram(s), "U" means unit(s), "bp" means base pair(s) and "kB" means kilobase(s).

**Nomenclature For Expression Cassettes**

The structure of an expression cassette is represented by a simple notation system of "X::Y::Z", wherein X describes the promoter fragment, Y describes the gene fragment, and Z describes the terminator fragment, which are all operably linked to one another.

**Transformation And Cultivation Of Yarrowia lipolytica**

*Y. lipolytica* strain ATCC #20362 was purchased from the American Type Culture Collection (Rockville, MD). *Y. lipolytica* strains were routinely grown at 28-30 °C in several media, according to the recipes shown below.

**Synthetic Complete Media ["SC"] Media (per liter):** 6.7 g Yeast Nitrogen base with ammonium sulfate and without amino acids; 20 g glucose; 1.9 g/L Yeast synthetic drop-out medium supplement without uracil

**High Glucose Media ["HGM"] (per liter):** 80 glucose, 2.58 g KH₂PO₄ and 5.36 g K₂HPO₄, pH 7.5 (do not need to adjust).

**Synthetic Dextrose Media ["SD"] (per liter):** 6.7 g Yeast Nitrogen base with ammonium sulfate and without amino acids; 20 g glucose.

**Fermentation Medium ["FM"] (per liter):** 6.7 g/L YNB without amino acids; 6 g/L KH₂PO₄; 2 g/L K₂HPO₄; 1.5 g/L MgSO₄-heptahydrate; 5 g/L yeast extract; 2% carbon source (wherein the carbon source is either glucose or sucrose).
Transformation of Y. lipolytica was performed as described in U.S. Pat. Appl. Pub. No. 2009-0093543-A1, hereby incorporated herein by reference.

**Fatty Acid Analysis Of Yarrowia lipolytica**

For fatty acid ["FA"] analysis, cells were collected by centrifugation and lipids were extracted as described in Bligh, E. G. & Dyer, W. J. (Can. J. Biochem. Physiol., 37:911-917 (1959)). Fatty acid methyl esters ["FAMEs"] were prepared by transesterification of the lipid extract with sodium methoxide (Roughan, G., and Nishida I., Arch Biochem Biophys., 276(1):38-46 (1990)) and subsequently analyzed with a Hewlett-Packard 6890 GC fitted with a 30-m X 0.25 mm (i.d.) HP-INNOWAX (Hewlett-Packard) column. The oven temperature was from 170 °C (25 min hold) to 185 °C at 3.5 °C/min.

For direct base transesterification, Yarrowia cells (0.5 mL culture) were harvested, washed once in distilled water, and dried under vacuum in a Speed-Vac for 5-10 min. Sodium methoxide (100 µl of 1%) and a known amount of C15:0 triacylglycerol (C15:0 TAG; Cat. No. T-145, Nu-Check Prep, Elysian, MN) was added to the sample, and then the sample was vortexed and rocked for 30 min at 50 °C. After adding 3 drops of 1 M NaCl and 400 µl hexane, the sample was vortexed and spun. The upper layer was removed and analyzed by GC.

Alternately, a modification of the base-catalysed transesterification method described in *Lipid Analysis*, William W. Christie, 2003 was used for routine analysis of the broth samples from either fermentation or flask samples. Specifically, broth samples were rapidly thawed in room temperature water, then weighed (to 0.1 mg) into a tarred 2 mL microcentrifuge tube with a 0.22 µm Corning® Costar® Spin-X® centrifuge tube filter (Cat. No. 8161). Sample (75 - 800 µl) was used, depending on the previously determined DCW. Using an Eppendorf 5430 centrifuge, samples are centrifuged for 5-7 min at 14,000 rpm or as long as necessary to remove the broth. The filter was removed, liquid was drained, and ~500 µl of deionized water was added to the filter to wash the sample. After
centrifugation to remove the water, the filter was again removed, the liquid drained and the filter re-inserted. The tube was then re-inserted into the centrifuge, this time with the top open, for ~3-5 min to dry. The filter was then cut approximately 1/2 way up the tube and inserted into a fresh 2 mL round bottom Eppendorf tube (Cat. No. 22 36 335-2).

The filter was pressed to the bottom of the tube with an appropriate tool that only touches the rim of the cut filter container and not the sample or filter material. A known amount of C15:0 TAG (above) in toluene was added and 500 µl of freshly made 1% sodium methoxide in methanol solution. The sample pellet was firmly broken up with the appropriate tool and the tubes were closed and placed in a 50 °C heat block (VWR Cat. No. 12621-088) for 30 min. The tubes were then allowed to cool for at least 5 min. Then, 400 µl of hexane and 500 µl of a 1 M NaCl in water solution were added, the tubes were vortexed for 2x 6 sec and centrifuged for 1 min. Approximately 150 µl of the top (organic) layer was placed into a GC vial with an insert and analyzed by GC.

FAME peaks recorded via GC analysis were identified by their retention times, when compared to that of known fatty acids, and quantitated by comparing the FAME peak areas with that of the internal standard (C15:0 TAG) of known amount. Thus, the approximate amount (µg) of any fatty acid FAME ['µg FAME'] is calculated according to the formula: (area of the FAME peak for the specified fatty acid/ area of the standard FAME peak) * (µg of the standard C15:0 TAG), while the amount (µg) of any fatty acid ['µg FA'] is calculated according to the formula: (area of the FAME peak for the specified fatty acid/area of the standard FAME peak) * (µg of the standard C15:0 TAG) * 0.9503, since 1 µg of C15:0 TAG is equal to 0.9503 µg fatty acids. Note that the 0.9503 conversion factor is an approximation of the value determined for most fatty acids, which range between 0.95 and 0.96.

The lipid profile, summarizing the amount of each individual fatty acid as a wt % of TFAs, was determined by dividing the individual FAME peak area by the sum of all FAME peak areas and multiplying by 100.
Analysis Of Total Lipid Content And Composition In *Yarrowia lipolytica* By Flask Assay

Flask assays were conducted as follows to analyze the total lipid content and composition in a particular strain of *Y. lipolytica*. Specifically, one loop of freshly streaked cells was inoculated into 3 mL FM medium and grown overnight at 250 rpm and 30 °C. The OD$_{600nm}$ was measured and an aliquot of the cells were added to a final OD$_{600nm}$ of 0.3 in 25 mL FM medium in a 125 mL flask. After 2 days in a shaking incubator at 250 rpm and at 30 °C, 6 mL of the culture was harvested by centrifugation and resuspended in 25 mL HGM in a 125 mL flask. After 5 days in a shaking incubator at 250 rpm and at 30 °C, a 1 mL aliquot was used for fatty acid analysis (above) and 10 mL dried for dry cell weight ["DCW"] determination.

For DCW determination, 10 mL culture was harvested by centrifugation for 5 min at 4000 rpm in a Beckman GH-3.8 rotor in a Beckman GS-6R centrifuge. The pellet was resuspended in 25 mL of water and re-harvested as above. The washed pellet was re-suspended in 20 mL of water and transferred to a pre-weighed aluminum pan. The cell suspension was dried overnight in a vacuum oven at 80 °C. The weight of the cells was determined.

Total lipid content of cells ["TFAs % DCW"] is calculated and considered in conjunction with data tabulating the concentration of each fatty acid as a weight percent of TFAs ["% TFAs"] and the EPA content as a percent of the dry cell weight ["EPA % DCW"].

**EXAMPLE 1**

Identification Of A *Yarrowia lipolytica* Gene Having Homology To The *Saccharomyces cerevisiae* YAP1

An ortholog to the S. cerevisiae Yap1 (GenBank Accession No. NM_001182362; SEQ ID NO:1) ["ScYap1"] was identified in *Yarrowia lipolytica* by conducting BLAST searches using ScYap1 as the query sequence against the public *Y. lipolytica* protein database of the "Yeast project Genolevures" (Center for Bioinformatics, LaBRI, Talence Cedex, France) (see also Dujon, B. et al., *Nature*, 430(6995):35-44 (2004)).
The protein sequence having the best homology (with an expectation value of 1.8e-18) to ScYap1 among all Y. lipolytica proteins, YALI0F03388p (GenBank Accession No. XP_504945; SEQ ID NO:4), was given the designation “YiYap”. YALI0F03388p was annotated therein as “weakly similar to uniprot|Q9P5L6 Neurospora crassa NCU03905.1 related to AP-1-like transcription factor”.

An alignment of ScYap1 and the putative YiYap1 is shown in FIG. 2. Both proteins have a basic leucine zipper (bZIP) motif, corresponding to a N-terminal basic region enriched in basic amino acids that is adjacent to a leucine zipper that is characterized by several leucine residues regularly spaced at seven-amino acid intervals. With respect to the figure, arginine and lysine amino acid residues in bold font and underlined correspond to the basic region; a star highlights each of the leucine residues within the leucine zipper. Vertical boxes highlight cysteine residues within the N-terminal Cys-rich domain of ScYap1 (i.e., corresponding to Cys303, Cys310 and Cys315 of SEQ ID NO:2) and the C-terminal Cys-rich domain (i.e., corresponding to Cys598, Cys620 and Cys629 of SEQ ID NO:2). Five of these residues are conserved in YiYap1. As discussed in Toone and Jones (Curr. Opin. Genet. Dev., 9: 55-61 (1999)), the bZIP domain and the cysteine rich domains are characteristics of AP-1 family proteins.

Using the protein sequence encoding YALI0F03388p (SEQ ID NO:4), National Center for Biotechnology Information ["NCBI"] BLASTP 2.2.26+ (Basic Local Alignment Search Tool; Altschul, S. F., et al., Nucleic Acids Res., 25:3389-3402 (1997); Altschul, S. F., et al., FEBS J., 272:5101-5109 (2005)) searches were conducted to identify sequences having similarity within the BLAST “nr” database (comprising all non-redundant GenBank CDS translations, RefSeq protein sequences from NCBI’s Reference Sequence Project, the Brookhaven Protein Data Bank ["PDB"] protein sequence database, the SWISS-PROT protein sequence database, the Protein Information Resource ["PIR"] protein sequence database and the Protein Research Foundation ["PRF"] protein sequence database).
The results of the BLASTP comparison summarizing the sequence to which SEQ ID NO:4 has the most similarity may be reported according to the % identity, % similarity and Expectation value. “% Identity” is defined as the percentage of amino acids that are identical between the two proteins. “% Similarity” is defined as the percentage of amino acids that are identical or conserved between the two proteins. “Expectation value” estimates the statistical significance of the match, specifying the number of matches, with a given score, that are expected in a search of a database of this size absolutely by chance.

A large number of proteins were identified as sharing similarity to YALI0F03388p (SEQ ID NO:4). Table 4 provides a partial summary of those hits having an Expectation value greater or equal to “2e-13” and annotation that specifically identified the protein (i.e., while hits to hypothetical proteins are excluded), although this should not be considered as limiting to the disclosure herein. The proteins in Table 4 shared between 13-87% query coverage with SEQ ID NO:4.

<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
<th>Query coverage</th>
<th>E value</th>
</tr>
</thead>
<tbody>
<tr>
<td>XP_002847259.1</td>
<td>Chap1 [Arthroderma otae CBS 113480]</td>
<td>86%</td>
<td>2e-39</td>
</tr>
<tr>
<td>EGS19655.1</td>
<td>putative Ap-1-like transcription factor [Chaetomium thermophilum var. thermophilum DSM 1495]</td>
<td>87%</td>
<td>1e-37</td>
</tr>
<tr>
<td>AAS64313.1</td>
<td>Chap1 [Cochliobolus heterostrophus]</td>
<td>28%</td>
<td>2e-27</td>
</tr>
<tr>
<td>XP_002145733.1</td>
<td>bZIP transcription factor AP-1/Yap1, putative [Penicillium marneffei ATCC 18224]</td>
<td>39%</td>
<td>4e-27</td>
</tr>
<tr>
<td>EGR47222.1</td>
<td>transcription factor [Trichoderma reesei QM6a]</td>
<td>30%</td>
<td>5e-27</td>
</tr>
<tr>
<td>EFQ30244.1</td>
<td>transcription factor PAP1 [Glomerella graminicola M1.001]</td>
<td>50%</td>
<td>6e-27</td>
</tr>
<tr>
<td>XP_001394766.2</td>
<td>bZIP transcription factor (AP-1) [Aspergillus niger CBS 513.88]</td>
<td>26%</td>
<td>1e-26</td>
</tr>
<tr>
<td>XP_001258114.1</td>
<td>bZIP transcription factor (AP-1), putative [Neosartorya fischeri NRRL 181]</td>
<td>38%</td>
<td>1e-26</td>
</tr>
<tr>
<td>EEH47458.1</td>
<td>bZIP transcription factor (AP-1) [Paracoccidioides brasiliensis Pb18]</td>
<td>38%</td>
<td>1e-26</td>
</tr>
<tr>
<td>XP_002477983.1</td>
<td>bZIP transcription factor AP-1/Yap1, putative</td>
<td>41%</td>
<td>2e-26</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Identity</td>
<td>Coverage</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>EGE77501.1</td>
<td>BZIP transcription factor AP-1/Yap1 [Ajiellozymes dermatitidis ATCC 18188]</td>
<td>41%</td>
<td>3e-26</td>
</tr>
<tr>
<td>XP_002627437.1</td>
<td>bZIP transcription factor AP-1/Yap1 [Ajiellozymes dermatitidis]</td>
<td>41%</td>
<td>3e-26</td>
</tr>
<tr>
<td>EFY90531.1</td>
<td>AP-1-like protein [Matarhizium acridum CQMa 102]</td>
<td>29%</td>
<td>8e-26</td>
</tr>
<tr>
<td>EFW21644.1</td>
<td>bZIP transcription factor [Coccidioides posadasii str. Silveira]</td>
<td>37%</td>
<td>1e-25</td>
</tr>
<tr>
<td>XP_003065816.1</td>
<td>bZIP family transcription factor [Coccidioides posadasii C735 delta SOWgp]</td>
<td>37%</td>
<td>2e-25</td>
</tr>
<tr>
<td>CAX15423.1</td>
<td>Ap1-like transcription factor [Botryotinia fuckelliana]</td>
<td>22%</td>
<td>2e-25</td>
</tr>
<tr>
<td>XP_001268032.1</td>
<td>bZIP transcription factor (AP-1), putative [Aspergillus clavatus NRRL 1]</td>
<td>26%</td>
<td>3e-25</td>
</tr>
<tr>
<td>EER41522.1</td>
<td>bZIP transcription factor [Ajiellozymes capsulatus H143]</td>
<td>40%</td>
<td>4e-25</td>
</tr>
<tr>
<td>EGC49388.1</td>
<td>bZIP transcription factor [Ajiellozymes capsulatus H88]</td>
<td>40%</td>
<td>4e-25</td>
</tr>
<tr>
<td>XP_002145732.1</td>
<td>bZIP transcription factor AP-1/Yap1, putative [Penicillium marneffei ATCC 18224]</td>
<td>41%</td>
<td>7e-25</td>
</tr>
<tr>
<td>EGX87755.1</td>
<td>bZIP transcription factor (AP-1), putative [Cordyceps militaris CM01]</td>
<td>22%</td>
<td>1e-24</td>
</tr>
<tr>
<td>CBX91516.1</td>
<td>similar to AP1-like transcription factor [Leptosphaeria maculans JN3]</td>
<td>28%</td>
<td>2e-24</td>
</tr>
<tr>
<td>EFX02671.1</td>
<td>AP-1-like, bzip transcription factor [Grosmannia clavigera kw1407]</td>
<td>43%</td>
<td>3e-24</td>
</tr>
<tr>
<td>XP_003017058.1</td>
<td>bZIP transcription factor AP-1/Yap1, putative [Arthrodema benhamiae CBS 112371]</td>
<td>36%</td>
<td>4e-24</td>
</tr>
<tr>
<td>XP_003021545.1</td>
<td>bZIP transcription factor AP-1/Yap1, putative [Trichophyton verrucosum HKI 0517]</td>
<td>27%</td>
<td>6e-24</td>
</tr>
<tr>
<td>EGE03693.1</td>
<td>bZIP transcription factor AP-1/Yap1 [Trichophyton equinum CBS 127.97]</td>
<td>27%</td>
<td>6e-24</td>
</tr>
<tr>
<td>XP_001931984.1</td>
<td>Chap1 [Pyrenophora tritici-repentis Pt-1C-BFP]</td>
<td>15%</td>
<td>8e-24</td>
</tr>
<tr>
<td>ACM50933.1</td>
<td>AP-1-like protein [Alternaria alternata]</td>
<td>57%</td>
<td>1e-23</td>
</tr>
<tr>
<td>EGO54582.1</td>
<td>PAP1-domain-containing protein [Neurospora tetrasperma FGSC 2509]</td>
<td>28%</td>
<td>1e-23</td>
</tr>
<tr>
<td>EGY17906.1</td>
<td>Chap1 [Verticillium dahliae VdLs.17]</td>
<td>24%</td>
<td>6e-23</td>
</tr>
<tr>
<td>EGP91344.1</td>
<td>bZIP transcription factor [Mycosphaerella graminicola IPO323]</td>
<td>40%</td>
<td>1e-22</td>
</tr>
<tr>
<td>EFZ02600.1</td>
<td>AP-1-like protein [Matarhizium anisopliae ARSEF 23]</td>
<td>38%</td>
<td>5e-22</td>
</tr>
</tbody>
</table>
Based on the BLASTP searches, YALI0F03388p (SEQ ID NO:4) shared the best similarity with hypothetical protein BC1G_14094 from Botryotinia fuckeliana (GenBank Accession No. XP_001547321), with 30% identity and 47% similarity, and an expectation value of 1e-41.

Among proteins with known function, the best hits were to: Chap1 from Arthroderma otae CBS 113480 (GenBank Accession No. XP_002847259.1), having 30% identity and 46% similarity, and an expectation value of 2e-39; the putative Ap-1-like transcription factor from
Chaetomium thermophilum var. thermophilum DSM 1495, having 31% identity and 46% similarity, and an expectation value of 1e-37; and, Chap1 from Cochliobulus heterostrophus (GenBank Accession No. AAS64313), having 48% identity and 64% similarity, and an expectation value of 2e-27. Chap1 is known as a functional homolog of S. cerevisiae Yap1 (S. Lev et al., Eukaryotic Cell, 4(2):443-454 (2005)).

Based on the above analyses, SEQ ID NO:3 was hypothesized to encode the Yap1 transcription factor of Y. lipolytica ("YlYap1"), wherein the protein sequence is set forth as SEQ ID NO:4.

It is not surprising that YlYap1 shares such relatively low percent identity and similarity with other bZIP transcription factors. For example, the Candida glabrata CgAP1p (GenBank Accession No. XP 446996) has been positively characterized as a functional ortholog of Yap1 (Chen, K.-H. et al., Gene, 386(1-2):63-72 (2007)). Despite shared functionality, Chen et al. reports that the Candida glabrata CgAP1p showed only 37% amino acid identity with S. cerevisiae Yap1p (GenBank Accession No. NP 013707), 30% identity with Kluyveromyces lactis KIAP1p (GenBank Accession No. P56095), 26% identity with Candida albicans CAP1p (GenBank Accession No. AAD00802), and 19% identity with Schizosaccharomyces pombe Pap1p (GenBank Accession No. CAB66170); notably, however, the identity between Candida glabrata CgAP1p and S. cerevisiae Yap1p was especially high in the bZip domain (73% identity), the N-terminal cysteine-rich domain (75% identity) and the C-terminal cysteine-rich domain (85% identity).

Thus, despite the sequence analyses described above, further functional analyses were necessary to confirm that YlYap1 functioned in a manner homologous to that of ScYap1.

**EXAMPLE 2**

**Increased Hydrogen Peroxide Sensitivity In Yarrowia lipolytica YAP1 Knockout Strain Y4184U (yap1Δ)**

The present Example describes the use of construct pYRH60 (FIG. 3A; SEQ ID NO:5) to down-regulate expression of chromosomal YAP1 gene from an EPA producing engineered strain of Yarrowia lipolytica,
Y4184U (Example 7, infra). Transformation of Y. lipolytica strain Y4184U with the YAP1 knockout construct fragment resulted in strain Y4184U (yap1Δ). The effect of the Yap1 knockout on oxidative stress sensitivity and on accumulated lipid level and EPA production was determined and compared. Specifically, knockout of YAP1 resulted in hyper-sensitivity against H₂O₂, as compared to cells whose native Yap1 had not been knocked out.

**Generation Of Strain Y4184U (yap1Δ)**

Plasmid pYRH60 was derived from plasmid pYPS161, which was described in U.S. Patent App. No. 2010-0062502 (Example 2, FIG. 5A, SEQ ID NO:40 therein) and contained the following components:

<table>
<thead>
<tr>
<th>RE Sites And Nucleotides Within SEQ ID NO:6</th>
<th>Description Of Fragment And Chimeric Gene Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ascl/BsiWI (1521-157)</td>
<td>1364 bp PEX10 knockout fragment #1 of Yarrowia PEX10 gene (GenBank Accession No. AB036770)</td>
</tr>
<tr>
<td>PacI/Sphi (5519-4229)</td>
<td>1290 bp PEX10 knockout fragment #2 of Yarrowia PEX10 gene (GenBank Accession No. AB036770)</td>
</tr>
<tr>
<td>SalI/EcoRI (7170-5551)</td>
<td>Yarrowia URA3 gene (GenBank Accession No. AJ306421)</td>
</tr>
<tr>
<td>2451-1571</td>
<td>ColE1 plasmid origin of replication</td>
</tr>
<tr>
<td>3369-2509</td>
<td>ampicillin-resistance gene (Amp&lt;sup&gt;R&lt;/sup&gt;) for selection in E. coli</td>
</tr>
<tr>
<td>3977-3577</td>
<td>E. coli f1 origin of replication</td>
</tr>
</tbody>
</table>

A 940 bp 5' promoter region (SEQ ID NO:7) of the Y. lipolytica YAP1 gene (“YIYAP1”; SEQ ID NO:3) replaced the Ascl/BsiWI fragment of pYPS161 (SEQ ID NO:6) and a 1164 bp 3' terminator region (SEQ ID NO:32) of the YIYAP1 gene replaced the PacI/SphiI fragment of pYPS161 to produce pYRH60 (SEQ ID NO:5; FIG. 3A).

*Y. lipolytica* strain Y4184U was transformed with the purified 4.7 kB Ascl/SphiI fragment of YAP1 knockout construct pYRH60 (SEQ ID NO:5) (General Methods).

To screen for cells having the yap1 deletion, quantitative real time PCR on YIYap1 was conducted, with the Yarrowia translation elongation
factor gene TEF1 (GenBank Accession No. AF054510) used as the control. Real time PCR primers and a TaqMan® probe targeting the YAP1 gene and the control TEF1 gene, respectively, were designed with Primer Express software version 2.0 (Applied Biosystems, Foster City, CA). Specifically, real time PCR primers Y1-EF-1214F (SEQ ID NO:8), Y1-EF-1270R (SEQ ID NO:9), YAP1-346F (SEQ ID NO:10) and YAP1-409R (SEQ ID NO:11) were designed, as well as YAP1-366T (i.e., 5’ 6-FAM™-CGGGCTGCCAAGGGCC-TAMRA™, wherein the nucleotide sequence is set forth as SEQ ID NO:13). The TaqMan probe YL-EF-MGB-1235T (i.e., 5’ 6-FAM™-CCTCACTGACTCC-TAMRA™, wherein the nucleotide sequence is set forth as SEQ ID NO:12) was obtained from Applied Biosystems. The 5’ end of the TaqMan fluorogenic probes have the 6-FAM™ fluorescent reporter dye bound, while the 3’ end comprises the TAMRA™ quencher. PCR primers and the YAP1 probe were obtained from Sigma-Genosys (Woodlands, TX).

Knockout candidate DNA was prepared by suspending 1 colony in 50 μl of water. Reactions for TEF1 and YAP1 were run in the same Real Time PCR well, in triplicate, for each sample. Real time PCR reactions included 10 pmoles each of forward and reverse primers (i.e., Y1-EF-1214F, Y1-EF-1270R, YAP1-346F and YAP1-409R, supra), and 2.5 pmoles TaqMan® probe (i.e., YL-EF-MGB-1235T and YAP1-366T, supra), 10 μl TaqMan® Universal PCR Master Mix--No AmpErase® Uracil-N-Glycosylase (UNG) (Catalog No. PN 4326614, Applied Biosystems), 1 μl colony suspension and 8.5 μl RNase/DNase free water for a total volume of 20 μl per reaction. Reactions were run on the ABI PRISM® 7900 Sequence Detection System under the following conditions: initial denaturation at 95 °C for 10 min, followed by 40 cycles of denaturation at 95 °C for 15 sec and annealing at 60 °C for 1 min.

Real time data was collected automatically during each cycle by monitoring 6-FAM™ fluorescence. Data analysis was performed using TEF1 gene threshold cycle (C_T) values for data normalization as per ABI PRISM® 7900 Sequence Detection System instruction manual (see ABI User Bulletin #2 "Relative Quantitation of Gene Expression"). Knockout
clones were identified as having no detectable signal for the YAP1 gene and a C_T value for TEF1 ≤ 30.

The methodology set forth above identified one of the colonies screened as a yap1 knockout. The Y. lipolytica yap1Δ mutant of Y4184U was designated RHY240.

H_2O_2 Sensitivity Assays With Knockout Strain Y4184U (yap1Δ)

In S. cerevisiae, strains lacking Yap1 are hypersensitive to killing by H_2O_2. This phenotype is related to Yap1’s role in controlling the induction of oxidative stress defense genes, such as TRR1 (cytoplasmic thioredoxin reductase), TRX2 (thioredoxin), GLR1 (glutathione reductase), and GSH1 (γ-glutamylcysteine synthetase). To test the function of the putative YILYap1 as an oxidative stress regulator, Y4184U (yap1Δ) was subjected to a H_2O_2 sensitivity assay.

Y4184U (yap1Δ) and Y4184 (control) cells were grown to an exponential phase (OD_600 of ~0.5) in SC medium and diluted to an OD_600 of 0.01 with fresh SC medium. Aliquots (100 μl) of the diluted cultures were incubated with fresh H_2O_2 at final concentrations from 0 to 50 mM at 30 °C for 1 hr, and 7 μl from each sample was spotted onto YPD plates. Cells were further grown at 30 °C for 2 days on the YPD plate.

Y4184U (yap1Δ) cells showed much higher sensitivity to H_2O_2 stress than the control strain Y4184 (FIG. 4A). This result supports the hypothesis that YILYap1, corresponding to YALI0F0338p, was important for oxidative stress defense in Y. lipolytica and was a functional homolog of ScYap1.

EXAMPLE 3
Overexpression Of Yarrowia lipolytica YAP1 In Saccharomyces cerevisiae

YAP1 Knockout Strain BY4743 (yap1Δ)

The present Example describes the use of centromeric plasmid pYRH61 (FIG. 3B; SEQ ID NO:14) to overexpress the putative YILYap1 (SEQ ID NO:4) in a S. cerevisiae yap1Δ strain, to evaluate the effect on oxidative stress sensitivity. Specifically, overexpression of YILYAP1 resulted in functional complementation of hyper-sensitivity against H_2O_2 in the S. cerevisiae yap1Δ strain.
Construction Of *S. cerevisiae* Overexpression Plasmid pYRH61

Plasmid pYRH61 was derived from plasmid pRS316 (Sikorski and Hieter, *Genetics*, 122:19-27 (1989)), a centromeric plasmid with *URA3* as a selective marker. The pYRH61 contained the following components:

<table>
<thead>
<tr>
<th>RE Sites And Nucleotides Within SEQ ID NO:14</th>
<th>Description Of Fragment And Chimeric Gene Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>*SalI/*SpeI (7442-8042)</td>
<td>601 bp <em>FBA1</em> promoter region of <em>S. cerevisiae</em> (GenBank Accession No. X15003)</td>
</tr>
<tr>
<td>*NotI/*SacI (1613-2634)</td>
<td>1022 bp <em>FBA1</em> terminator region of <em>S. cerevisiae</em> (GenBank Accession No. X15003)</td>
</tr>
<tr>
<td>*SpeI/*NotI (1-1612)</td>
<td><em>YIIAP1</em> (YALI0F03388g: SEQ ID NO:3) (GenBank Accession No. XP504945)</td>
</tr>
<tr>
<td>*SalI/*SacI (2635-7436)</td>
<td>pRS316 vector backbone</td>
</tr>
</tbody>
</table>

Specifically, a 1.6 kb fragment of the *YIIAP1* gene was amplified by PCR from the *Y. lipolytica* genome using primers *Yl.Yap1-F-Spel* (SEQ ID NO:15) and *Yap1-R* (SEQ ID NO:16). The reaction mixture contained 1 μl of the genomic DNA, 1 μl each of the primers (from 20 μM stocks), 2 μl water, and 45 μl AccuPrime Pfx SuperMix from Invitrogen. Amplification was carried out as follows: initial denaturation at 95 °C for 5 min, followed by 35 cycles of denaturation at 95 °C for 60 sec, annealing at 55 °C for 60 sec, and elongation at 68 °C for 180 sec. A final elongation cycle at 72 °C for 7 min was carried out, followed by reaction termination at 4 °C. A 1.6 kb DNA fragment was obtained from the PCR reaction.

The amplified gene was digested with *SpeI/*NotI and cloned with a 601 bp 5' promoter region (SEQ ID NO:17) of the *S. cerevisiae FBA1* gene ["ScFBA1"] and a 1022 bp 3' terminator region of ScFBA1 (SEQ ID NO:18) into pRS316 (SEQ ID NO:19) to produce pYRH61 (SEQ ID NO:14; FIG. 3B). Thus, pYRH61 contained a chimeric ScFBA1::YIIAP1::ScFBA1 gene.

**H$_2$O$_2$ Sensitivity Assays With *S. cerevisiae* Strains Expressing pYRH61**
S. cerevisiae strains BY4743 (MATa/α his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0) and its isogenic yap1Δ strain BY4743 (yap1Δ) (obtained from Invitrogen, Carlsbad, CA) were transformed with pRS316 (vector control) or pYRH61 to evaluate the effect of YIYap1 overexpression on oxidative stress sensitivity.

Cells were grown to an exponential phase (OD_600 of ~0.5) in SC medium lacking uracil and diluted to an OD_600 of 0.01 with fresh SC medium. Aliquots (100 μl) of the diluted cultures were incubated with fresh H_2O_2 at the final concentrations from 0 to 50 mM at 30 °C for 1 hr, and 7 μl from each sample was spotted onto YPD plates. Spotted cells were further grown at 30 °C for 2 days on the YPD plate.

FIG. 4B shows the results of the H_2O_2 sensitivity assay. Specifically, the top two rows are BY4743 transformants (i.e., with either the control vector or pYRH61, respectively), while the bottom two rows are BY4743 (yap1Δ) transformants (i.e., with either the control vector or pYRH61, respectively).

As shown in FIG. 4B, the BY4743 yap1Δ strain transformed with control plasmid pRS316 showed higher sensitivity to H_2O_2 stress than its isogenic BY4743 wild type strain with either the control or pYRH61. When YIYap1 was over-expressed in the BY4743 yap1Δ strain, cells become much more resistant to the oxidative stress than BY4743 yap1Δ transformants with the control plasmid, suggesting the YIYap1 (SEQ ID NO:4) conferred the resistance against oxidative stress.

The results herein support the hypothesis that YIYap1, corresponding to YAL10F03388p, was a functional homolog of ScYap1 and was associated with oxidative stress defense.

EXAMPLE 4
Overexpression Of Yarrowia lipolytica YAP1 In Y. lipolytica

Strains Y4184 And Y9502

The present Example describes synthesis of overexpression construct pYRH43 (FIG. 5A; SEQ ID NO:20) and its transformation into Y. lipolytica strains Y4184U (Example 7) and Y9502U (Example 8). The
effect of YIYAP1 overexpression on accumulated lipid level was
determined and compared. Specifically, YIYAP1 overexpression resulted
in increased total lipid (measured as total fatty acids as a percent of the
total dry cell weight ["TFAs % DCW"]) as compared to cells whose native
Yap1 level had not been manipulated.

**Construction Of Y. lipolytica Overexpression Plasmid pYRH43**

Plasmid pYRH43 was derived from plasmid pZuFmEaD5s
(described in Example 6 of U.S. Pat. 7,943,365, hereby incorporated
herein by reference). Plasmid pZuFmEaD5s contained a chimeric
FBAINm::EaD5S::PEX20 gene, wherein: (i) FBAINm is a Y. lipolytica
promoter upstream of the fba1 gene encoding a fructose-bisphosphate
aldolase enzyme (E.C. 4.1.2.13) (U.S. Pat. 7,202,356); (ii) EaD5S is a
synthetic delta-5 desaturase derived from *Euglena anabaena* and codon-
optimized for expression in *Yarrowia*, flanked by *Ncol/NotI* restriction
enzyme sites; and, (iii) PEX20 is a *PEX20* terminator sequence from the
*Yarrowia PEX20* gene (GenBank Accession No. AF054613).

A 1.6 kb fragment of the YIYAP1 gene was amplified by PCR from
the Y. lipolytica genome using primers Yap1-F (SEQ ID NO:21) and Yap1-
R (SEQ ID NO:16). The reaction mixture contained 1 μl of the genomic
DNA, 1 μl each of the primers (from 20 μM stocks), 2 μl water, and 45 μl
AccuPrime Pfx Supermix from Invitrogen. Amplification was carried out as
follows: initial denaturation at 95 °C for 5 min, followed by 35 cycles of
denaturation at 95 °C for 1 sec, annealing at 55 °C for 1 sec, and
elongation at 68 °C for 3 sec. A final elongation cycle at 72 °C for 7 min
was carried out, followed by reaction termination at 4 °C. A 1.6 kb DNA
fragment was obtained from the PCR reaction.

The amplified gene was digested with *PciI/NotI* and used to replace
the *Ncol/NotI* fragment of pZuFmEaD5s to produce pYRH43. Thus,
pYRH43 contained a chimeric FBAINm::YIYAP1::PEX20 gene.

**Identification Of Transformant Strains Y4184U+YAP1 And Y9502U+Yap1**

By Quantitative Real Time PCR

Plasmid pYRH43 was cut with *BsiWI/PacI* and a 4.4 kb fragment
was isolated and used for transformation (General Methods) into Y.
lipolytica strains Y4184U (Example 7) and Y9502U (Example 8), thereby producing strains Y4184U+YAP1 and Y9502U+Yap1.

Overexpression of YIYAP1 was confirmed by performing quantitative real time RT-PCR, using the Yarrowia TEF1 gene as the control in a manner similar to that described in Example 2.

Primers were qualified for real time quantitation using a dilution series of genomic DNA and the PCR conditions detailed below. Linear regression analysis was performed using the obtained C_T values versus log ng DNA for each primer and probe set and the efficiencies were confirmed to be within 90-110%.

cDNA from strains Y4184U+YAP1 and Y9502U+Yap1 was prepared by first isolating RNA using a Qiagen RNeasy™ kit (Valencia, CA). Residual genomic DNA was then eliminated by treating 2 µg of RNA with DNase (Catalog No. PN79254, Qiagen) for 15 min at room temperature, followed by inactivation for 5 min at 75 °C. The cDNA was generated from 1 µg of treated RNA using the High Capacity cDNA Reverse Transcription Kit from Applied Biosystems (Catalog No. PN 4368813), according to the manufacturer’s recommended protocol.

Real time PCR reactions for YITEF1 and YIYAP1 were run separately in triplicate for each sample. Real time PCR reactions included 0.2 µl each of forward and reverse primers (100 µM) (i.e., ef-324F [SEQ ID NO:22], ef-392R [SEQ ID NO:23], YAP1-346F [SEQ ID NO:10] and YAP1-409R [SEQ ID NO:11]), 0.05 µl of each TaqMan® probe (100 µM) (i.e., ef-345T [i.e., 5’ 6-FAM™-TGCTGGTGTTGGTGAGTT-TAMRA™, wherein the nucleotide sequence is set forth as SEQ ID NO:24] and YAP1-366T [i.e., 5’ 6-FAM™-CGGGCTGCCAAAGGGC-TAMRA™, wherein the nucleotide sequence is set forth as SEQ ID NO:13]), 10 µl TaqMan® Universal PCR Master Mix--No AmpErase® Uracil-N-Glycosylase (UNG) (Catalog No. PN 4326614, Applied Biosystems), 1 µl diluted cDNA (1:10), and 8.55 µl RNase/DNase free water for a total volume of 20 µl per reaction. Reactions were run on the ABI PRISM 7900 Sequence Detection System under the following conditions: initial denaturation at 95 °C for 10 min, followed by 40 cycles of denaturation at
95 °C for 15 sec and annealing at 60 °C for 1 min. A negative reverse transcription RNA control of each sample was run with the TEF1 primer set to confirm the absence of genomic DNA. Real time data was collected as described in Example 2.

Based on this analysis, it was concluded that the Y4184U+Yap1 strain showed approximately 2.9-fold higher expression level of the YIYAP1 gene, as compared to that of the Y4184U (Ura+) control strain, thereby confirming functionality of plasmid pYRH43.

**Lipid Content And Composition In Transformant Strain Y4184U+YAP1**

*Y. lipolytica* strain Y4184U (Ura+) (control) and strain Y4184U+Yap1 were grown under comparable oleaginous conditions. More specifically, oleaginous conditions were achieved by first growing the cultures aerobically in 25 mL of SD medium (starting OD$_{600}$ of ~0.3) at 30 °C for 48 h, and then harvesting the cells by centrifugation. The pellets were then resuspended in 25 mL of HGM and further incubated for 5 days in a shaker incubator at 250 rpm and 30 °C.

The DCW, total lipid content of cells ["TFAs % DCW"], the concentration of each fatty acid as a weight percent of TFAs ["% TFAs"] and the EPA productivity (i.e., EPA content as its percent of the dry cell weight ["EPA % DCW"])) for *Y. lipolytica* Y4184U (Ura+) control and Y4184U+Yap1 strains are shown below in Table 7, while averages are highlighted in gray and indicated as “Ave”. Abbreviations for fatty acids are as follows: stearic acid (18:0), oleic acid (18:1), linoleic acid (18:2), and eicosapentaenoic acid ("EPA", 20:5).

<table>
<thead>
<tr>
<th>Strains</th>
<th>DCW (g/L)</th>
<th>TFAs % DCW</th>
<th>% TFAs</th>
<th>EPA % DCW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>18:0</td>
<td>18:1</td>
</tr>
<tr>
<td>Y4184U (Ura+)</td>
<td>5.48</td>
<td>14.1</td>
<td>2.0</td>
<td>8.1</td>
</tr>
<tr>
<td></td>
<td>4.38</td>
<td>15.5</td>
<td>1.7</td>
<td>7.8</td>
</tr>
<tr>
<td><strong>AVE</strong></td>
<td><strong>4.93</strong></td>
<td><strong>14.8</strong></td>
<td><strong>1.9</strong></td>
<td><strong>8.0</strong></td>
</tr>
<tr>
<td>Y4184U +Yap1</td>
<td>3.68</td>
<td>16.5</td>
<td>2.1</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>4.62</td>
<td>14.7</td>
<td>2.0</td>
<td>8.6</td>
</tr>
</tbody>
</table>
Overexpression of YIYAP1 (SEQ ID NO:4), corresponding to locus YALI0F03388p, in Y4184U increased lipid content ["TFAs % DCW"] by approximately 12% and increased average EPA titer ["EPA % DCW"] approximately 15%, as compared to that of strain Y4184U (Ura+).

Lipid Content And Composition In Transformant Strain Y9502U+YAP1

*Y. lipolytica* strain Y9502 (control) and strain Y9502U+Yap1 (three isolates) were grown in duplicate under comparable oleaginous conditions, *supra*. Table 8 summarizes the DCW, TFAs % DCW, the concentration of each fatty acid as % TFAs, and EPA % DCW, in a format similar to that used in Table 7.

**Table 8. Lipid Content And Composition In *Y. lipolytica* Strains Y9502 And Y9502U+Yap1**

<table>
<thead>
<tr>
<th>Strains</th>
<th>DCW (g/L)</th>
<th>TFAs % DCW</th>
<th>% TFAs</th>
<th>EPA % DCW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>18:0</td>
<td>18:1</td>
</tr>
<tr>
<td>Y9502</td>
<td>3.7</td>
<td>31.3</td>
<td>2.3</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>3.6</td>
<td>30.7</td>
<td>2.3</td>
<td>4.7</td>
</tr>
<tr>
<td><strong>AVE</strong></td>
<td><strong>3.6</strong></td>
<td><strong>31.0</strong></td>
<td><strong>2.3</strong></td>
<td><strong>4.7</strong></td>
</tr>
<tr>
<td>Y9502U+Yap1</td>
<td>4.2</td>
<td>37.0</td>
<td>2.2</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td>4.2</td>
<td>35.8</td>
<td>2.2</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td>4.1</td>
<td>35.5</td>
<td>2.2</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td>4.2</td>
<td>35.7</td>
<td>2.2</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td>4.2</td>
<td>36.1</td>
<td>2.2</td>
<td>5.3</td>
</tr>
<tr>
<td><strong>AVE</strong></td>
<td><strong>4.2</strong></td>
<td><strong>36.0</strong></td>
<td><strong>2.2</strong></td>
<td><strong>5.3</strong></td>
</tr>
</tbody>
</table>

Overexpression of YIYAP1 (SEQ ID NO:4), corresponding to locus YALI0F03388p, in Y9502U increased lipid content ["TFAs % DCW"] by
approximately 16% and increased average EPA titer ["EPA % DCW"]
approximately 15%, as compared to that of Y9502.

Thus, it appears that overexpression of Y1YAP1 in a PUFA-
producing strain of Yarrowia lipolytica provided increased resistance to
oxidative stresses. One beneficial outcome of this increased resistance to
oxidative stresses is increased protection against lipid peroxidation, which
thereby resulted in increased lipid and PUFA content.

EXAMPLE 5

Identification Of A Yarrowia lipolytica Gene Having Homology To The
Saccharomyces cerevisiae GPX3

An ortholog to the S. cerevisiae Gpx3 (GenBank® Accession No.
NM_001179559; SEQ ID NO:26) ["ScGpx3"] was identified in Yarrowia
lipolytica by conducting BLAST searches using ScGpx3 as the query
sequence against the public Y. lipolytica protein database of the “Yeast
project Genolevures” (Center for Bioinformatics, LaBRI, Talence Cedex,
France) (see also Dujon, B. et al., Nature, 430 (6995):35-44 (2004)).

The protein sequence having the best homology (with an
expectation value of 4e-68) to ScGpx3 among all Y. lipolytica proteins,
YALI0E02310p (GenBank Accession No. XP_503454; SEQ ID NO:28),
was given the designation “YIGpx3”. YALI0E02310p was annotated
therein as “highly similar to uniprotP40581 Saccharomyces cerevisiae
YIR037w HYR1 (ohnolog of YKL026C) Thiol peroxidase that functions as
a hydroperoxide receptor to sense intracellular hydroperoxide levels and
transduce a redox signal to the Yap1p transcription factor”.

An alignment of ScGpx3 and the putative YIGpx3 is shown in FIG.
6. Vertical boxes highlight Cys36 and Cys82 of ScGpx3, important for
inter- and intra-molecular interactions (Delaunay, A., et al., Cell, 111:471-
481 (2002)). These residues are conserved in YIGpx3.

Using the protein sequence encoding YALI0E02310p (SEQ ID
NO:28), NCBI BLASTP 2.2.26+ searches were conducted to identify
sequences having similarity within the BLAST "nr" database, according to
the methodology set forth in Example 1.
A large number of proteins were identified as sharing significant similarity to YALI0E02310p (SEQ ID NO:28). Table 9 provides a partial summary of those hits having an Expectation value greater or equal to “8e-72” and annotation that specifically identified the protein (i.e., while hits to hypothetical proteins and proteins from *Saccharomyces cerevisiae* are excluded), although this should not be considered as limiting to the disclosure herein. The proteins in Table 9 shared between 93-95% query coverage with SEQ ID NO:28.

<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
<th>Query coverage</th>
<th>E value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP_985509.1</td>
<td>AFL039Cp [Ashbya gossypii ATCC 10895]</td>
<td>95%</td>
<td>1e-85</td>
</tr>
<tr>
<td>XP_002491803.1</td>
<td>Thiol peroxidase that functions as a hydroperoxide receptor [Komagataella pastoris GS115]</td>
<td>94%</td>
<td>8e-85</td>
</tr>
<tr>
<td>XP_002548683.1</td>
<td>peroxiredoxin HYR1 [Candida tropicalis MYA-3404]</td>
<td>94%</td>
<td>2e-79</td>
</tr>
<tr>
<td>XP_002548650.1</td>
<td>peroxiredoxin HYR1 [Candida tropicalis MYA-3404]</td>
<td>94%</td>
<td>1e-77</td>
</tr>
<tr>
<td>EGV62163.1</td>
<td>glutathione peroxidase [Candida tenuis ATCC 10573]</td>
<td>94%</td>
<td>3e-76</td>
</tr>
<tr>
<td>XP_714295.1</td>
<td>potential glutathione peroxidase/redox transducer [Candida albicans SC5314]</td>
<td>94%</td>
<td>3e-76</td>
</tr>
<tr>
<td>XP_002420878.1</td>
<td>hydrogen peroxide resistance protein, putative; peroxiredoxin, putative; thiol peroxidase, putative [Candida dubliniensis CD36]</td>
<td>94%</td>
<td>8e-76</td>
</tr>
<tr>
<td>NP_596146.1</td>
<td>glutathione peroxidase Gpx1 [Schizosaccharomyces pombe 972h-]</td>
<td>93%</td>
<td>3e-75</td>
</tr>
<tr>
<td>EFW96327.1</td>
<td>Glutathione-Dependent Phospholipid Peroxidase Hyr1 [Ogataea parapomorpha DL-1]</td>
<td>95%</td>
<td>2e-73</td>
</tr>
<tr>
<td>XP_002172470.1</td>
<td>glutathione peroxidase Gpx1 [Schizosaccharomyces japonicus yFS275]</td>
<td>93%</td>
<td>2e-73</td>
</tr>
<tr>
<td>XP_001384693.1</td>
<td>glutathione peroxidase [Schefteromyces stipitis CBS 6054]</td>
<td>94%</td>
<td>7e-72</td>
</tr>
<tr>
<td>XP_001698575.1</td>
<td>glutathione peroxidase [Chlamydomonas reinhardtii]</td>
<td>94%</td>
<td>8e-72</td>
</tr>
</tbody>
</table>
Based on the BLASTP searches, YAL10E02310p (SEQ ID NO:28) shared the best similarity with a hypothetical protein from *Ashbya gossypii* (GenBank® Accession No. NP_985509), with 73% identity and 86% similarity, and an expectation value of 1e-85. Among proteins with known function, the best hit was the thiol peroxidase from *Pichia pastoris* (GenBank® Accession No. XP_002491803, renamed as *Komagataella pastoris*), with 71% identity and 89% similarity with an expectation value of 8e-85, followed by ScGPX3 with 72% identity and 86% similarity, and an expectation value of 7e-84.

Based on the above analyses, SEQ ID NO:27 was hypothesized to encode the Gpx3 thiol peroxidase of *Y. lipolytica* ("YIGpx3"), wherein the protein sequence is set forth as SEQ ID NO:28.

EXAMPLE 6

**Overexpression Of Yarrowia lipolytica GPX3 In Y. lipolytica Strain Y4184**

The present Example describes synthesis of overexpression construct pYRH65 (FIG. 5B; SEQ ID NO:29) and its transformation into *Y. lipolytica* strain Y4184U (Example 7). The effect of YIGPX3 overexpression on accumulated lipid level was determined and compared. Specifically, YIGPX3 overexpression resulted in increased total lipid (measured as total fatty acids as a percent of the total dry cell weight ["TFAs % DCW"]) as compared to cells whose native Gpx3 level had not been manipulated.

**Construction Of Y. lipolytica Overexpression Plasmid pYRH65**

A 510 bp fragment encoding the YAL10E02310g was amplified from genomic DNA of *Yarrowia lipolytica* ATCC #20362 using primers GPX3-F (SEQ ID NO:30) and GPX3-R (SEQ ID NO:31). The reaction mixture contained 1 µl of the genomic DNA, 1 µl each of the primers (from 20 µM stocks), 2 µl water, and 45 µl AccuPrime Pfx SuperMix from Invitrogen. Amplification was carried out as follows: initial denaturation at 95 °C for 5 min, followed by 35 cycles of denaturation at 95 °C for 60 sec, annealing at 55 °C for 60 sec, and elongation at 68 °C for 60 sec. A final elongation cycle at 72 °C for 7 min was carried out, followed by reaction termination at 4 °C. A 0.51 kb DNA fragment was obtained from the PCR reaction.
The amplified gene was then cut with Ncol/NotI and used to produce pYRH65 (FIG. 5B; SEQ ID NO:29), containing the following components:

<table>
<thead>
<tr>
<th>RE Sites And Nucleotides Within SEQ ID NO:29</th>
<th>Description Of Fragment And Chimeric Gene Components</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pmel BsiWI</strong> (6183–317)</td>
<td>FBAINm::YIGPX3::PEX20, comprising:</td>
</tr>
<tr>
<td></td>
<td>• FBAINm: Yarrowia lipolytica FBAINm promoter Pmel /Ncol (U.S. Pat. No. 7,202,356);</td>
</tr>
<tr>
<td></td>
<td>• YIGPX: Yarrowia lipolytica GPX3 (Ncol/NotI)</td>
</tr>
<tr>
<td></td>
<td>• PEX20: Pex20 terminator sequence from Yarrowia PEX20 gene NotI/ BsiWI (GenBank Accession No. AF054613)</td>
</tr>
<tr>
<td><strong>BsiWI/Ascl</strong> (318-1211)</td>
<td>894 bp 5’ portion of Yarrowia Lip7 gene (labeled as “LipY-5” in Figure; GenBank Accession No. AJ549519)</td>
</tr>
<tr>
<td><strong>Pacl/SphI</strong> (3920/4681)</td>
<td>762 bp 3’ portion of Yarrowia Lip7 gene (labeled as “LipY-3” in Figure; GenBank Accession No. AJ549519)</td>
</tr>
<tr>
<td><strong>Pacl/Pmel</strong> (4682-6182)</td>
<td>Yarrowia URA3 gene (GenBank Accession No. AJ306421)</td>
</tr>
<tr>
<td>2200-3060</td>
<td>Ampicillin-resistance gene (AmpR) for selection in E. coli</td>
</tr>
</tbody>
</table>

**Lipid Content And Composition In Transformant Strain Y4184U+Gpx3**

Plasmid pYRH65 was cut with BsiWI/Pacl and a 3.3 kB fragment was isolated and used for transformation of Y. lipolytica strain Y4184U, thereby producing strain Y4184U+Gpx3.

Y. lipolytica strain Y4184U (Ura+) (control) and strain Y4184U+Gpx3 were grown under comparable oleaginous conditions (as described in Example 4). Table 11 summarizes the DCW, TFAs % DCW, the concentration of each fatty acid as % TFAs, and EPA % DCW, in a format similar to that used in Table 7.

<table>
<thead>
<tr>
<th>Strains</th>
<th>DCW (g/L)</th>
<th>TFAs % DCW</th>
<th>% TFAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y4184U (Ura+) And Y4184U+Gpx3</td>
<td>18:1</td>
<td>18:2</td>
<td>20:5</td>
</tr>
</tbody>
</table>

69
<table>
<thead>
<tr>
<th></th>
<th>6.40</th>
<th>11.2</th>
<th>10.1</th>
<th>28.2</th>
<th>24.8</th>
<th>2.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y4184U</td>
<td>5.34</td>
<td>15.5</td>
<td>8.8</td>
<td>28.7</td>
<td>27.5</td>
<td>4.3</td>
</tr>
<tr>
<td>(Ura+)</td>
<td>6.20</td>
<td>13.3</td>
<td>9.8</td>
<td>28.3</td>
<td>25.5</td>
<td>3.4</td>
</tr>
<tr>
<td>AVE</td>
<td>5.98</td>
<td>13.3</td>
<td>9.6</td>
<td>28.4</td>
<td>25.9</td>
<td>3.5</td>
</tr>
<tr>
<td>Y4184U</td>
<td>4.92</td>
<td>16.6</td>
<td>10.7</td>
<td>27.1</td>
<td>25.5</td>
<td>4.2</td>
</tr>
<tr>
<td>+Gpx3</td>
<td>4.28</td>
<td>20.0</td>
<td>11.2</td>
<td>26.2</td>
<td>25.7</td>
<td>5.1</td>
</tr>
<tr>
<td>AVE</td>
<td>4.34</td>
<td>19.5</td>
<td>11.0</td>
<td>26.6</td>
<td>25.7</td>
<td>5.6</td>
</tr>
</tbody>
</table>

Overexpression of YIGPX3 (SEQ ID NO:27), corresponding to locus YALI0E02310g, in Y4184U increased lipid content ['TFAs % DCW'] by approximately 47% and increased average EPA titer ['EPA % DCW'] approximately 40%, as compared to that of strain Y4184U (Ura+).

Thus, it appears that overexpression of YIGpx3 in a PUFA-producing strain of *Yarrowia lipolytica* provided increased resistance to oxidative stresses. One beneficial outcome of this increased resistance to oxidative stresses was increased protection against lipid peroxidation, which thereby resulted in increased lipid and PUFA content.

**EXAMPLE 7**

**Generation Of Yarrowia lipolytica Strains Y4184 And Y4184U For High EPA Production**

*Y. lipolytica* strain Y4184U was used as a host in Examples 4 and 6. Strain Y4184U was derived from *Y. lipolytica* ATCC #20362 and is capable of producing high EPA relative to the total lipids via expression of a delta-9 elongase/ delta-8 desaturase pathway. The strain has a *Ura*-phenotype and its construction is described in Example 7 of PCT Publication No. WO 2008/073367, hereby incorporated herein by reference.

The development of strain Y4184U required the construction of strains Y2224, Y4001, Y4001U, Y4036, Y4036U, Y4069, Y4084, Y4084U1, Y4127 (deposited with the American Type Culture Collection on November 29, 2007, under accession number ATCC PTA-8802), Y4127U2, Y4158, Y4158U1 and Y4184.
The final genotype of strain Y4184 (producing 30.7% EPA of total lipids) with respect to wildtype *Yarrowia lipolytica* ATCC #20362 was unknown 1-, unknown 2-, unknown 4-, unknown 5-, unknown 6-, unknown 7-, YAT1::ME3S::Pex16, EXP1::ME3S::Pex20 (2 copies), GPAT::EgD9e::Lip2, FBAInm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, FBA::EgD9eS::Pex20, YAT1::EgD9eS::Lip2, GPD::EgD9eS::Lip2, GPDIN::EgD8M::Lip1, YAT1::EgD8M::Aco, EXP1::EgD8M::Pex16, FBAInm::EgD8M::Pex20, FBAIn::EgD8M::Lip1 (2 copies), GPM/FBAIn::FmD12S::Oct, EXP1::FmD12S::Aco, YAT1::FmD12::Oct, GPD::FmD12::Pex20, EXP1::EgD5S::Pex20, YAT1::EgD5S::Aco, YAT1::Rd5S::Oct, FBAIn::EgD5::Aco, FBAInm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::YICPT1::Aco, GPD::YICPT1::Aco.

Abbreviations above are as follows: ME3S is a codon-optimized C\textsubscript{16/18} elongase gene, derived from *Mortierella alpina* [U.S. Pat. No. 7,470,532]; EgD9e is a *Euglena gracilis* delta-9 elongase gene [U.S. Pat. No. 7,645,604]; EgD9eS is a codon-optimized delta-9 elongase gene, derived from *Euglena gracilis* [U.S. Pat. No. 7,645,604]; EgD8M is a synthetic mutant delta-8 desaturase [U.S. Pat. No. 7,709,239], derived from *Euglena gracilis* [U.S. Pat. No. 7,256,033]; FmD12 is a *Fusarium moniliforme* delta-12 desaturase gene [U.S. Pat. No. 7,504,259]; FmD12S is a codon-optimized delta-12 desaturase gene, derived from *Fusarium moniliforme* [U.S. Pat. No. 7,504,259]; EgD5 is a *Euglena gracilis* delta-5 desaturase [U.S. Pat. No. 7,678,560]; EgD5S is a codon-optimized delta-5 desaturase gene, derived from *Euglena gracilis* [U.S. Pat. No. 7,678,560]; RD5S is a codon-optimized delta-5 desaturase, derived from *Peridinium* sp. CCMP626 [U.S. Pat. No. 7,695,950]; PaD17 is a *Pythium aphanidermatum* delta-17 desaturase [U.S. Pat. No. 7,556,949]; PaD17S is a codon-optimized delta-17 desaturase, derived from *Pythium aphanidermatum* [U.S. Pat. No. 7,556,949]; and, YICPT1 is a *Yarrowia lipolytica* diacetylglycerol cholinephosphotransferase gene [U.S. Pat. No. 7,932,077].
Finally, in order to disrupt the Ura3 gene in strain Y4184, construct pZKUE3S (PCT Publication No. WO 2008/073367, SEQ ID NO:78 therein) was used to integrate a EXP1::ME3S::Pex20 chimeric gene into the Ura3 gene of strain Y4184 to result in strains Y4184U1 (11.2% EPA of total lipids), Y4184U2 (10.6% EPA of total lipids) and Y4184U4 (15.5% EPA of total lipids), respectively (collectively, Y4184U).

It is noted that PCT Publication No. WO 2008/073367 describes a discrepancy in the EPA % TFAs quantified in Y4184 (30.7%) versus Y4184U (average 12.4%) due to differing growth conditions.

**EXAMPLE 8**

**Generation Of Yarrowia lipolytica Strains Y9502 And Y9502U For High EPA Production**

*Y. lipolytica* strain Y9502U was used as a host in Example 4. Strain Y9502U was derived from *Y. lipolytica* ATCC #20362 and is capable of producing high EPA relative to the total lipids via expression of a delta-9 elongase/delta-8 desaturase pathway. The strain has a *Ura*- phenotype.

**Genotype Of Yarrowia lipolytica Strain Y9502**

The generation of strain Y9502 is described in U.S. Pat. Appl. Pub. No. 2010-0317072-A1. Strain Y9502, derived from *Y. lipolytica* ATCC #20362, was capable of producing about 57.0% EPA relative to the total lipids via expression of a delta-9 elongase/delta-8 desaturase pathway.

The final genotype of strain Y9502 with respect to wildtype Yarrowia lipolytica ATCC #20362 was *Ura+, Pex3-, unknown 1-, unknown 2-, unknown 3-, unknown 4-, unknown 5-, unknown6-, unknown 7-, unknown 8-, unknown9-, unknown 10-, YAT1::ME3S::Pex16, GPD::ME3S::Pex20, YAT1::ME3S::Lip1, FBAINm::EgD9eS::Lip2, EXP1::EgD9eS::Lip1, GPAT::EgD9e::Lip2, YAT1::EgD9eS::Lip2, FBAINm::EgD8M::Pex20, EXP1::EgD8M::Pex16, FBAIN::EgD8M::Lip1, GPD::EaD8S::Pex16 (2 copies), YAT1::E389D9eS/EgD8M::Lip1, YAT1::EgD9eS/EgD8M::Aco, FBAINm::EaD9eS/EaD8S::Lip2, GPD::FmD12::Pex20, YAT1::FmD12::Oct, EXP1::FmD12::Aco, GPDIN::FmD12::Pex16, EXP1::EgD5M::Pex16, FBAIN::EgD5SM::Pex20, GPDIN::EgD5SM::Aco, GPM::EgD5SM::Oct, EXP1::EgD5SM::Lip1,
YAT1::EaD5SM::Oct, FBAINm::PaD17::Aco, EXP1::PaD17::Pex16, YAT1::PaD17S::Lip1, YAT1::YICPT1::Aco, YAT1::MCS::Lip1, FBA::MCS::Lip1, YAT1::MaLPAAAT1S::Pex16.

Abbreviations used above and not set forth in Example 7 are as follows: EaD8S is a codon-optimized delta-8 desaturase gene, derived from *Euglena anabaena* [U.S. Pat. 7,790,156]; E389D9eS/EgD8M is a DGLA synthase created by linking a codon-optimized delta-9 elongase gene (“E389D9eS”), derived from *Eutreptiella sp.* CCMP389 (U.S. Pat. 7,645,604), to the delta-8 desaturase “EgD8M” *supra* [U.S. Pat. Appl. Pub. No. 2008-0254191-A1]; EgD9eS/EgD8M is a DGLA synthase created by linking the delta-9 elongase “EgD9eS” *supra* to the delta-8 desaturase “EgD8M” *supra* [U.S. Pat. Appl. Pub. No. 2008-0254191-A1]; EaD9eS/EgD8M is a DGLA synthase created by linking a codon-optimized delta-9 elongase gene (“EaD9eS”), derived from *Euglena anabaena* [U.S. Pat. 7,794,701], to the delta-8 desaturase “EgD8M” *supra* [U.S. Pat. Appl. Pub. No. 2008-0254191-A1]; EgD5M and EgD5SM are synthetic mutant delta-5 desaturase genes comprising a mutant HPGs (SEQ ID NO:41) motif [U.S. Pat. App. Pub. 2010-0075386-A1], derived from *Euglena gracilis* [U.S. Pat. 7,678,560]; EaD5SM is a synthetic mutant delta-5 desaturase gene comprising a mutant HaGG (SEQ ID NO:42) motif [U.S. Pat. App. Pub. 2010-0075386-A1], derived from *Euglena anabaena* [U.S. Pat. 7,943,365]; MCS is a codon-optimized malonyl-CoA synthetase gene, derived from *Rhizobium leguminosarum* bv. *viciae* 3841 [U.S. Pat. App. Pub. 2010-0159558-A1], and, MaLPAAAT1S is a codon-optimized lysophosphaticid acid acyltransferase gene, derived from *Mortierella alpina* [U.S. Pat. 7,879,591].

For a detailed analysis of the total lipid content and composition in strain Y9502, a flask assay was conducted wherein cells were grown in 2 stages for a total of 7 days. Based on analyses, strain Y9502 produced 3.8 g/L DCW, 37.1 TFAs % DCW, 21.3 EPA % DCW, and the lipid profile was as follows, wherein the concentration of each fatty acid is as a weight percent of TFAs ["% TFAs"]: 16:0 (palmitate)—2.5, 16:1 (palmitoleic acid)—0.5, 18:0 (stearic acid)—2.9, 18:1 (oleic acid)—5.0, 18:2 (LA)—
12.7, ALA—0.9, EDA—3.5, DGLA—3.3, ARA—0.8, ETrA—0.7, ETA—2.4, EPA—57.0, other—7.5.

**Genotype Of Yarrowia lipolytica Strain Y9502U**

To disrupt the *Ura3* gene in strain Y9502, *SalI/PacI*-digested construct pZKUM (see U.S. Pat. Appl. Pub. No. 2009-0093543-A1, Table 15, SEQ ID NO:133 and FIG. 8A therein) was used to integrate an *Ura3* mutant gene into the *Ura3* gene of strain Y9502, according to the General Methods. A total of 27 transformants (selected from a first group comprising 8 transformants, a second group comprising 8 transformants, and a third group comprising 11 transformants) were grown on Minimal Media + 5-fluoroorotic acid [“MM+ 5-FOA”] selection plates and maintained at 30°C for 2 to 5 days. MM+ 5-FOA comprises (per liter): 20 g glucose, 6.7 g Yeast Nitrogen base, 75 mg uracil, 75 mg uridine and an appropriate amount of FOA (Zymo Research Corp., Orange, CA), based on FOA activity testing against a range of concentrations from 100 mg/L to 1000 mg/L (since variation occurs within each batch received from the supplier).

Further experiments determined that only the third group of transformants possessed a real *ura*- phenotype.

The *ura*- cells were scraped from the MM + 5-FOA plates and subjected to fatty acid analysis, according to the General Methods. In this way, GC analyses showed that there were 28.5%, 28.5%, 27.4%, 28.6%, 29.2%, 30.3% and 29.6% EPA of TFAs in pZKUM-transformants #1, #3, #6, #7, #8, #10 and #11 grown on MM + 5-FOA plates of group 3, respectively. These seven strains were designated as strains Y9502U12, Y9502U14, Y9502U17, Y9502U18, Y9502U19, Y9502U21 and Y9502U22, respectively (collectively, Y9502U).

**EXAMPLE 9**

**Identification Of A Yarrowia lipolytica Gene Having Homology To The Saccharomyces cerevisiae Tsa1 Gene**

An ortholog to the *S. cerevisiae* Tsa1 (GenBank® Accession No. NP_013684; SEQ ID NO:34) [“ScTsa1”] was identified in *Yarrowia lipolytica* by conducting BLAST searches using ScTsa1 as the query
sequence against the public *Y. lipolytica* protein database of the “Yeast project Genolevures” (Center for Bioinformatics, LaBRI, Talence Cedex, France) (see also Dujon, B. et al., *Nature*, 430 (6995):35-44 (2004)).

The protein sequence having the best homology (with an expectation value of 1e-82) to ScTsa1 among all *Y. lipolytica* proteins, YALI0B15125g (GenBank Accession No. XP_500915.1; SEQ ID NO:36), was given the designation “YITsa1”. YALI0B15125g was annotated therein as “highly similar to uniprot|P34760 Saccharomyces cerevisiae YML028w TSA1 (ohnolog of YDR453C) Thioredoxin peroxidase, acts as both a ribosome-associated and free cytoplasmic antioxidant”.

An alignment of ScTsa1 and the putative YITsa1 is shown in FIG. 7. There are only two Cys residues both in ScTsa1 and YITsa1. Vertical boxes highlight Cys48 and Cys171 of ScTsa1, important for inter- and intra-molecular interactions (Tachibana, T. et al., *J. Biol. Chem.*, 284:4464-4472 (2009)). The former Cys residue is conserved in YITsa1, while the latter is shifted two amino acids upstream in YITsa1 when compared to ScTsa1.

Using the protein sequence encoding YALI0B15125g (SEQ ID NO:36), NCBI BLASTP 2.2.26+ searches were conducted to identify sequences having similarity within the BLAST “nr” database, according to the methodology set forth in Example 1.

A large number of proteins were identified as sharing significant similarity to YALI0B15125g (SEQ ID NO:36). Table 12 provides a partial summary of those hits having an Expectation value greater or equal to “2e-102” and annotation that specifically identified the protein (i.e., while hits to hypothetical proteins are excluded), although this should not be considered as limiting to the disclosure herein. The proteins in Table 12 shared between 95-100% query coverage with SEQ ID NO:36.

Table 12. Genes Sharing Similarity To YITsa1 (SEQ ID NO:36)

<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
<th>Query coverage</th>
<th>E value</th>
</tr>
</thead>
<tbody>
<tr>
<td>XP_500915.1</td>
<td>YALI0B15125p [Yarrowia lipolytica] &gt;emb</td>
<td>CAG83166.1</td>
<td>YALI0B15125p</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Query coverage</td>
<td>E value</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>XP_002616355.1</td>
<td>peroxiredoxin TSA1 [Clavispora lusitaniae ATCC 42720]</td>
<td>100%</td>
<td>4e-117</td>
</tr>
<tr>
<td>XP_001485052.1</td>
<td>peroxiredoxin TSA1 [Meyerozyma guillermondii ATCC 6260]</td>
<td>100%</td>
<td>3e-115</td>
</tr>
<tr>
<td>EGW31724.1</td>
<td>peroxiredoxin TSA1 [Spathaspora passalidarum NRRL Y-27907]</td>
<td>100%</td>
<td>1e-114</td>
</tr>
<tr>
<td>XP_001382622.1</td>
<td>Peroxiredoxin TSA1 [Scheffersomyces stipitis CBS 6054]</td>
<td>100%</td>
<td>2e-114</td>
</tr>
<tr>
<td>XP_002491977.1</td>
<td>Thioredoxin peroxidase, acts as both a ribosome-associated and free cytoplasmic antioxidant [Komagataella pastoris GS115]</td>
<td>99%</td>
<td>4e-113</td>
</tr>
<tr>
<td>XP_001526168.1</td>
<td>peroxiredoxin TSA1 [Lodderomyces elongisporus NRRL YB-4239]</td>
<td>100%</td>
<td>4e-112</td>
</tr>
<tr>
<td>EFW97887.1</td>
<td>putative peroxiredoxin [Ogataea parapolymorpha DL-1]</td>
<td>98%</td>
<td>5e-111</td>
</tr>
<tr>
<td>ACV49765.1</td>
<td>putative peroxiredoxin [Ogataea angusta]</td>
<td>98%</td>
<td>1e-110</td>
</tr>
<tr>
<td>BAH80187.1</td>
<td>thioredoxin peroxidase 1 [Komagataella pastoris]</td>
<td>95%</td>
<td>2e-110</td>
</tr>
<tr>
<td>XP_002547929.1</td>
<td>peroxiredoxin TSA1 [Candida tropicalis MYA-3404]</td>
<td>100%</td>
<td>2e-110</td>
</tr>
<tr>
<td>XP_716082.1</td>
<td>likely thioredoxin peroxidase [Candida albicans]</td>
<td>100%</td>
<td>2e-109</td>
</tr>
<tr>
<td>XP_002419517.1</td>
<td>thioredoxin peroxidoxin, putative; [Candida dubliniensis CD36]</td>
<td>100%</td>
<td>8e-109</td>
</tr>
<tr>
<td>EEU06015.1</td>
<td>Tsa1p [Saccharomyces cerevisiae JAY291]</td>
<td>100%</td>
<td>1e-104</td>
</tr>
<tr>
<td>NP_013684.1</td>
<td>Peroxiredoxin TSA1 (also Cytoplasmic thiol peroxidase 1) [Saccharomyces cerevisiae]</td>
<td>100%</td>
<td>3e-104</td>
</tr>
<tr>
<td>EGA57449.1</td>
<td>Tsa1p [Saccharomyces cerevisiae FostersB]</td>
<td>100%</td>
<td>2e-102</td>
</tr>
</tbody>
</table>

Based on the BLASTP searches, YALI0B15125g (SEQ ID NO:36) shared the best similarity with the Tsa1 peroxiredoxin from Clavispora lusitaniae ATCC 42720 (GenBank® Accession No. XP_002616355.1), with 81% identity and 92% similarity with an expectation value of 4e-117, followed by the TSA1 peroxiredoxin from Meyerozyma guillermondii ATCC 6260 with 80% identity and 91% similarity, and an expectation value of 3e-115.
Based on the above analyses, SEQ ID NO:35 was hypothesized to encode the TSA1 peroxiredoxin of *Y. lipolytica* ("YITsa1"), wherein the protein sequence is set forth as SEQ ID NO:36.

It is hypothesized herein that overexpression of YITsa1 in a PUFA-producing strain of *Yarrowia lipolytica* will provide increased resistance to oxidative stresses. One beneficial outcome of this increased resistance to oxidative stresses will be increased protection against lipid peroxidation, which will thereby result in increased lipid and PUFA content.

**EXAMPLE 10**

**Identification Of A *Yarrowia lipolytica* Gene Having Homology To The *Saccharomyces cerevisiae* Ybp1 Gene**

An ortholog to the *S. cerevisiae* Ybp1 (GenBank® Accession No. NP_009775.1; SEQ ID NO:38) ["ScYbp1"] was identified in *Yarrowia lipolytica* by conducting BLAST searches using ScYbp1 as the query sequence against the public *Y. lipolytica* protein database of the "Yeast project Genolevures" (Center for Bioinformatics, LaBRI, Talence Cedex, France) (see also Dujon, B. et al., *Nature*, 430 (6995):35-44 (2004)).

The protein sequence having the best homology (with an expectation value of 5e-22) to ScYbp1 among all *Y. lipolytica* proteins, YALI0B03762g (GenBank Accession No. XP_500469.1; SEQ ID NO:40), was given the designation "YLYbp1". YALI0B03762g was annotated therein as "weakly similar to uniprot|P53169 Saccharomyces cerevisiae YGL060w YBP2 (ohnolog of YBR216C) Central kinetochore associated protein that mediates mitotic progression".

An alignment of ScYbp1 and the putative YLYbp1 is shown in FIG. 8, although little sequence conservation between the proteins is noted.

Using the protein sequence encoding YALI0B03762g (SEQ ID NO:40), NCBI BLASTP 2.2.26+ searches were conducted to identify sequences having similarity within the BLAST "nr" database, according to the methodology set forth in Example 1.

Several proteins were identified as sharing similarity to YALI0B03762g (SEQ ID NO:40). Table 13 provides a partial summary of those hits having an Expectation value greater or equal to "2e-37" and
annotation that specifically identified the protein (i.e., while hits to hypothetical proteins and proteins from *Saccharomyces cerevisiae* are excluded), although this should not be considered as limiting to the disclosure herein. The proteins in Table 13 shared between 74-93% query coverage with SEQ ID NO:40.

<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
<th>Query coverage</th>
<th>E value</th>
</tr>
</thead>
<tbody>
<tr>
<td>XP_001386941.2</td>
<td>YAP1 binding protein 2 (YBP2) <em>Scheffersomyces stipitis</em> CBS 6054</td>
<td>75%</td>
<td>6e-53</td>
</tr>
<tr>
<td>EGV63342.1</td>
<td>YAP1 binding protein 2 <em>Candida tenuis</em> ATCC 10573</td>
<td>74%</td>
<td>3e-48</td>
</tr>
<tr>
<td>EGW32572.1</td>
<td>YAP1 binding protein 2 <em>Spathaspora passalidarum</em> NRRL Y-27907</td>
<td>79%</td>
<td>3e-46</td>
</tr>
<tr>
<td>XP_002492586.1</td>
<td>YAP1-binding protein 1 <em>Komagataella pastoris</em> CBS 7435</td>
<td>93%</td>
<td>5e-46</td>
</tr>
<tr>
<td>XP_002417933.1</td>
<td>redox regulator, putative <em>Candida dubliniensis</em> CD36</td>
<td>88%</td>
<td>9e-38</td>
</tr>
<tr>
<td>XP_722350.1</td>
<td>potential redox regulator <em>Candida albicans</em> SC5314</td>
<td>88%</td>
<td>2e-37</td>
</tr>
</tbody>
</table>

Based on the BLASTP searches, YALI0B03762g (SEQ ID NO:40) shared the best similarity with hypothetical protein CLUG_00080 from *Clavispora lusitaniae* ATCC 42720 (GenBank® Accession No. XP_002618921.1), with 28% identity and 48% similarity, and an expectation value of 1e-54.

Among proteins with known function, the best hit was the YAP1 binding protein 2 from *Scheffersomyces stipitis* CBS 6054 (GenBank® Accession No. XP_001386941.2), with 30% identity and 49% similarity with an expectation value of 6e-53, followed by YAP1 binding protein 2 from *Candida tenuis* ATCC 10573 (GenBank® Accession No. EGV63342.1) with 28% identity and 47% similarity, and an expectation value of 3e-48.

Based on the above analyses, SEQ ID NO:39 was hypothesized to encode the YAP1 binding protein of *Y. lipolytica* ("YIYbp1"), wherein the protein sequence is set forth as SEQ ID NO:40.
The protein sequence set forth in SEQ ID NO:40 was aligned with the following proteins set forth in Table 14, using a CLUSTAL W (1.81) multiple sequence alignment (FIG. 9; Thompson J.D., et al., *Nucleic Acids Res.* 22:4673-4680 (1994)) to further evaluate YLYbp1. It is hypothesized that each of these proteins encode a homolog of Ybp1.

**Table 14. Proteins Aligned With YLYbp1 (SEQ ID NO:40)**

<table>
<thead>
<tr>
<th>Protein</th>
<th>Annotation</th>
<th>GenBank Accession No.</th>
<th>SEQ ID NO</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Saccharomyces cerevisiae</em> S288c Ybp1</td>
<td>Protein required for oxidation of specific cysteine residues of the transcription factor Yap1p, resulting in the nuclear localization of Yap1p in response to stress</td>
<td>NP_009775.1</td>
<td>38</td>
</tr>
<tr>
<td><em>Candida glabrata</em> unnamed protein product CAGL0K06743g</td>
<td>similar to uniprot</td>
<td>P38315 S. cerevisiae YBR216c</td>
<td>CAG61477.1</td>
</tr>
<tr>
<td><em>Kluyveromyces lactis</em> NRRL Y-1140 hypothetical protein KLLA-ORF8035</td>
<td>similar to uniprot</td>
<td>P38315 S. cerevisiae YBR216C YBP1 and to uniprot</td>
<td>P53169 S. cerevisiae YGL060W YBP2</td>
</tr>
<tr>
<td><em>Scheffersomyces stipitis</em> CBS 6054 (Pichia stipitis CBS 6054) YAP1 binding protein 2 (YBP2)</td>
<td>required for the oxidative stress response to peroxides via the Yap1p transcription factor</td>
<td>XP_001386941.2</td>
<td>45</td>
</tr>
<tr>
<td><em>Zygosaccharomyces rouxi</em> CBS 732 Hypothetical protein ZYRO-ORF6798</td>
<td>--</td>
<td>XP_002495870.1</td>
<td>46</td>
</tr>
<tr>
<td><em>Candida albicans</em> SC5314 YBP1 (CaO19.5034)</td>
<td>similar to S. cerevisiae YBP1 (YBR216C) redox regulator of thioredoxin transcriptional regulatory factor YAP1</td>
<td>XP_722236.1</td>
<td>47</td>
</tr>
</tbody>
</table>

Relatively few regions of sequence conservation were observed between the proteins upon visual inspection of the alignment. However, each of the seven proteins was included within the kinetochor_Ybp2 super family (Pfam08568; described as a family of proteins integrally involved in the central kinetochore) upon analysis using the "Identify Conserved Domains" tool of National Center for Biotechnology Information ["NCBI"] to view conserved domains detected within the protein sequence using a

It is hypothesized herein that overexpression of YlYbp1 in a PUFA-producing strain of *Yarrowia lipolytica* will provide increased resistance to oxidative stresses. One beneficial outcome of this increased resistance to oxidative stresses will be increased protection against lipid peroxidation, which will thereby result in increased lipid and PUFA content.
CLAIMS

What is claimed is:

1. A transgenic oleaginous yeast having increased oil content comprising increased Yap1 transcription factor activity wherein the increased oil content is compared to the oil content of a non-transgenic oleaginous yeast.

2. The transgenic oleaginous yeast of claim 1 wherein the increased Yap1 transcription factor activity results from overexpressing the Yap1 transcription factor, by increasing the interaction between the transcription factor and a protein that is capable of activating the transcription factor, or by a combination thereof.

3. The transgenic oleaginous yeast of claim 2 wherein the protein that is capable of activating the transcription factor is selected from the group consisting of: Gpx3, Ybp1 and Tsa1.

4. The transgenic oleaginous yeast of claim 2, wherein the Yap1 transcription factor comprises a nucleotide sequence encoding a polypeptide having transcription factor activity and comprising:
   a) a bZIP leucine zipper motif;
   b) an N-terminal Cys-rich domain comprising a sequence of at least two cysteine residues that are separated by at least 6 amino acids; and,
   c) a C-terminal Cys-rich domain comprising a sequence of at least two cysteine residues that are separated by at least 8 amino acids.

5. The transgenic oleaginous yeast of claim 4, wherein the sequence of the Yap1 transcription factor is selected from the group consisting of SEQ ID NO:2 and SEQ ID NO:4.
6. The transgenic oleaginous yeast of claim 3, wherein the Gpx3 protein comprises:
   a) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the polypeptide has at least 70% amino acid identity, based on the BLASTP method of alignment, when compared to a sequence selected from the group consisting of SEQ ID NO:26 [ScGpx3] or SEQ ID NO:28 [YIGpx3];
   b) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the nucleotide sequence has at least 70% sequence identity, based on the BLASTN method of alignment, when compared to a sequence selected from the group consisting of SEQ ID NO:25 [ScGpx3] or SEQ ID NO:27 [YIGpx3]; or,
   c) a complement of the nucleotide sequence of (a) or (b), wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

7. The transgenic oleaginous yeast of claim 3, wherein the Tsa1 protein comprises:
   a) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the polypeptide has at least 70% amino acid identity, based on the BLASTP method of alignment, when compared to a sequence selected from the group consisting of SEQ ID NO:34 [ScTsa1] or SEQ ID NO:36 [YITsa1];
   b) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the nucleotide sequence has at least 70% sequence identity, based on the BLASTN
method of alignment, when compared to a sequence selected from the group consisting of SEQ ID NO:33 [ScTsa1] or SEQ ID NO:35 [YITsa1]; or,
c) a complement of the nucleotide sequence of (a) or (b), wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

8. The transgenic oleaginous yeast of claim 3, wherein the Ybp1 protein comprises:
a) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the polypeptide is selected from the group consisting of SEQ ID NO:38 [ScYbp1] or SEQ ID NO:40 [YIYbp1]; or,
b) a nucleotide sequence encoding a polypeptide capable of interacting with the Yap1 transcription factor to increase Yap1 transcription factor activity, wherein the polypeptide sequence is classified within a kinetochor_Ybp2 super family, based on a conserved domain method of analysis; or,
c) a complement of the nucleotide sequence of (a) or (b), wherein the complement and the nucleotide sequence consist of the same number of nucleotides and are 100% complementary.

9. The transgenic oleaginous yeast of claim 1, wherein the transgenic oleaginous yeast is from a genus selected from the group consisting of: Yarrowia, Candida, Rhodotorula, Rhodosporidium, Cryptococcus, Trichosporon, and Lipomyces.

10. The transgenic oleaginous yeast of claim 8, wherein the transgenic oleaginous yeast is Yarrowia lipolytica.
11. The transgenic oleaginous yeast of claim 1, wherein the transgenic oleaginous yeast produces at least one polyunsaturated fatty acid.

12. A method of increasing oil content in an oleaginous yeast comprising:
   a) engineering the oleaginous yeast to overexpress a protein selected from the group consisting of:
      (i) a Yap1 transcription factor;
      (ii) a protein that is capable of activating the transcription factor;
      (iii) a combination of (a) and (b); and,
   b) growing the oleaginous yeast under suitable conditions to result in increased oil content when compared to the oil content of a non-transgenic oleaginous yeast.
Application number / Numéro de demande: 2823210

Figures: 4A, 4B

Pages: 6/17

Unscannable items received with this application
(Request original documents in File Prep. Section on the 10th floor)

Documents reçu avec cette demande ne pouvant être balayés
(Commander les documents originaux dans la section de la préparation des dossiers au 10ième étage)
ScYap1 (SEQ ID NO:2)
YlYap1 (SEQ ID NO:4)

---

ScYap1 (SEQ ID NO:2)
YlYap1 (SEQ ID NO:4)

---

ScYap1 (SEQ ID NO:2)
YlYap1 (SEQ ID NO:4)

---

ScYap1 (SEQ ID NO:2)
YlYap1 (SEQ ID NO:4)

---

ScYap1 (SEQ ID NO:2)
YlYap1 (SEQ ID NO:4)

---

ScYap1 (SEQ ID NO:2)
YlYap1 (SEQ ID NO:4)

---

FIG. 2A
ScYap1 (SEQ ID NO:2)  CGPDKPIALDKKEVPAISSILSSNSPAITNTWBHSNITDNPANVIATDATKYE
YlYap1 (SEQ ID NO:4)  ---KAELLQRSATASASPPTTNTVPSAAAGSTQQSAGQPFSSTPS---

ScYap1 (SEQ ID NO:2)  SGFGRGFDMSANHYVNNDSTGSTDSTGKTNKMNNSDDVLPFISESFPDMQVT
YlYap1 (SEQ ID NO:4)  ---TTAPPQLSASVAAGLDLPSGDFS---FLDPDDMFMDMFS

ScYap1 (SEQ ID NO:2)  NFFPSGPGTIGNNAASNTNPSLQQSSKEIDIPFINANLAFDDNSTNIQLQPPSESQ8QNK
YlYap1 (SEQ ID NO:4)  ------------------------YRDVFSETAHLD

ScYap1 (SEQ ID NO:2)  FDYDMPFPRDDSSKECNLNGPEFLEDDDDKKAANMSESSSLIKNLQNEPPELPQYLQS
YlYap1 (SEQ ID NO:4)  FSLPELTETTS----------MFDPDPFDHSSVDVGKPY------LST

ScYap1 (SEQ ID NO:2)  VPGNESSEIQNNSLGQNDNGNNDNDNVPSKEGSSLRCSEIWDRITTHPKYSID
YlYap1 (SEQ ID NO:4)  MGATHSGVNGGQSGAPEVKE------EDEDLLMFSKPTLMGCTAVNDRTSHPKF8GD

ScYap1 (SEQ ID NO:2)  IDVGEGISELMKAKCSEGQVINAEDVQLALNKHIM
YlYap1 (SEQ ID NO:4)  IDIGEGILERNKAKCSEGVVTLTEDVDVCVLSTFQ-

* - single, fully conserved residue  . - conservation of weak groups
: - conservation of strong groups  - no consensus

FIG. 2B
FIG. 3B
FIG. 5A
FIG. 5B
**FIG. 6**

* - single, fully conserved residue  
: - conservation of strong groups  
- - conservation of weak groups  
- - no consensus
ScTsai (SEQ ID NO:34)  MVAQVQKPATFKTKAVDVGFVDEVSLDKYKGYVVLAFIPMTFVCFPTEIIAFSEA AK
YlTsai (SEQ ID NO:36)  MVATVQHPAPDFFKKTAVGSGYVFEVEVSLDQPKGWVVLAFIPMTFVCFPTEIIAYSDAVS

ScTsai (SEQ ID NO:34)  KPEEQGAWQLFASDSEYSSLLAWTNIPRKEGGLGIPINPLADTNHSLSRDYGVLIEEEG
YlTsai (SEQ ID NO:36)  QPKERGAEVLFASDSEYSSLLAWTVPARKDGGLGPVINPLADTNHTLSDKYGVILPEAG

ScTsai (SEQ ID NO:34)  VALRGLFIIDPKGVHRITINDLPVRGNVDEALRLVEAFQWTDKNTVHL[CNWTPCAATI
YlTsai (SEQ ID NO:36)  VALRGIIFIDPKGVVRQITINDLPVRGTSVEELRLIDAFQTEKHGEV[DAANQKGSDTI

ScTsai (SEQ ID NO:34)  KPTVEDSKYFYEAANK
YlTsai (SEQ ID NO:36)  KADPVNAKEYFEKANK

* - single, fully conserved residue
: - conservation of strong groups
: - conservation of weak groups
- - no consensus

FIG. 7
FIG. 8B
CaYbp1 (SEQ ID NO:47) EYLQRRLLTGFLTQVTYLHANINTGEGYSIEHFSWLQOQSKKSKIKVFERDG---AFCDFRFV
SsYbp1 (SEQ ID NO:45) EYLQRKLLTGFISQLIQMSMDNLINFYTLHLSFLQVPHRGQKLYFEYSVNLVFMDRLA
CgYbp1 (SEQ ID NO:43) IELQVRLLLNLCTFSSVAYCVKLNDKTEVVYFIKLI---NKDLQPFYRSVHDIIRYY
ScYbp1 (SEQ ID NO:38) SELLQVRLLRLRLCTFGRISTP1KTVTNNADVYKYYCALN---QQFELSAYTEYLEDFTCYY
ZrYbp1 (SEQ ID NO:46) SI1QQLRSLSTFPAVEGCLKNKASRDLRBYFRHLR---NTEFHLSENDEEELVLSKRFY
KlYbp1 (SEQ ID NO:44) EVLQRKLLCNI1LTSALHQ1LKLRTCS1LLNTHSLQG-IPTLSTSEYLQGLTDL0LSRYY
YlYbp1 (SEQ ID NO:40) QSLSTYFFYSWFHRVAVRWSSSNLFRQKIHSTHELPRAEARKYDNSKNKSAYTVTTYNW

CaYbp1 (SEQ ID NO:47) ESSLFIDTTLLKCFQ-GFITDSHKLHIGYKKNKASDIHIIELIIFERVVVDYQKNVLTST
SsYbp1 (SEQ ID NO:45) ELALSYDINLTQHFK-SMVADSHTLILRSFDYS---IDRDELSAQIFKEKVYQVDYQKLTAMS
CgYbp1 (SEQ ID NO:43) QIAFDIDLNDFEN-DILRETRGIYEDVTKRINETNNTDKNASDIILLKAGYEVQK-QLMPSFLDVQIEGEQV-NIKECR11YKSVFPEISAVNDEAKLVLRMAYTYFQVCK-QLMPSFLDVQIEGEQV-
ScYbp1 (SEQ ID NO:38) ZrYbp1 (SEQ ID NO:46) KLDBFSLDKDSEQKLDSCVRESVYRSLPKDSEIKELKEITTNPVQLAYYTEVEK-TLALSLLDPSQYFL---PLQBEIQEDVQEATKGGFLDLGFLAGFSKSSLS38ASP1AFDY
YlYbp1 (SEQ ID NO:40) *

CaYbp1 (SEQ ID NO:47) IVDSDAKA1KDSII1GELILFTSHTAGKNNFAKPTMS1HDSLVMLLRTTIPQMVNPKFINA
SsYbp1 (SEQ ID NO:45) TINSDBAKRDSPLGFYLLTHALS1HIRTPELRLYFSDAVLVLTVLYVPE1IQSTTFVK--TAREKEINPD1TG111Sgefny1ENGDDH1-DIDIAADAL1YLRFA1SESLPSPTC1NV
CgYbp1 (SEQ ID NO:43) --AASKMV1GDLVYFGI1TYELNQ1LVKEMN1TAIAL1YLYRFTTP11YSKYYNV
ScYbp1 (SEQ ID NO:48) --L1KLWHENHSGNF1ILG11H11ITYETQKHLPE1111Y11CRT1PS1FSTTYT1NL
ZrYbp1 (SEQ ID NO:46) --1ANVKE111LDPAG1ILHLSNSF1NLPS1AT1LQLQ8AT1MYLRFVT11PSM1FSA1FL1ENR
KlYbp1 (SEQ ID NO:44) SMYDASEIF1PSQEG1GLMLATQYMMENRDSH-LNT1RDQLV1LTVLHVRSSPK1F1LP-F
YlYbp1 (SEQ ID NO:40) * * * * *

FIG. 9C
CaYbp1 (SEQ ID NO:47) GNVHDVVFVWWFWALYQQQIINSKNLQEILSIPKPLLRTFQCLFLFIIVKSEGKPNFKYM
SsYbp1 (SEQ ID NO:45) GVEDATITFHTWALYQTSPLNN-KSVEIIIAAIIFPVLITIYYQVIVPPVTVITNSRPNFKYA
CgYbp1 (SEQ ID NO:43) TIEGVARYWIWALTTDDNN---ILKEKLAELSPLVILHSVILNLLVKNQHQVNE-E1RMI
Scvbp1 (SEQ ID NO:38) AVESVSRYWLLYATTEPB---DVKKEKLKNISVFVVTKTLHVLQKNCIQNVQ-OCLRI
ZrYbp1 (SEQ ID NO:46) YAEQTARYWVWAVTNNKVQ---KLKBEKELPSYIVTFVFLQMNQSCNQPNP-BARMH
Kllybp1 (SEQ ID NO:44) SSHDLARTWLYFALTTNSH---DLMDLSKLDLPSYIVTLYLQTELRACLQIND-NLRTT
YlYbp1 (SEQ ID NO:40) AITDDLIFWGWTTLKROMP---------EVRQLEAFYVKLYQFLVFTSASSLP-E1RNI

CaYbp1 (SEQ ID NO:47) LLLTLLTKLTLTLSP-DTGVEFIKDSLNNCPYEVSVPYSLIGVGLYQNLIN---EKWDVNUSEL
SsYbp1 (SEQ ID NO:45) VTLTLVRVLALSPEDSLDSVFKDSLHNCPYESEKPIMIGVKKELITK---DKSSSTSDTT
CgYbp1 (SEQ ID NO:43) TFFTLTIRLCCLPENCSYFEPMLMEDLDNCAVFGKCHGLRLDLV1K---VDHSVSSNNT
Scvbp1 (SEQ ID NO:38) TFFTLTRLLCLPKEPVAPEPIIDLVKTSLPPLAKTSALCVFKDLSSR---R-IKNDKDS
ZrYbp1 (SEQ ID NO:46) SFFLLTRRLMPETSPEVFVLDLTLLTCFPFTNNLAGIKCGLKDLTRRNCQNKQSLQPS
Kllybp1 (SEQ ID NO:44) QFSILTRLCLPREDPAFNPDRFTDSSLCPYEQAKCCCAILKDMMQH---ERKVPQKS
YlYbp1 (SEQ ID NO:40) AYTLCGLRLYLYQHESVPSAFIATDDACPFENAGYLAQVNGIKLRRMPIS---E1SDDLQ

CaYbp1 (SEQ ID NO:47) EKINISSSS-SNTPPKLPVRNG-IKRKHPSLTNSLNDLVLINNNSKNKAYEDNS---KIG
SsYbp1 (SEQ ID NO:45) EALANSEKTVPLPTLPPASSRASSRYPFTLTKARLVEDILALQVAQDSAPVTHESTVAI
CgYbp1 (SEQ ID NO:43) DTEDELSESAQLKINNEKRAKKT---FITLDPKRAEDLHAIKTLKETKKS---EM
Scvbp1 (SEQ ID NO:38) BTTDFKVEKLQVDSNKAQSNHVRQODSKMKAVHDCQLQTIQDSPTAD---A
ZrYbp1 (SEQ ID NO:46) NMNITKSDDKRRSTSTSEPPLP---RAYIDINERMDAASIAAMTFQDQKAK---
Kllybp1 (SEQ ID NO:44) DDELDLAKDMELKIKNSPPPLPS---RAYMLNDRDIATLHSITLAIACSAD---P
YlYbp1 (SEQ ID NO:40) SKLRIIPDVTRENGVEHQKASQTT---IPTTPEHVDIHKSLCNAALE---

FIG. 9D
<table>
<thead>
<tr>
<th>Protein</th>
<th>SEQ ID NO:</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CaYbp1</td>
<td>47</td>
<td>DPSKLSTIAAYLNLLVAI</td>
</tr>
<tr>
<td>SsYbp1</td>
<td>45</td>
<td>DPSKLSTI.SAYLNLLVI.</td>
</tr>
<tr>
<td>CgYbp1</td>
<td>43</td>
<td>KDYILLVLNYIKFFISTFAH</td>
</tr>
<tr>
<td>ScYbp1</td>
<td>38</td>
<td>KKSDILLILLYTLYNLIFIVL</td>
</tr>
<tr>
<td>ZrYbp1</td>
<td>46</td>
<td>GKDHIILINFLNFFNGLSQK</td>
</tr>
<tr>
<td>KLYbp1</td>
<td>44</td>
<td>ESKKVKTLLYTLYNFLNAFLT</td>
</tr>
<tr>
<td>YLYbp1</td>
<td>40</td>
<td>-QENTHVITWLNLPHSTVK-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein</th>
<th>SEQ ID NO:</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>CaYbp1</td>
<td>47</td>
<td>NAAGMLEBITIERFNE</td>
</tr>
<tr>
<td>SsYbp1</td>
<td>45</td>
<td>NAAGILEBITIDRIKS</td>
</tr>
<tr>
<td>CgYbp1</td>
<td>43</td>
<td>DANEKLRSLTE---------</td>
</tr>
<tr>
<td>ScYbp1</td>
<td>38</td>
<td>DKNESLNDYFTGIK-</td>
</tr>
<tr>
<td>ZrYbp1</td>
<td>46</td>
<td>IANETLGGKHL---------</td>
</tr>
<tr>
<td>KLYbp1</td>
<td>44</td>
<td>STVASISSKL---------</td>
</tr>
<tr>
<td>YLYbp1</td>
<td>40</td>
<td>LALDVIAKTV---------</td>
</tr>
</tbody>
</table>

* - Single, fully conserved residue
: - Conservation of strong groups
- - Conservation of weak groups
- - No consensus

FIG. 9E
FIG. 1

Gpx3

\[ \text{H}_2\text{O}_2 \]

\[ \text{H}_2\text{O} \]

YAP1 (reduced)

Yap1 (oxidized)