MIT Libraries | DSpace@MIT

MIT Open Access Articles

Functional overexpression and characterization of lipogenesis-related genes in the oleaginous yeast Yarrowia lipolytica

The MIT Faculty has made this article openly available. ***Please share*** how this access benefits you. Your story matters.

Citation: Silverman, Andrew M., Kangjian Qiao, Peng Xu, and Gregory Stephanopoulos. "Functional Overexpression and Characterization of Lipogenesis-Related Genes in the Oleaginous Yeast Yarrowia Lipolytica." Applied Microbiology and Biotechnology 100, no. 8 (February 26, 2016): 3781–3798.

As Published: http://dx.doi.org/10.1007/s00253-016-7376-0

Publisher: Springer Berlin Heidelberg

Persistent URL: http://hdl.handle.net/1721.1/107181

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike

Massachusetts Institute of Technology

Functional Overexpression and Characterization of Lipogenesis-Related Genes in the Oleaginous Yeast *Yarrowia lipolytica*

Andrew M. Silverman, Kangjian Qiao, Peng Xu, and Gregory Stephanopoulos#

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA

Keywords: lipogenesis, metabolic engineering, gene expression, fermentation

#Address correspondence to Gregory Stephanopoulos, gregstep@mit.edu

A.M.S. and K.Q. contributed equally to this work.

Abstract

Single Cell Oil (SCO) is an attractive energy source due to scalability, utilization of low-cost renewable feedstocks and type of product(s) made. Engineering strains capable of producing high lipid titers and yields is crucial to the economic viability of these processes. However, lipid synthesis in cells is a complex phenomenon subject to multiple layers of regulation, making gene target identification a challenging task. In this study, we aimed to identify genes in the oleaginous yeast *Yarrowia lipolytica* whose overexpression enhances lipid production by this organism. To this end, we examined the effect of the overexpression of a set of 44 native genes on lipid production in *Y. lipolytica*, including those involved in glycerolipid synthesis, fatty acid synthesis, central carbon metabolism, NADPH generation, regulation, and metabolite transport, and characterized each resulting strain's ability to produce lipids growing on both glucose and acetate as a sole carbon source.

Our results suggest that a diverse subset of genes was effective at individually influencing lipid production in *Y. lipolytica*, sometimes in a substrate-dependent manner. The most productive strain on glucose overexpressed the diacylglycerol acyltransferase DGA2 gene, increasing lipid titer, cellular content, and yield by 236%, 165%, and 246%, respectively, over our control strain. On acetate, our most productive strain overexpressed the acylglycerol-phosphate acyltransferase SLC1 gene, with a lipid titer, cellular content, and yield increase of 99%, 91%, and 151%, respectively, over the control strain. Aside from genes encoding enzymes that directly catalyze the reactions of lipid synthesis, other ways by which lipogenesis was increased in these cells include overexpressing the glycerol-3-phosphate dehydrogenase GPD1 gene to increase production of glycerol head groups, and overexpressing the 6-phosphogluconolactonase SOL3 gene from the oxidative pentose phosphate pathway to increase

NADPH availability for fatty acid synthesis. Taken together, our study demonstrates that the overall kinetics of microbial lipid synthesis is sensitive to a wide variety of factors. Fully optimizing a strain for single cell oil processes could involve manipulating and balancing many of these factors, and, due to mechanistic differences by which each gene product investigated here impacts lipid synthesis, there is a high likelihood that many of these genes will work synergistically to further increase lipid production when simultaneously overexpressed.

Introduction

Concerns over energy security and global climate change have necessitated the development of alternative sources of liquid fuels from renewable and domestic feedstocks (Asif and Muneer 2007; Solomon et al. 2009). Liquid biofuels such as biodiesel are particularly attractive for use in transportation due to their high energy density and compatibility with existing infrastructure (Demirbas 2007; Huber et al. 2006). Single Cell Oil (SCO) from fungi as a biodiesel precursor has many attractive features relative to other sources, as SCO processes allow for a smaller physical footprint and can utilize low-cost raw materials as feedstocks (Meng et al. 2009).

Oleaginous yeasts have been extensively studied as promising candidates for use in SCO processes, as they can readily accumulate lipids as a significant (over 20%) fraction of their biomass. Specifically, *Y. lipolytica* has emerged both as a convenient host for industrial processes and as a model organism for investigating lipid synthesis and accumulation in microbes and higher organisms. *Y. lipolytica* can achieve a lipid content of up to 36% (grams of lipids per gram of total dry cell weight, or DCW) in some laboratory settings (Beopoulos et al. 2009), with over 90% of those lipids stored in the form of triacylglycerols (TAGs). Additionally, *Y. lipolytica* is capable of metabolizing a wide variety of substrates in many industrially relevant contexts, including glucose from plant cellulose and hemicellulose hydrolysates (Papanikolaou et al. 2009; Stephanopoulos 2007), acetate and other volatile fatty acids (VFAs) from agro-industrial and municipal wastes (Fontanille et al. 2012), glycerol from industrial waste streams (including recycle streams from biodiesel manufacturing) (Papanikolaou et al. 2003; Papanikolaou and Aggelis 2002), and also very hydrophobic substrates such as fatty acids and TAGs derived from animal fat and alkanes from petroleum sludge (Papanikolaou et al. 2002; Fickers et al. 2005; Rahman et al. 2003). Conversion to biodiesel involves the transesterification

of the TAGs into fatty acid alkyl esters (Beopoulos et al. 2011), which can be performed in multiple ways (Folch et al 1957; Griffiths et al. 2010). *Y. lipolytica* is also the most genetically tractable oleaginous yeast, with a fully sequenced (albeit roughly annotated) genome (Beopoulos et al. 2009), and an assortment of genetic tools that have been developed to allow for insertion and overexpression of genes as well as genetic locus targeting to create gene knockouts (Nicaud et al. 1989; Fickers et al. 2003; Blazeck et al. 2011; Tai and Stephanopoulos 2013; Verbeke et al. 2013).

In this study, we sought to improve the production of biodiesel precursor lipids in *Y. lipolytica* when growing on two of the industrially relevant substrates mentioned above, glucose and acetate. The processes by which *Y. lipolytica* converts these substrates to lipids are very different, and summarized in **Figure 1**. In the case of glucose, lipogenesis in oleaginous organisms is triggered in response to the depletion in the medium of nitrogen required for normal growth and cell division, (Ratledge 2002), but also phosphate (Wu et al. 2010), sulfur (Wu et al. 2011), or certain metal ions required for proper enzyme function (Rossi et al. 2011). Upon nitrogen depletion, AMP deaminase becomes activated, converting AMP into IMP (inosine monophosphate) and producing ammonium ions. AMP is an allosteric activator of the isocitrate dehydrogenase complex in mitochondria, so that AMP depletion causes citrate and isocitrate to accumulate and exit from the mitochondrion to the cytosol where it is cleaved into acetyl-CoA and oxaloacetate (OxAc) by an enzyme called ATP:citrate lyase (ACL) (Ratledge 2002). The acetyl-CoA produced by ACL is then converted to malonyl-CoA by acetyl-CoA carboxylase (Acc1) in the first committed step of fatty acid synthesis (Wakil et al. 1983). Fatty acid synthesis begins with the binding of the two carbon acetyl unit of acetyl-CoA to a cysteine residue of the β-ketoacyl-ACP synthase domain of the fatty acid synthase (FAS) complex. Then the attached

carbon chain is elongated through successive rounds of simultaneous decarboxylation and condensation of malonyl-CoA with the growing carbon chain followed by reduction by 2 NADPH molecules to form a saturated acyl-ACP chain. In *Y. lipolytica*, the product of the FAS complex, palmitate (C16:0), is then shuttled to the endoplasmic reticulum (ER) to be directly incorporated into a lipid molecule in the Kennedy Pathway (Coleman and Lee 2004), or it can first undergo NADPH-dependent desaturation and/or a two-carbon elongation to form palmitoleate (C16:1), stearate (C18:0), oleate (C18:1), or linoleate (C18:2) (Xue et al. 2013). The supply of NADPH for fatty acid synthesis in oleaginous organisms can be provided by a variety of sources, including the oxidative pentose phosphate pathway (oxPPP), NADP+-dependent isocitrate dehydrogenase, or malic enzyme (Ratledge 2002). Previous research has determined that malic enzyme in *Y. lipolytica* is not responsible for producing appreciable amounts of NADPH for facilitating lipid synthesis (Zhang et al. 2013).

Biosynthesis of TAGs in the Kennedy pathway begins with a glycerol-3-phosphate acyltransferase (GPAT) catalyzing the transfer of a fatty acyl group from CoA to glycerol-*sn*-3-phosphate (G3P) to form 1-acylglycerol-*sn*-3-phosphate (lysophosphatidate or LPA) (Coleman and Lee 2004). Next, acylglycerol-phosphate acyltransferase (AGPAT or LPAT) catalyzes the acylation of the LPA at the second carbon to form 1,2-diacylglycerol-*sn*-3-phosphate (phosphatidate or PA). PA is then dephosphorylated to form diacylglycerol (DAG) by phosphatidate phosphatase (PAP or lipin). Lastly, an acyl group is transferred to the third carbon of DAG to form TAG by diacylglycerol acyltransferase (DGAT). DAGs can also be acted on by CDS to form CDP-diacylglycerol, which is the precursor to non-phosphatidate phospholipids, and fatty acids from phospholipids can be transferred to DAGs to form TAGs by phospholipid:diacylglyerol acyltransferase (or PDAT) enzymes. TAGs are stored in lipid bodies

(or lipid particles), which are bounded in membranes that are formed from many types of lipids, but mainly phospholipids, and these membranes must change in lipid composition to allow for increases in particle size as TAGs accumulate in the cell (Thiam et al. 2013).

The pathway for converting acetate to lipids differs significantly from that of glucose. Yeasts normally uptake acetate trough symport with protons at higher pH (Casal et al. 1996) or passive diffusion (as acetic acid) at lower pH (Augstein et al. 2003). Once in the cell, dissociated acetate can be converted by the acetyl-CoA synthetase (ACS2) enzyme to acetyl-CoA. The acetyl-CoA pool can be utilized in three ways: 1) carbon can go directly into fatty acid synthesis through Acc1, 2) the carbon can condense with OxAc in the mitochondrion to form citrate and be oxidized to carbon dioxide in the tricarboxylic acid (TCA) cycle to make energy (ATP) for the cell, or 3) the acetyl-CoA units can be converted in the glyoxylate cycle to succinate, which can be used downstream in anapleurotic, gluconeogenic, or other anabolic pathways (De la Peña Mattozzi et al. 2010).

Overexpressing genes involved in relevant metabolic pathways is one of the most common methods by which metabolic engineers can increase cellular production of a chemical of interest. Interest in engineering oleaginous organisms in this way to increase lipid production has grown in response to the recent development of genetic tools and rough genome annotations (Beopoulos et al. 2009). In *Mucor circinelloides,* overexpression of the native or a heterologous (from *Mortierella alpina*) malic enzyme increased the lipid content of the engineered cells by 2.5-fold over wild type (Zhang et al. 2007). In *Y. lipolytica*, overexpression of genes directly involved in lipid synthesis such as ACC1 and DGA1 conferred significant increases in the lipid titers (17.5% and 176.6%, respectively) and cellular oil content (roughly 2-fold and 4-fold, respectively) achieved in high nitrogen (atomic C:N = 20) fermentations (Tai and Stephanopoulos 2013).

Overexpressions of the native NAD+-dependent glycerol-3-phopshate dehydrogenase GPD1 (Dulermo and Nicaud 2011) and hexokinase HXK1 (Lazar et al. 2014) in *Y. lipolytica* were also found to situationally increase lipid production and accumulation depending on other genetic modifications in the strains when grown on glucose/oleic acid mixtures and fructose, respectively. Another study showed that, in addition to DGA1-overexpression increasing total lipid content by 3-fold, *Y. lipolytica* can be engineered to express heterologous fatty acid desaturases and elongases to produce significant titers of the important dietary supplement omega-3 eicosapentaenoic acid (EPA, C20:5), which is not naturally produced in these cells (Xue et al. 2013).

Regardless of carbon source, TAG production requires many different inputs, including fatty acids and glycerol head groups, reducing equivalents in NADPH, and energy (as ATP), and experiments on other oleaginous organisms strongly suggest that acetyl-CoA availability for flux through the fatty acid synthesis pathway is not necessarily the only major limiting factor for lipid production (Zhang et al. 2007). The reactions to produce these different inputs come from many different metabolic pathways that occur across several different cellular compartments and are subject to regulations of varying tightness and mechanism (Ratledge 2004; Seip et al. 2013). As *de novo* lipid production and storage is a complex phenomenon, here we attempt to gain a better understanding of the effects of many genes that are potentially important to this process by engineering several strains of *Y. lipolytica*, with each one overexpressing a single native enzyme with an additional copy under the strong and previously characterized TEF-1α promoter (Müller et al. 1998; Tai and Stephanopoulos 2013), and evaluating each strain's ability to produce lipids in fermentations of glucose and acetate. The genes tested here represent many different areas potentially important to lipid production, including the Kennedy pathway, fatty acid synthesis,

central carbon metabolism, NADPH generation, regulation, and metabolite transport. In addition to these groups, because some genes are poorly annotated or are unique to *Y. lipolytica* and similar oleaginous yeasts, we thought it potentially useful to include some genes from *Y. lipolytica* that are upregulated during lipogenesis, or that have high homology to those genes whose upregulation coincides with lipogenesis in other organisms, even if their functions are not entirely clear (Liu et al. 2009; Zhu et al. 2012; Schmidt et al. 2012; Morin et al. 2011; Liu et al. 2011). By overexpressing a single gene at a time, we aimed to isolate the influence of each one's specific expression on lipid production, although it is important to note that the lack of improvement in lipid production from a gene's overexpression does not necessarily prove that it is not somehow critical to the overall process. Only native genes were used in this study in order to avoid potential issues with codon optimization, or requirement of non-native protein folding chaperones or post-translational modifications for activity. We tested a total of 44 genes, summarized in **Table 1**, and compared their effect on fermentation performance to that of a control strain that expresses *E. coli* LacZ (beta-galactosidase) under the TEF-1α promoter; as LacZ has no substrate to act on in these fermentations, this strain is meant to approximate wild-type but with the additional burden of overexpressing a gene in a similar manner to our experimental strains (Tai and Stephanopoulos 2013). The genes in **Table 1** are not meant to comprise an exhaustive list of potential gene targets for enhancing lipid synthesis; some genes critically important to lipid synthesis were not included in this study, as they have already been sufficiently characterized elsewhere. Our results indicate that while many of the tested gene overexpressions were ineffective at influencing lipid synthesis in *Y. lipolytica* on both glucose and acetate, production was increased significantly in a diverse subset of our gene targets on at least one substrate. Taken together, the results in this study suggest that there is no single

limiting factor in lipid synthesis, and the kinetics of lipid production can be influenced by increasing the concentrations of many different metabolic intermediates and enzymes relevant to the process.

Materials and Methods

Strains, growth media, and fermentation conditions

All *Y. lipolytica* strains used in this study resulted from transformations of strain Po1g (Yeastern Biotech, Taipei, Taiwan), a Leu^- auxotrophic derivative of wild-type strain W29 (ATCC20460). All *Escherichia coli* strains developed in this study are plasmid carrying transformants of the cloning strain DH10β (New England Biolabs, Ipswich, MA). The *Y. lipolytica* and *E. coli* strains used in this study are listed in **Table S1**.

Standard growth conditions for *E. coli* for the purposes of plasmid construction and generation have been previously described (Sambrook and Russell 2001). Rich YPD medium for *Y. lipolytica* strains contains 10 g/L yeast extract (BD Biosciences, New Jersey), 20 g/L Bacto™ Peptone (BD), and 20 g/L glucose (Sigma-Aldrich, St. Louis, MO). Rich YPA medium for *Y. lipolytica* strains contains 10 g/L yeast extract, 20 g/L Bacto™ Peptone, and 20 g/L anhydrous sodium acetate (Avantor Performance Materials, Center Valley, PA). Yeast synthetic minimal medium containing glucose (YSMg) used for fermentations contains 1.7 g/L Yeast Nitrogen Base (YNB, without amino acids or ammonium sulfate) (Amresco, Solon, OH), 50 g/L glucose, and 1.84 g/L ammonium sulfate (Sigma) for an atomic C:N ratio of 60. Yeast synthetic minimal medium containing acetate (YSMa) used for fermentations contains 1.7 g/L Yeast Nitrogen Base (YNB, without amino acids or ammonium sulfate), 50 g/L anhydrous sodium acetate, and 2.02 g/L ammonium sulfate for an atomic C:N ratio of 40.

For fermentations, strains are inoculated into 3 mL rich (YPD or YPA) medium in a 14 mL polypropylene tube with a vented cap (VWR, Radnor, PA) and shaken for 24 hours at 200 rpm and 30°C for the first pre-culture. Yeast from the first pre-culture was then transferred into 14 mL tubes containing 3 mL of YSMg or YSMa to achieve an initial optical density (OD_{600}) of 0.05 and shaken for 36 hours at 200 rpm and 30°C for the second pre-culture. Yeast from the second pre-culture was then centrifuged and washed once with YSMg or YSMa medium and inoculated into 10 mL 24 well-plates (Axygen, Union City, CA) containing 3 mL of YSMg or YSMa per well to an initial OD_{600} of 0.05. Well plates were covered by AeraSeal™ hydrophobic sealing films (Excel Scientific, Victorville, CA) and placed in a Multitron Pro Infors incubator (Infors AG, Bottmingen, Switzerland) and shaken for 120 hours (for glucose fermentations) or 144 hours (for acetate fermentations) at 800 rpm, 30°C, and 75% humidity. Residual glucose or acetate in the medium at the end of the fermentation was measured by HPLC (Agilent, Santa Clara, CA), and dry cell weight measurements were carried out gravimetrically using 1.5 mL samples washed in sterile deionized water and dried at 60°C for 48 hours. Three biological replicate wells were used for each strain.

Plasmid and yeast strain construction and verification

The coding sequences of all genes tested in this study were integrated into plasmid pMT015, described previously (Tai and Stephanopoulos 2013). Plasmid pMT015 contains a cloning site that is flanked upstream by the strong *Y. lipolytica* Transcription Elongation Factor-1α (TEF-1α) promoter and associated intron and downstream by the XPR2 terminator. Overexpression plasmids were constructed using the Gibson Assembly method (Gibson et al. 2009) and Cloning Kit (New England Biolabs). Forward primers for polymerase chain reaction (PCR) amplification of coding sequences do not contain the first ATG codon (as an ATG already exists before the

intron in pMT015) and instead contain the bases TAACCGCAG (containing the branch point and 3' splice site of the TEF-1α intron) between the vector overlap site and the coding sequence binding site. The primers used to PCR amplify pMT015 and the coding sequence for construction of each plasmid are listed in **Table S2**.

Gibson Assembly products were transformed into DH10β using standard techniques (Sambrook and Russell 2001) and transformants were recovered and grown on selective Luria-Bertani medium containing carbenicillin. Confirmation of the correct coding sequence in frame with the promoter was carried out by PCR of the appropriate site followed by sequencing of purified plasmids obtained with the Qiagen Miniprep Kit (Qiagen, Valencia, CA).

For construction of *Y. lipolytica* strains, purified plasmids were digested with *Not*I or *Sac*II restriction enzymes (NEB) and transformed into yeast using the lithium acetate method (Chen et al. 1997). Yeast transformants were recovered on plates of YSM made with 15 g/L Bacto™ Agar (BD) and supplemented with 0.69 g/L CSM-Leu (Sunrise Science Products, San Diego, CA). Integration of the gene overexpression cassettes in each yeast strain was confirmed by PCR of genomic DNA purified with the Zymo YeaStar Genomic DNA Kit (Zymo Research, Irvine, CA).

RNA extraction and qRT-PCR Analysis for gene expression quantification

The methods for RNA extraction and gene expression analysis used in this study have been previously described (Tai and Stephanopoulos 2013), with some modification. Samples from 50 mL shake flask cultures of each strain (including the LacZ-expressing control strain) growing on YSMg media were harvested at 36 hours after inoculation and pelleted. The pellets were resuspended in 1 mL Trizol reagent (Life Technologies, New York), vortexed with acid-washed 425-600 μm glass beads (Sigma) for 20 minutes at 4°C, and centrifuged. 1 mL of the

supernatant was transferred to a new tube and 200 µL of chloroform (Sigma) was then added. The tubes were then briefly vortexed and centrifuged. 400 µL of the top aqueous phase of the resulting mixture was then added to 400 µL of acid phenol-chloroform-isoamyl alcohol (125:24:1) and briefly vortexed and centrifuged. 250 µL of the top phase of this mixture was then added to 1 mL of a cold 25:1 mixture of ethanol and 3 M sodium acetate (pH 5.2) in nuclease-free water to precipitate the RNA. After centrifugation, the RNA pellet is then washed twice in cold 70% ethanol, dried, and resuspended in nuclease-free water. All centrifugation steps were performed at 12,000 rcf for 15 minutes at 4°C.

Confirmation and quantification of overexpression of each gene was performed using the iScript One-Step RT-PCR Kit with SYBR Green (Bio-Rad, Hercules, CA) in 96-well plates with the Bio-Rad iCycler iQ Real-Time PCR Detection System. The RNA concentration in each well was approximately 400 pg/µL. Raw fluorescence data was then input into the Real Time PCR Miner program (Zhao and Fernald 2005) to obtain PCR amplification efficiencies and cycle threshold (Ct) values for each well. These results were then input into the REST 2009 program (Qiagen) to obtain fold overexpressions with confidence intervals for the upregulated gene in each strain relative to the LacZ control strain. The *Y. lipolytica* actin gene (YALI0D08272g) was used as a reference. Two replicate wells were analyzed for each gene and strain combination.

Lipid quantification: TAG derivatization and gas chromatograph analysis

Derivatization of fatty acids from lipids into fatty acid methyl esters (FAMEs) was achieved with a modified form of the direct transesterification method (Griffiths et al. 2010). Samples of 200 µL were centrifuged and the media aspirated. For the transesterification reaction, 100 µL of hexane containing known concentrations of methyl tridecanoate (C13-FAME) (Sigma) and

glyceryl triheptadecanoate (C17-TAG) (Sigma) standards was added to the pellet, followed by the addition of 500 µL of 0.5 N sodium hydroxide in methanol. The resulting mixture was vortexed at room temperature for one hour. Following transesterification, the reaction mixture is neutralized by the slow addition of 40 µL sulfuric acid (Sigma). To extract the FAMEs for GC analysis, 500 µL of hexane was then added and the mixture was vortexed at room temperature for 30 minutes. After extraction, the mixture is centrifuged at 8,000 rcf for one minute, and the top hydrophobic phase is transferred to a vial for Gas Chromatography (GC) analysis. Samples were analyzed in a Bruker 450-GC with a Flame Ionization Detector (FID) and an HP-INNOWax (30 m x 0.25 mm x 0.15 µm) GC column (Agilent). The GC method used helium as a mobile phase at a flow rate of 1.5 mL/min, an injector temperature of 260°C with a 10:1 split ratio, a column temperature held constant at 200°C, and an FID temperature of 260°C with gas flow rates of helium, hydrogen, and air of 25 mL/min, 30 mL/min, and 300 mL/min, respectively. The total concentration of FAMEs in the GC sample was calculated as the sum of concentrations of methyl palmitate (C16:0), methyl palmitoleate (16:1), methyl stearate (C18:0), methyl oleate (C18:1), and methyl linoleate (C18:2). Lipid titers of fermentation cultures calculated from GC results were normalized to the average C13:0 peak area and again using the maximum ratio of C17:0 peak area to C13:0 peak area for all samples derivatized and analyzed at the same time in order to control for errors introduced by solvent evaporation, differences in injection volume, and differences in transesterification efficiency. For each biological replicate in the fermentations, two lipid samples were analyzed.

Results

Verification and quantification of gene overexpression in all strains by qRT-PCR

Previous work in our lab characterized the full *Y. lipolytica* TEF-1α promoter and demonstrated its effectiveness in driving high expression levels of its target gene (Tai and Stephanopoulos 2013). We placed an additional copy of each of the 44 genes described in **Table 1** under the control of this promoter in their respective strains. After verifying the presence of our overexpression cassette in each strain's genomic DNA, we wanted to quantify the increase in expression of each gene relative to that of the LacZ-overexpressing control strain to confirm our gene cassettes were functional. RNA from each strain was harvested from glucose shake flask cultures at 36 hours and qRT-PCR analyses were performed. The resulting increases in expressions of each gene under control of TEF-1α are shown in **Figure 2**.

From **Figure 2** it is clear the expression platform has worked, as all genes studied are overexpressed in their respective strains. The magnitudes of overexpression observed here range from small but statistically significant fold-changes as seen in the IDP1 (2.1), REG1 (2.3), and FAS2 (2.6) overexpression strains, to larger fold-changes in expression as seen in the SLD (18.4), SOL3 (18.4), ABC2 (25.6), AMD1 (43.9), PSD (52.6), and YAS1 (71.0) overexpression strains. This data could potentially indicate that genes from the former group are already highly expressed in *Y. lipolytica* under the specific conditions tested (growth phase in glucose minimal media), while genes from the latter group are not normally as highly expressed.

Effects of gene overexpression on fermentation performance

After confirming that the overexpression cassettes were functional, we performed two sets of fermentations for all strains: one on a glucose minimal medium, and one on an acetate minimal medium. Fermentations were carried out in 24-well plates with a reaction volume of 3 mL per well. Lipid production of each strain was assessed in terms of lipid titer (grams/liter lipid in the medium, a product of specific cellular production and growth), cellular lipid content (grams

lipids per gram dry cell weight), and lipid yield (grams lipids per gram of carbon substrate consumed). Our LacZ control strain produced an average lipid titer, content, and yield on glucose of 0.71 g/L, 0.21 g lipid/g DCW, and 0.042 g lipid/g glucose consumed, and on acetate, produced an average titer, content, and yield of 0.42 g/L, 0.23 g lipid/g DCW, and 0.053 g lipid/g acetate consumed.

<u>Kennedy Pathway/Glycerolipid Synthesis Genes:</u> Fermentation results for strains overexpressing Kennedy Pathway and glycerolipid synthesis related genes are shown in **Figure 3**. On glucose, the largest increases in lipid production over the LacZ control strain were achieved by overexpressing the DGA2 gene (increases in titer of 236%, in content of 165%, and in yield of 246%), followed by the SLC1 gene (86%, 73%, and 87%, respectively). These two gene products catalyze the third (final) and second acyltransferase reaction, respectively, involved in TAG synthesis. The effect of SLC1 overexpression when growing on acetate (increase in t/c/y of 99%, 91%, and 154%) was consistent with the results from glucose, however the DGA2 overexpression strain showed very poor performance in acetate fermentations. A microscopic evaluation revealed that the DGA2 strain underwent a transition to a hyphal morphological state, a stress response characterized by long, filamentous cells with generally slower growth rate and low capacity for lipid storage. Hyphal growth was also seen in the PDR16 overexpression strain when grown on acetate.

On glucose, the LRO1 and GPA gene overexpression strains showed small but significant increases in lipid titer (19% and 26% respectively) and yield (19% and 28%, respectively), and while both had higher average lipid contents than the LacZ strain, statistical significance was not achieved for the lipid content of the LRO1 strain. Interestingly, the lipase/esterase TGL4 (TAG lipase) and TGL1 (steryl ester hydrolase) overexpressions increased lipid production

significantly, with titer increases of 29% and 27%, respectively, yield increases of 45% and 29%, respectively, and a lipid content increase of 19% for TGL1 overexpressing cells. The lipid content increase for TGL4 overexpressing cells was not significant. None of these genes showed significant improvement in lipid production on acetate.

In contrast to the results above, there were two Kennedy pathway/glycerolipid synthesis genes that showed greater lipid production on acetate but not glucose. The PSD (phosphatidylserine decarboxylase) strain showed increases in titer and yield (39% and 45%, respectively) compared to the LacZ control strain. In addition to PSD, overexpression of the CSR1 gene (determined to be so by homology to the *S. cerevisiae* CSR1 gene), coding for a phosphatidylinositol transfer protein (PITP), led to a 69% increase in lipid titer and a large increase in biomass (see **Table S3**) over the control strain, but increases in content or yield were not statistically significant.

Central Carbon Metabolism Genes: From the results in **Figure 4**, the only gene in this group that promoted lipogenesis when overexpressed was GPD1, or glycerol-3-phosphate dehydrogenase. GPD1 overexpression caused an increase in titer, content, and yield of 40%, 31%, and 55%, respectively, over the control strain on glucose and also an increase in titer and yield of 30% and 38%, respectively, on acetate, with a non-significant increase in lipid content. These results demonstrate that the availability of glycerol-3-phosphate in the cell is important contributor to the overall kinetics of lipid synthesis, with glycerol-3-P concentrations in the cell being normally suboptimal for achieving the highest possible rates of lipid synthesis.

Fatty Acid Synthesis Genes: **Figure 5** shows that none of the genes in the fatty acid synthesis group were effective at increasing lipid production on glucose when overexpressed. On acetate, only the FAS1 overexpressing strain produced significantly more lipids, giving a 33% increase in

titer, but no significant increase in content or yield. Higher dry cell weight measurements (as shown in **Table S3**) in this strain compared to the control suggest that FAS1 overexpression can contribute to a higher growth rate on acetate.

NADPH Generating Genes: Three potential sources of NADPH that could contribute to fatty acid synthesis were tested in this study: the oxidative pentose phosphate pathway (oxPPP), a cytosolic NADP+-specific isocitrate dehydrogenase, and an NADP+-specific glutamate dehydrogenase. Malic enzyme activity is a major source of NADPH and a limiting factor for lipid synthesis in many oleaginous fungi (Zhang et al. 2007), but previous work of ours (data not shown) and others (Zhang et al. 2013) showed no effect of overexpression of the only native malic enzyme isoform, MAE1, on lipid synthesis in *Y. lipolytica*. It has been suggested elsewhere (Zhang et al. 2013) that MAE1 in *Y. lipolytica* may be mostly specific for NAD+ and not NADP+. Results for fermentations of strains overexpressing the other potentially important NADPH-producing genes are given in **Figure 6**.

The only gene of this group to significantly increase lipogenesis when overexpressed was SOL3 (6-phosphogluconolactonase), catalyzing the intermediate step of the oxPPP. SOL3 overexpression increased the lipid titer, content, and yield on glucose by 23%, 18%, and 27% (all statistically significant) over the control strain. SOL3 does not produce NADPH itself, but the steps in the oxPPP immediately before and after do, which are catalyzed by ZWF1 and GND1 gene products. On acetate, SOL3 overexpression greatly increased the lipid content (97% over control), but caused a severe reduction in growth rate, resulting in lower overall titer, and suggesting that on acetate, SOL3 expression levels can cause a diversion of too much carbon flux towards synthesis of lipids and away from non-lipid biomass. Curiously, neither NADPH-producing gene in the oxPPP significantly increased lipid production when overexpressed.

Regulators of Lipid Synthesis Genes: We tested six genes that are involved in regulating enzymes involved in fatty acid and lipid synthesis. The results from these fermentations are shown in **Figure 7**. The two strains in this group that did show some increase in lipogenesis were the YAS1 and YAS2 overexpression strains, but only on acetate. Increases in lipid production were relatively mild in the YAS1 strain (14% in titer and 22% in yield, respectively), and were more significant in the YAS2 overexpressing strain (24% in titer, 24% in content, and 47% in yield). None of the SAP190, TOR2, REG1, or GLC7 genes contributed to increased lipid synthesis when overexpressed.

Transporters and Miscellaneous Genes: The genes in this group included five genes (SLY1, ABC2, PSR1, SKY, and YIP3) that are the closest *Y. lipolytica* homologues to genes that are upregulated in the oleaginous yeast *Rhodosporidium toruloides* during lipogenesis (Liu et al. 2009; Zhu et al. 2012) as well as the AMD1 and SCP2 genes. From the results of the fermentations of strains overexpressing these genes in **Figure 8**, none of the seven genes in this group increase lipid synthesis on glucose when overexpressed. Two genes in this group, PSR1 and YIP3, conferred a strongly lipogenic phenotype to their overexpression strains when grown on acetate. The strains overexpressing these genes had increases in lipid content of 51% and 46% over the control strain, respectively, but otherwise grew poorly, suggesting that a high fatty acid flux in these strains prevented the normal development of biomass. PSR1 likely codes for a transmembrane protein phosphatase involved in stress responses, and YIP3 likely codes for a Rab-GDI dissociation factor which is involved in the maintenance of intracellular membranes; aside from their apparent upregulation coinciding with the onset of lipogenesis in *R. toruloides*, their roles in driving lipogenesis in *Y. lipolytica* when growing on acetate are very much unclear. The other three genes in this group whose transcription correlates to lipogenesis in this study did

not increase lipogenesis in *Y. lipolytica* when overexpressed and therefore do not demonstrate a causal link between expression and lipid synthesis within the ranges of expression tested in this study.

Effects of gene overexpression on fatty acid composition

In addition to observing the impact of their overexpression on total fatty acid titer, content, and yield, we were also interested in assessing if some of the native genes we overexpressed in *Y. lipolytica* affected their respective strains' fatty acid profiles, or the distributions of chain lengths and degrees of unsaturation among the fatty acids produced by the cells. The results of that analysis performed on the glucose fermentation data for a select few fatty acid acyltransferase and desaturase genes are shown in **Figure 9**.

Our results show that, on glucose, *Y. lipolytica* normally produces a lipid profile highest in oleate (C18:1, 48.8% by weight of total fatty acid concentration), lower in palmitate (C16:0, 21.1%) and stearate (C18:0, 15.4%), and much lower in linoleate (C18:2, 10.7%) and palmitoleate (C16:1, 4.0%). Our DGA2-overexpression strain showed little difference in fatty acid profile, with the only large change being a 41% reduction in the fraction of C18:2 compared to the control, which was accounted for by slight increases in C16:0, C16:1, and C18:0. It is possible that this could be explained by Dga2 having no fatty acid preference, but rather it simply outcompetes the elongases and desaturases for fatty acids before they can become the heavily modified C18:2 molecules.

Overexpression of LRO1, a PDAT, increased the C16:1 fraction in its overexpression strain by 105%, and overexpression of SLC1, an AGPAT, increased C18:0 fraction by 38%. TGL1 overexpression caused small decreases in saturated fatty acid fraction and increases in monounsaturated fatty acid fraction, including an 85% increase in C16:1 content. These results

could indicate significant preferences of these enzymes towards specific substrates. While the percentage changes of C16:1 fraction in some of these strains are striking, it is important to note the normally low baseline fraction of C16:1 in these cells, and even in the strain with the highest C16:1 content, the LRO1 overexpressing strain, C16:1 is still a very minor product (8.3% of total FA).

The most dramatic change in fatty acid profile did not come from overexpressing an acyltransferase gene, but rather it came from a more obvious desaturase gene. The Fad2 enzyme is a Δ12-desaturase, which in *Y. lipolytica* catalyzes the conversion of C18:1 to C18:2. Although FAD2 overexpression did not affect the total amount of lipids produced (see **Figure 5**), a very large increase (180%) in C18:2 fraction and a commensurate decrease in C18:1 fraction was observed, to the point where the two are produced in roughly equal amounts in the FAD2 overexpression strain. This finding demonstrates that enzymes that directly modify fatty acids themselves can more strongly determine a strain's profile of fatty acids than the preferences of the acyltransferases that act on them after modification.

Discussion

In this study, we constructed 44 strains of *Y. lipolytica* that each overproduce a single gene product that either has activity in directly catalyzing the reactions of lipid synthesis or may be involved in increasing the activity of those enzymes through regulation or enhancing the availabilities of their substrates or cofactors. We first demonstrated that our TEF-1α promoter expression platform worked to significantly enhance expression of each strain's respective gene. In addition to confirming overexpression, it is possible that the magnitude of overexpression of each strain's respective gene in our study can provide information as to the baseline expression level of that gene in wild-type *Y. lipolytica* during growth phase on glucose. Strong inverse

correlations have previously been observed between fold overexpression and baseline expression level when comparing several genes overexpressed in the same manner in other organisms, through ectopic inductions of transcription factors in *M. tuberculosis* (Rustad et al. 2014) and through transcriptional activator-bound Cas9-mediated upregulation of genes in human 293FT cells (Konermann et al. 2014). The highly overexpressed genes in our study could therefore belong to pathways that are ordinarily not highly active during growth phase on glucose. Alternatively, the native promoters of these genes may not be sensitive to the overabundance of their mRNA product.

Some gene products in this study are involved in either mainly linear pathways or are subunits of a complex with a specific stoichiometric makeup of proteins. We noted the apparent possible differences in baseline expression level for some of these genes relative to the others in their respective pathways or complexes. This difference can be seen in the effect of TEF-1α overexpression on SOL3 (18.4-fold increase), the middle step in the linear oxidative pentose phosphate pathway, compared to both ZWF1 (6.7-fold increase) and GND1 (5.9-fold increase), which are the steps before and after SOL3, respectively. FAS1 (5.1-fold increase) also may have a lower baseline expression level than FAS2 (2.6-fold increase), despite both enzymes being present in equal stoichiometric amounts in the fatty acid synthase complex. In both cases, the genes with higher-fold overexpressions in our study could be differentially transcriptionally regulated compared to the rest of the genes in their pathways and therefore transcript levels of these specific genes could be important in controlling the flux through these pathways.

Our results demonstrated that a diverse subset of the genes tested here can confer greater lipid production onto their overexpression strains. The most dramatic increase in lipid synthesis from glucose came from overexpressing the DGA2 gene, and our largest improvements on

acetate came from the SLC1 overexpressing strain. As these genes are directly involved in the process of TAG synthesis, these results are not surprising. Previously, the overexpression of the other diacylglycerol acyltransferase gene (DGA1, or YALI0E32769g) has been shown to be similarly effective at achieving large increases in lipid production (Tai and Stephanopoulos 2013; Xue et al. 2013), and deletion of any DGA gene significantly decreased lipid content (by 43% for DGA2) in *Y. lipolytica* (Zhang et al. 2012). Two possible explanations for this include overexpressed DGA outcompeting CDS for DAG, which is the first step in the synthesis of phospholipids required for plasma membranes (Tai and Stephanopoulos 2013; Courchesne et al. 2009), or that creating a "sink" for fatty acids in general stimulates their production further (Tai and Stephanopoulos 2013; Thelen and Ohlrogge 2002). While the effects of DGA2 overexpression can be explained by either or both of these hypotheses, the effects of SLC1 overexpression may only be explained by the latter, as SLC1 acts prior to DAG synthesis. One study in which a yeast SLC1 was expressed in plants found the transgenic plants had up to a 48% increased oil content in their seeds (Zou et al. 1997), further giving support to this theory.

The cause of the observed hyphal growth of the DGA2 strain on acetate is unclear; it is possible that, due to the proximity of acetate to the fatty acid synthesis pathway, cells overexpressing DGA2 could divert too much flux away from pathways that would generate biomass or energy. However, *Y. lipolytica* cells overexpressing the DGA1 isoform have shown large increases in lipid production on acetate similar to those on glucose (Tai 2012). The reason for the PDR16 strain's hyphal growth may not the same as the reason for the DGA2 strain's hyphal growth, as PDR16 overexpression did not cause increased lipogenesis on glucose.

Other genes directly involved in glycerolipid synthesis had more modest effects. LRO1 codes for an enzyme that catalyzes the transfer of fatty acids from phospholipids to DAGs to

form TAGs, but is mainly active in yeast during the exponential growth phases, and not during the (non-dividing) lipid accumulation phase (Müllner and Daum 2004), suggesting that this pathway is likely only a minor contributor to overall TAG synthesis in *Y. lipolytica*. It is possible that GPA overexpression increases lipid synthesis on glucose by the same mechanism as SLC1, although to a much smaller extent. Regarding TGL4 and TGL1, although these genes exhibit catabolic activity towards neutral lipids like TAGs and steryl esters (SE), TGL4 in *Saccharomyces cerevisiae* is a multifunctional enzyme, also exhibiting AGPAT activity (as well as steryl ester hydrolase activity like TGL1), and it is possible that TGL1 is similarly multifunctional. Expression of the *S. cerevisiae* TGL4 gene into *P. pastoris* was shown to result in an increase in phospholipid biosynthesis (Rajakumari and Daum 2010). Contributions to lipid production from TGL4 overexpression in *Y. lipolytica* therefore may come mostly from increasing the growth rate by providing more phospholipids for membrane synthesis, which also explains the lack of a statistically significant increase in the lipid content of this strain; indeed, we observed significant increases in total biomass generated by the end of the fermentation (see **Table S3**). As PSD is a key enzyme in the biosynthesis of phospholipids like phosphatidylethanolamine and phosphatidylcholine, overexpression of this gene also likely contributes an increase in phospholipid synthesis. Just like in the TGL4 and TGL1 overexpression strains, a small but significant increase in total biomass was observed in the PSD strain, providing more evidence of a causal link between phospholipid biosynthesis and growth rate. To explain the increase in titer and biomass in the CSR1 strain without a consistently large increase in yield, it is possible that TCA cycle flux is higher in these cells.

Our data indicated that lipid production in *Y. lipolytica* was positively influenced by increasing glycerol-3-phosphate availability. From our results, we can speculate as to the form

taken by the additional lipids produced by the GPD1 overexpression strain; the lipid content increased when growing on glucose, suggesting additional lipids in the form of TAGs, while for acetate, the lack of an increase in lipid content combined with an increase in biomass formed (see **Table S3**) suggests that the additional lipids were directed more towards phospholipid synthesis. Aside from glycerol-3-P availability promoting lipid synthesis, another study indicates that higher glycerol-3-P availability achieved by knocking out the mitochondrial glycerol-3-P dehydrogenase GUT2, which mainly catalyzes the reverse reaction (glycerol-3-phosphate to dihydroxyacetone phosphate), prevents the β-oxidation of fatty acids accumulated when *Y. lipolytica* is grown on oleate (Beopoulos et al. 2008). From another study where the oil content in the seeds of *B. napus* was increased by expressing a yeast GPD1, it was hypothesized that glycerol-3-P availability may be an important signal for upregulation of fatty acid synthesis genes in some organisms (Vigeolas et al. 2007).

Neither ACS2 nor PYC1 overexpression contributed significantly lipid synthesis. Although it has been speculated that ACS2 overexpression can contribute to lipid synthesis (Courchesne et al. 2009) because it was found that overexpressing ACS in *E. coli* led to an increase in the rate of acetate utilization (Lin et al. 2006), in our study, we found that ACS2 overexpression only increased biomass significantly (by 49%, see **Table S3**) on acetate. Lipid production in our ACS2 overexpressing strain was too inconsistent to achieve statistical significance in the results, suggesting that acetyl-CoA availability by itself does not drive lipid synthesis in a monotonic manner. The reaction catalyzed by the PYC1 gene product (pyruvate to oxaloacetate) is implicated in lipogenesis and gluconeogenesis in humans, and is a component of the NADPH-producing transhydrogenase (malate-pyruvate-oxaloacetate) cycle (Mackall and Lane 1977), however, the lack of contribution of malic enzyme to NADPH production and lipid synthesis in

Y. lipolytica (Zhang et al. 2013) likely explains the ineffectiveness of overexpressing this gene in increasing lipid production in our strain.

With respect to the genes of the FAS complex proteins and the overexpression of FAS1 leading to higher lipid titers on acetate, it is interesting to note that the fold change of FAS1 overexpression in its respective strain compared to the control shown in **Figure 2** was approximately twice that of FAS2 overexpression in its strain (5.1 vs. 2.6), which could be evidence of a lower baseline of expression of FAS1 in wild-type *Y. lipolytica*, although it is important to note that our gene overexpression quantification experiment was performed when the strains were growing on glucose and not acetate. As the FAS complex contains equal amounts of FAS1 and FAS2 subunits, it is possible that FAS1 is not produced in stoichiometric amounts needed to complex with every copy of FAS2 that is translated; excess copies of FAS subunits are rapidly degraded *in vitro* (Schweizer and Hofmann 2004). In *S. cerevisiae*, the FAS1 and FAS2 genes contain different numbers of both constitutive general transcription factor binding sites and inositol/choline response elements (ICRE), which are bound by the INO2/INO4 complex to upregulate phospholipid synthesis genes under conditions of low inositol and choline (Schweizer and Hofmann 2004). Research on *S. cerevisiae* FAS genes also shows that increasing the expression level of FAS1 can influence the expression level of FAS2 by interfering with the function of a "downstream repressive element" contained in the FAS2 gene, while increasing FAS2 expression level does not have the same effect on FAS1 (Wenz et al. 2001).

To help explain the curious result of SOL3 overexpression affecting lipogenesis while that of either of the two NADPH-producing gene products does not, although the ZWF1 gene product catalyzes the first committed step of the oxPPP, the intermediate it produces, 6-

phosphogluconolactone, is highly reactive and thought to be toxic to the cells if accumulated (Miclet et al. 2001). In addition, our gene expression analysis results (**Figure 2**) showed that fold change in respective gene expression compared to control in the SOL3 strain was approximately three times higher than those of the ZWF1 and GND1 strains, which could suggest that SOL3 has a lower baseline of expression and its transcription is a limiting factor for oxPPP activity. SOL3 overexpression therefore would allow for increased flux through the oxPPP without causing side-effects detrimental to growth.

Recent metabolic flux analysis (MFA) studies on *Y. lipolytica* reported that a high-oil producing ACC1- and DGA1-overexpressing strain has a greater than two-fold higher oxPPP flux than a lower oil wild-type strain when both are grown on glucose during their lipid-production (nitrogen-limited) phases, and oxPPP flux potentially accounts for all NADPH generation consumed by fatty acid synthesis (Wasylenko et al. 2015). Similar studies also suggest that oxPPP flux in the high-oil strain can potentially produce all of the NADPH for fatty acid synthesis when grown on acetate, as well [unpublished results from our lab]. Because all other potential NADPH-producing pathways did not increase lipid production, then between our results and the MFA studies, we can conclude that the oxidative pentose phosphate pathway is the only major source of NADPH for *Y. lipolytica* growing on both glucose and acetate, and other native enzymes are largely ineffective at producing NADPH in significant quantities for use in lipid synthesis.

Aside from those coding for the subunits of the Yas1/Yas2 complex, the genes coding for regulators of fatty acid and lipid synthesis enzymes in this study were largely ineffective at driving lipogenesis in *Y. lipolytica* when overexpressed. Both Sap190 and the Reg1/Glc7 complex are involved in, among other things, dephosphorylating and inactivating the Snf1

protein, which is a homologue of AMPK in mammals and can inactivate Acc1 (Ruiz et al 2011). The Tor2 protein is a master regulator that can activate lipid production in response to nutrient availability with a signal cascade that ultimately ends in inactivating Snf1 (Raught et al. 2001). From our results, it is clear that higher transcription levels of these genes alone cannot cause the regulatory changes needed to increase activity of lipid synthesis related genes, and it is likely that post-translational modifications to these gene products or allosteric activators are critical for their functionality.

While the Yas1/Yas2 complex has been found to be involved in upregulating genes for alkane metabolism (Endoh-Yamagami et al. 2007), these two proteins are also the closest homologues in *Y. lipolytica* to the Ino2/Ino4 complex proteins from *S. cerevisiae*, and may transcriptionally upregulate phospholipid (Hirakawa et al. 2009) and fatty acid synthesis (Schweizer and Hofmann 2004) genes when concentrations of intracellular inositol are low. Inositol drives the repression of the Ino2/Ino4 complex by Opi1 (closest homologue in *Y. lipolytica*: YALI0C14784p), and ICRE-containing genes are significantly upregulated in *S. cerevisiae* ΔOPI1 mutants (Henry et al. 2012). Our results suggest that overproducing the enzymes of this complex does not allow for its complete repression, and inositol may be present in lower intracellular concentrations when *Y. lipolytica* is grown on acetate as opposed to glucose, allowing for differences in the effect of their overproduction between both substrates tested.

Amd1, or AMP deaminase, is thought to be the enzyme responsible for beginning the process of neutral lipid production and accumulation in *Y. lipolytica* on glucose by removing the AMP needed for TCA cycle activity (Ratledge 2002), but overexpression of this enzyme was not effective at promoting lipid synthesis beyond wild-type levels. It has been shown that the

products of the reaction catalyzed by yeast Amd1, IMP and NH3, strongly inhibit Amd1 activity, the latter of which is present in significant amounts in the growth phase; AMD1 is also already natively constitutively expressed in yeast (Merkler et al. 1989). This suggests that the intracellular environment is critical to Amd1 activity, while the exact transcription level of the gene is not.

The titers and yields of lipids are not the only important outputs to a biodiesel manufacturing process; the specific fatty acid profile has an important effect on the fuel's performance and stability. The cetane number of a fuel, a measure of ignition delay and quality of combustion, is higher in biodiesel with longer chains and lower degrees of unsaturation (Ramos et al. 2009). Higher degrees of unsaturation make also biodiesel more susceptible to oxidation (Ramos et al. 2009). However, the lubricity of the fuel is improved with higher degrees of unsaturation (Knothe 2005) and the cold-flow properties improve with higher unsaturation and shorter chain length (Dunn 2011). Therefore, the optimal fatty acid composition for biodiesel may be context dependent (Hoekman et al 2012).

It has been observed that acyltransferase enzymes have different preferences (including the lack thereof) for different fatty acids to incorporate into lipids, even among close homologues from different species. For example, the Dga2 and Lro1 enzymes in *S. cerevisiae* demonstrate a preference for unsaturated fatty acids, especially C18:1 (Grillitsch et al. 2011), while the Dga2 enzyme in *P. pastoris* showed strong preference to unsaturated C16:0 and C18:0, and the Lro1 enzyme showed no preference at all (Ivashov et al. 2013). *Y. lipolytica* Dga1, when singly overproduced, showed a significant preference for C18:0 above all other fatty acids. It has also been observed that increasing overall fatty acid availability by overexpressing ACC1 can have an impact on the fatty acid profile (Tai and Stephanopoulos 2013). While there were some

noticeable differences in this study in fatty acid fractions between acyltransferase-overexpressing strains and the control LacZ-expressing strain, we observed that overexpressing the oleoyl-CoA Δ12-desaturase FAD2 gene led to the most dramatic shift in fatty acid profile, causing the conversion of a large fraction of C18:1 to C18:2 in the resulting cells. Due to its relative reactivity and the inverse correlation between cetane number and unsaturation degree, it is unlikely that biodiesel made from this strain alone would be preferable to that made from any of the other strains mentioned here. However, native production of linoleic acid in *Y. lipolytica* has been used as a starting point for synthesis of other fatty acids like the diet supplement EPA (Xue et al. 2013), so this finding could potentially be very useful for other applications.

In conclusion, lipid synthesis is a complex phenomenon, requiring not only carbon in the form of acetyl-CoA for fatty acids, but also other inputs like energy, NADPH, and glycerol head groups. The many enzymes involved in converting a substrate to lipids act in many different cellular compartments, and are subject to many layers of transcriptional, post-translational, and feedback inhibition control. In our study, lipid synthesis was increased in response to not only greater concentrations of directly related enzymes (either through their direct overexpression or upregulation by overexpressed transcriptional activators) and availability of fatty acids, but also greater availability of other substrates in those reactions such as glycerol head groups (from GPD1) and NADPH (from the oxidative pentose phosphate pathway), and the effects are sometimes substrate-dependent. Hence, despite being extensively regulated, the overall kinetics of lipid synthesis are still sensitive to a variety of factors, with changes in each one individually having influences of varying magnitude. In addition to the total amount of lipids produced, we also demonstrate here that the specific chain length and degree of unsaturation of the biodiesel produced by these SCO processes can be readily altered.

Although the overexpressions of many genes tested here had relatively modest effects, their potential importance should not be discounted. Due to the vast differences in these genes' respective modes of action in regards to influencing lipid synthesis, overexpressions of multiple genes with non-redundant functions will likely work synergistically to further improve the process when simultaneously expressed (Tai and Stephanopoulos 2013). Because commodities like fuels generally have high trading volumes and demand and low profit margins, even small increases in the yield of a biodiesel production process can have large effects on profitability and could mean the difference between becoming economically viable or not in competition with traditional, non-renewable energy sources (Hill et al. 2006).

Acknowledgements: This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science program, under Award Number DE- SC0008744.

The authors declare that they have no conflict of interest. This article does not contain any studies with human participants or animals performed by any of the authors.

References

Asif M., Muneer T. (2007). Energy supply, its demand and security issues for developed and emerging economies. Renew Sust Energ Rev, 11(7), 1388–1413. doi:10.1016/j.rser.2005.12.004

Augstein, A., Barth, K., Gentsch, M., Kohlwein, S. D., & Barth, G. (2003). Characterization, localization and functional analysis of Gpr1p, a protein affecting sensitivity to acetic acid in the yeast *Yarrowia lipolytica*. Microbiology, 149(3), 589–600. doi:10.1099/mic.0.25917-0

Beopoulos, A., Mrozova, Z., Thevenieau, F., Le Dall, M.-T., Hapala, I., Papanikolaou, S., Chardot, T., Nicaud, J.-M. (2008). Control of lipid accumulation in the yeast *Yarrowia lipolytica*. Appl Environ Microb, 74(24), 7779–7789. doi:10.1128/AEM.01412-08

Beopoulos, A., Cescut, J., Haddouche, R., Uribelarrea, J.-L., Molina-Jouve, C., & Nicaud, J.-M. (2009). *Yarrowia lipolytica* as a model for bio-oil production. Prog Lipid Res, 48(6), 375–87. doi:10.1016/j.plipres.2009.08.005

Beopoulos, A., Chardot, T., & Nicaud, J.-M. (2009). *Yarrowia lipolytica*: A model and a tool to understand the mechanisms implicated in lipid accumulation. Biochimie, 91(6), 692–6. doi:10.1016/j.biochi.2009.02.004

Beopoulos, A., Nicaud, J. M., & Gaillardin, C. (2011). An overview of lipid metabolism in yeasts and its impact on biotechnological processes. Appl Microbiol Biot, 90(4), 1193-1206.

Blazeck, J., Liu, L., Redden, H., & Alper, H. (2011). Tuning gene expression in *Yarrowia lipolytica* by a hybrid promoter approach. Appl Environ Microb, 77(22), 7905–14. doi:10.1128/AEM.05763-11

Casal, M., Cardoso, H., & Leao, C. (1996). Mechanisms regulating the transport of acetic acid in *Saccharomyces cerevisiae*. Microbiology, (142), 1385–1390.

Chen, D. C., Beckerich, J. M., & Gaillardin, C. (1997). One-step transformation of the dimorphic yeast *Yarrowia lipolytica*. Appl Microbiol Biot, 48, 232–235. doi:10.1007/s002530051043

Coleman, R. A., & Lee, D. P. (2004). Enzymes of triacylglycerol synthesis and their regulation. Prog Lipid Res, 43(2), 134–176. doi:10.1016/S0163-7827(03)00051-1

Courchesne, N. M. D., Parisien, A., Wang, B., & Lan, C. Q. (2009). Enhancement of lipid production using biochemical, genetic and transcription factor engineering approaches. J Biotechnol, 141(1-2), 31–41. doi:10.1016/j.jbiotec.2009.02.018

De la Peña Mattozzi, M., Kang, Y., & Keasling, J. D. (2010). Feast: Choking on Acetyl-CoA, the Glyoxylate Shunt, and Acetyl-CoA-Driven Metabolism. Handb Hydrocarb Lipid Microbiol, 1649–1660. doi:10.1007/978-3-540-77587-4

Demirbas, A. (2007). Importance of biodiesel as transportation fuel. Energ Policy, 35(9), 4661–4670. doi:10.1016/j.enpol.2007.04.003

Dulermo, T., & Nicaud, J.-M. (2011). Involvement of the G3P shuttle and β-oxidation pathway in the control of TAG synthesis and lipid accumulation in *Yarrowia lipolytica*. Metab Eng, 13(5), 482–91. doi:10.1016/j.ymben.2011.05.002

Dunn, R. O. (2011). Improving the Cold Flow Properties of Biodiesel by Fractionation. Soybean - Application and Technology, 975(2). doi:10.5772/14624

Endoh-Yamagami, S., Hirakawa, K., Morioka, D., Fukuda, R., & Ohta, A. (2007). Basic helix-loop-helix transcription factor heterocomplex of Yas1p and Yas2p regulates cytochrome P450 expression in response to alkanes in the yeast *Yarrowia lipolytica*. Eukaryot Cell, 6(4), 734–743. doi:10.1128/EC.00412-06

Fickers, P., Le Dall, M. T., Gaillardin, C., Thonart, P., & Nicaud, J. M. (2003). New disruption cassettes for rapid gene disruption and marker rescue in the yeast *Yarrowia lipolytica*. J Microbiol Meth, 55(3), 727–737. doi:10.1016/j.mimet.2003.07.003

Fickers, P., Benetti, P.-H., Waché, Y., Marty, a, Mauersberger, S., Smit, M. S., & Nicaud, J.-M. (2005). Hydrophobic substrate utilization by the yeast *Yarrowia lipolytica*, and its potential applications. FEMS Yeast Res, 5(6-7), 527–43. doi:10.1016/j.femsyr.2004.09.004

Folch, J., Lees, M., & Sloane-Stanley, G. H. (1957). A simple method for the isolation and purification of total lipids from animal tissues. J Biol Chem, 226(1), 497-509.

Fontanille, P., Kumar, V., Christophe, G., Nouaille, R., & Larroche, C. (2012). Bioconversion of volatile fatty acids into lipids by the oleaginous yeast *Yarrowia lipolytica*. Bioresource Technol, 114, 443–9. doi:10.1016/j.biortech.2012.02.091

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. a, & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods, 6(5), 343–345. doi:10.1038/nmeth.1318

Griffiths, M. J., Van Hille, R. P., & Harrison, S. T. L. (2010). Selection of direct transesterification as the preferred method for assay of fatty acid content of microalgae. Lipids, 45(11), 1053-1060.

Grillitsch, K., Connerth, M., Köfeler, H., Arrey, T. N., Rietschel, B., Wagner, B., Karas, M., Daum, G. (2011). Lipid particles/droplets of the yeast *Saccharomyces cerevisiae* revisited: lipidome meets proteome. BBA-Mol Cell Biol L, 1811(12), 1165–1175. doi:10.1016/j.bbalip.2011.07.015

Henry, S. a., Kohlwein, S. D., & Carman, G. M. (2012). Metabolism and regulation of glycerolipids in the yeast *Saccharomyces cerevisiae*. Genetics, 190(February), 317–349. doi:10.1534/genetics.111.130286

Hill, J., Nelson, E., Tilman, D., Polasky, S., & Tiffany, D. (2006). Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. P Natl Acad Sci USA, 103, 11206–11210. doi:10.1073/pnas.0604600103

Hirakawa, K., Kobayashi, S., Inoue, T., Endoh-Yamagami, S., Fukuda, R., & Ohta, A. (2009). Yas3p, an Opi1 family transcription factor, regulates cytochrome P450 expression in response to n-alkanes in *Yarrowia lipolytica*. J Biol Chem, 284(11), 7126–7137. doi:10.1074/jbc.M806864200

Hoekman, S. K., Broch, A., Robbins, C., Ceniceros, E., & Natarajan, M. (2012). Review of biodiesel composition, properties, and specifications. Renew Sust Energ Rev, 16(1), 143–169. doi:10.1016/j.rser.2011.07.143

Huber, G. W., Iborra, S., & Corma, A. (2006). Synthesis of transportation fuels from biomass: chemistry, catalysts, and engineering. Chem Rev, 106(9), 4044–98. doi:10.1021/cr068360d

Ivashov, V. a., Zellnig, G., Grillitsch, K., & Daum, G. (2013). Identification of triacylglycerol and steryl ester synthases of the methylotrophic yeast *Pichia pastoris*. BBA-Mol Cell Biol L, 1831(6), 1158–1166. doi:10.1016/j.bbalip.2013.03.004

Knothe, G. (2005). Dependence of biodiesel fuel properties on the structure of fatty acid alkyl esters. Fuel Process Technol, 86, 1059–1070. doi:10.1016/j.fuproc.2004.11.002

Konermann, S., Brigham, M. D., Trevino, A. E., Joung, J., Abudayyeh, O. O., Barcena, C., Hsu, P. D., Habib, N., Gootenberg, J. S., Nishimasu, H., Nureki, O., Zhang, F. (2014). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature, 517(7536), 583–588. doi:10.1038/nature14136

Lazar, Z., Dulermo, T., Neuvéglise, C., Crutz-Le Coq, A.-M., & Nicaud, J.-M. (2014). Hexokinase-a limiting factor in lipid production from fructose in *Yarrowia lipolytica*. Metab Eng, 26, 89–99. doi:10.1016/j.ymben.2014.09.008

Lin, H., Castro, N. M., Bennett, G. N., & San, K. Y. (2006). Acetyl-CoA synthetase overexpression in *Escherichia coli* demonstrates more efficient acetate assimilation and lower acetate accumulation: A potential tool in metabolic engineering. Appl Microbiol Biot, 71, 870–874. doi:10.1007/s00253-005-0230-4

Liu, H., Zhao, X., Wang, F., Li, Y., Jiang, X., & Ye, M. (2009). Comparative proteomic analysis of *Rhodosporidium toruloides* during lipid accumulation. FEMS Yeast Res, 2011(11), 553–566.

Liu, H., Zhao, X., Wang, F., Jiang, X., Zhang, S., Ye, M., Zhao, Z. K., Zou, H. (2011). The proteome analysis of oleaginous yeast *Lipomyces starkeyi*. FEMS Yeast Res, 11(1), 42–51. doi:10.1111/j.1567-1364.2010.00687.x

Mackall, J. C., & Lane, M. D. (1977). Role of pyruvate carboxylase in fatty acid synthesis: alterations during preadipocyte differentiation. Biochem Bioph Res Co, 79(3), 720–725. doi:10.1016/0006-291X(77)91171-8

Meng, X., Yang, J., Xu, X., Zhang, L., Nie, Q., & Xian, M. (2009). Biodiesel production from oleaginous microorganisms. Renew Energ, 34(1), 1–5. doi:10.1016/j.renene.2008.04.014

Merkler, D. J., Wali, A. S., Taylor, J., & Schramm, V. L. (1989). AMP deaminase from yeast: role in AMP degradation, large scale purification, and properties of the native and proteolyzed enzyme. J Biol Chem, 264(35), 21422–21430.

Miclet, E., Stoven, V., Michels, P. a M., Opperdoes, F. R., Lallemand, J. Y., & Duffieux, F. (2001). NMR spectroscopic analysis of the first two steps of the pentose-phosphate pathway elucidates the role of 6-phosphogluconolactonase. J Biol Chem, 276(37), 34840–34846. doi:10.1074/jbc.M105174200

Morin, N., Cescut, J., Beopoulos, A., Lelandais, G., Le Berre, V., Uribelarrea, J.-L., Molina-Jouve, C., Nicaud, J.-M. (2011). Transcriptomic analyses during the transition from biomass production to lipid accumulation in the oleaginous yeast *Yarrowia lipolytica*. PLOS ONE, 6(11), e27966. doi:10.1371/journal.pone.0027966

Müller, S., Sandal, T., Kamp-Hansen, P., & Dalbøge, H. (1998). Comparison of Expression Systems in the Yeasts *Klyveromyces lactis*, *Hansenula polymorpha*, *Schizosaccharomyces pombe*, and *Yarrowia lipolytica*. Cloning of Two Novel Promoters from *Yarrowia lipolytica*. Yeast, 14, 1267–1283.

Müllner, H., & Daum, G. (2004). Dynamics of neutral lipid storage in yeast. Acta Biochim Pol, 51(2), 323–347. doi:035001323

Nicaud, J. M., Fabre, E., & Gaillardin, C. (1989). Expression of invertase activity in *Yarrowia lipolytica* and its use as a selective marker. Curr Genet, 16(4), 253-260.

Papanikolaou, S., & Aggelis, G. (2002). Lipid production by *Yarrowia lipolytica* growing on industrial glycerol in a single-stage continuous culture, Bioresource Technol, 82(July 2001), 43–49.

Papanikolaou, S., Chevalot, I., Komaitis, M., Marc, I., & Aggelis, G. (2002). Single cell oil production by *Yarrowia lipolytica* growing on an industrial derivative of animal fat in batch cultures. Appl Microbiol Biot, *58*(3), 308–12. doi:10.1007/s00253-001-0897-0

Papanikolaou, S., Muniglia, L., Chevalot, I., Aggelis, G., & Marc, I. (2003). Accumulation of a cocoa-butter-like lipid by *Yarrowia lipolytica* cultivated on agro-industrial residues. Curr Microbiol, 46(2), 0124-0130.

Papanikolaou, S., Chatzifragkou, A., Fakas, S., Galiotou-Panayotou, M., Komaitis, M., Nicaud, J. M., & Aggelis, G. (2009). Biosynthesis of lipids and organic acids by *Yarrowia lipolytica* strains cultivated on glucose. Eur J Lipid Sci Tech, 111(12), 1221-1232.

Rahman, K. S. M., Rahman, T. J., Kourkoutas, Y., Petsas, I., Marchant, R., & Banat, I. M. (2003). Enhanced bioremediation of n-alkane in petroleum sludge using bacterial consortium

amended with rhamnolipid and micronutrients. Bioresource Technol, 90(2), 159–168. doi:10.1016/S0960-8524(03)00114-7

Rajakumari, S., & Daum, G. (2010). Multiple functions as lipase, steryl ester hydrolase, phospholipase, and acyltransferase of Tgl4p from the yeast *Saccharomyces cerevisiae*. J Biol Chem, 285(21), 15769–15776. doi:10.1074/jbc.M109.076331

Ramos, M. J., Fernández, C. M., Casas, A., Rodríguez, L., & Pérez, Á. (2009). Influence of fatty acid composition of raw materials on biodiesel properties. Bioresource Technol, 100, 261–268. doi:10.1016/j.biortech.2008.06.039

Ratledge, C. (2002). Regulation of lipid accumulation in oleaginous micro-organisms. Biochem Soc T, 30(6), 47–50.

Ratledge, C. (2004). Fatty acid biosynthesis in microorganisms being used for Single Cell Oil production. Biochimie, 86(11), 807–15. doi:10.1016/j.biochi.2004.09.017

Raught, B., Gingras, a C., & Sonenberg, N. (2001). The target of rapamycin (TOR) proteins. P Natl Acad Sci USA, 98(13), 7037–7044. doi:10.1073/pnas.121145898

Rossi, M., Amaretti, A., Raimondi, S., & Leonardi, A. (2011). Getting lipids for biodiesel production from oleaginous fungi. Biodiesel - Feedstocks and Processing Technologies. ISBN: 978-953-307-713-0, InTech

Ruiz, A., Xu, X., & Carlson, M. (2011). Roles of two protein phosphatases, Reg1-Glc7 and Sit4, and glycogen synthesis in regulation of SNF1 protein kinase. P Natl Acad Sci USA, 108(16), 6349–6354. doi:10.1073/pnas.1102758108

Rustad, T. R., Minch, K. J., Ma, S., Winkler, J. K., Hobbs, S., Hickey, M., Brabant, W., Turkarslan, S., Price, N. D., Baliga, N. S., Sherman, D. R. (2014). Mapping and manipulating

the *Mycobacterium tuberculosis* transcriptome using a transcription factor overexpression-derived regulatory network. Genome Biol, 15, 1–11. doi:10.1186/s13059-014-0502-3

Sambrook, J., & Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. CSHL press, Cold Spring Harbor, NY, USA.

Schmidt, C., Ploier, B., Koch, B., & Daum, G. (2012). Analysis of yeast lipid droplet proteome and lipidome. Method Cell Biol, 116, 15-37.

Schweizer, E., & Hofmann, J. (2004). Microbial type I fatty acid synthases (FAS): major players in a network of cellular FAS systems. Microbiol Mol Biol R, 68(3), 501–517. doi:10.1128/MMBR.68.3.501-517.2004

Seip, J., Jackson, R., He, H., Zhu, Q., & Hong, S.-P. (2013). Snf1 is a regulator of lipid accumulation in *Yarrowia lipolytica*. Appl Environ Microb, *79*(23), 7360–70. doi:10.1128/AEM.02079-13

Solomon S, Plattner, G.-K., Knutti, R., & Friedlingstein, P. (2009). Irreversible climate change due to carbon dioxide emissions. P Natl Acad Sci USA, 106(6), 1704–9. doi:10.1073/pnas.0812721106

Stephanopoulos, G. (2007). Challenges in engineering microbes for biofuels production. Science, (February), 801–805.

Tai, M. (2012). Metabolic Engineering of Oleaginous Yeast for the Production of Biofuels. Ph.D. Thesis. Massachusetts Institute of Technology.

Tai, M., & Stephanopoulos, G. (2013). Engineering the push and pull of lipid biosynthesis in oleaginous yeast *Yarrowia lipolytica* for biofuel production. Metab Eng, 15, 1–9. doi:10.1016/j.ymben.2012.08.007

Thelen, J. J., & Ohlrogge, J. B. (2002). Metabolic engineering of fatty acid biosynthesis in plants. Metab Eng, 4, 12–21. doi:10.1006/mben.2001.0204

Thiam, A. R., Farese, R. V, & Walther, T. C. (2013). The biophysics and cell biology of lipid droplets. Nat Rev Mol Cell Bio, 14(12), 775–86. doi:10.1038/nrm3699

Verbeke, J., Beopoulos, A., & Nicaud, J.-M. (2013). Efficient homologous recombination with short length flanking fragments in Ku70 deficient *Yarrowia lipolytica* strains. Biotechnol Lett, 35(4), 571–6. doi:10.1007/s10529-012-1107-0

Vigeolas, H., Waldeck, P., Zank, T. and Geigenberger, P. (2007), Increasing seed oil content in oil-seed rape (*Brassica napus* L.) by over-expression of a yeast glycerol-3-phosphate dehydrogenase under the control of a seed-specific promoter. Plant Biotechnol J, 5:431–441. doi: 10.1111/j.1467-7652.2007.00252.x

Wakil, S. J., Stoops, J. K., & Joshi, V. C. (1983). Fatty acid synthesis and its regulation. Annu Rev Biochem, 52, 537–79. doi:10.1146/annurev.bi.52.070183.002541

Wasylenko, T. M., Ahn, W. S., & Stephanopoulos, G. (2015). The oxidative pentose phosphate pathway is the primary source of NADPH for lipid overproduction from Glucose in *Yarrowia lipolytica*. Metab Eng, 30, 27–39. doi:10.1016/j.ymben.2015.02.007

Wenz, P., Schwank, S., Hoja, U., & Schüller, H. J. (2001). A downstream regulatory element located within the coding sequence mediates autoregulated expression of the yeast fatty acid synthase gene FAS2 by the FAS1 gene product. Nucleic Acids Res, 29(22), 4625–4632. doi:10.1093/nar/29.22.4625

Wu, S., Hu, C., Jin, G., Zhao, X., & Zhao, Z. K. (2010). Phosphate-limitation mediated lipid production by *Rhodosporidium toruloides*. Bioresource Technol, 101(15), 6124–9. doi:10.1016/j.biortech.2010.02.111

Wu, S., Zhao, X., Shen, H., Wang, Q., & Zhao, Z. K. (2011). Microbial lipid production by *Rhodosporidium toruloides* under sulfate-limited conditions. Bioresource Technol, 102(2), 1803–7. doi:10.1016/j.biortech.2010.09.033

Xue, Z., Sharpe, P. L., Hong, S.-P., Yadav, N. S., Xie, D., Short, D. R., Damude, H. G., Rupert, R. A., Seip, J. E., Wang, J., Pollak, D. W., Bostick, M. W., Bosak, M. D., Macool, D. J., Hollerbach, D. H., Zhang H., Arcilla, D. M., Bledsoe, S. A., Croker, K., McCord, E. F., Tyreus, B. D., Jackson, E. N., Zhu, Q. (2013). Production of omega-3 eicosapentaenoic acid by metabolic engineering of *Yarrowia lipolytica*. Nat Biotechnol, 31(8), 734–40. doi:10.1038/nbt.2622

Zhang, H., Damude, H. G., & Yadav, N. S. (2012). Three diacylglycerol acyltransferases contribute to oil biosynthesis and normal growth in *Yarrowia lipolytica*. Yeast, 29, 25–38. doi:10.1002/yea.1914

Zhang, H., Zhang, L., Chen, H., Chen, Y. Q., Ratledge, C., Song, Y., & Chen, W. (2013). Regulatory properties of malic enzyme in the oleaginous yeast, *Yarrowia lipolytica*, and its non-involvement in lipid accumulation. Biotechnol Lett, 35(12), 2091–8. doi:10.1007/s10529-013-1302-7

Zhang, Y., Adams, I. P., & Ratledge, C. (2007). Malic enzyme: the controlling activity for lipid production? Overexpression of malic enzyme in *Mucor circinelloides* leads to a 2.5-fold increase in lipid accumulation. Microbiology, 153(Pt 7), 2013–25. doi:10.1099/mic.0.2006/002683-0

Zhao, S., Fernald, R. D. (2005). Comprehensive algorithm for quantitative real-time polymerase chain reaction. J Comput Biol, Oct;12(8):1045-62.

Zou, J., Katavic, V., Giblin, E. M., Barton, D. L., MacKenzie, S. L., Keller, W. A., Hu, X., Taylor, D. C. (1997). Modification of seed oil content and acyl composition in the brassicaceae by expression of a yeast sn-2 acyltransferase gene. Plant Cell, 9(June), 909–923. doi:10.1105/tpc.9.6.909

Zhu, Z., Zhang, S., Liu, H., Shen, H., Lin, X., Yang, F., Zhou, Y. J., Jin, G., Ye, M., Zou, H., Zhao, Z. K. (2012). A multi-omic map of the lipid-producing yeast *Rhodosporidium toruloides*. Nat Comm, 3, 1112. doi:10.1038/ncomms2112

Table 1: Genes Characterized in this Study

Gene Number	NCBI Locus Tag	Gene Name	Deduced/Hypothetical Function
Control			
LZ	b0344	LacZ	beta-galactosidase from *E. coli*; strains overexpressing this gene are control strains
Kennedy Pathway/Glycerolipid Synthesis			
OP1	YALI0D07986g	DGA2	diacylglycerol acyltransferase
OP2	YALI0E16797g	LRO1	phospholipid:diacylglycerol acyltransferase
OP3	YALI0A10362g	GPA	glyerol-3-phosphate acyltransferase
OP4	YALI0D27016g	PAH1	lipin, phosphatidate phosphatase
OP5	YALI0F06578g	ARE1	acyl-CoA sterol acyltransferase
OP6	YALI0E18964g	SLC1	acylglycerol-phosphate acyltransferase
OP7	YALI0F10010g	TGL4	triacylglycerol lipase; may have some acylglycerol-phosphate acyltransferase activity
OP8	YALI0F19514g	ALE1	lysophospholipid acyltransferase
OP9	YALI0C00209g	SCT1	glycerol-3-phosphate acyltransferase; dihydroxyacetone phosphate acyltransferase
OP10	YALI0E14443g	CDS1	CDP-diacylglycerol synthase
OP11	YALI0F23837g	PGS1	cardiolipin synthase
OP12	YALI0B18854g	EEB	putative carboxylic ester hydrolase; may have some fatty acid ethyl ester synthesis activity
OP13	YALI0D03480g	PSD	phosphatidylserine decarboxylase
OP14	YALI0A08448g	PDR16	phosphatidylinositol transfer protein; may be involved in regulation of lipid biosynthesis or transport
OP15	YALI0E32035g	TGL1	steryl ester hydrolase
OP16	YALI0C17545g	CSR1	phosphatidylinositol transfer protein; may be involved in regulation of lipid biosynthesis or transport
OP17	YALI0D05995g	SAC1	phosphatidylinositol phosphate phosphatase
OP18	YALI0C14014g	LOA	lysophosphatidic acid acyltransferase
Central Carbon Metabolism			

OP19	YALI0F05962g	ACS2	acetate-CoA ligase
OP20	YALI0B02948g	GPD1	NAD+-dependent glycerol-3-phosphate dehydrogenase
OP21	YALI0C24101g	PYC1	cytosolic pyruvate carboxylase, involved in transhydrogenase cycle with malic enzyme
Fatty Acid Synthesis			
OP22	YALI0E30591g	BPL1	acetyl-CoA carboxylase-biotin ligase
OP23	YALI0B15059g	FAS1	fatty acid synthase subunit beta
OP24	YALI0B19382g	FAS2	fatty acid synthase subunit alpha
OP25	YALI0B10153g	FAD2	Oleoyl-CoA delta-12 desaturase
OP26	YALI0E21131g	SLD	delta-8 desaturase; may have desaturase activity at other sites
NADPH Generation			
OP27	YALI0E22649g	ZWF1	NADP+-dependent glucose-6-phosphate dehydrogenase
OP28	YALI0B15598g	GND1	NADP+-dependent 6-phosphogluconate dehydrogenase
OP29	YALI0E11671g	SOL3	6-phosphogluconolactonase
OP30	YALI0F04095g	IDP1	cytosolic NADP+-specific isocitrate dehydrogenase
OP31	YALI0F17820g	GDH3	NADP+-specific glutamate dehydrogenase
Regulators of Lipid Synthesis			
OP32	YALI0F11869g	SAP190	Associated with Sit4 protein phosphatase, which can inactivate SNF1, an ACC1-inactivating protein kinase
OP33	YALI0F07084g	TOR2	target of rapamycin protein, phosphoinositide 3-kinase-related protein kinase, master regulator involved in promoting growth (including lipid synthesis) in response to nutrient availability
OP34	YALI0C02387g	YAS1	HLH transcription factor, upregulates cytochrome p450 genes in response to alkanes with YAS2; some similarity to INO4/INO2 from *Saccharomyces cerevisiae*, which upregulates phospholipid biosynthesis genes
OP35	YALI0E32417g	YAS2	HLH transcription factor, upregulates cytochrome p450 genes in response to alkanes with YAS1; some similarity to INO4/INO2 from *Saccharomyces cerevisiae*
OP36	YALI0B16808g	REG1	regulatory subunit of Reg1-Glc7 protein phosphatase 1 (PP1) complex, PP1 can inactivate SNF1, which can inactivate Acc1
OP37	YALI0A08077g	GLC7	catalytic subunit of Reg1-Glc7 protein phosphatase 1 (PP1) complex, PP1 can inactivate SNF1, which can inactivate Acc1
Transporters			
OP38	YALI0E01298g	SCP2	sterol carrier protein; may be involved in fatty acid transport
OP39	YALI0D20416g	SLY1	Golgi transport protein; similar gene upregulated in *R. toruloides* during lipogenesis
OP40	YALI0C20265g	ABC2	ABC (ATP-binding cassette) transporter; similar gene upregulated in *R. toruloides* during lipogenesis
Miscellaneous Genes			
OP41	YALI0E11495g	AMD1	adenosine monophosphate (AMP) deaminase, depletes cellular AMP to divert carbon from TCA cycle towards lipogenesis
OP42	YALI0D11726g	PSR1	probable transmembrane protein phosphatase for stress response; similar gene upregulated in *R. toruloides* during lipogenesis
OP43	YALI0A18590g	DSK1	Probable Dis1-suppressing protein kinase, involved in regulation of mitosis and possibly pre-mRNA splicing; similar gene upregulated in *R. toruloides* during lipogenesis
OP44	YALI0A09845g	YIP3	probable Rab-GDI dissociation factor, involved in maintenance of intracellular membranes; similar gene upregulated in *R. toruloides* during lipogenesis

Figure 1: Map of the metabolic pathways from substrate (glucose or acetate) to lipids. Enzymes are written in italics, and those studied in this work are also underlined. Arrows with circular ends denote activation of the reaction or enzyme they point to. G6P: glucose-6-phosphate; 6PGL: 6-phosphogluconolactone; 6PG: 6-phosphogluconate; Rib5P: ribulose-5-phosphate; R5P: ribose-5-phosphate; X5P: xylulose-5-phosphate; F6P: fructose-6-phosphate; G3P: glyceraldehyde-3-phosphate; DHAP: dihydroxyacetone phosphate; PEP: phosphoenolpyruvate; Gly3P: glycerol-3-phosphate; Ac-CoA: acetyl-CoA; Mal-CoA: malonyl-CoA; FA: fatty acid; AGP: acylglyerol-phosphate (lysophosphatidate); PA: phosphatidate; CDP-DAG: cytidine diphosphate diacylglycerol; PL: glycerophospholipids; DAG: diacylglycerol; TAG: triacylglycerol.

Figure 2: Expression of genes in this study with an additional copy under control of the TEF-1α promoter relative to expression solely under their respective native promoters. Actin was used as the reference gene. Samples were taken from strains growing in glucose-containing media in shake flasks. Samples were analyzed with two technical replicates; error bars represent a 95% confidence interval for fold-overexpression. All samples are statistically significant for increased expression ($\alpha = 0.05$).

Figure 3: Fermentation performance of strains overexpressing glycerolipid synthesis genes on (A) 50 g/L glucose at an atomic C:N of 60 after 120 hours and (B) 50 g/L sodium acetate at an atomic C:N of 40 after 144 hours. Performance metrics used to evaluate strains include lipid titer, cellular lipid content, and lipid yield. An "H" denotes extreme growth defect due to hyphal formation in this strain under this condition. Error bars represent standard deviations. Fermentations were carried out in triplicate for all strains.

Figure 4: Fermentation performance of strains overexpressing central carbon metabolism genes on (A) 50 g/L glucose at an atomic C:N of 60 after 120 hours and (B) 50 g/L sodium acetate at an atomic C:N of 40 after 144 hours. Performance metrics used to evaluate

strains include lipid titer, cellular lipid content, and lipid yield. Error bars represent standard deviations. Fermentations were carried out in triplicate for all strains.

Figure 5: Fermentation performance of strains overexpressing fatty acid synthesis genes on (A) 50 g/L glucose at an atomic C:N of 60 after 120 hours and (B) 50 g/L sodium acetate at an atomic C:N of 40 after 144 hours. Performance metrics used to evaluate strains include lipid titer, cellular lipid content, and lipid yield. Error bars represent standard deviations. Fermentations were carried out in triplicate for all strains.

Figure 6: Fermentation performance of strains overexpressing NADPH generation genes on (A) 50 g/L glucose at an atomic C:N of 60 after 120 hours and (B) 50 g/L sodium acetate at an atomic C:N of 40 after 144 hours. Performance metrics used to evaluate strains include lipid titer, cellular lipid content, and lipid yield. Error bars represent standard deviations. Fermentations were carried out in triplicate for all strains.

Figure 7: Fermentation performance of strains overexpressing regulator genes on (A) 50 g/L glucose at an atomic C:N of 60 after 120 hours and (B) 50 g/L sodium acetate at an atomic C:N of 40 after 144 hours. Performance metrics used to evaluate strains include lipid titer, cellular lipid content, and lipid yield. Error bars represent standard deviations. Fermentations were carried out in triplicate for all strains.

Figure 8: Fermentation performance of strains overexpressing metabolite transport and other genes on (A) 50 g/L glucose at an atomic C:N of 60 after 120 hours and (B) 50 g/L sodium acetate at an atomic C:N of 40 after 144 hours. Performance metrics used to evaluate strains include lipid titer, cellular lipid content, and lipid yield. Error bars represent standard deviations. Fermentations were carried out in triplicate for all strains.

Figure 9: Fatty acid profiles of selected strains growing on glucose. Percentages shown are the mean values of six measurements of lipid wt% of C16:0, C16:1, C18:0, C18:1, and C18:2 analyzed by GC-FID. C16:0: palmitate; C16:1: palmitoleate; C18:0: stearate; C18:1: oleate; C18:2: linoleate.

Figure 1

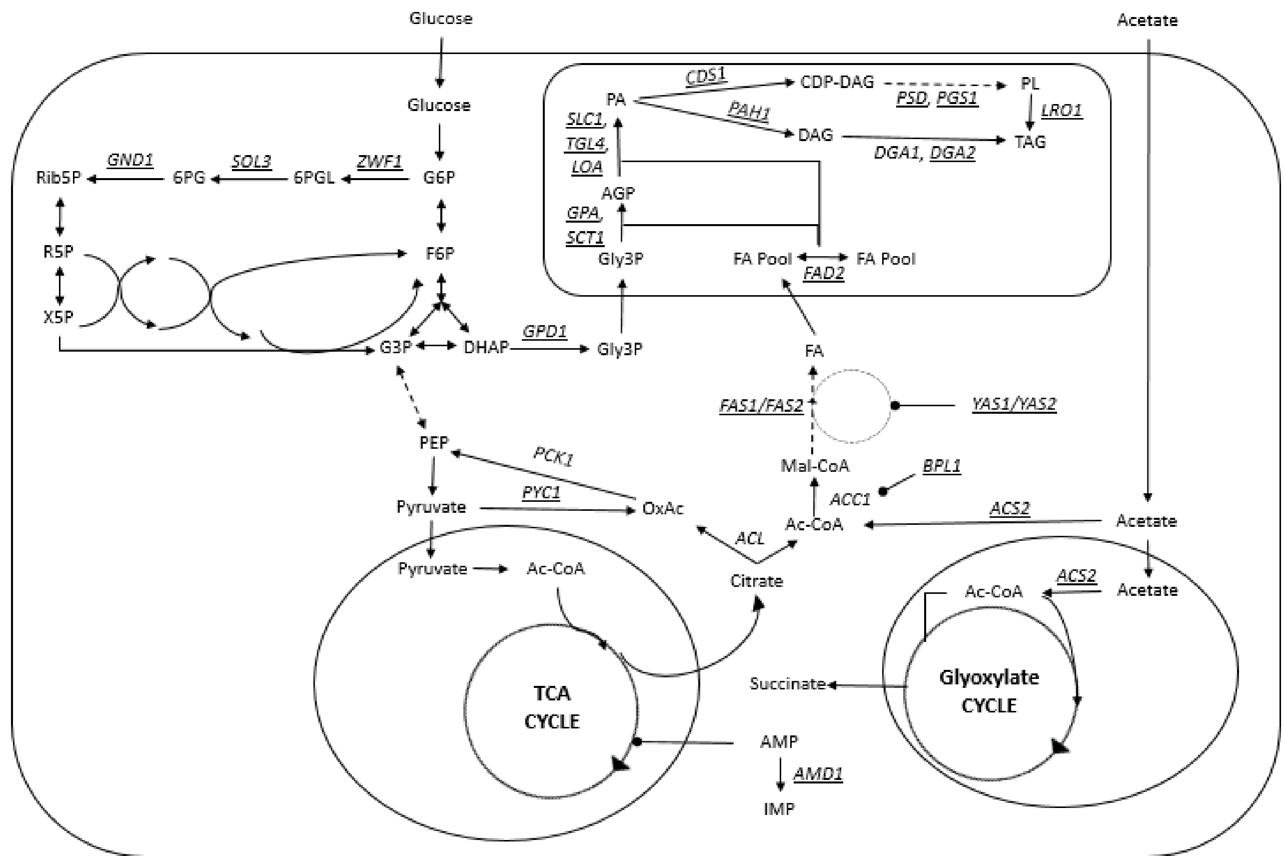


Figure 2

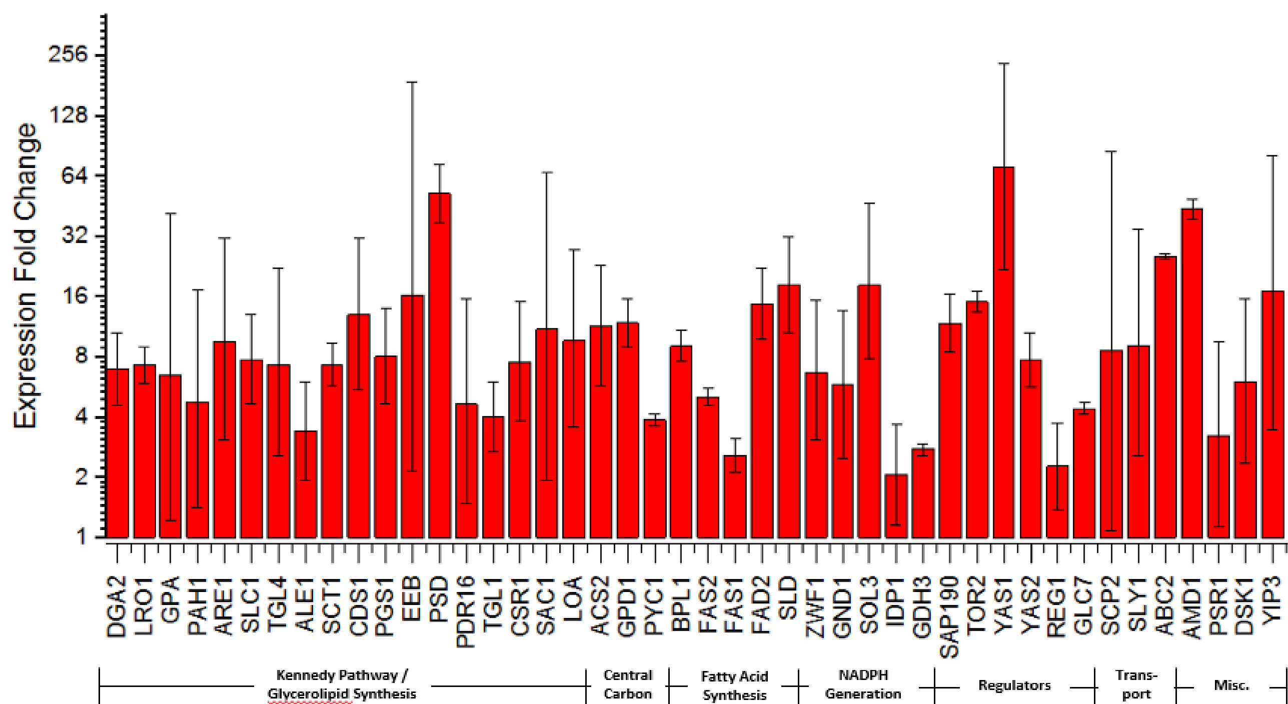


Figure 3

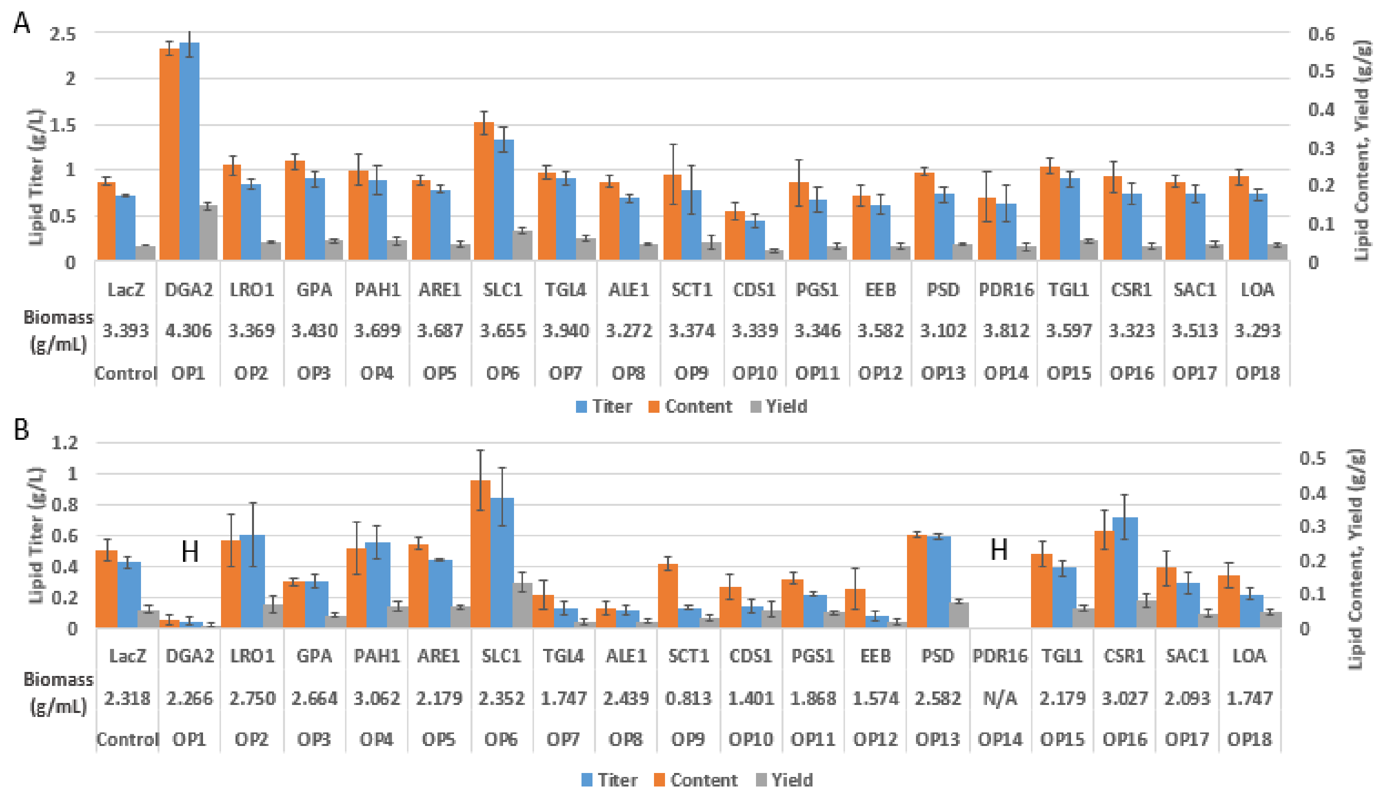


Figure 4

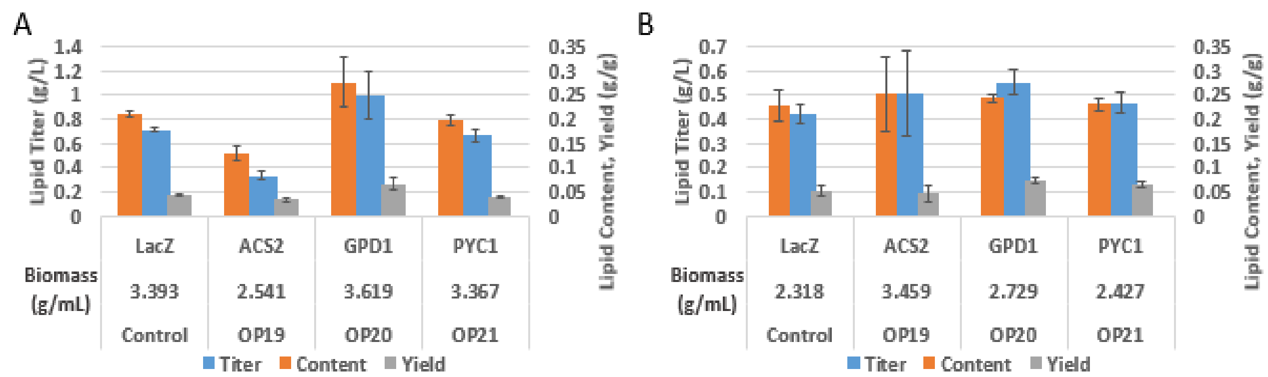


Figure 5

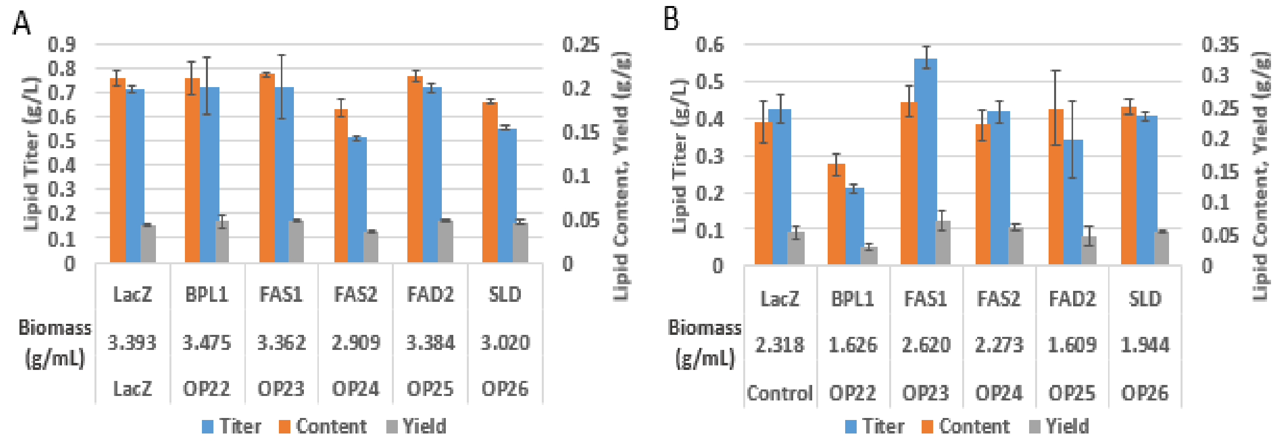


Figure 6

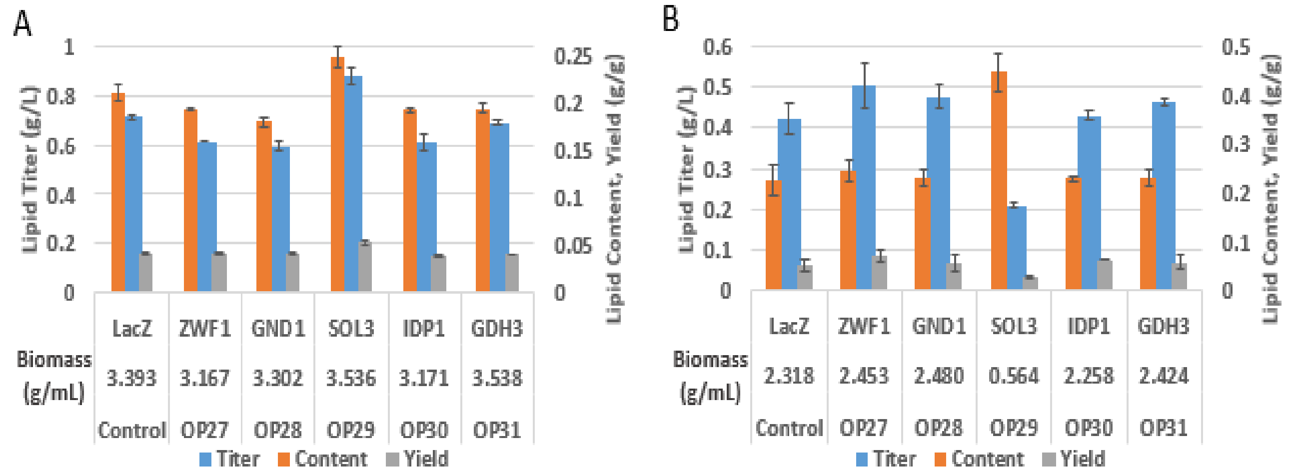


Figure 7

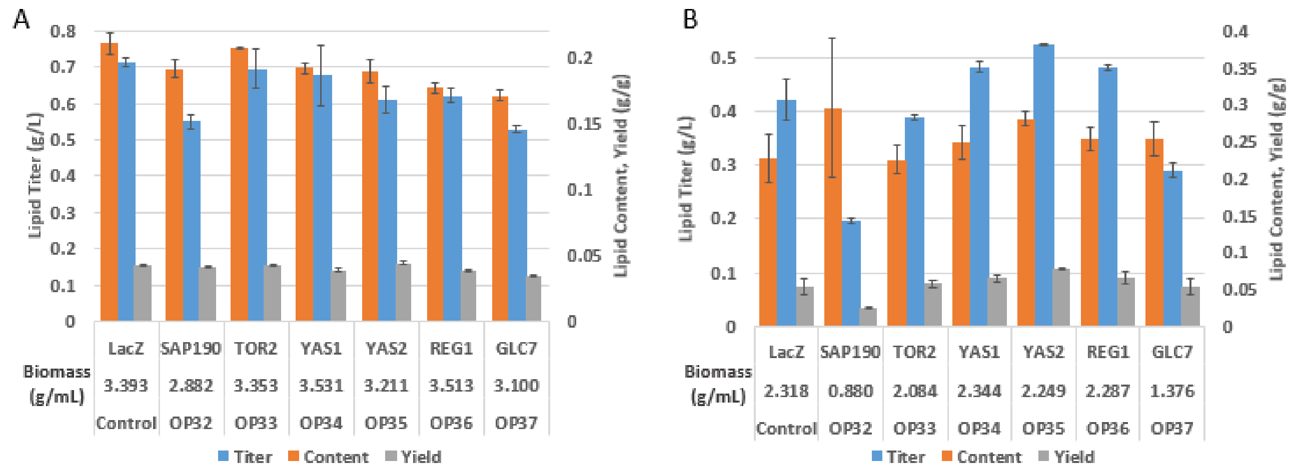


Figure 8

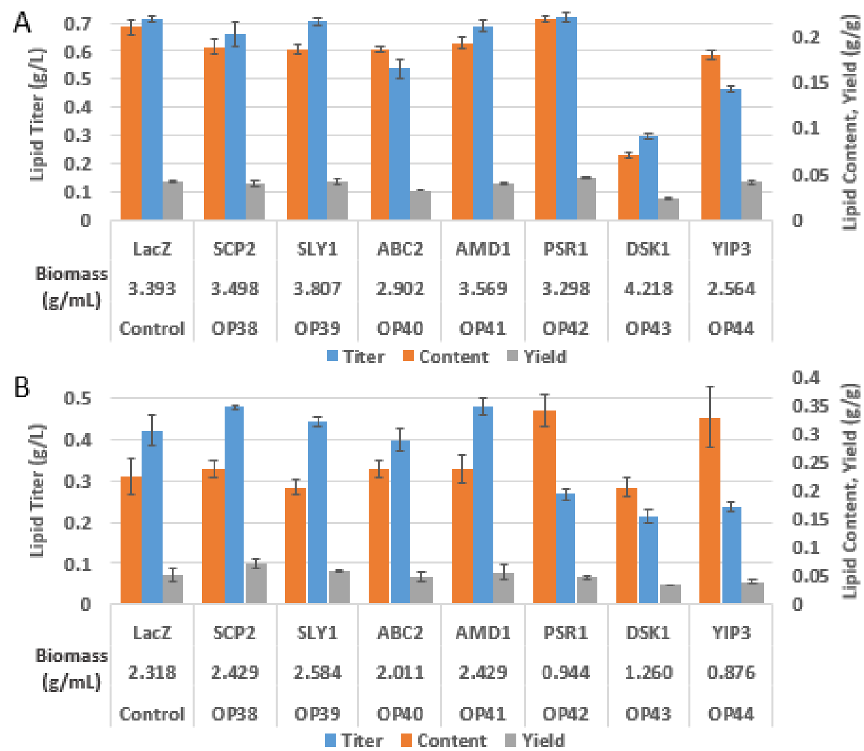
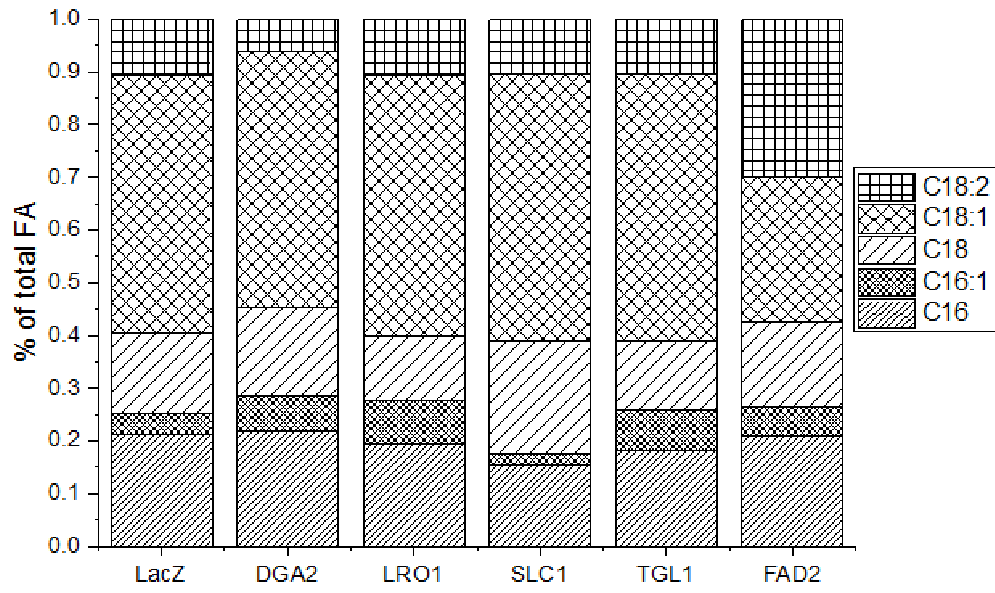


Figure 9



1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
% of total FA
LacZ
DGA2
LRO1
SLC1
TGL1
FAD2
C18:2
C18:1
C18
C16:1
C16