Metabolic Engineering Strategies for Increasing Lipid Production in Oleaginous Yeast

by

Andrew Michael Silverman

M.S. Chemical Engineering Practice, Massachusetts Institute of Technology (2011)
B.S. Biochemistry and Molecular Biology, University of Florida (2008)
B.S. Chemical Engineering, University of Florida (2008)

Submitted to the Department of Chemical Engineering
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2015

Signature of Author..... Signature redacted
Department of Chemical Engineering
September 4, 2015

Certified by..... Signature redacted .....
Gregory N. Stephanopoulos
Willard Henry Dow Professor of Chemical Engineering and Biotechnology
Thesis Supervisor

Accepted by..... Signature redacted
Richard D. Braatz
Professor of Chemical Engineering
Chairman, Committee for Graduate Students

Metabolic Engineering Strategies for Increasing Lipid Production in Oleaginous Yeast

by

Andrew Michael Silverman

Submitted to the Department of Chemical Engineering on September 4, 2015,
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Chemical Engineering

Abstract

Although petroleum and other fossil fuels have traditionally been used to fulfill our energy needs, rising concerns over energy security and the climate-changing effects of our continual greenhouse gas emissions have led to great interest in developing a domestic source of renewable fuel with low net carbon emissions. Biodiesel is an attractive option for replacing petroleum-based fuels used in the transportation sector due to its compatibility with existing infrastructure. Single cell oils from heterotrophic oleaginous microorganisms as a source of biodiesel allow for high productivity from a wide array of potential feedstocks, including agro-industrial and municipal waste streams.

The goal of this work is to use the tools of rational metabolic engineering to improve lipid production in the non-conventional oleaginous yeast *Yarrowia lipolytica* on two representative carbon sources, glucose and acetate. Previous work in this area achieved considerable success with the simultaneous overexpression of the native acetyl-CoA carboxylase (ACC1) and diacylglycerol acyltransferase (DGA2) genes; the resulting strain was used as a benchmark to evaluate our own efforts.

We began with the compilation of a set of 44 genes and evaluated the effects of the individual overexpression of each gene on the ability of the resulting strain to produce lipids in fermentations of glucose and acetate. The genes tested here represent many different functions potentially important to lipid production, including the Kennedy pathway, fatty acid synthesis, central carbon metabolism, NADPH generation, regulation, and metabolite transport. Our results demonstrate that a diverse subset of gene overexpressions led to significant improvements in lipid production on at least one substrate. The largest improvements unsurprisingly came from overexpressing genes directly related to triacylglycerol synthesis, such as diacylglycerol acyltransferase DGA1, which on glucose increased the lipid titer, content and yield by 236%, 165%, and 246%, respectively, over our wild-type control strain, and the acylglycerolphosphate acyltransferase SLC1 gene, which increased titer/content/yield on glucose by 86%/73%/87% and on acetate by 99%/91%/151%. Significant improvements were also detected from genes that more indirectly effect lipogenesis, such as glycerol-3-phosphate dehydrogenase GPD (which produces head groups for triacylglycerol molecules) and the 6-phosphogluconolactoase SOL3 (catalyzing the middle step of the NADPH-producing oxidative pentose phosphate pathway).

We next chose the aforementioned SLC1, GPD, and SOL3 genes for use in continued rational engineering of our benchmark strain due to the significance of their effects and the lack of redundancy in their likely mechanism of improving lipogenesis when overexpressed along with ACC1 and DGA2. The results of this investigation indicate that the strain overexpressing ACC1, DGA2, and SLC1 may be superior to our benchmark strain, increasing lipid content and yield by 24% and 20%, respectively, with a statistically equivalent titer on acetate. This strain produces the highest reported overall lipid yield of an oleaginous yeast on acetate, at 0.207 g lipids/g acetate.

In addition to gene overexpressions, we explored the use of gene knockouts by removing the SNF1 gene, which codes for a protein kinase that can inhibit Acc1. This led to an improvement in the lipid content and yield on glucose and acetate, with the effect much greater on acetate. We then complemented this ΔSNF1 strain with the previously validated overexpressions from our benchmark strain. We achieved the notable result of a highly lipogenic phenotype on acetate with a strain that overexpresses DGA2 in the ΔSNF1 background; this strain produced a lipid content of 0.767 g lipids/g DCW, the highest lipid content reported for an oleaginous yeast growing on acetate as a sole carbon source, and a lipid yield of 0.200 g lipids/g acetate.

Lastly, we discuss a couple projects that did not result in any strain improvement. We used differential expression analysis to identify lipase genes that could be responsible for TAG degradation in *Yarrowia* by determining which genes are more highly expressed during lipostatic phases than during growth and lipid production phases, however, the knockout of the genes identified using this method did not result in the increase in lipid production. A project involving heterologous pathway construction in our benchmark strain for improving lipid yield and productivity on acetate by conserving ATP in activating acetate and providing a shorter metabolic pathway for generating NADPH also resulted in a lack of significant improvements.

As our understanding of the intricacies of the metabolic network becomes clearer, and the functions of previously uncharacterized genes and genetic elements in *Yarrowia lipolytica* are investigated, the potential for the use of this organism in industrial-scale production of biodiesel grows. Incremental improvements in lipid production, even marginal ones, that are typical of the latter stages of host strain optimization, could make the difference between a single cell oil

process for biodiesel becoming economically viable or not in competition with traditional, non-renewable energy sources.

Thesis Supervisor: Gregory Stephanopoulos

Title: Willard Henry Dow Professor of Chemical Engineering and Biotechnology

Acknowledgements

First and foremost, I must thank my advisor, Greg Stephanopoulos, for giving me the opportunity to work in his lab and in a field that I knew I wanted to be a part of since before I started grad school. As a research mentor, Greg struck a perfect balance between providing me with clear objectives, ideas, and resources for my work and allowing me the freedom to independently design and execute my own plans. In the face of rather long stretches of adversity throughout my time here and through many failed projects, he taught me to trust myself as a researcher and quickly move on to new ideas when the results did not match our expectations; my work progressed much more smoothly once I began to internalize this concept. His incredible patience for when things did not go right and infectious enthusiasm for when they did provided me with a constant source of motivation for me throughout my time here. Similarly, I must thank my thesis committee members, Hadley Sikes and Chris Love, who were also always very supportive and encouraging. They were immensely helpful in distilling the scattered, half-formed ideas I would bring up at every meeting down to a more reasonable, coherent plan, and their advice had a large influence in shaping this work.

Next, I would like to thank the members of the Stephanopoulos group, who taught me so much and helped create a working environment in the lab that gave me something to look forward to almost every day. Mitchell Tai taught me how to work with *Yarrowia*, and showed me the value of being exceptionally well-organized; I never had a problem locating a strain or plasmid he made or replicating the results of any of his experiments, and he made it incredibly easy for me to get started on my own work because of it. Sagar Chakraborty taught me how to run a bioreactor and generously assisted me in preparing my bioreactor experiments. Kangjian Qiao worked with me on a couple projects and introduced me to a few molecular biology

techniques that were extremely useful and made my cloning efforts substantially more productive. Tom Wasylenko also contributed significantly to my work in providing me with interesting ideas for some projects and helping me to explain the results of some others. I also would like to thank Brian Pereira, Adel Ghaderi, Ben Woolston, Steven Edgar, Mark Keibler, and Haoran Zhang, who along with the rest of the group, were always willing to commiserate with me when projects failed (and congratulate me when they worked) and were never lacking in solid advice to give at the same time. In addition to those in our group, I must thank Scott Baker and his research group at the Pacific Northwest National Laboratory for creating and providing us with a *Yarrowia* strain that was more amenable to genetic disruption; without this, much of the work in this thesis could not have been done.

I need to thank my family for the immense amount of moral support over these last several years. My mother JoAnn, my father Richard, my brother Josh, my sister Rachel, and brother-in-law Sergio showed an unshakable faith that I would eventually finish this long process. They were a constant source of encouragement through the many stretches of unexpected obstacles and disappointments. Lastly, I would like to thank my friends and classmates here for the supportive environment and for helping me to enjoy my time away from the lab.

Contents

List of Figures

List of Tables

Chapter 1

Single Cell Oil as an Alternative Energy Source

1.1 Motivations for the development of alternative energy sources

Although petroleum and other fossil fuels have traditionally been used to fulfill our energy needs, rising concerns over energy security and the climate-changing effects of our continual greenhouse gas emissions have led to great interest in developing a domestic source of renewable energy with low net carbon emissions [1]. Roughly 60% of the net petroleum imported by the U.S. in 2014 came from members of the Organization of Petroleum Exporting Countries (OPEC) [2], many of which are relatively economically or politically unstable. In addition to this, an overwhelming majority of climate scientists believe that anthropogenic forces, most notably increases in greenhouse gases such as carbon dioxide, are largely if not entirely responsible for the 0.6°C rise in average global temperature from the period of 1951-2010. This increase in temperature is predicted to rapidly intensify, rising by a similar amount over the next 20 years even under scenarios with the most aggressive reductions in carbon emissions [3], leading to the rising of sea levels and increased frequency of heat waves and other extreme weather phenomena. Rising carbon dioxide concentrations are also predicted to contribute to the acidification of the oceans that is already beginning to be observed.

A significant contributor to greenhouse gas emissions comes from the transportation sector. Transportation accounted for approximately 27% of greenhouse gas emissions in the U.S. in 2013, and 13% of global emissions in 2007 [4]. Approximately 95% of the fuel for transportation globally is derived from petroleum-based sources.

1.2 Single cell oil (SCO) processes for production of biodiesel

Biodiesel is an attractive option for replacing petroleum-based fuels used in the transportation sector due to its compatibility with existing infrastructure while significantly reducing greenhouse gas emissions. Currently, most biodiesel is made from plant oils, but in

order to approach meeting the demand for fuels, this process requires prohibitively large amounts of arable land to be devoted to relatively slow growing feedstocks that are expensive to maintain instead of food crops [5, 6]. To prevent this food-fuel competition in land use, research into single cell oil (SCO) from microorganisms as a method of producing biodiesel has grown considerably in recent years.

"Oleaginous" microorganisms are defined as being capable of producing lipids to make up at least 20% of their total biomass, mostly stored in the form of triacylglycerol (TAG) molecules; these microorganisms grow quickly and produce a distribution of fatty acids that are very similar to the plant oils currently used for biodiesel production. Perhaps most importantly, many of the heterotrophic species that are candidates for use in SCO processes can grow on and produce lipids from a wide variety of organic molecules from diverse sources, allowing for a means of valorizing agricultural, municipal, or industrial waste streams with minimal pretreatment [7]. Some oleaginous microorganisms, such as *Yarrowia lipolytica*, have been shown to be genetically tractable using current methods in molecular biology and are able to be engineered for improved lipid productivity and the ability to metabolize more substrates [8]. With high productivity and the ability to effectively utilize a diverse set of inexpensive raw inputs, the use of SCO processes to produce biodiesel can help to meet the rising demand for domestic, renewable, low-carbon emitting alternatives to traditional fossil fuels.

1.3 Main objectives of this work and thesis overview

Biodiesel from single cell oil processes has the potential to be a widely used alternative energy source for reducing carbon emissions and dependence on oil imports, but the economic landscape is not currently favorable towards its implementation. Important parameters that greatly influence the economic viability of SCO processes include the lipid productivities and

yields of the host organism when growing on different carbon substrates. As there have previously been successful attempts at engineering one oleaginous organism in particular, *Yarrowia lipolytica*, the main focus of this work is to build upon those previous successes and engineer strains with increased ability to produce lipids on two representative substrates, glucose and acetate. To accomplish this, the studies reported herein involve the use of a diverse set of metabolic engineering tools, including native gene overexpression, gene knockouts, and heterologous pathway expression.

Chapter 2 provides some more background on the oleaginous organism *Yarrowia lipolytica*. As a "non-conventional" yeast, *Yarrowia* significantly differs in physiology from other model yeasts such as *Saccharomyces cerevisiae*. This Chapter begins with a description of *Yarrowia lipolytica* and a discussion of the advantages of using this species as a host organism for SCO processes, then provides a primer on the biochemistry of *de novo* lipid production from glucose and acetate, complete with theoretical maximum yields for fatty acids on each substrate, and ends with a brief overview of prior metabolic engineering efforts to improve lipid production in this organism and diversify the potential inputs and output distributions of these processes.

In **Chapter 3**, we profile 44 native different genes potentially related to lipogenesis in *Yarrowia* by overexpressing them individually and comparing the fermentation performances of the resulting strains in glucose and acetate containing media. The goal of this project is to identify genes that may be useful for making additional improvements to lipid productivity or yield on these substrates in a background that already overexpress other genes previously validated to be effective for this purpose. In **Chapter 4**, we select three of the genes from the screening study in **Chapter 3**, overexpress them in this high-oil background, and evaluate the lipid-production ability of the resulting strains on glucose and acetate.

Chapter 5 involves the development of a strain and method for performing gene knockouts in *Yarrowia*, a technique that had previously only been reliable for *Yarrowia* in our lab for the marker gene URA3, due to the availability of an effective screening method (recovery plates complemented with 5'fluoroorotic acid) for null strains of this specific gene. Using a *Yarrowia* strain with enhanced genetic tractability, we successfully knocked out the SNF1 gene, which codes for a negative regulator of fatty acid synthesis. After assessing the lipogenic phenotype of our ΔSNF1 strain on glucose and acetate, we attempt to optimize this strain by adding other modifications in the form of native gene overexpressions that have been previously validated as contributing to a lipogeneic phenotype in *Yarrowia*.

Chapter 6 is an investigation into which lipases in *Yarrowia* may be responsible for the previously observed phenomenon of triacylglycerol degradation in this organism during fermentations. We used a differential expression analysis method to identify lipase genes with increased expression in a high-oil producing strain in an acetate bioreactor fermentation during the lipostatic late-stationary phase compared to the growth phase. We knocked out the genes of interest in the high-oil background and assessed the resulting strains' fermentation performances.

In **Chapter 7**, we present two heterologous pathways that we hypothesized would improve either the yield or productivity of lipids on acetate by conserving energy (ATP) or providing a shorter pathway for producing reducing equivalents for fatty acid synthesis. We engineer high-oil strains with each of these pathways and test for improvements in lipid productivity and yield on acetate.

Lastly, **Chapter 8** summarizes the work done in this thesis, highlighting the most promising results. We also include a brief discussion of a couple general themes recurring throughout the

work that was reported here and that which was left out of this thesis, and conclude with recommendations for future work in this area.

References:

1) Viju, C., & Kerr, W. a. (2013). Taking an option on the future: Subsidizing biofuels for energy security or reducing global warming. *Energy Policy*, *56*, 543–548. doi:10.1016/j.enpol.2013.01.020
2) U.S. Energy Information Association. *FAQ: How much petroleum does the United States import and from where?* http://www.eia.gov/tools/faqs/faq.cfm?id=727&t=6. Accessed: August 2015.
3) IPCC. 2014 Climate Change Synthesis Report: Fifth Assessment Report (AR5). http://www.ipcc.ch/pdf/assessment-report/ar5/syr/AR5_SYR_FINAL_SPM.pdf. Accessed: August 2015.
4) EPA. Greenhouse Gas Emissions. http://www.epa.gov/climatechange/ghgemissions. Accessed: August 2015.
5) Rossi, M., Amaretti, A., Raimondi, S., & Leonardi, A. (2011). Getting Lipids for Biodiesel Production from Oleaginous Fungi. *Biodiesel – Feedstocks and Processing Technologies*. doi:10.5772/25864
6) Meng, X., Yang, J., Xu, X., Zhang, L., Nie, Q., & Xian, M. (2009). Biodiesel production from oleaginous microorganisms. *Renewable Energy*, *34*(1), 1–5. doi:10.1016/j.renene.2008.04.014
7) 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 Technology*, *114*, 443–9. doi:10.1016/j.biortech.2012.02.091
8) Tai, M. (2012). Metabolic Engineering of Oleaginous Yeast for the Production of Biofuels. Ph.D. Thesis. Massachusetts Institute of Technology.

Chapter 2

The Use of *Yarrowia lipolytica* in Single Cell Oil Processes

2.1 *Yarrowia lipolytica* as a model organism for the study of lipid biosynthesis and host for industrial processes

The choice of a host organism to use as the biocatalyst in a single cell oil (SCO) process is a critical parameter that greatly influences many aspects of process design and economics. So-called "oleaginous" microbes offer a clear advantage in that they can naturally accumulate in excess of 20% of their total biomass in the form of lipids [1] without any genetic modifications. Oleaginous microbes can include both phototrophic organisms such as microalgae, and heterotrophic organisms in yeasts and bacteria. While the ability of phototrophic oleaginous organisms to directly fix carbon dioxide to produce biodiesel precursors seems highly advantageous, the growth rates (and therefore lipid productivity) of these organisms are generally poor, and the need for light requires large land areas for maximum exposure [2]. One heterotrophic oleaginous yeast, *Yarrowia lipolytica*, has recently emerged as a model organism for the study of lipogenesis and as a platform for SCO production.

Yarrowia lipolytica is a "non-conventional" yeast, in that it is physiologically and phylogenetically very distinct from the typical model yeasts used in basic biological or biotechnological research, such as *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Yarrowia* is a dimorphic fungus, with two physiological forms: a round "yeast" form, and a true hyphal form (as opposed to the pseudo-hyphal form found in *Saccharomyces*), characterized by rod-shaped, slowly growing filamentous cells; transition from yeast to hyphal form is generally in response to stress [3]. It is an obligate aerobe, and a Crabtree-negative yeast, unlike *Saccharomyces*. As an oleaginous organism, wild-type *Yarrowia* has been found to accumulate up to 36% of its dry cell weight (DCW) in the form of lipids [4], over 90% of which are in the form of triacylglycerols (TAGs). These TAGs are stored in membrane-bound lipid bodies that originate from the endoplasmic reticulum. Aside from use in SCO processes, *Yarrowia* is also

used in the production of organic acids like citric acid [5] and can secrete enzymes such as lipases for use in food and other industries [3]; *Yarrowia* is classified as a GRAS (Generally Recognized As Safe) organism by the U.S. Food and Drug Administration. To make these products, *Yarrowia* has the ability to metabolize a wide variety of substrates in many industrially relevant contexts, including glucose from plant cellulose and hemicellulose hydrolysates [6, 7], acetate and other volatile fatty acids (VFAs) from agro-industrial and municipal wastes [8], glycerol from industrial waste streams (including recycle streams from biodiesel manufacturing) [9 ,10], and also very hydrophobic substrates such as fatty acids and TAGs derived from animal fat and alkanes from petroleum sludge [11, 12].

The most unique advantage of using *Yarrowia lipolytica* for SCO processes has to be the prior research done on its genetics. Such efforts have yielded the complete sequence to and rough annotation of its haploid genome [13] and the development of some genetic tools [14-16] for achieving high levels of constitutive expression of native and heterologous genes. Transformation methods [17] and auxotrophic and antimycotic markers for integrating DNA cassettes into the genome have been devised for adding genes in this organism, and especially genetically tractable strains have been developed [18] that allow for high frequency of homologous recombination for the purposes of directly modifying or knocking out genes for gene function research or metabolic engineering applications. This has led to the genetic modification of strains that can synthesize lipids at higher rates, utilize non-native metabolic pathways to consume previously unusable substrates (e.g. xylose) [19], and produce customizable distributions of fatty acids of varying length and degree of unsaturation [20, 21]. As *Yarrowia* has many genes that have little to no similarity to those in other well-studied organisms, improving the gene annotations through loss-of-function studies or protein-protein or

protein-nucleotide interaction assays could continue to provide us with insight into the phenomena that are more unique to *Yarrowia*, such as hydrophobic substrate uptake, and those that are more ubiquitous or relevant to higher organisms, such as obesity, and will expand the industrial potential of this versatile organism.

2.2 Biochemistry of TAG biosynthesis in *Yarrowia lipolytica*

2.2.1 The fatty acid biosynthetic pathway

Fatty acids are the component of lipids that are esterified (if free) or transesterified (if in the form of glycerolipids or steryl esters) with alcohols to form the fatty acid alkyl esters that make up biodiesel. The biosynthesis pathway for fatty acids is pictured in **Figure 2.1**. In the first committed step of fatty acid biosynthesis, cytosolic acetyl-CoA, a central carbon metabolite used as a substrate in many biosynthetic pathways, is carboxylated by the enzyme acetyl-CoA carboxylase (Acc1) to form malonyl-CoA at the cost of one ATP [22]. Acc1 uses biotin as a cofactor, and the enzyme Acc-biotin ligase (Acc2) is required for its activity. Fatty acid synthesis begins with the binding of acetyl-CoA to a particular cysteine residue of the β-ketoacyl-ACP (Acyl-Carrier Protein) 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. For each chain, the elongation and desaturation process is performed seven times to form the 16 carbon palmitoyl-ACP, which in yeast is then transferred from the ACP to a free CoA-SH by palmitoyl transferase to form palmitoyl-CoA [23] (in animals, free palmitate is cleaved from ACP using a thiolase).

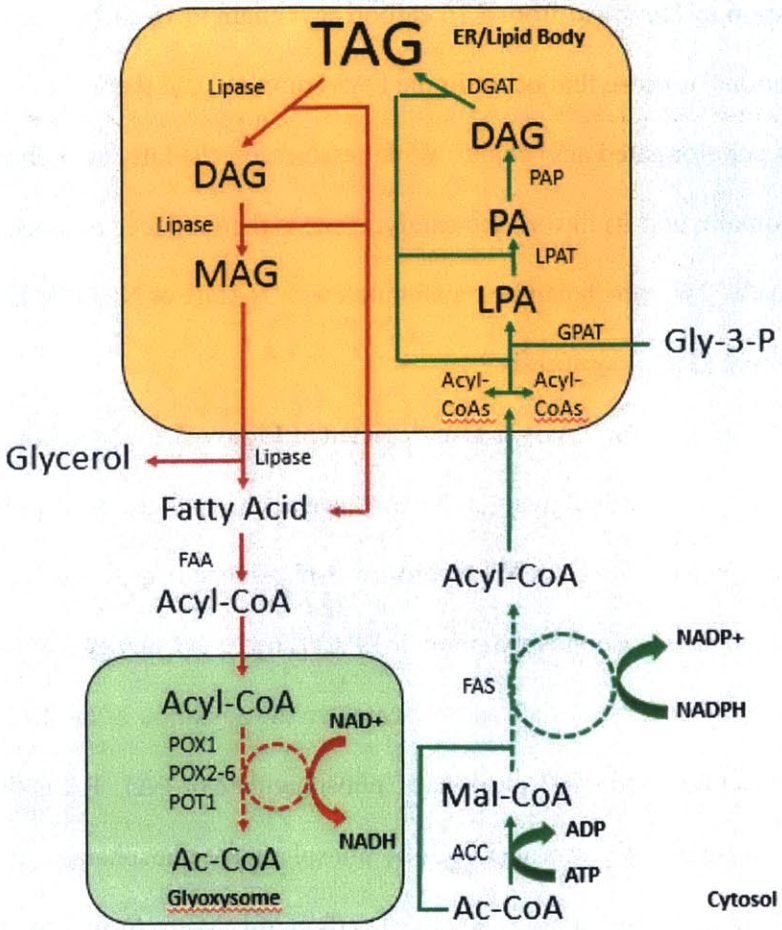

Figure 2.1: Pathways of fatty acid synthesis, TAG synthesis, and lipid degradation. Fatty acid synthesis is shown by the green arrows in the cytosol, TAG synthesis by the green arrows in the ER/Lipid Body, and lipid degradation by the red arrows.

Palmitoyl-CoA is transported into the endoplasmic reticulum, where it can be acted on by elongases and/or desaturases to form the other four fatty acyl-CoAs that *Yarrowia* incorporates into glycerolipids: along with palmitoyl-CoA (C16:0), the cells also make palmitoleoyl-CoA (C16:1), stearoyl-CoA (C18:0), oleoyl-CoA (18:1), and linoleoyl-CoA (C18:2). *Yarrowia* has two genes that may code for putative fatty acid elongases: GNS1 (NCBI Locus Tag: YALI0B20196g), and a currently unnamed gene that codes for a protein with the same GNS elongase domain (NCBI Locus Tag: YALI0F06754g). The fatty acid desaturases in *Yarrowia* include SCD1, a stearoyl-CoA delta-9 desaturase, and ODE1, an oleoyl-CoA delta-12

desaturase. Elongation in *Yarrowia* from a 16 carbon acyl chain to 18 carbons proceeds by the same chemical reactions as those that occur in the FAS complex, and therefore also requires 2 NADPH molecules per elongated acyl chain. With desaturases, the fatty acyl chain is acted on by the desaturase domain, and its diiron-oxo catalytic site is then reduced and regenerated by the cytochrome b5 domain [24], which can be regenerated with NADH or NADPH [25].

2.2.2 Triacylglycerol (TAG) synthesis

The biosynthesis pathway for TAGs is also depicted in **Figure 2.1**. 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 (Gly3P) to form 1-acylglycerol-*sn*-3-phosphate (lysophosphatidate or LPA) [26]. Next, acylglycerol-phosphate acyltransferases (AGPAT or LPAT) such as Slc1 catalyze 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, another acyl group is transferred to the third carbon of DAG to form TAG by diacylglycerol acyltransferase (DGAT). There are two DGAT genes in *Yarrowia*: DGA1 (YALI0D07986g) and DGA2 (YALI0E32769g) [27]. 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 such as Lro1. TAGs are stored in lipid bodies (or lipid particles), which are bounded in membranes that contain high fractions of phospholipids and smaller fractions of sterols; these membranes change in lipid composition to allow for increases in particle size as TAGs accumulate in the cell [28].

2.2.3 NADPH generation for fatty acid biosynthesis

For each two-carbon elongation step in fatty acid biosynthesis, two NADPH molecules are needed to reduce the growing acyl chain. The question of which pathway(s) provide(s) the main source(s) of cytosolic NADPH for fatty acid synthesis in *Yarrowia lipolytica* has received a lot of interest recently. Generally, there are three NADPH-producing pathways that are thought to provide cytosolic reducing equivalents for elongation and desaturation of acyl-CoAs, with even phylogenetically similar organisms sometimes having dramatically different flux distributions among the three: 1) malic enzyme [29], 2) NADP+-specific isocitrate dehydrogenase [29], and 3) glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase from the oxidative pentose phosphate pathway (oxPPP) [30]. Previously, it has been determined that the overexpression of the native malic enzyme (MAE1) [30] gene does not lead to improved lipid synthesis, and Mae1 in *Yarrowia* was determined to be mitochondrial with a much higher affinity for NAD+ than NADP+ [30], while Idp2, the putative NADP+-specific isocitrate dehydrogenase in *Yarrowia*, has been speculated by some to be mitochondrial as well [31]. By process of elimination, it seems that the oxPPP is most likely the main, if not the only, source of cytosolic NADPH for fatty acid synthesis. To further back this claim, the oxPPP has been shown through metabolic flux analysis (MFA) to be the main NADPH producing pathway in *Yarrowia* for lipid synthesis on glucose, as strains engineered for high lipid flux automatically increase their oxPPP flux by an amount commensurate to the extra fatty acid synthesis flux [32].

2.2.4 TAG and fatty acid degradation

The process of TAG and fatty acid degradation is also shown in **Figure 2.1** (red arrows). TAG degradation involves the sequential removal of free fatty acids from the glycerol head group by enzymes called lipases [33]. The free fatty acids are released from the lipid bodies into

the cytosol, and are then activated by fatty acyl-CoA ligase (Faa) to form fatty acyl-CoA [34]. Fatty acyl-CoA is then transported into peroxisomes or glyoxysomes, and undergoes a process called β-oxidation, the first step of which involves oxidizing the fatty acyl-CoA to a trans-2-enoyl-CoA by acyl-CoA oxidase (Pox1), which transfers electrons to FAD; in glyoxysomal β-oxidation, the $FADH_2$ formed by this step does not produce ATP as in the mitochondrial respiratory electron transport chain, but is oxidized directly by molecular oxygen [35]. Water is then added across the carbon-carbon double bond of the enoyl-CoA, and the resulting hydroxyl group is oxidized with NAD+ by one of the other Pox enzymes (Pox2-6, depending on the chain length [36]) to form a 3-ketoacyl-CoA. Lastly, a peroxisomal oxoacyl thiolase (Pot1) enzyme catalyzes the replacement of a free CoA with the first two carbons in the acyl chain, producing acetyl-CoA. Successive rounds of this β-oxidation process for are needed for each two carbon unit to be freed from the acyl-CoA chain. The enzymes of the glyoxylate cycle, such as isocitrate lyase (Icl1) and malate synthase (Mls1) are needed to produce biomass from the acetyl-CoA produced by fatty acid degradation. There is evidence to suggest that TAG degradation in *Yarrowia* readily occurs when nitrogen in the media has been depleted, even in the presence of excess carbon [6]; in many yeasts, including *Saccharomyces cerevisiae*, ICL1 and MLS1 are repressed by glucose, however, glucose repression in *Yarrowia lipolytica* is not absolute, allowing for significant fatty acid mobilization while glucose is left in the media.

2.2.5 Cytosolic acetyl-CoA generation from glucose

The pathway through which extracellular glucose becomes the cytosolic acetyl-CoA precursors for lipid synthesis is shown in **Figure 2.2**. Glucose enters the cell through one of several hexose transporters in *Yarrowia*, which vary in their substrate specificity and binding affinity. Once inside the cell, it is phosphorylated to glucose-6-phosphate (G6P) at the cost of

one ATP by one of two hexokinases (Hxk1: YALI0B22308p, Hxk2: YLAI0E20207p) or glucokinase (Glk1: YALI0E15488p). G6P can then enter glycolysis to make two pyruvate molecules, 3 ATP, and 2 NADH, or it can enter the pentose phosphate pathway and lose one carbon to decarboxylation to make 2 NADPH per pass through the pathway. DHAP from glycolysis can be also be converted to glycerol-3-phosphate (Gly3P) by NADH-consuming Gly3P dehydrogenase (Gpd), producing the head groups needed for glycerolipid synthesis. Pyruvate enters the mitochondria and is then undergoes oxidative decarboxylation by the pyruvate dehydrogenase complex (Pdh) to form acetyl-CoA, which then enters the TCA cycle.

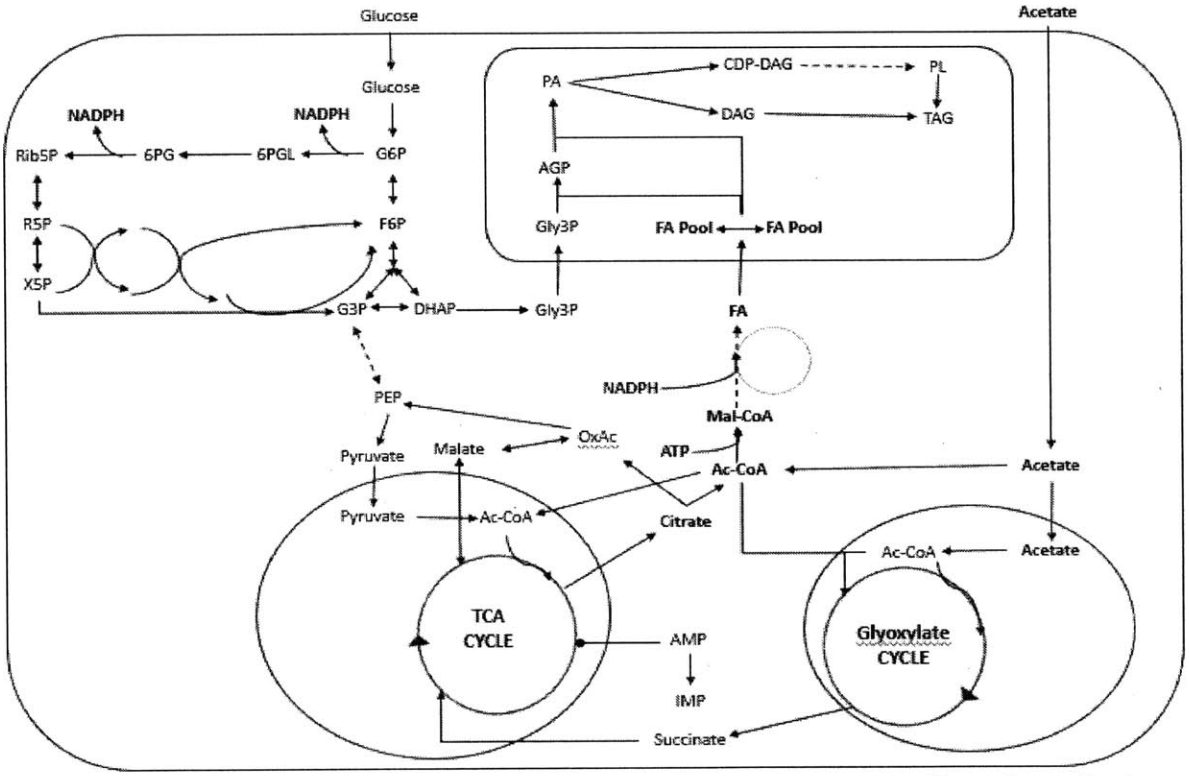

Figure 2.2: Pathways from extracellular glucose and acetate to TAGs and required NADPH and ATP.

The onset of lipogenesis in oleaginous organisms coincides with the depletion of some nutrient in the medium required for normal growth and cell division, typically nitrogen [37], but also potentially phosphate [38], sulfur [39], or certain metal ions required for proper enzyme function [40]. The process of lipogenesis begins with the conversion of AMP to IMP (inosine monophosphate) by AMP deaminase (Amd1). AMD1 is constitutively expressed, and it has been found that ATP is an allosteric activator of the enzyme [41], so it is possible that the increase in activity in this enzyme after nitrogen depletion is merely due to the increase in energy charge when the cell is unable to use ATP for biomass because of other limiting factors. AMP is an activator of the mitochondrial isocitrate dehydrogenase (Idh), so in response to the removal of AMP by AMP deaminase, the TCA cycle ceases activity, causing citrate to accumulate in and then exit the mitochondria. Once in the cytosol, citrate is cleaved by ATP:citrate lyase (Acl) to form oxaloacetate and acetyl-CoA. It is believed that the presence of ACL genes in an organism is a necessary condition for being oleaginous [37]; for example, *Saccharomyces cerevisiae*, a non-oleaginous yeast, does not have this enzyme. The acetyl-CoA produced by the Acl complex is available for fatty acid synthesis, while the oxaloacetate produced can be shuttled back into the mitochondrion using the enzymes and transporters of the malate/aspartate/oxaloacetate cycle [42].

2.2.6 Cytosolic acetyl-CoA generation from acetate

Figure 2.2 also shows the metabolic pathways from extracellular acetate to make both lipids and the required NADPH for fatty acid biosynthesis. Acetate enters the cell either in the form of acetic acid [43], or through symport with a hydrogen ion [44]; protons that get imported to the cell in this process are exported out by an (H+)-ATPase, the ATP cost of which is generally assumed to be one ATP per proton, but there is evidence that this exchange rate can vary

considerably [45-47]. Acetate is converted to acetyl-CoA by the acetyl-CoA synthetase (Acs) enzyme; this enzyme is localized in many compartments in the cell, including the cytosol, peroxisomes/glyoxysomes, and the nucleus, but there is only one isoform of Acs in *Yarrowia* (YALI0F05962p). Yeast Acs enzymes form AMP from ATP when converting acetate to acetyl-CoA [48], which puts the cost of "activating" acetate for use in downstream metabolic pathways at effectively 2 ATP per molecule when accounting for the equilibrium of ATP, ADP, and AMP.

While converting extracellular acetate to cytosolic acetyl-CoA is a very simple pathway, a fraction of this acetyl-CoA generated when growing on acetate must be used to produce NADPH from the oxPPP and ATP for acetyl-CoA carboxylase (and to cover the cost of importing and activating each acetate molecule upon entering the cell). To produce NADPH, acetyl-CoA is either synthesized in the glyoxysome or imported into it using the glyoxysomal carnitine acyltransferase system [49]. Two glyoxysomal acetyl-CoA molecules enter the glyoxylate cycle, which results in the net production of a succinate molecule and the reduction of one NAD+ to NADH. Succinate from the glyoxylate cycle can then migrate to the mitochondria, where it can be converted to malate by the reactions of the TCA cycle, producing $FADH_2$. The malate produced here then exits the mitochondrion and is converted to oxaloacetate by a cytosolic malate dehydrogenase (Mdh), producing another NADH in the process. Cytosolic oxaloacetate is then converted to phosphoenolpyruvate (PEP), an intermediate of glycolysis, by the enzyme PEP carboxykinase, at the cost of one ATP (or other nucleotide triphosphate). PEP can then go through either the glyceroneogenesis pathway to produce glycerol-3-phosphate at a cost of two NADH molecules and one ATP per PEP, or two molecules of PEP can be converted to one G6P molecule through gluconeogenesis at the cost of two ATP and two NADH. The G6P converted in this way can then be used to produce NADPH in the pentose phosphate pathway. Any

additional ATP needed for lipid synthesis or acetate metabolism may come from the TCA cycle, which requires transport of acetyl-CoA into mitochondria using the mitochondrial carnitine acetyltransferase system.

2.2.7 Regulation of neutral lipid synthesis

As a highly energy and reducing-equivalent intensive pathway, fatty acid and storage lipid biosynthesis is extensively regulated at multiple levels and at many nodes. The first committed step to fatty acid synthesis, catalyzed by Acc1, is subject to transcriptional controls (by several regulators), post-translational modifications, and effector-level feedback inhibition. Transcriptional control of ACC1 and the FAS1 and FAS2 genes in yeasts are controlled at least in part in response to inositol and choline, two precursors involved in synthesis of various phospholipids, by the transcriptional activating Ino2/Ino4 complex and by the repressor Opi1. Although Ino2 and Ino4 are not annotated as existing in *Yarrowia*, there is some evidence to suggest that the Yas1 and Yas2 proteins, the closest homologs to the Ino2/Ino4 proteins, could have similar activity [50]. ACC1, FAS1, and FAS2 are also transcriptionally upregulated during the G1 phase of the cell cycle to allow for phospholipid synthesis in growing and proliferating cells, but it is unclear if the aforementioned transcriptional regulators are involved in this process [51]. As for post-translational modifications, Acc1 is phosphorylated and inactivated by the Ser/Thr kinase Snf1, most likely on the Ser1178 residue [52], in response to a low cellular energy charge (ATP:AMP ratio). The product of the FAS complex, palmitoyl-CoA, and other fatty acyl-CoAs, also cause feedback inhibition of Acc1 activity [53].

Studies of regulation of neutral lipid synthesis enzymes seem to be limited. There is evidence to suggest that DGAT enzymes are upregulated in *Yarrowia* in response to nitrogen depletion [54]. The distribution of flux from the phosphatidic acid branch point node between

phospholipid and TAG synthesis is in part regulated in response to the availability of several phospholipid precursors; phospholipid synthesis is downregulated in response to inositol and choline, and activity of Dgk1, which catalyzes the phosphorylation of DAG to PA, is stimulated by the availability of CTP [55]. Post-translational modifications of constitutively expressed intracellular lipases such as Tgl4 by Cdc28 (cyclin-dependent kinase) may also be needed for their activity, and they occur in *Saccharomyces* as part of the transition from stationary phase in response to sensing the nutrients required for growth [56].

2.3 Pathway-based analysis for calculation of theoretical yield of fatty acids

The theoretical maximum yield is a critical parameter needed for evaluating the utility and economic viability of any chemical process. Estimates of maximum yield can be obtained through simple balances of elemental components and degrees of reduction between substrate and product. However, for processes involving microbial biocatalysis, for many substrate-product pairings, the most likely metabolic pathways that will be utilized by the host organism are known. In many cases, including the knowledge of these pathways and using a system of equations that take into account the intermediates, unavoidable byproducts, and the need to balance the production of energy and reducing equivalent cofactors in the metabolic process leads to a significantly lower, and necessarily more realistic, theoretical maximum yield.

2.3.1 Maximum yield of palmitate from glucose

From the description of the fatty acid biosynthesis pathway in **Section 2.2.1** and the TAG biosynthesis pathway in **Section 2.2.2**, we can determine that the biochemical equation for the production of three moles of palmitate in the form of one mole of glyceryl tripalmitate (C16:0-TAG) is the following:

24 Acetyl-CoA + Gly3P + 21 ATP + 42 NADPH → Glyceryl tripalmitate

Using this equation, the knowledge that the oxidative pentose phosphate pathway is the major NADPH producing pathway in *Yarrowia*, and the information about the most direct path from glucose to cytosolic acetyl-CoA presented in **Section 2.2.5**, we can construct a system of equations describing the intermediates and cofactors created in the process, all of which must be balanced. A condensed form of this system of equations is shown in **Table 2.1**. In writing the system of equations in this way, determining the coefficients multiplying each equation (signifying the relative flux through each equation's pathway) to achieve no net consumption or generation of any intermediate or cofactor (aside from excess ATP, which can be removed by the cell using one of many futile cycles) is a relatively trivial process. **Table 2.2** shows the correct linear combination of equations from **Table 2.1** needed to calculate the theoretical yield palmitate in glyceryl tripalmitate from glucose, which is 0.266 g palmitate/g glucose.

Table 2.1: Pathway equations for determining theoretical yield of palmitate from glucose.

Rxn Product	Glucose	G6P	Ac-CoA.c	Ac-CoA.m	ATP	NADH	NADPH	FADH2	G3P	Palmitate
Overall TAG Synthesis	0	0	-24	0	-21	0	-42	0	-1	3
TCA Cycle	0	0	0	-1	1	3	0	1	0	0
NADPH (PPP)	0	-1	0	0	0	0	12	0	0	0
Glycolysis (G3P)	0	-1	0	0	-1	-2	0	0	2	0
Glycolysis/PyrDH	0	-1	0	2	3	4	0	0	0	0
Hexokinase	-1	1	0	0	-1	0	0	0	0	0
ETC (NADH)	0	0	0	0	2.5	-1	0	0	0	0
ETC (FADH2)	0	0	0	0	1.5	0	0	-1	0	0
TCA-Citrate-ACL	0	0	1	-1	-1	0	0	0	0	0

Table 2.2: Solved system of equations for determining theoretical yield of palmitate from glucose.

Rxn Product	Glucose	G6P	Ac-CoA.c	Ac-CoA.m	ATP	NADH	NADPH	FADH2	G3P	Palmitate
Overall TAG Synthesis	0	0	-24	0	-21	0	-42	0	-1	3
NADPH (PPP)	0	-3.5	0	0	0	0	42	0	0	0
Glycolysis (G3P)	0	-0.5	0	0	-0.5	-1	0	0	1	0
Glycolysis/PyrDH	0	-12	0	24	36	48	0	0	0	0
TCA-Citrate-ACL	0	0	24	-24	-24	0	0	0	0	0
Total:	0	-16	0	0	-9.5	47	0	0	0	3
ATP Value:	0	-16	0	0	-9.5	117.5	0	0	0	0
		92	Excess ATP			16	Glucose to 3 Palmitate			
							MW	Mass	Theoretical Yield	
						Glucose	180.16	2882.56		
						Palmitate	255.42	766.26	0.266	

2.3.2 Maximum yield of palmitate from acetate

Using the same basis equation and general methods, the theoretical yield of palmitate in C16:0-TAG from acetate can be calculated. As discussed in **Section 2.2.6**, while the pathway from extracellular acetate to cytosolic acetyl-CoA is simple, involving the expenditure of two ATP through the Acs enzyme, most of the pathways described in the equations used to determine theoretical yield are for the process of generating NADPH and Gly3P from acetyl-CoA. This system of equations is given in **Table 2.3**. In taking intracellular pH maintenance during acetate import into account in these theoretical yield calculations, it is important to note that the amount of protons pumped out of the cell per ATP hydrolyzed can vary based on the cell's environment [47]; as the pH of the medium increases sharply at the beginning of our acetate fermentations with *Yarrowia*, it is not unreasonable to assume that the lower (or more negative) free energy associated with pumping protons across the membrane into the more alkaline medium will not require an equivalent number of ATP molecules hydrolyzed. As the ATP cost of intracellular pH

maintenance is variable and unknown in our specific process, we did not include this as a constraint to the theoretical yield in our calculations.

Table 2.3: Pathway equations for determining theoretical yield of palmitate from acetate.

Rxn Product	G6P	Ac-CoA	ATP	NADH	NADPH	FADH2	Succ	OxAc	PEP	Gly3P	Palmitate
Overall TAG-Palmitate	0	-24	-21	0	-42	0	0	0	0	-1	3
Pentose Phosphate Pathway Cycle	-1	0	0	0	12	0	0	0	0	0	0
Glyoxylate Cycle	0	-2	0	1	0	0	1	0	0	0	0
Succinate to OxAc.c	0	0	0	1	0	1	-1	1	0	0	0
PEPCK	0	0	-1	0	0	0	0	-1	1	0	0
Ac-CoA to PEP (Sum of Previous 3 lines)	0	-2	-1	2	0	1	0	0	1	0	0
Gluconeogenesis	1	0	-2	-2	0	0	0	0	-2	0	0
Glyceroneogenesis	0	0	-1	-2	0	0	0	0	-1	1	0
Acetate Activation	0	1	-2	0	0	0	0	0	0	0	0
TCA Cycle - Electron Transport	0	-1	10	0	0	0	0	0	0	0	0

Table 2.4 shows the correct linear combination of equations from **Table 2.3** needed to solve for the maximum possible yield of palmitate in C16:0-TAG from acetate. Due to the large ATP cost in activating the acetate used by the cell, the cell runs a significant energy deficit by producing the cytosolic acetyl-CoA and NADPH needed for fatty acid synthesis. The most efficient way to produce ATP to make up this energy deficit is through the TCA cycle, which produces 1 ATP, 1 FADH (which can be oxidized to produce 1.5 ATP, and 3 NADH (which can each be oxidized to produce 2.5 ATP); each additional acetate needed by the cell to replenish ATP in this way produces eight net ATP. In adding the 24 mol acetate needed to make the carbon in 3 mol palmitate, the 16 mol acetate needed to make the Gly3P and NADPH, and the 10.9375 mol of acetate needed to make up the ATP deficit, **Table 2.4** shows that the theoretical maximum yield of palmitate in *Yarrowia* on acetate is approximately 0.255 g palmitate/g acetate.

Table 2.4: Solved system of equations for determining theoretical yield of palmitate from glucose.

Rxn Product	G6P	Ac-CoA	ATP	NADH	NADPH	FADH2	Succ	OxAc	PEP	Gly3P	Palm
Basis	0	-24	-21	0	-42	0	0	0	0	-1	3
Glyceroneogenesis	0	0	-1	-2	0	0	0	0	-1	1	0
Pentose Phosphate Pathway Cycle	-3.5	0	0	0	42	0	0	0	0	0	0
Gluconeogenesis	3.5	0	-7	-7	0	0	0	0	-7	0	0
Ac-CoA to PEP (Sum of Previous 3 lines)	0	-16	-8	16	0	8	0	0	8	0	0
Total:	0	-40	-37	7	0	8	0	0	0	0	3
ATP Value:		-80	-37	17.5		12					
		-87.5	ATP Deficit			50.9375	Ac for 3 Palmitate				
		10.9375	More Ac Needed				MW	Mass	Theoretical Yield		
						Ac	59.04	3007.35			
						Palm	255.42	766.26	0.255		

2.4 Previous efforts in metabolic engineering of *Yarrowia lipolytica* for SCO processes

While wild-type *Yarrowia lipolytica* is capable of producing very significant quantities of fatty acids that can be used towards the production of biodiesel, the biggest advantage in using this specific organism over other oleaginous species is the development of genetic tools to further increase lipid productivity or yield, or produce different distributions of fatty acids, including those that are not natively synthesized. With the development and characterization of several promoters, genes in *Yarrowia* can be overexpressed constitutively with varying magnitude [14-16], or induced in the presence of certain carbon sources or additives to growth media [19]. Heterologous genes and pathways can also be added using these promoters and given tunable expression profiles. By transforming DNA cassettes with large (> 1 kb) homology

regions, it is possible (in some strains) to target loci and disrupt genes of interest with auxotrophic or antimycotic marker genes.

Previous work in our lab involved the overexpression of the native acetyl-CoA carboxylase ACC1 gene (YALI0C11407g) under the constitutive hp4d promoter and the diacylglycerol acyltransferase DGA2 gene (YALI0E32769g) under the stronger constitutive TEF1α promoter to increase TAG production in *Yarrowia lipolytica* [16]. Overexpression of ACC1 led to a roughly twofold increase in lipid content and threefold increase in lipid yield over wild-type cells, while overexpression of DGA2 led to a roughly fourfold increase in content and yield in low C:N (~20) glucose media; overexpression of both genes simultaneously revealed that these modifications' effects were additive, causing a 4.7-fold increase in lipid content and a nearly fivefold increase in lipid yield over wild-type. The latter strain, called MTYL065, produced a 61.7% lipid content and a lipid yield of 0.195 g lipid/g glucose in a bioreactor with 90 g/L glucose at a C:N of 100. This strain performs similarly well when using acetate as the carbon source in a bioreactor, capable of achieving the same lipid content (~62%) and high yields (0.152 g lipids/g acetate) [19]. Due to the high productivity and versatility of this strain, the ACC1 and DGA2 overexpression modifications are used throughout this work as a basis for further strain engineering.

Heterologous pathway engineering can also be used to improve lipid productivity in *Yarrowia* or add to the diversity of products that can be synthesized or substrates that can be metabolized. Significant quantities of eicosapentaenoic acid (EPA, C20:5), an important nutritional supplement, were synthesized in *Yarrowia* by adding several copies of four genes from various algae, euglenoids, and other fungi: a Δ9-elongase, a Δ8-desaturase, a Δ5-desaturase, and a Δ17-desaturase [21]. In previous work in our lab, *Yarrowia* was engineered to

consume xylose, a five carbon sugar that is normally unable to be metabolized by *Yarrowia*, through heterologous expression of the xylose reductase XYL1 and xylitol dehydrogenase XYL2 genes from *Scheffersomyces stipitis* [19].

Gene knockouts can be an effective way to remove metabolic pathways that compete with or reverse the pathways of interest, and can also be used to remove regulatory elements that could otherwise limit flux through product-forming pathways. However, achieving gene targeting through homologous recombination of transformed DNA cassettes is difficult in many strains; fortunately, studies have shown that the removal of the KU70 gene in *Yarrowia* leads to a strain that is significantly more genetically tractable and more frequently uses homologous recombination for DNA cassette integration [18]. As *Yarrowia* has been shown to have appreciable TAG mobilization activity during stationary phase, accumulation and *de novo* synthesis of lipids has been shown to improve when the intracellular lipases TGL3 and TGL4 [56] or the POX genes of the fatty acid β-oxidation pathway [36] are disrupted. Removal of SNF1, a gene encoding a protein kinase and inhibitor of Acc1, has been shown to increase the lipid content of *Yarrowia* [57].

Using these tools, significant progress in lipid production in *Yarrowia lipolytica* has already been achieved. As our understanding of the intricacies of the metabolic network becomes clearer, and the functions of previously uncharacterized genes and genetic elements are investigated, the potential for the use of this organism in industrial-scale production of lipid-derived chemicals for use as fuels, nutritional supplements, or in a variety of other applications gets closer to becoming realized.

References:

1) Meng, X., Yang, J., Xu, X., Zhang, L., Nie, Q., & Xian, M. (2009). Biodiesel production from oleaginous microorganisms. *Renewable Energy*, *34*(1), 1–5. doi:10.1016/j.renene.2008.04.014
2) Rude, M. a., & Schirmer, A. (2009). New microbial fuels: a biotech perspective. *Current Opinion in Microbiology*, *12*(3), 274–281. doi:10.1016/j.mib.2009.04.004
3) Coelho, M. a Z., Amaral, P. F. F., & Belo, I. (2010). Yarrowia lipolytica : an industrial workhorse. *Applied Microbiology and Microbial Biotechnology*, 930–944.
4) 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. *Progress in Lipid Research*, *48*(6), 375–87. doi:10.1016/j.plipres.2009.08.005
5) Papanikolaou, S., Galiotou-Panayotou, M., Fakas, S., Komaitis, M., & Aggelis, G. (2008). Citric acid production by Yarrowia lipolytica cultivated on olive-mill wastewater-based media. *Bioresource Technology*, *99*(7), 2419–2428. doi:10.1016/j.biortech.2007.05.005
6) 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. *European journal of lipid science and technology*, *111*(12), 1221-1232.
7) Stephanopoulos, G. (2007). Challenges in engineering microbes for biofuels production. *Science*, (February), 801–805.
8) 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 Technology*, *114*, 443–9. doi:10.1016/j.biortech.2012.02.091
9) 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. *Current Microbiology*, *46*(2), 0124-0130.
10) Papanikolaou, S., & Aggelis, G. (2002). Lipid production by Yarrowia lipolytica growing on industrial glycerol in a single-stage continuous culture, *82*(July 2001), 43–49.
11) 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. *Applied Microbiology and Biotechnology*, *58*(3), 308–12. doi:10.1007/s00253-001-0897-0
12) Fickers, P., Benetti, P.-H., Waché, Y., Marty, a, Mauersberger, S., Smit, M. S., & Nicaud, J.-M. (2005). Hydrophobic substrate utilisation by the yeast Yarrowia lipolytica, and its potential applications. *FEMS Yeast Research*, *5*(6-7), 527–43. doi:10.1016/j.femsyr.2004.09.004
13) Dujon, B., Sherman, D., Fischer, G., Durrens, P., Casaregola, S., Lafontaine, I., ... Souciet, J.-L. (2004). Genome evolution in yeasts. *Nature*, *430*(6995), 35–44. doi:10.1038/nature02579
14) Madzak, C., Tréton, B., & Blanchin-Roland, S. (2000). Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous

proteins in the yeast Yarrowia lipolytica. *Journal of Molecular Microbiology and Biotechnology*, *2*(2), 207–216.
15) Blazeck, J., Liu, L., Redden, H., & Alper, H. (2011). Tuning gene expression in yarrowia lipolytica by a hybrid promoter approach. *Applied and Environmental Microbiology*, *77*(22), 7905–7914. doi:10.1128/AEM.05763-11
16) Tai, M., & Stephanopoulos, G. (2013). Engineering the push and pull of lipid biosynthesis in oleaginous yeast Yarrowia lipolytica for biofuel production. *Metabolic Engineering*, *15*, 1–9. doi:10.1016/j.ymben.2012.08.007
17) Chen, D. C., Beckerich, J. M., & Gaillardin, C. (1997). One-step transformation of the dimorphic yeast Yarrowia lipolytica. *Applied Microbiology and Biotechnology*, *48*, 232–235. doi:10.1007/s002530051043
18) Verbeke, J., Beopoulos, A., & Nicaud, J.-M. (2013). Efficient homologous recombination with short length flanking fragments in Ku70 deficient Yarrowia lipolytica strains. *Biotechnology Letters*, *35*(4), 571–6. doi:10.1007/s10529-012-1107-0
19) Tai, M. (2012). Metabolic Engineering of Oleaginous Yeast for the Production of Biofuels. Ph.D. Thesis. Massachusetts Institute of Technology.
20) Zhang, B., Chen, H., Li, M., Gu, Z., Song, Y., Ratledge, C., ... Chen, W. (2013). Genetic engineering of Yarrowia lipolytica for enhanced production of trans-10, cis-12 conjugated linoleic acid. *Microbial Cell Factories*, *12*(1), 70. doi:10.1186/1475-2859-12-70
21) Xue, Z., Sharpe, P. L., Hong, S.-P., Yadav, N. S., Xie, D., Short, D. R., ... Zhu, Q. (2013). Production of omega-3 eicosapentaenoic acid by metabolic engineering of Yarrowia lipolytica. *Nature Biotechnology*, *31*(8), 734–40. doi:10.1038/nbt.262
22) Wakil, S. J., Stoops, J. K., & Joshi, V. C. (1983). Fatty acid synthesis and its regulation. *Annual Review of Biochemistry*, *52*, 537–79. doi:10.1146/annurev.bi.52.070183.002541
23) Engeser, H., Hübner, K., Straub, J., & Lynen, F. (1979). Identity of malonyl and palmitoyl transferase of fatty acid synthetase from yeast. *European Journal of Biochemistry / FEBS*, *101*(2), 413–422.
24) Martin, C. E., Oh, C.-S., & Jiang, Y. (2007). Regulation of long chain unsaturated fatty acid synthesis in yeast. *Biochimica et Biophysica Acta*, *1771*(3), 271–285. doi:10.1016/j.bbalip.2006.06.010
25) Vergéres, G., & Waskell, L. (1995). Cytochrome b5, its functions, structure and membrane topology. *Biochimie*, *77*(7-8), 604–620. doi:10.1016/0300-9084(96)88176-4
26) Coleman, R. A., & Lee, D. P. (2004). Enzymes of triacylglycerol synthesis and their regulation. *Progress in Lipid Research*, *43*(2), 134–176. doi:10.1016/S0163-7827(03)00051-1
27) 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
28) Thiam, A. R., Farese, R. V, & Walther, T. C. (2013). The biophysics and cell biology of lipid droplets. *Nature Reviews. Molecular Cell Biology*, *14*(12), 775–86. doi:10.1038/nrm369
29) 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
30) 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. *Biotechnology Letters*, *35*(12), 2091–8. doi:10.1007/s10529-013-1302-7

31) Ratledge, C. (2014). The role of malic enzyme as the provider of NADPH in oleaginous microorganisms: A reappraisal and unsolved problems. *Biotechnology Letters*, *36*(8), 1557–1568. doi:10.1007/s10529-014-1532-3

32) 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. *Metabolic Engineering*, *30*, 27–39. doi:10.1016/j.ymben.2015.02.007

33) Czabany, T., Athenstaedt, K., & Daum, G. (2007). Synthesis, storage and degradation of neutral lipids in yeast. *Biochimica et Biophysica Acta*, *1771*(3), 299–309. doi:10.1016/j.bbalip.2006.07.001

34) Hiltunen, J. K., Mursula, A. M., Rottensteiner, H., Wierenga, R. K., Kastaniotis, A. J., & Gurvitz, A. (2003). The biochemistry of peroxisomal β-oxidation in the yeast Saccharomyces cerevisiae. *FEMS Microbiology Reviews*, *27*(1), 35–64. doi:10.1016/S0168-6445(03)00017-2

35) Nelson, D. L. & Cox, M. M. (2005). *Lehninger principles of biochemistry* (4th ed.). New York: W.H. Freeman.

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

37) Ratledge, C. (2002). Regulation of lipid accumulation in oleaginous micro-organisms. *Biochemical Society Transactions*, *30*(6), 47–50.

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

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

40) Rossi, M., Amaretti, A., Raimondi, S., & Leonardi, A. (2011). Getting lipids for biodiesel production from oleaginous fungi. *Feedstocks and Processing Technologies*.

41) 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. *The Journal of Biological Chemistry*, *264*(35), 21422–21430.

42) Minárik, P., Tomásková, N., Kollárová, M., & Antalík, M. (2002). Malate dehydrogenases--structure and function. *General Physiology and Biophysics*, *21*(3), 257–265.

43) 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

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

45) Ambesi, A., Miranda, M., Petrov, V. V, & Slayman, C. W. (2000). Biogenesis and function of the yeast plasma-membrane H(+)-ATPase. *The Journal of Experimental Biology*, *203*(Pt 1), 155–160.
46) Stephanopoulos, G. N., Aristidou, A. A., Nielsen, J. (1998). *Metabolic Engineering: Principles and Methodologies*. Academic Press, San Diego, CA.
47) Venema, K., & Palmgren, M. G. (1995). Metabolic modulation of transport coupling ratio in yeast plasma membrane H+-ATPase. *Journal of Biological Chemistry*. doi:10.1074/jbc.270.33.19659
48) Jogl, G., & Tong, L. (2004). Crystal Structure of Yeast Acetyl-Coenzyme A Synthetase in Complex with AMP. *Biochemistry*, *43*(6), 1425–1431. doi:10.1021/bi035911a
49) Papanikolaou, S., & Aggelis, G. (2010). Yarrowia lipolytica: A model microorganism used for the production of tailor-made lipids. *European Journal of Lipid Science and Technology*, *112*(6), 639–654. doi:10.1002/ejlt.200900197
50) 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. *The Journal of Biological Chemistry*, *284*(11), 7126–7137. doi:10.1074/jbc.M806864200
51) Tehlivets, O., Scheuringer, K., & Kohlwein, S. D. (2007). Fatty acid synthesis and elongation in yeast. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1771*(3), 255–270. doi:10.1016/j.bbalip.2006.07.004
52) Choi, J. W., & Da Silva, N. a. (2014). Improving polyketide and fatty acid synthesis by engineering of the yeast acetyl-CoA carboxylase. *Journal of Biotechnology*, *187*, 56–59. doi:10.1016/j.jbiotec.2014.07.430
53) Shirra, M. K., Patton-Vogt, J., Ulrich, a, Liuta-Tehlivets, O., Kohlwein, S. D., Henry, S. a, & Arndt, K. M. (2001). Inhibition of acetyl coenzyme A carboxylase activity restores expression of the INO1 gene in a snf1 mutant strain of Saccharomyces cerevisiae. *Molecular and Cellular Biology*, *21*(17), 5710–5722. doi:10.1128/MCB.21.17.5710-5722.2001
54) Morin, N., Cescut, J., Beopoulos, A., Lelandais, G., Le Berre, V., Uribelarrea, J.-L., … 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
55) 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
56) Dulermo, T., Tréton, B., Beopoulos, A., Gnankon, A. P. K., Haddouche, R., & Nicaud, J. M. (2013). Characterization of the two intracellular lipases of Y. lipolytica encoded by TGL3 and TGL4 genes: New insights into the role of intracellular lipases and lipid body organisation. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1831*(9), 1486–1495. doi:10.1016/j.bbalip.2013.07.001
57) Seip, J., Jackson, R., He, H., Zhu, Q., & Hong, S. P. (2013). Snf1 is a regulator of lipid accumulation in Yarrowia lipolytica. *Applied and Environmental Microbiology*, *79*(23), 7360–7370. doi:10.1128/AEM.02079-13

Chapter 3

Functional Overexpression and Characterization of Individual Lipogenesis-Related Genes in *Yarrowia lipolytica*

Adapted from:

Silverman, A. M.*, Qiao, K.*, Xu, P., and Stephanopoulos, G. Functional overexpression and characterization of lipogenesis-related genes in the oleaginous yeast *Yarrowia lipolytica.* Submitted.

3.1 Introduction

Concerns over energy security and global climate change have necessitated the development of alternative sources of liquid fuels from renewable and domestic feedstocks [1,2]. Liquid biofuels such as biodiesel are particularly attractive for use in transportation due to their high energy density and compatibility with existing infrastructure [3,4]. 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 [5].

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 [6], 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 [7,8], acetate and other volatile fatty acids (VFAs) from agro-industrial and municipal wastes [9], glycerol from industrial waste streams (including recycle streams from biodiesel manufacturing) [10,11], and also very hydrophobic substrates such as fatty acids and TAGs derived from animal fat and alkanes from petroleum sludge [12, 13]. Conversion to biodiesel involves the transesterification of the TAGs into fatty acid alkyl esters [14], which can be performed in multiple ways [15, 16]. *Y. lipolytica* is also the most genetically tractable oleaginous yeast, with a fully sequenced (albeit roughly annotated) genome [17], and a good

arsenal 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 [18-21].

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 *Yarrowia* converts these substrates to lipids are very different, and summarized in **Figure 3.1**. 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, [22], but also phosphate [23], sulfur [24], or certain metal ions required for proper enzyme function [25]. 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) [22]. 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 [26]. 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 [27], 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) [28]. 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 [22]. Previous research has determined that malic enzyme in *Yarrowia* is not responsible for producing appreciable amounts of NADPH for facilitating lipid synthesis [29].

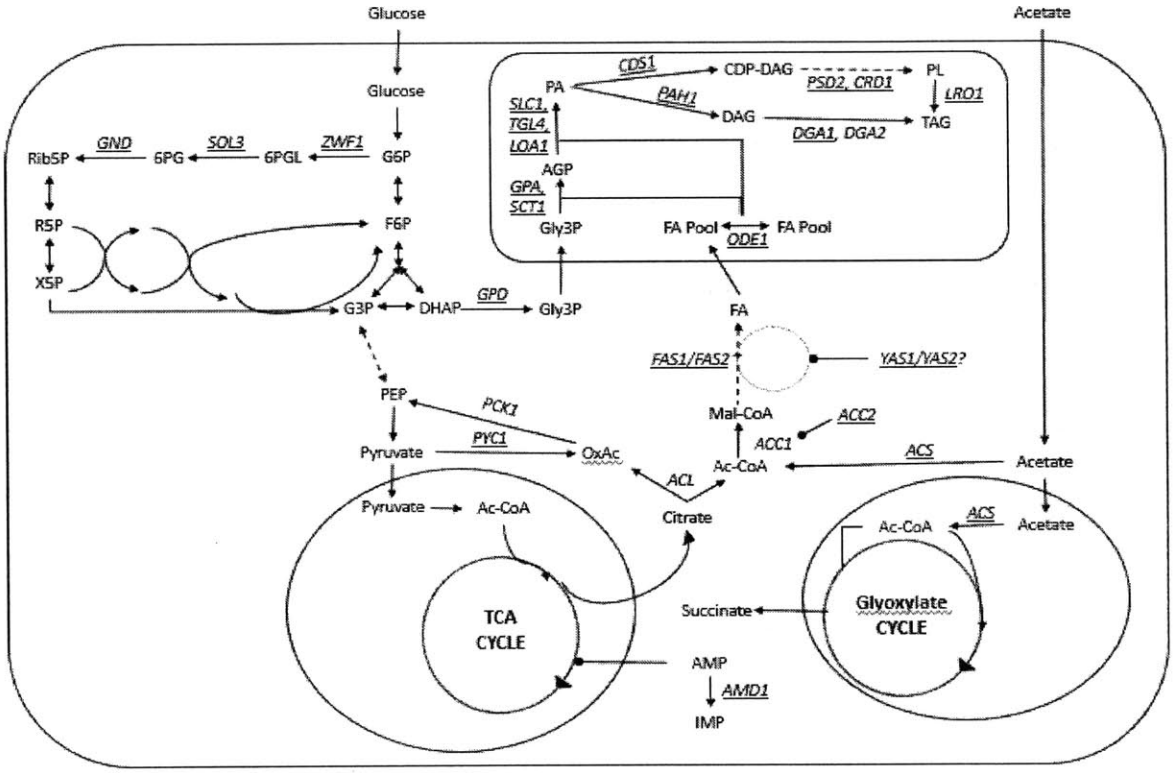

Figure 3.1: Map of the metabolic pathways from substrate (glucose or acetate) to lipids. Enzymes are written in italics, and those studied are also underlined. Arrows with circular ends denote activation of the reaction or enzyme they point to. G6P: glucose-6-phosphate; 6PGL: 6-phosphogluconolactone; 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.

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) [27]. 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 [30].

The pathway for converting acetate to lipids differs significantly from that of glucose. Yeasts normally uptake acetate trough symport with protons at higher pH [31] or passive diffusion (as acetic acid) at lower pH [32]. Once in the cell, dissociated acetate can be converted by the Acetyl-CoA Synthase (ACS) 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 OAA 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 [33].

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 [17]. 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 [34]. In *Y. lipolytica,* overexpression of enzymes directly involved in lipid synthesis such as ACC1 and DGA2 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 [20]. Overexpressions of the native NAD+-dependent glycerol-3-phopshate dehydrogenase GPD1 [35] and hexokinase HXK1 [36] in *Yarrowia* 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 DGA2-overexpression increasing total lipid content by 3-fold, *Yarrowia* 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 [28].

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 [34]. 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 [37, 38]. As *de novo* lipid production and storage is a complex phenomenon, here we attempt to gain a better understanding of the effect of genes deemed most important to this process by engineering several strains of *Y. lipolytica* such that that each overexpress a single native enzyme with an additional copy under the strong TEF-1α promoter (previously characterized [20]) 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 *Yarrowia* and similar oleaginous yeasts, we thought it potentially useful to include some genes from *Yarrowia* 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 [39-43]. 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 3.1**, 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 [20]. 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.

Table 3.1: Genes Characterized in this Study

Gene Number	NCBI Locus Tag	Gene Name	Likely Function
Control			
LZ	b0344	LacZ	beta-galactosidase from *E. coli*; strains overexpressing this gene are control strains as in [20]
Kennedy Pathway Genes/Glycerolipid Synthesis			
OP1	YALI0D07986g	DGA1	diacylglycerol acyltransferase
OP2	YALI0E16797g	LRO1	phospholipid:diacylglycerol acyltransferase
OP3	YALI0A10362g	GPA	glyerol-3-phosphate acyltransferase
OP4	YALI0D27016g	PAH1	lipin, phosphatidate phosphatase
OP5	YALI0F06578g	ARE	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	CRD1	cardiolipin synthase
OP12	YALI0B18854g	EEB1	putative carboxylic ester hydrolase; may have some fatty acid ethyl ester synthesis activity
OP13	YALI0D03480g	PSD2	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	LOA1	lysophosphatidic acid acyltransferase
Central Carbon Metabolism			
OP19	YALI0F05962g	ACS	acetate-CoA ligase
OP20	YALI0B02948g	GPD	cytosolic NAD+-dependent glycerol-3-phosphate dehydrogenase
OP21	YALI0C24101g	PYC1	cytosolic pyruvate carboxylase, involved in transhydrogenase cycle with malic enzyme
Fatty Acid Synthesis			
OP22	YALI0E30591g	ACC2	acetyl-CoA carboxylase-biotin ligase
OP23	YALI0B15059g	FAS1	fatty acid synthase subunit beta
OP24	YALI0B19382g	FAS2	fatty acid synthase subunit alpha
OP25	YALI0B10153g	ODE1	oleate 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	GND	NADP+-dependent 6-phosphogluconate dehydrogenase
OP29	YALI0E11671g	SOL3	6-phosphogluconolactonase
OP30	YALI0F04095g	IDP	cytosolic NADP+-specific isocitrate dehydrogenase
OP31	YALI0F17820g	GDH	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 [44]
OP33	YALI0F07084g	TOR	target of rapamycin protein, phosphoinositide 3-kinase-related protein kinase, master regulator involved in promoting growth (including lipid synthesis) in response to nutrient availabilty [45]
OP34	YALI0C02387g	YAS1	HLH transcription factor, upregulates cytochrome p450 genes in response to alkanes with YAS2 [46]; 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 [46]; 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 [44]
OP37	YALI0A08077g	Glc7	catalytic subunit of Reg1-Glc7 protein phosphatase 1 (PP1) complex, PP1 can inactivate SNF1, which can inactivate ACC1 [44]
Transporters			
OP38	YALI0E01298g	SCP	sterol carrier protein; may be involved in fatty acid transport
OP39	YALI0D20416g	SLY1	Golgi transport protein; similar gene upregulated in *R. toruloides* during lipogenesis [39, 40]
OP40	YALI0C20265g	ABC2	ABC (ATP-binding cassette) transporter; similar gene upregulated in *R. toruloides* during lipogenesis [39, 40]
Miscellaneous Genes			
OP41	YALI0E11495g	AMD1	adenosine monophosphate (AMP) deaminase, depletes cellular AMP to divert carbon from TCA cycle towards lipogenesis [22]
OP42	YALI0D11726g	PSR	probable transmembrane protein phosphatase for stress response; similar gene upregulated in *R. toruloides* during lipogenesis [39, 40]
OP43	YALI0A18590g	SKY1	probable protein kinase regulating RNA metabolism; similar gene upregulated in *R. toruloides* during lipogenesis [39, 40]
OP44	YALI0A09845g	YIP3	probable Rab-GDI dissociation factor, involved in maintenance of intracellular membranes; similar gene upregulated in *R. toruloides* during lipogenesis [39, 40]

Table 3.2: Plasmids and Strains Used in this Study

Gene Number	Gene Name	Expression Plasmid	*Corresponding Yarrowia* Strain
LZ	LacZ	pMT037	MTYL037
OP1	DGA1	pOP1	YLOP1
OP2	LRO1	pOP2	YLOP2
OP3	GPA	pOP3	YLOP3
OP4	PAH1	pOP4	YLOP4
OP5	ARE	pOP5	YLOP5
OP6	SLC1	pOP6	YLOP6
OP7	TGL4	pOP7	YLOP7
OP8	ALE1	pOP8	YLOP8
OP9	SCT1	pOP9	YLOP9
OP10	CDS1	pOP10	YLOP10
OP11	CRD1	pOP11	YLOP11
OP12	EEB1	pOP12	YLOP12
OP13	PSD2	pOP13	YLOP13
OP14	PDR16	pOP14	YLOP14
OP15	TGL1	pOP15	YLOP15
OP16	CSR1	pOP16	YLOP16
OP17	SAC1	pOP17	YLOP17
OP18	LOA1	pOP18	YLOP18
OP19	ACS	pOP19	YLOP19
OP20	GPD	pOP20	YLOP20
OP21	PYC1	pOP21	YLOP21
OP22	ACC2	pOP22	YLOP22
OP23	FAS1	pOP23	YLOP23
OP24	FAS2	pOP24	YLOP24
OP25	ODE1	pOP25	YLOP25
OP26	SLD	pOP26	YLOP26
OP27	ZWF1	pOP27	YLOP27
OP28	GND	pOP28	YLOP28
OP29	SOL3	pOP29	YLOP29
OP30	IDP	pOP30	YLOP30
OP31	GDH	pOP31	YLOP31
OP32	SAP190	pOP32	YLOP32
OP33	TOR	pOP33	YLOP33
OP34	YAS1	pOP34	YLOP34
OP35	YAS2	pOP35	YLOP35
OP36	Reg1	pOP36	YLOP36
OP37	Glc7	pOP37	YLOP37
OP38	SCP	pOP38	YLOP38
OP39	SLY1	pOP39	YLOP39
OP40	ABC2	pOP40	YLOP40
OP41	AMD1	pOP41	YLOP41
OP42	PSR	pOP42	YLOP42
OP43	SKY1	pOP43	YLOP43
OP44	YIP3	pOP44	YLOP44

3.2 Materials and Methods

3.2.1 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 3.2**.

Standard growth conditions for *E. coli* for the purposes of plasmid construction and generation have been previously described [47]. Rich YPD medium for *Y. lipolytica* strains contains 10 g/L yeast extract (BD Biosciences, New Jersey), 20 g/L Bacto^{TM} 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^{TM} 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.

3.2.2 Plasmid and yeast strain construction and verification

The coding sequences of all genes tested in this study were integrated into plasmid pMT015, described previously [20]. 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 [48] 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.

Gibson Assembly products were transformed into DH10β using standard techniques [47] 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 NotI or SacII restriction enzymes (NEB) and transformed into yeast using the lithium acetate method [49]. Yeast transformants were recovered on plates of YSMg 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).

3.2.3 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 [20], 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, Grand Island, NY), 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 [50] 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 *Yarrowia* actin gene (YALI0D08272g) was used as a reference. Two replicate wells were analyzed for each gene and strain combination.

3.2.4 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 [16]. 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.

3.3 Results and Discussion

3.3.1 Overexpression of each strain's respective gene was verified and quantified by qRT-PCR

Previous work in our lab characterized the full *Yarrowia* TEF-1α promoter and demonstrated its effectiveness in driving high expression levels of its target gene [20]. We placed each of the 44 genes described in **Table 3.1** 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 3.2**.

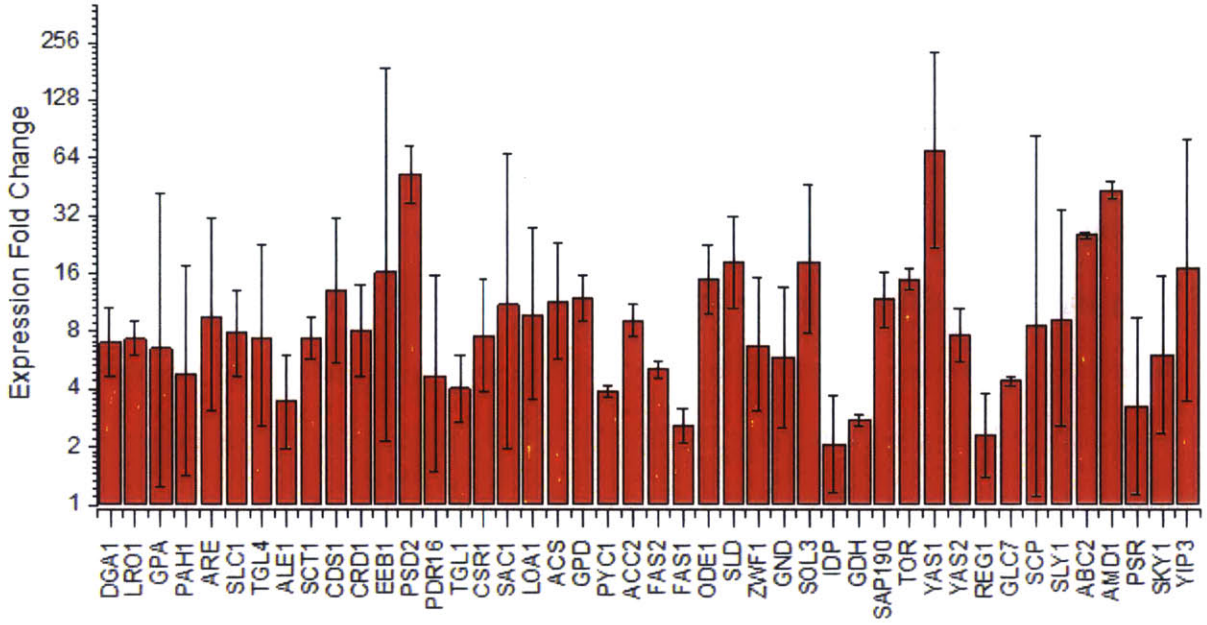

Figure 3.2: Expression of genes in this study with an additional copy under control of the TEF-1α promoter relative to expression solely under their respective native promoters. Actin was used as the reference gene. Samples were taken from strains growing in glucose-containing media in shake flasks. 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 ($\alpha = 0.05$).

From **Figure 3.2** it is clear the expression platform has worked, as all genes studied are significantly overexpressed in their respective strains. Especially large fold-changes in expression were seen in EEB1 (16.3), PSD2 (52.6), ODE1 (14.8), SLD (18.4), SOL3 (18.4), YAS1 (71.0), ABC2 (25.6), AMD1 (43.9), and YIP3 (17.1) overexpression strains, possibly indicating a lower baseline expression level for these genes during growth phase when the RNA samples were harvested. Strong inverse correlations have previously been observed between fold overexpression and baseline expression level when comparing several genes overexpressed in the same manner, through ectopic inductions of transcription factors in *M. tuberculosis* [51] and through transcriptional activator-bound Cas9-mediated upregulation of genes in human 293FT

cells [52]. 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 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 GND (5.9-fold increase), which are the steps before and after SOL3, respectively. FAS2 (5.1-fold increase) also may have a lower baseline expression level than FAS1 (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.

3.3.2 Overexpression of many gene targets influences lipid production of *Yarrowia* to varying degrees

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.

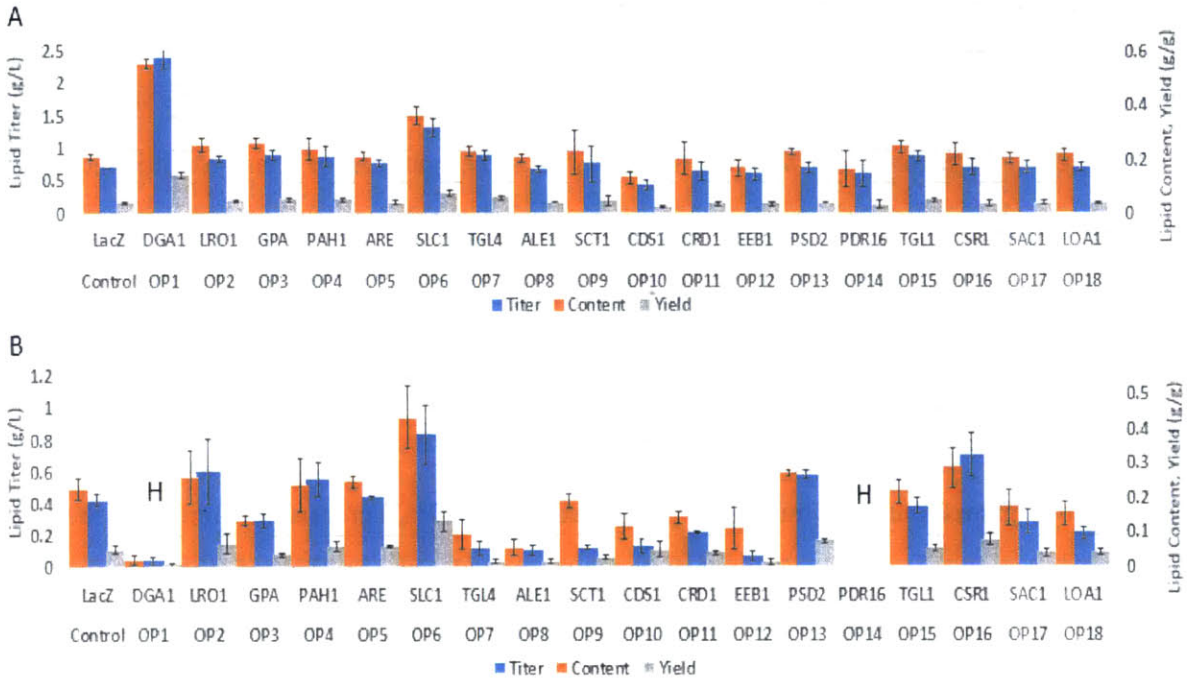

Figure 3.3: Fermentation performance of strains overexpressing glycerolipid synthesis genes on (A) 50 g/L glucose at an atomic C:N of 60 after 120 hours and (B) 50 g/L sodium acetate at an atomic C:N of 40 after 144 hours. Performance metrics used to evaluate strains include lipid titer, cellular lipid 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.

Kennedy Pathway/Glycerolipid Synthesis Genes: Fermentation results for strains overexpressing Kennedy Pathway and glycerolipid synthesis related genes are shown in **Figure 3.3**. On glucose, the largest increases in lipid production over the LacZ control strain were achieved by overexpressing the DGA1 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. Overexpression of the other diacylglycerol acyltransferase gene (YALI0E32769g) has been shown to be similarly effective at achieving large increases in lipid production [20, 28], and deletion of any DGA gene significantly decreased lipid content (by 43% for DGA1) in *Yarrowia* [53]. 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 [20, 54], or that creating a "sink" for fatty acids in general stimulates their production further [20, 55]. While the effects of DGA1 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; however, 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 [56], further giving support to this theory.

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 DGA1 overexpression strain showed very poor performance in acetate fermentations. A microscopic evaluation revealed that the DGA1 strain underwent a transition to a hyphal morphological state, a stress response characterized by long, filamentous cells with low growth rate and capacity for lipid storage. While the reason for this is unclear, it is possible that, due to the proximity of acetate to the fatty acid synthesis pathway, cells overexpressing DGA1 could divert too much flux away from pathways that would generate biomass or energy; however, *Yarrowia* cells overexpressing the other DGA isoform have shown large increases in lipid production on acetate similar to those on glucose [57]. In addition to DGA1, hyphal growth was also seen in the PDR16 overexpression strain when grown on acetate, however the reason for this strain's hyphal growth may not the same as the

reason for the DGA1 strain's hyphal growth, as PDR16 overexpression did not cause increased lipogenesis on glucose.

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. LRO1 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 stationary (non-dividing) phase [58], suggesting that this pathway is likely only a minor contributor to overall TAG synthesis in *Yarrowia*. It is possible that GPA overexpression increases lipid synthesis on glucose by the same mechanism as SLC1, although to a much smaller extent. Neither gene showed significant improvement in lipid production on acetate.

Interestingly, the lipase/esterase TGL4 (TAG lipase) and TGL1 (steryl ester hydrolase) overexpressions increased lipid production on glucose 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. 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 expression of the *S. cerevisiae* TGL4 gene into *P. pastoris* resulted in an increase in phospholipid biosynthesis [59]. It is possible that TGL1 is similarly multifunctional. Contributions to lipid production from TGL4 overexpression therefore likely comes mostly from increasing the growth rate by providing more phospholipids for membrane synthesis, which 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 3.3**). Neither gene overexpression conferred increased lipid production when grown on acetate.

In contrast to the results discussed above, there were two Kennedy pathway/glycerolipid synthesis genes that showed greater lipid production on acetate but not glucose. PSD2 showed increases in titer and yield (39% and 45%, respectively) compared to the LacZ control strain. As PSD2 is a key enzyme in the biosynthesis of phospholipids like phosphatidylethanolamine and phosphatidylcholine, overexpression of this gene 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 PSD2 strain, providing more evidence of a causal link between phospholipid biosynthesis and growth rate. In addition to PSD2, 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 over the control strain, but increases in content or yield were not statistically significant. To explain the increase in titer and biomass without a consistently large increase in yield, it is possible that TCA cycle flux is higher in these cells.

Central Carbon Metabolism Genes: From the results in **Figure 3.4**, the only gene in this group that promoted lipogenesis when overexpressed was GPD, or glycerol-3-phosphate dehydrogenase. GPD 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. From our results, we can speculate as to the form taken by the additional lipids produced by this 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 3.3**) 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 [60]. 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 [61].

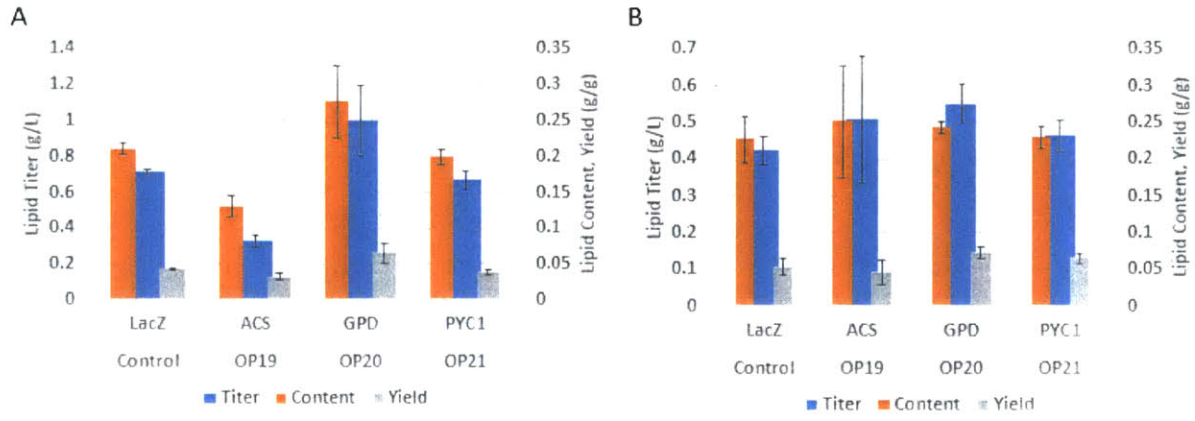

Figure 3.4: Fermentation performance of strains overexpressing central carbon metabolism genes on (A) 50 g/L glucose at an atomic C:N of 60 after 120 hours and (B) 50 g/L sodium acetate at an atomic C:N of 40 after 144 hours. Performance metrics used to evaluate strains include lipid titer, cellular lipid content, and lipid yield. Error bars represent standard deviations. Fermentations were carried out in triplicate for all strains.

Neither ACS nor PYC1 overexpression contributed significantly lipid synthesis. Although it has been speculated that ACS overexpression can contribute to lipid synthesis [54] because it was found that overexpressing ACS in *E. coli* led to an increase in the rate of acetate utilization [62], in our study, we found that ACS overexpression only increased biomass significantly (by 49%, see **Table 3.3**) on acetate. Lipid production in our ACS 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 PYC1 (pyruvate to oxaloacetate) is implicated in lipogenesis and gluconeogenesis in humans, and is a component of the NADPH-producing transhydrogenase (malate-pyruvate-oxaloacetate) cycle [63], however, the lack of contribution of malic enzyme to NADPH production and lipid synthesis in *Y. lipolytica* [29] likely explains the ineffectiveness of overexpressing this gene in increasing lipid production in our strain.

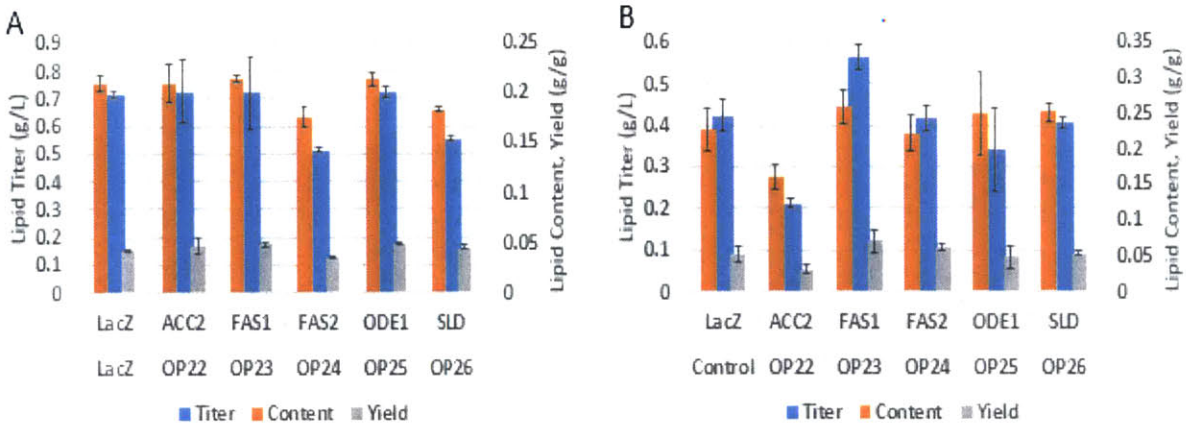

Figure 3.5: Fermentation performance of strains overexpressing fatty acid synthesis genes on (A) 50 g/L glucose at an atomic C:N of 60 after 120 hours and (B) 50 g/L sodium acetate at an atomic C:N of 40 after 144 hours. Performance metrics used to evaluate strains include lipid titer, cellular lipid content, and lipid yield. Error bars represent standard deviations. Fermentations were carried out in triplicate for all strains.

Fatty Acid Synthesis Genes: **Figure 3.5** 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 DCW measurements (as shown in **Table 3.3**) in this strain compared to the control suggest that FAS1 overexpression can contribute to a higher growth rate. It is interesting to note that the fold change of FAS1 overexpression in its respective strain compared to the control shown in **Figure 3.2** 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* [64]. 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 [64]. 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 [65].

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 a limiting factor for lipid synthesis in many oleaginous fungi [34], but previous work of ours (data not shown) and others [29] showed no effect of overexpression of the only native malic enzyme isoform, MAE1, on lipid synthesis in *Y. lipolytica*. It has been suggested elsewhere [29] that MAE1 in *Y. lipolytica* may be mostly specific for NAD+ and not NADP+. Results for fermentations of strains overexpressing potentially important NADPH-producing genes are given in **Figure 3.6**.

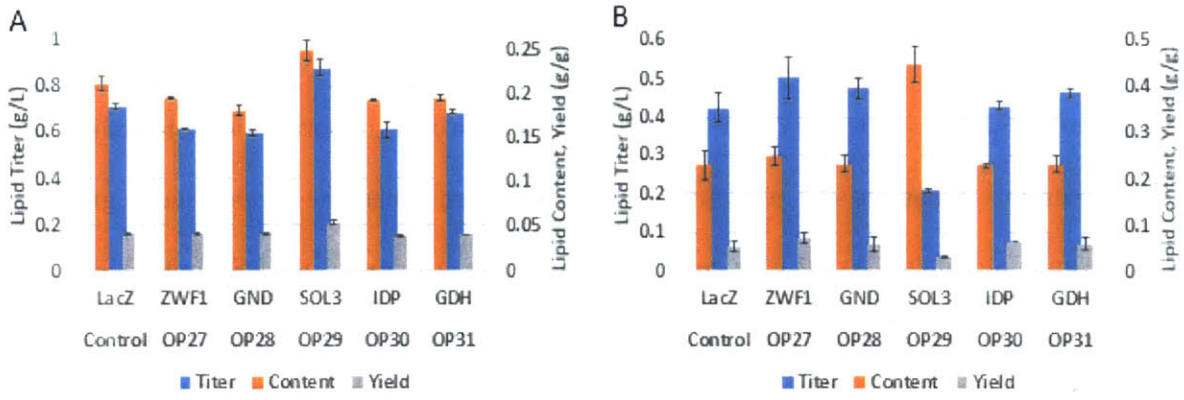

Figure 3.6: Fermentation performance of strains overexpressing NADPH generation genes on (A) 50 g/L glucose at an atomic C:N of 60 after 120 hours and (B) 50 g/L sodium acetate at an atomic C:N of 40 after 144 hours. Performance metrics used to evaluate strains include lipid titer, cellular lipid content, and lipid yield. Error bars represent standard deviations. Fermentations were carried out in triplicate for all strains.

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 GND gene products. Curiously, neither NADPH-producing gene in the oxPPP significantly increased lipid production. 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 [66]. In addition, our gene expression analysis results (**Figure 3.2**) 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 GND 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. 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.

Recent metabolic flux analysis (MFA) studies on *Y. lipolytica* reported that a high-oil producing ACC1- and DGA2-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 [67]. 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.

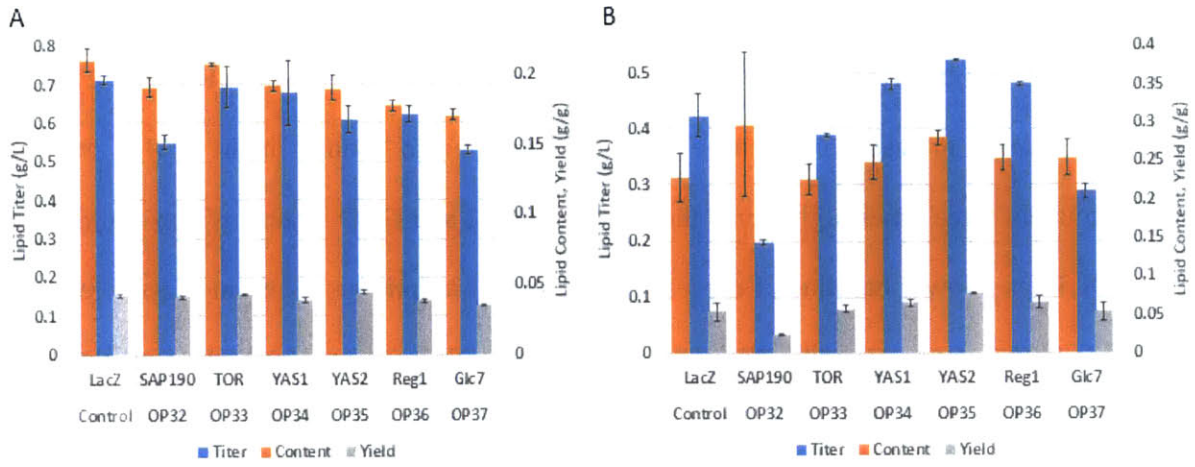

Figure 3.7: Fermentation performance of strains overexpressing regulator genes on (A) 50 g/L glucose at an atomic C:N of 60 after 120 hours and (B) 50 g/L sodium acetate at an atomic C:N of 40 after 144 hours. Performance metrics used to evaluate strains include lipid titer, cellular lipid content, and lipid yield. Error bars represent standard deviations. Fermentations were carried out in triplicate for all strains.

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 3.7**. None of the SAP190, TOR, REG1, or GLC7 genes contributed to increased lipid synthesis 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 [44]. The Tor 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 [45]. 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.

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). While the Yas1/Yas2 complex has been found to be involved in upregulating genes for alkane metabolism [46], 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 [68] and fatty acid synthesis [64] genes when concentrations of intracellular inositol are low. Inositol drives the repression of the Ino2/Ino4 complex by Opi1 (closest homologue in *Y. lipoltyica*: YALI0C14784p), and ICRE-containing genes are significantly upregulated in *S. cerevisiae* ΔOPI1 mutants [69]. 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.

Transporters and Miscellaneous Genes: The genes in this group included five genes (SLY1, ABC2, PSR, SKY, and YIP3) that are the closest *Yarrowia* homologues to genes that are upregulated in the oleaginous yeast *Rhodosporidium toruloides* during lipogenesis [39, 40] as well as the AMD1 and SCP genes. From the results of the fermentations of strains overexpressing these genes in **Figure 3.8**, none of these genes increase lipid synthesis on glucose when overexpressed. Amd1, or AMP deaminase, is thought to be the enzyme responsible for beginning the lipogenic phase on glucose by removing the AMP needed for TCA cycle activity [22], however, 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 [70]. This suggests that the intracellular environment is critical to Amd1 activity, while the exact transcription level of the gene is not.

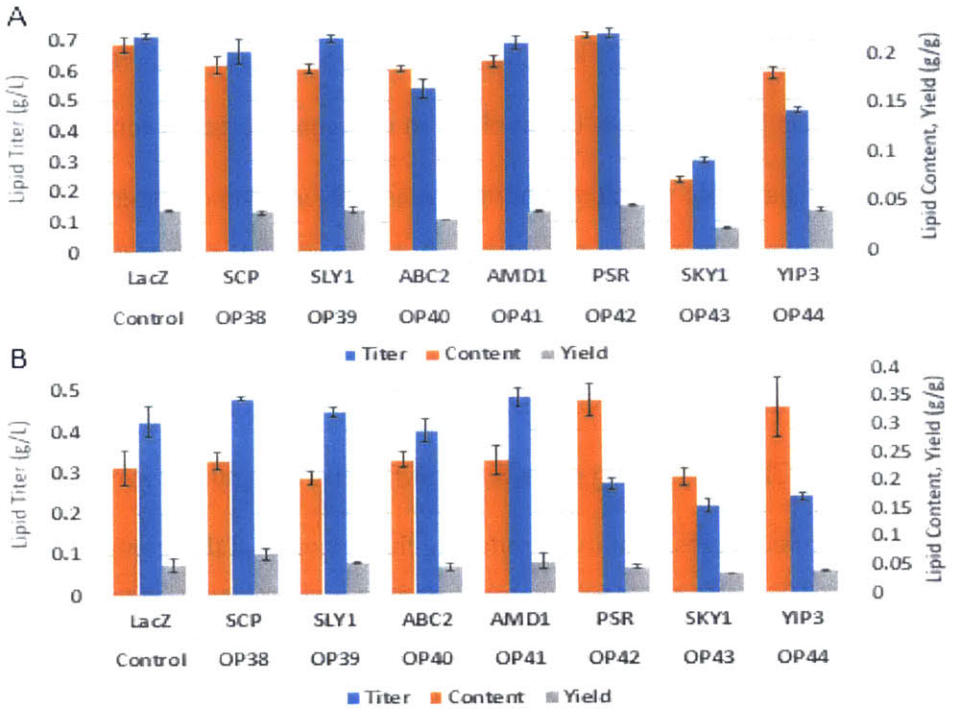

Figure 3.8: Fermentation performance of strains overexpressing metabolite transport and "miscellaneous" genes on (A) 50 g/L glucose at an atomic C:N of 60 after 120 hours and (B) 50 g/L sodium acetate at an atomic C:N of 40 after 144 hours. Performance metrics used to evaluate strains include lipid titer, cellular lipid content, and lipid yield. Error bars represent standard deviations. Fermentations were carried out in triplicate for all strains.

Two genes in this group, PSR 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. PSR 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 *Yarrowia* when growing on acetate are very much unclear. The other three genes in this group whose transcription correlates to lipogenesis in this study do not demonstrate a causal link between expression and lipid synthesis from this study.

Table 3.3: Dry Cell Weight per mL (and Standard Deviation) of Strains at the End of Fermentations

Strain	Overexpressed Gene	Glucose		Acetate	
		DCW/mL	DCW SD	DCW/mL	DCW SD
MTYL037	LacZ	3.393	0.079	2.318	0.130
YLOP1	DGA1	4.306	0.395	2.266	0.286
YLOP2	LRO1	3.369	0.127	2.750	0.363
YLOP3	GPA	3.430	0.093	2.664	0.396
YLOP4	PAH1	3.699	0.189	3.062	0.650
YLOP5	ARE	3.687	0.116	2.179	0.208
YLOP6	SLC1	3.655	0.067	2.352	0.108
YLOP7	TGL4	3.940	0.264	1.747	0.512
YLOP8	ALE1	3.272	0.211	2.439	0.363
YLOP9	SCT1	3.374	0.169	0.813	0.079
YLOP10	CDS1	3.339	0.144	1.401	0.104
YLOP11	CRD1	3.346	0.461	1.868	0.275
YLOP12	EEB1	3.582	0.007	1.574	0.917
YLOP13	PSD2	3.102	0.259	2.582	0.057
YLOP14	PDR16	3.812	0.248	N/A	N/A
YLOP15	TGL1	3.597	0.079	2.179	0.187
YLOP16	CSR1	3.323	0.177	3.027	0.196
YLOP17	SAC1	3.513	0.153	2.093	0.530
YLOP18	LOA1	3.293	0.277	1.747	0.286
YLOP19	ACS	2.541	0.119	3.459	0.245
YLOP20	GPD	3.619	0.051	2.729	0.200
YLOP21	PYC1	3.367	0.118	2.427	0.085
YLOP22	ACC2	3.475	0.197	1.626	0.234
YLOP23	FAS1	3.362	0.597	2.620	0.204
YLOP24	FAS2	2.909	0.125	2.273	0.121
YLOP25	ODE1	3.384	0.065	1.609	0.363
YLOP26	SLD	3.020	0.070	1.944	0.120
YLOP27	ZWF1	3.167	0.035	2.453	0.085
YLOP28	GND	3.302	0.108	2.480	0.122
YLOP29	SOL3	3.536	0.017	0.564	0.063
YLOP30	IDP	3.171	0.196	2.258	0.050
YLOP31	GDH	3.538	0.064	2.424	0.144
YLOP32	SAP190	2.882	0.075	0.880	0.364
YLOP33	TOR	3.353	0.256	2.084	0.184
YLOP34	YAS1	3.531	0.377	2.344	0.207
YLOP35	YAS2	3.211	0.087	2.249	0.068
YLOP36	Reg1	3.513	0.179	2.287	0.133
YLOP37	Glc7	3.100	0.012	1.376	0.080
YLOP38	SCP	3.498	0.129	2.429	0.160
YLOP39	SLY1	3.807	0.053	2.584	0.101
YLOP40	ABC2	2.902	0.178	2.011	0.090
YLOP41	AMD1	3.569	0.015	2.429	0.149
YLOP42	PSR	3.298	0.108	0.944	0.030
YLOP43	SKY1	4.218	0.187	1.260	0.179
YLOP44	YIP3	2.564	0.087	0.876	0.111

3.3.3 Overexpression of some genes can alter fatty acid composition

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 [71]. Higher degrees of unsaturation make also biodiesel more susceptible to oxidation [71]. However, the lubricity of the fuel is improved with higher degrees of unsaturation [72] and the cold-flow properties improve with higher unsaturation and shorter chain length [73]. Therefore, the optimal fatty acid composition for biodiesel may be context dependent [74].

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 Dga1 and Lro1 enzymes in *S. cerevisiae* demonstrate a preference for unsaturated fatty acids, especially C18:1 [75], while the Dga1 enzyme in *P. pastoris* showed strong preference to unsaturated C16:0 and C18:0, and the Lro1 enzyme showed no preference at all [76]. *Yarrowia* Dga2, 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 [20]. We therefore were interested to see if some of the native genes we overexpressed in *Yarrowia* had an impact on the fatty acid profile. Results of that analysis performed on the glucose fermentation data are shown in **Figure 3.9**.

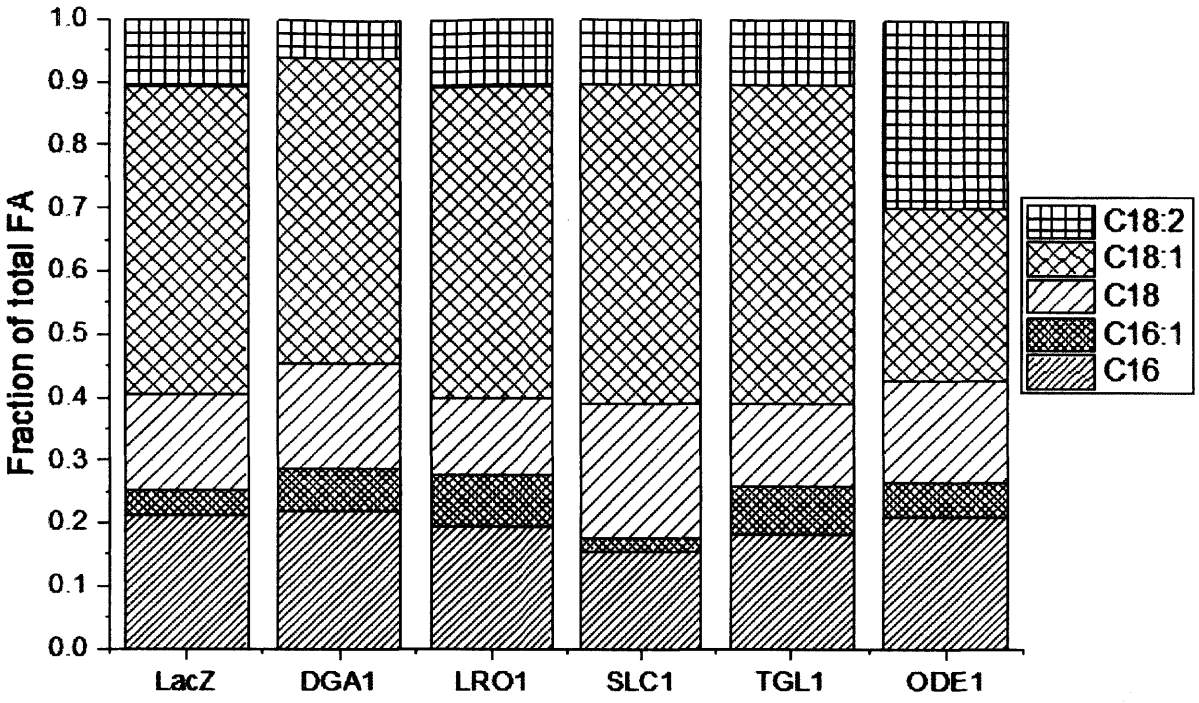

Figure 3.9: Fatty acid profiles of selected strains growing on glucose. Percentages shown are the mean values of six measurements of lipid wt% of C16:0, C16:1, C18:0, C18:1, and C18:2 analyzed by GC-FID. C16:0: palmitate; C16:1: palmitoleate; C18:0: stearate; C18:1: oleate; C18:2: linoleate.

Our results show that, on glucose, *Yarrowia* 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%). In contrast to the DGA2 overexpression strain mentioned above, our DGA1-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 Dga1 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 fraction, the LRO1 overexpressing strain, it 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 Ode1 enzyme is a Δ12-desaturase, which in *Yarrowia* catalyzes the conversion of C18:1 to C18:2. Although ODE1 overexpression did not affect the total amount of lipids produced (see **Figure 3.5**), 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 ODE1 overexpression strain. 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. Native production of linoleic acid in *Yarrowia* has however been used as a starting point for synthesis of other fatty acids like EPA [28], so this finding could potentially be very useful for other applications. 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.

3.4 Conclusions

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 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. Our results demonstrated that a diverse subset of the genes tested here can confer greater lipid production onto their overexpression strains. 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 GPD) 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 individual one 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.

The most dramatic increase in lipid synthesis from glucose came from overexpressing the DGA1 gene, producing lipid titers, contents, and yields that were 236%, 165%, and 246% greater than those of the control strain; our largest improvements on acetate came from the SLC1 overexpressing strain, which produced lipid liters, contents, and yields that were 99%, 91%, and

151% greater than those of our control. As these genes are directly involved in the process of TAG synthesis, these results are not surprising. The overexpressions of genes affecting other inputs to the process had comparatively smaller effects, however, 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 [20]. 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 [77].

References:

1. Asif, M., & Muneer, T. (2007). Energy supply, its demand and security issues for developed and emerging economies. *Renewable and Sustainable Energy Reviews*, *11*(7), 1388–1413. doi:10.1016/j.rser.2005.12.004
2. Solomon, S., Plattner, G.-K., Knutti, R., & Friedlingstein, P. (2009). Irreversible climate change due to carbon dioxide emissions. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(6), 1704–9. doi:10.1073/pnas.0812721106
3. Demirbas, A. (2007). Importance of biodiesel as transportation fuel. *Energy Policy*, *35*(9), 4661–4670. doi:10.1016/j.enpol.2007.04.003
4. Huber, G. W., Iborra, S., & Corma, A. (2006). Synthesis of transportation fuels from biomass: chemistry, catalysts, and engineering. *Chemical Reviews*, *106*(9), 4044–98. doi:10.1021/cr068360d
5. Meng, X., Yang, J., Xu, X., Zhang, L., Nie, Q., & Xian, M. (2009). Biodiesel production from oleaginous microorganisms. *Renewable Energy*, *34*(1), 1–5. doi:10.1016/j.renene.2008.04.014
6. 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. *Progress in Lipid Research*, *48*(6), 375–87. doi:10.1016/j.plipres.2009.08.005
7. 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. *European journal of lipid science and technology*, *111*(12), 1221-1232.
8. Stephanopoulos, G. (2007). Challenges in engineering microbes for biofuels production. *Science*, (February), 801–805.
9. 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 Technology*, *114*, 443–9. doi:10.1016/j.biortech.2012.02.091
10. 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. *Current Microbiology*, *46*(2), 0124-0130.
11. Papanikolaou, S., & Aggelis, G. (2002). Lipid production by Yarrowia lipolytica growing on industrial glycerol in a single-stage continuous culture, *82*(July 2001), 43–49.
12. 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. *Applied Microbiology and Biotechnology*, *58*(3), 308–12. doi:10.1007/s00253-001-0897-0
13. Fickers, P., Benetti, P.-H., Waché, Y., Marty, a, Mauersberger, S., Smit, M. S., & Nicaud, J.-M. (2005). Hydrophobic substrate utilisation by the yeast Yarrowia lipolytica, and its potential applications. *FEMS Yeast Research*, *5*(6-7), 527–43. doi:10.1016/j.femsyr.2004.09.004
14. Beopoulos, A., Nicaud, J. M., & Gaillardin, C. (2011). An overview of lipid metabolism in yeasts and its impact on biotechnological processes. *Applied microbiology and biotechnology*, *90*(4), 1193-1206.

15. 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.
16. 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.
17. 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
18. Blazeck, J., Liu, L., Redden, H., & Alper, H. (2011). Tuning gene expression in Yarrowia lipolytica by a hybrid promoter approach. *Applied and Environmental Microbiology*, *77*(22), 7905–14. doi:10.1128/AEM.05763-11
19. Nicaud, J. M., Fabre, E., & Gaillardin, C. (1989). Expression of invertase activity in Yarrowia lipolytica and its use as a selective marker. *Current genetics*, *16*(4), 253-260.
20. Tai, M., & Stephanopoulos, G. (2013). Engineering the push and pull of lipid biosynthesis in oleaginous yeast Yarrowia lipolytica for biofuel production. *Metabolic Engineering*, *15*, 1–9. doi:10.1016/j.ymben.2012.08.007
21. Verbeke, J., Beopoulos, A., & Nicaud, J.-M. (2013). Efficient homologous recombination with short length flanking fragments in Ku70 deficient Yarrowia lipolytica strains. *Biotechnology Letters*, *35*(4), 571–6. doi:10.1007/s10529-012-1107-0
22. Ratledge, C. (2002). Regulation of lipid accumulation in oleaginous micro-organisms. *Biochemical Society Transactions*, *30*(6), 47–50.
23. Wu, S., Hu, C., Jin, G., Zhao, X., & Zhao, Z. K. (2010). Phosphate-limitation mediated lipid production by Rhodosporidium toruloides. *Bioresource Technology*, *101*(15), 6124–9. doi:10.1016/j.biortech.2010.02.111
24. Wu, S., Zhao, X., Shen, H., Wang, Q., & Zhao, Z. K. (2011). Microbial lipid production by Rhodosporidium toruloides under sulfate-limited conditions. *Bioresource Technology*, *102*(2), 1803–7. doi:10.1016/j.biortech.2010.09.033
25. Rossi, M., Amaretti, A., Raimondi, S., & Leonardi, A. (2011). Getting lipids for biodiesel production from oleaginous fungi. *Feedstocks and Processing Technologies*.
26. Wakil, S. J., Stoops, J. K., & Joshi, V. C. (1983). Fatty acid synthesis and its regulation. *Annual Review of Biochemistry*, *52*, 537–79. doi:10.1146/annurev.bi.52.070183.002541
27. Coleman, R. A., & Lee, D. P. (2004). Enzymes of triacylglycerol synthesis and their regulation. *Progress in Lipid Research*, *43*(2), 134–176. doi:10.1016/S0163-7827(03)00051-1
28. Xue, Z., Sharpe, P. L., Hong, S.-P., Yadav, N. S., Xie, D., Short, D. R., ... Zhu, Q. (2013). Production of omega-3 eicosapentaenoic acid by metabolic engineering of Yarrowia lipolytica. *Nature Biotechnology*, *31*(8), 734–40. doi:10.1038/nbt.2622
29. 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. *Biotechnology Letters*, *35*(12), 2091–8. doi:10.1007/s10529-013-1302-7
30. Thiam, A. R., Farese, R. V, & Walther, T. C. (2013). The biophysics and cell biology of lipid droplets. *Nature Reviews. Molecular Cell Biology*, *14*(12), 775–86. doi:10.1038/nrm3699
31. Casal, M., Cardoso, H., & Leao, C. (1996). Mechanisms regulating the transport of acetic acid in Saccharomyces cerevisiae. *Microbiology*, (142), 1385–1390.

32. 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
33. De la Peña Mattozzi, M., Kang, Y., & Keasling, J. D. (2010). Feast: Choking on Acetyl-CoA, the Glyoxylate Shunt, and Acetyl-CoA-Driven Metabolism. *Handbook of Hydrocarbon and Lipid Microbiology*, 1649–1660. doi:10.1007/978-3-540-77587-4
34. 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 (Reading, England)*, *153*(Pt 7), 2013–25. doi:10.1099/mic.0.2006/002683-0
35. 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. *Metabolic Engineering*, *13*(5), 482–91. doi:10.1016/j.ymben.2011.05.002
36. 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. *Metabolic Engineering*, *26*, 89–99. doi:10.1016/j.ymben.2014.09.008
37. 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
38. Seip, J., Jackson, R., He, H., Zhu, Q., & Hong, S.-P. (2013). Snf1 is a regulator of lipid accumulation in Yarrowia lipolytica. *Applied and Environmental Microbiology*, *79*(23), 7360–70. doi:10.1128/AEM.02079-13
39. Liu, H., Zhao, X., Wang, F., Li, Y., Jiang, X., & Ye, M. (2009). Comparative proteomic analysis of Rhodosporidium toruloides during lipid accumulation. *FEMS Yeast Research*, *2011*(11), 553–566. doi:10.1002/yea
40. Zhu, Z., Zhang, S., Liu, H., Shen, H., Lin, X., Yang, F., ... Zhao, Z. K. (2012). A multi-omic map of the lipid-producing yeast Rhodosporidium toruloides. *Nature Communications*, *3*, 1112. doi:10.1038/ncomms2112
41. Schmidt, C., Ploier, B., Koch, B., & Daum, G. (2012). Analysis of yeast lipid droplet proteome and lipidome. *Methods in cell biology*, *116*, 15-37.
42. Morin, N., Cescut, J., Beopoulos, A., Lelandais, G., Le Berre, V., Uribelarrea, J.-L., ... 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
43. Liu, H., Zhao, X., Wang, F., Jiang, X., Zhang, S., Ye, M., ... Zou, H. (2011). The proteome analysis of oleaginous yeast Lipomyces starkeyi. *FEMS Yeast Research*, *11*(1), 42–51. doi:10.1111/j.1567-1364.2010.00687.x
44. 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. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(16), 6349–6354. doi:10.1073/pnas.1102758108
45. Raught, B., Gingras, a C., & Sonenberg, N. (2001). The target of rapamycin (TOR) proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(13), 7037–7044. doi:10.1073/pnas.121145898
46. 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*. *Eukaryotic Cell*, *6*(4), 734–743. doi:10.1128/EC.00412-06
47. Sambrook, J., & Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. CSHL press, Cold Spring Harbor, NY, USA.
48. 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. *Nature Methods*, *6*(5), 343–345. doi:10.1038/nmeth.1318
49. Chen, D. C., Beckerich, J. M., & Gaillardin, C. (1997). One-step transformation of the dimorphic yeast Yarrowia lipolytica. *Applied Microbiology and Biotechnology*, *48*, 232–235. doi:10.1007/s002530051043
50. Sheng Zhao, Russell D. Fernald. Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J. Comput. Biol.* 2005 Oct;12(8):1045-62.
51. Rustad, T. R., Minch, K. J., Ma, S., Winkler, J. K., Hobbs, S., Hickey, M., ... Sherman, D. R. (2014). Mapping and manipulating the Mycobacterium tuberculosis transcriptome using a transcription factor overexpression-derived regulatory network. *Genome Biology*, *15*, 1–11. doi:10.1186/s13059-014-0502-3
52. Konermann, S., Brigham, M. D., Trevino, a E., Joung, J., Abudayyeh, O. O., Barcena, C., ... Zhang, F. (2014). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, *517*(7536), 583–588. doi:10.1038/nature14136
53. 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
54. Courchesne, N. M. D., Parisien, A., Wang, B., & Lan, C. Q. (2009). Enhancement of lipid production using biochemical, genetic and transcription factor engineering approaches. *Journal of Biotechnology*, *141*(1-2), 31–41. doi:10.1016/j.jbiotec.2009.02.018
55. Thelen, J. J., & Ohlrogge, J. B. (2002). Metabolic engineering of fatty acid biosynthesis in plants. *Metabolic Engineering*, *4*, 12–21. doi:10.1006/mben.2001.0204
56. Zou, J., Katavic, V., Giblin, E. M., Barton, D. L., MacKenzie, S. L., Keller, W. a, ... Taylor, D. C. (1997). Modification of seed oil content and acyl composition in the brassicaceae by expression of a yeast sn-2 acyltransferase gene. *The Plant Cell*, *9*(June), 909–923. doi:10.1105/tpc.9.6.909
57. Tai, M. (2012). Metabolic Engineering of Oleaginous Yeast for the Production of Biofuels. Ph.D. Thesis. Massachusetts Institute of Technology.
58. Müllner, H., & Daum, G. (2004). Dynamics of neutral lipid storage in yeast. *Acta Biochimica Polonica*, *51*(2), 323–347. doi:035001323
59. Rajakumari, S., & Daum, G. (2010). Multiple functions as lipase, steryl ester hydrolase, phospholipase, and acyltransferase of Tgl4p from the yeast Saccharomyces cerevisiae. *Journal of Biological Chemistry*, *285*(21), 15769–15776. doi:10.1074/jbc.M109.076331
60. Beopoulos, A., Mrozova, Z., Thevenieau, F., Le Dall, M.-T., Hapala, I., Papanikolaou, S., ... Nicaud, J.-M. (2008). Control of lipid accumulation in the yeast Yarrowia lipolytica. *Applied and Environmental Microbiology*, *74*(24), 7779–7789. doi:10.1128/AEM.01412-08
61. 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 Biotechnology Journal, 5: 431–441. doi: 10.1111/j.1467-7652.2007.00252.x
62. 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. *Applied Microbiology and Biotechnology*, *71*, 870–874. doi:10.1007/s00253-005-0230-4
63. Mackall, J. C., & Lane, M. D. (1977). Role of pyruvate carboxylase in fatty acid synthesis: alterations during preadipocyte differentiation. *Biochemical and Biophysical Research Communications*, *79*(3), 720–725. doi:10.1016/0006-291X(77)91171-8
64. Schweizer, E., & Hofmann, J. (2004). Microbial type I fatty acid synthases (FAS): major players in a network of cellular FAS systems. *Microbiology and Molecular Biology Reviews : MMBR*, *68*(3), 501–517. doi:10.1128/MMBR.68.3.501-517.2004
65. 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 Research*, *29*(22), 4625–4632. doi:10.1093/nar/29.22.4625
66. 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. *Journal of Biological Chemistry*, *276*(37), 34840–34846. doi:10.1074/jbc.M105174200
67. 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. *Metabolic Engineering*, *30*, 27–39. doi:10.1016/j.ymben.2015.02.007
68. 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. *The Journal of Biological Chemistry*, *284*(11), 7126–7137. doi:10.1074/jbc.M806864200
69. 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
70. 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. *The Journal of Biological Chemistry*, *264*(35), 21422–21430.
71. 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 Technology*, *100*, 261–268. doi:10.1016/j.biortech.2008.06.039
72. Knothe, G. (2005). Dependence of biodiesel fuel properties on the structure of fatty acid alkyl esters. *Fuel Processing Technology*, *86*, 1059–1070. doi:10.1016/j.fuproc.2004.11.002
73. Dunn, R. O. (2011). Improving the Cold Flow Properties of Biodiesel by Fractionation. *Soybean - Application and Technology*, *975*(2). doi:10.5772/14624
74. Hoekman, S. K., Broch, A., Robbins, C., Ceniceros, E., & Natarajan, M. (2012). Review of biodiesel composition, properties, and specifications. *Renewable and Sustainable Energy Reviews*, *16*(1), 143–169. doi:10.1016/j.rser.2011.07.143

75. Grillitsch, K., Connerth, M., Köfeler, H., Arrey, T. N., Rietschel, B., Wagner, B., … Daum, G. (2011). Lipid particles/droplets of the yeast Saccharomyces cerevisiae revisited: Lipidome meets Proteome. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1811*(12), 1165–1175. doi:10.1016/j.bbalip.2011.07.015
76. Ivashov, V. a., Zellnig, G., Grillitsch, K., & Daum, G. (2013). Identification of triacylglycerol and steryl ester synthases of the methylotrophic yeast Pichia pastoris. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1831*(6), 1158–1166. doi:10.1016/j.bbalip.2013.03.004
77. Hill, J., Nelson, E., Tilman, D., Polasky, S., & Tiffany, D. (2006). Environmental, economic, and energetic costs and benefits of biodiesel and ethanol biofuels. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 11206–11210. doi:10.1073/pnas.0604600103

Chapter 4

Construction and Characterization of Strains Combining Validated Gene Targets

4.1 Introduction

Single cell oil (SCO) processes for the production of biodiesel are a promising alternative to traditional fossil fuels, particularly for use in transportation. Oleaginous yeasts are convenient host organisms for such processes due to their natural tendency to accumulate lipids as a high fraction of their dry cell weight. *Yarrowia lipolytica* is one such yeast that is becoming more widely used in SCO process research due to its sequenced genome and genetic tractability [1]. *Yarrowia lipolytica* is also naturally capable of metabolizing a wide range of carbon substrates from large-scale industrially relevant feedstocks and converting them into intracellular lipids (over 90% of which are in the form of triacylglycerols, or TAGs), such as hexoses from plant biomass hydrolysate [2], volatile fatty acids, such as acetate, from municipal waste [3], glycerol and long-chain fatty acids derived from animal fat [4] and even very hydrophobic substrates like alkanes from petroleum sludge waste streams [5].

In many cases, these feedstocks for SCO processes can be acquired for a low or even negative cost. However, metabolic engineering of strains capable of quickly producing high lipid titers and yields is still crucial to the economic viability of these processes [6]. Many recent engineering efforts to improve strain performance have taken the form of overexpressing certain native genes or introducing non-native isoforms of genes that are involved in the overall process of lipid synthesis. The overexpression of the native DGA2 diacylglycerol acyltransferase gene in *Yarrowia lipolytica* caused a threefold increase in the lipid content of the resulting strain [7]. Increasing the supply of NADPH for fatty acid synthesis in *Mucor circincelloides* by either overexpressing its native NADP+-dependent malic enzyme or expressing a similar isoform from *Mortierella alpina* led to a 2.5-fold increase in lipid content and a similar increase in lipid titer [8].

By adding multiple genes to a single DNA cassette, or using multiple cassettes containing different auxotrophic or antifungal markers, many genes can be upregulated or introduced in the same strain. In this way, the overall rate of a metabolic pathway of interest can be increased by upregulating the activities of multiple steps that would otherwise be rate-limiting, or by enhancing the availability of multiple inputs needed in the pathway. Typically, individual gene overexpressions are first validated for their effectiveness at increasing the output of the desired product, and are later combined to determine whether their individual effects are additive. An example in plants involves increasing the neutral lipid content of *N. benthamiana* leaves (wild type neutral lipid content of 0.025 g/g) by expressing the glycolysis and fatty acid synthesis upregulating [9] WRI1 transcription factor (lipid content: 0.57 g/g) from *A. thaliana*, and the *A. thaliana* DGA1 gene (lipid content: 0.45 g/g) first individually, and then simultaneously (lipid content: 2.48 g/g) [10]. A previous finding in *Yarrowia lipolytica* demonstrated additive modifications by overexpression of the native acetyl-CoA carboxylase ACC1 gene in tandem with the native diacylglycerol acyltransferase DGA2 gene; ACC1 overexpression alone caused the lipid content to increase twofold in high nitrogen glucose containing media, while DGA2 overexpression alone caused a fourfold increase in lipid content [11]. Combined, ACC1 and DGA2 overexpression worked in a synergistic manner to give the resulting strain a lipid content 4.7 times that of the wild-type control.

Despite the relative conceptual simplicity of this tandem modification approach and some reported instances of success, the architecture of cellular metabolic networks is complex and comprised of highly connected pathways that can be either highly sensitive to perturbations or subject to multiple layers of local and global regulation. In many cases, enhancing the activities of multiple enzymes that individually contribute positively to the rate of a particular biosynthetic

pathway can lead to a lack of stackable effects in the overall pathway rate (which could be the result of a feedback inhibition regulation [12] or of product or byproduct toxicity [13]), or even a decrease in the rate of production from the base strain. In the latter case, this negative effect could be due to imbalances at the level of metabolite concentrations, i.e. diverting too many resources to the pathway of interest and away from other pathways needed for the cells to divide or produce energy or reducing equivalents [14], or at the level of transcription and translation (describing the concept of "metabolic burden" [15]), in which the cell decreases production of protein from other genes needed for proper cell function due to an over-dedication of transcription or translational machinery towards producing the artificially expressed genes. In the aforementioned example with *Yarrowia lipolytica*, while ACC1 and DGA2 overexpressions showed synergistic effects on lipid productivity and yield, and simultaneous overexpressing DGA2 along with the native stearoyl-CoA delta-9 desaturase SCD1 gene led to higher lipid productivity and yield than overexpression of DGA2 alone, overexpression of all three genes simultaneously led to significantly lower productivity and yield than the ACC1 and DGA2 only strain [16]. The lack of a synergistic effect in this case likely comes from a diversion of too many resources away from other pathways necessary for cell function, and illustrates that a delicate balance between metabolic pathways must be achieved for optimal product output.

The previously developed *Yarrowia lipolytica* ACC1 and DGA2 overexpression strain represents a major advancement in host organism engineering for more efficient SCO processes. In this study, we attempt to build on the success of this previous effort by combining these two modifications with other native gene overexpressions that were previously validated in the work detailed in Chapter 3. The genes overexpressed in tandem with ACC1 and DGA2 in this study include the acylglycerol phosphate acyltransferase (AGPAT) SLC1, the glycerol-3-phosphate

dehydrogenase GPD, and the 6-phosphogluconolactonase SOL3. SLC1 was chosen for further study due to the dramatic increases in lipid production obtained from its individual overexpression. SLC1 is an intermediate step in the Kennedy pathway for lipid synthesis. In transferring fatty acids from CoA groups onto acylglycerol phosphate molecules, SLC1 overexpression likely increases the rate of lipid production by lessening the pool of fatty acyl-CoAs, which are feedback inhibitors of Acc1 [11, 17]. GPD and SOL3 were chosen due to the differences in the mechanisms by which they increase lipid production when overexpressed in comparison with ACC1 and DGA2; instead of directly catalyzing reactions of the fatty acid and lipid synthesis pathways, GPD and SOL3 are part of different pathways which produce some of the inputs required for the overall lipid synthesis pathway, which we hypothesized would make synergy between these genes and ACC1 and DGA2 more likely. Gpd catalyzes the conversion of dihydroxyacetone phosphate (DHAP) from glycolysis into glycerol-3-phosphate (Gly3P), which is the backbone on which fatty acids are added to by transesterification to form lipid molecules. In *B. napus*, expression of a yeast GPD1 that led to a twofold increase in Gly3P dehydrogenase activity caused a 40% increase in seed oil content [18]. SOL3 is the middle step in the pentose phosphate pathway (PPP), which a previous study has demonstrated is the major NADPH producing pathway for lipid synthesis in *Yarrowia* [19]. Although this enzyme does not itself produce NADPH, its overproduction may lead to higher lipid synthesis due to the faster removal of toxic intermediates [20] of the PPP, or due to this enzyme's production being

transcriptionally limited compared to others in the PPP (see **Chapter 3**).

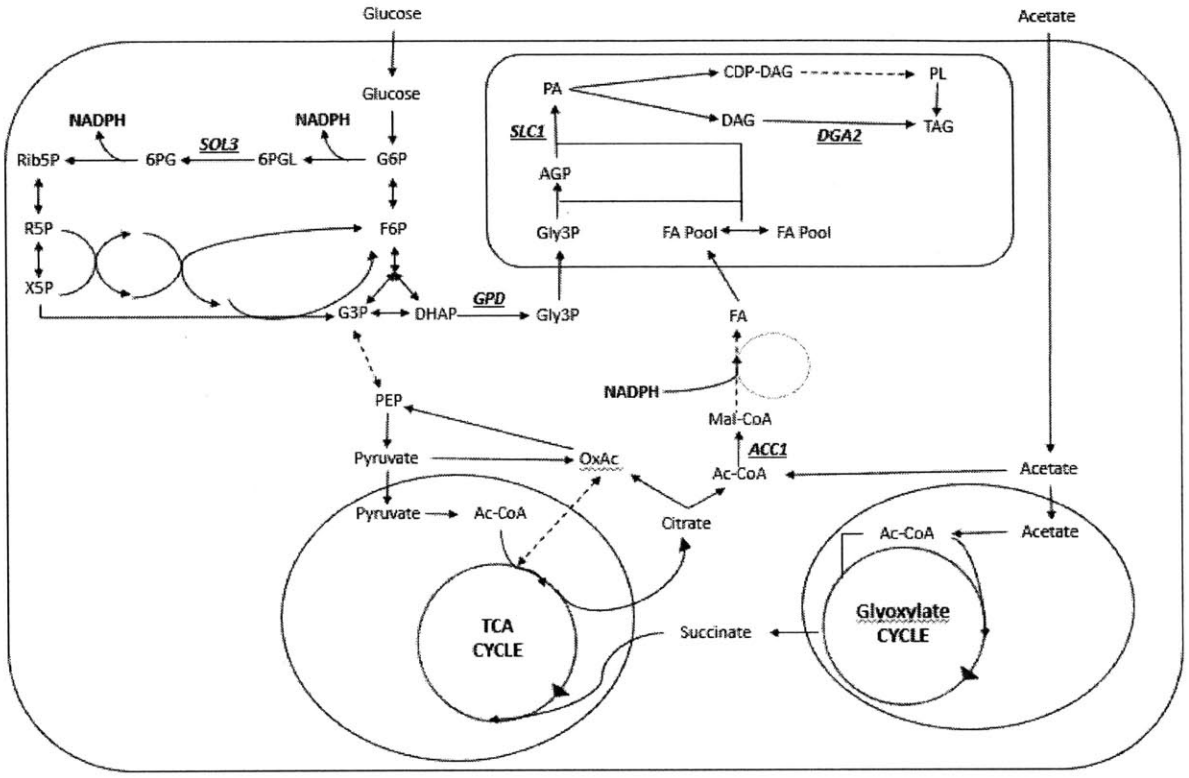

Figure 4.1: Map of the metabolic pathways from glucose and acetate to lipids. G6P: glucose-6-phosphate; 6PGL: 6-phosphogluconolactone; 6PG: 6-phosphogluconate; Rib5P: ribulose-5-phosphate; R5P: ribose-5-phosphate; X5P: xylulose-5-phosphate; F6P: fructose-6-phosphate; G3P: glyceraldehyde-3-phosphate; 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.

After the construction of the three strains, each overexpressing both ACC1 and DGA2 and one of either SLC1 (henceforth called strain "AD6"), GPD (strain "AD20"), or SOL3 (strain "AD29"), all three strains and the original ACC1- and DGA2- overexpression strain (strain "ADC") were grown in two fermentations, each with either glucose or acetate as the sole carbon source. **Figure 4.1** shows the pathways each carbon source takes from its uptake into the cell to

becoming a finished TAG molecule, with the five genes overexpressed in the constructs tested in this study highlighted. Due to the vast differences in the paths each carbon source must take, we expected that each construct may respond differently to each carbon source. Our results suggest that the ACC1, DGA2, and SLC1 overexpression strain may potentially be a more useful strain than the predecessor ACC1, DGA2 strain for both substrates.

4.2 Materials and Methods

4.2.1 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). Standard growth conditions for *E. coli* for the purposes of plasmid construction and generation have been previously described [21]. Rich YPD medium for *Y. lipolytica* strains contains 10 g/L yeast extract (BD Biosciences, New Jersey), 20 g/L Bacto^{TM} 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^{TM} 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 18 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 31 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 250 mL shake flasks containing 50 mL of YSMg or YSMa to an initial OD_{600} of 0.02. Flasks were closed with 34mm-45mm foam plugs (Jaece Industries Inc., North Tonawanda, NY) and placed in an Innova 43 Incubator Shaker (New Brunswick Scientific, Edison, NJ) and shaken for 155 hours (for glucose fermentations) or 178 hours (for acetate fermentations) at 250 rpm and 30°C. 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 flasks were used for each strain for glucose fermentations, and two biological replicate flasks each were used for acetate fermentations.

4.2.2 Plasmid and yeast strain construction and verification

Plasmid pMT065, containing the native ACC1 gene under control of the hp4d promoter and the native DGA2 gene under the full TEF-1α promoter (with intron), along with the LEU2 marker, was described previously [11]. The plasmid p91T contains a URA3 marker and an insertion site for gene coding sequences flanked by the full TEF-1α promoter upstream and by the XPR2 terminator downstream, all between two sites homologous to regions in the *Yarrowia* genome around the extracellular lipase LIP2 gene. Plasmid p91T was created by polymerase chain reaction

(PCR) amplifying pMT092, a previously described [16] derivative of the pACYCDUET-1 plasmid that contains the DGA2 gene in frame with the full TEF-1α promoter, with primers rAS396 and rAS397, which contain both binding sites and 5' extensions that overlap with each other, and performing a Gibson Assembly [22] on the single amplicon, closing the DNA molecule and excising out the DGA2 coding sequence. To construct the plasmids containing SLC1, GPD, and SOL3 overexpression genes ("pOP6U," "pOP20U," and "pOP29U," respectively), the coding sequences of all three genes without their ATG start codons were amplified by PCR and each was inserted into p91T that had been amplified by PCR with primers rAS429 and rAS317 by Gibson Assembly. The plasmids and primers used in this study are listed in **Table 4.1** and **Table 4.2**, respectively.

Table 4.1: Plasmids Used in this Study

Name	Marker	Gene(s) [Promoter-Coding]
pMT065	LEU2	hp4d-ACC1, TEF1α-DGA2
pMT092	URA3	TEF1α-DGA2
p91T	URA3	TEF1α-[Empty]
pOP6U	URA3	TEF1α-SLC1
pOP20U	URA3	TEF1α-GPD
pOP29U	URA3	TEF1α-SOL3

Gibson Assembly products were transformed into DH10β using standard techniques [21] and transformants were recovered and grown on selective Luria-Bertani medium containing chloramphenicol. Confirmation of gene insertion into pMT091 was carried out by PCR of the appropriate site followed by sequencing of purified plasmids obtained with the Qiagen Miniprep Kit (Qiagen, Valencia, CA).

Table 4.2: Primers Used in this Study

Name	Amplicon	Sequence (lowercase = extension, CAPS = BINDING)
rAS396	pMT092	gaggtacctccatggcctgtCCCCACGTTGCCGGTCTT
rAS397	pMT092	acaggccatggaggtacctcCTGCGGTTAGTACTGCAAAAAGTGCT
rAS429	p91T	catcttGAGGTACCTCCATGGC
rAS317	p91T	ctgcggttaGTACTGCAAAAAGTGCTGGTC
rAS470	SLC1	agcacttttgcagtactaaccgcagTCCGTTGCATCCAAGCTCGT
rAS471	SLC1	gaggtacctcaagatgCTACTGAGTCTTCTGGCCAGCGTAG
rAS419	GPD	agcacttttgcagtactaaccgcagAGCGCTCTACTTCGATCGTCC
rAS420	GPD	gaggtacctcaagatgCTAGTTGGCGTGGTAAAGAATCTCG
rAS421	SOL3	agcacttttgcagtactaaccgcagCCCAAGGTCATCTCTAAGAACGAATC
rAS422	SOL3	gaggtacctcaagatgCTAGATCTTGGAGAGGTTGACGCC

Strain ADC (ACC1 and DG2 overexpressing) was constructed by transforming pMT065 into strain po1g, as described previously [11]. For construction of new *Y. lipolytica* strains, strain Po1g (Leu$^-$) was first converted to strain Po1z (Leu$^-$ Ura$^-$) by knocking out the URA3 gene using a knockout cassette described previously [16] and screening for colonies on agar plates containing YSMg with 1 g/L 5'-fluoroorotic acid (5-FOA). Purified plasmids (pMT065, pOP6U, pOP20U, and pOP29U) were digested with the SacII restriction enzyme (NEB) and transformed into *Yarrowia* using the lithium acetate method [23]. 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). Po1z was transformed with pMT065 and colonies were recovered on plates of YSMg made with 15 g/L Bacto™ Agar (BD) and supplemented with 0.69 g/L CSM-Leu (Sunrise Science Products, San Diego, CA) to make strain po165. Strain po165 was then transformed with one of pOP6U, pOP20U, or pOP29U to form strains AD6, AD20, and AD29 after selection on YSMg agar plates supplemented with 0.67 g/L CSM-Leu-Ura (Sunrise). The strains used in this study and their genotypic descriptions are listed in **Table 4.3**.

Table 4.3: Strains Used in this Study

Name	Genotype	Source	Comments
po1g	MATa, leu2-270, ura3-302::URA3, xpr2-332, axp-2	Yeastern Biotech	Base strain, Leu auxotroph
po1z	As po1g, with ΔURA3	[16]	Leu and Ura auxotroph
po165	As po1z, with hp4d-ACC1, TEF1α-DGA2, LEU2+	[16], Remade for this work	ACC1, DGA2 overexpression, Ura auxotroph
ADC	As po1g, with hp4d-ACC1, TEF1α-DGA2, LEU2+	[11, 16], Remade for this work	ACC1, DGA2 overexpression
AD6	As po165, with TEF1α-SLC1, URA3+	This work	ACC1, DGA2, SLC1 overexpression
AD20	As po165, with TEF1α-GPD, URA3+	This work	ACC1, DGA2, GPD overexpression
AD29	As po165, with TEF1α-SOL3, URA3+	This work	ACC1, DGA2, SOL3 overexpression

4.2.3 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 [24]. Samples of 200 μL from glucose fermentations and 400 uL from acetate fermentations were taken at the end time point (155 hours for glucose, 178 hours for acetate) and 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.

4.3 Results and Discussion

4.3.1 Fermentation performance of multi-gene overexpression strains on glucose

The native SLC1, GPD, and SOL3 genes have previously been demonstrated (see **Chapter 3** [25]) to increase lipogenesis in *Yarrowia lipolytica* when individually overexpressed under the full Transcription Elongation Factor 1α (TEF-1α) promoter in a wild-type-approximate strain. To determine whether overexpressing any of these genes is useful in a closer to industrially relevant context (that is to say, in a high-oil producing strain), we transformed an additional copy of each of these genes individually into a previously developed highly oleaginous background that already overexpresses the native ACC1 and DGA2 genes. After confirming the presence of the additional copies of all genes needed to make each strain, all four strains (one for each gene

tested plus one control without an additional gene) were grown in triplicate shake flask fermentations with 50 mL of YSMg media for 155 hours. Fermentation performance of the strains was compared using the metrics of lipid titer (g lipids/L, this metric is also a proxy for average lipid productivity over the entire fermentation period), lipid content (g lipids/g dry cell weight), and lipid yield (g lipids/g glucose consumed). The results of the fermentation in glucose media are given in **Figure 4.2** and **Table 4.4**.

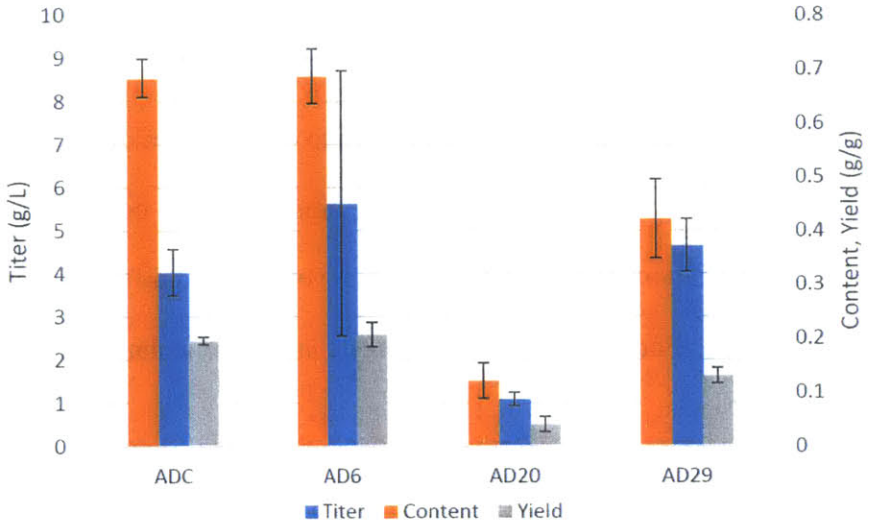

Figure 4.2: Fermentation performance of strains on glucose. Error bars represent the standard deviation of three biological replicates.

Table 4.4: Lipid Titers, Contents, Yields, and Total Biomass from Glucose Fermentations

Strain	Lip. Titer (g/L)	Lip. Content (g/g)	Lip. Yield (g/g)	DCW (g/L)
ADC	4.03	0.684	0.195	5.88
AD6	5.63	0.687	0.207	8.02
AD20	1.09	0.121	0.041	9.17
AD29	4.65	0.422	0.130	11.31

The ADC control strain produced a lipid titer of 4.03 g/L on average, making up 68.4% of dry cell weight, at a yield of 0.195 grams of lipids per gram of glucose. Our results show that strain AD6 performed the best of all strains in this experiment. The overexpression of SLC1 in this strain led to a 40% increase on average in lipid titer over the control strain, with the same lipid content, and a very slight improvement in yield. SLC1 is an acylglycerol phosphate acyltransferase, catalyzing the addition of a second fatty acyl group onto the glycerol backbone of a lipid molecule. This enzyme therefore acts before the branch point in lipid synthesis pathways between phospholipid and neutral lipid production. Phospholipids make up the outer and intracellular membranes of cells, and as such, are required for overall cell growth and division [26], while neutral lipids are mostly stored in lipid bodies in the endoplasmic reticulum for use as energy or a source of fatty acids for phospholipid synthesis during starvation conditions [27]. As the lipid titer increase of the AD6 strain was commensurate with the increase in its DCW (36%) over the ADC strain, it is likely that the SLC1 overexpression promoted higher growth by effectively increasing the production of phospholipids.

Surprisingly, strain AD20 performed extremely poorly in this experiment, with GPD overexpression leading to decreases in titer, content, and yield of 73%, 82%, and 79%. However, it is not likely that these three overexpressions led to unhealthy cells, as the DCW of this strain was on average 56% greater than that of the control strain. GPD is a cytosolic glycerol-3-phoshate dehydrogenase, catalyzing the reduction of dihydroxyacetone phosphate into glycerol-3-phosphate and oxidizing NADH to NAD+ in the process. GPD in the yeast Saccharomyces cerevisiae can be used to recycle excess reducing equivalents from NADH by producing glycerol from glycolytic intermediates [28]. Producing lipids at high rates represents an excess of carbon and energy in the cell, and therefore requires a commensurately high energy

input; with the TCA cycle becoming largely deactivated at the onset of lipogenesis in oleaginous organisms after nitrogen depletion [29], most of the ATP to make lipids must therefore come from NADH generated during glycolysis. The deleterious effects of GPD overexpression in this ACC1 and DGA2 overexpressing background on lipid production may come from reducing the supply of NADH too much to keep up with the energy demand for high lipid production, and the high biomass productivity exhibited by this strain may also be the result of high availability of the oxidized NAD+ allowing for high rates of glycolytic flux.

Strain AD29 showed an increase in lipid titer of 16%, but decreases in content and yield of 38% and 33%, respectively, compared to the control strain. This strain also produced an almost twofold higher (92%) total biomass than the control strain. SOL3, 6-phosphogluconolactonase, is the intermediate (but possibly rate-limiting in *Yarrowia* [25]) step in the NADPH producing pentose phosphate pathway. Aside from being the source for reducing electrons during fatty acid synthesis, NADPH is an important cofactor for many anabolic reactions in amino acid and nucleic acid biosynthetic pathways [30]. High-oil *Yarrowia* strains growing on glucose have been found to automatically increase their PPP flux in response to high rates of fatty acid synthesis [19], so in the case of the AD29 strain, it is likely that NADPH production was not a limiting factor in lipid biosynthesis. Excess NADPH likely increased the rates of the other anabolic pathways, leading to the extreme increase in biomass productivity relative to lipid synthesis.

Unfortunately, although a couple of our new strains on average outperformed the control strain in at least one metric, it is important to note that none of the increases in any metric we observed were found to be statistically significant at the $\alpha = 0.05$ level. However, it is probable

that with a higher number of biological replicates tested, the certainty in the differences observed here can be magnified to the point of achieving statistical significance.

In addition to the three metrics for fermentation performance described above, the fatty acid profiles of each strain were observed, and are illustrated in **Figure 4.3**. SLC1 overexpression did not seem to alter the fatty acid composition at all. Strain AD20 has a very different profile from the control strain, with a significantly lower fraction of stearate (C18) and higher fractions of palmitate (C16), palmitoleate (C16:1), and linoleate (C18:2). There does not seem to be an apparent explanation for this deviation in fatty acid profile. Strain AD29 demonstrated a striking increase in palmitoleate and oleate (C18:1) fractions compared to the saturated fatty acids. This may be more evidence of the SOL3 overexpression in this strain being effective at producing excess NADPH, as it is required for regenerating the cofactors used by fatty acid desaturases [31].

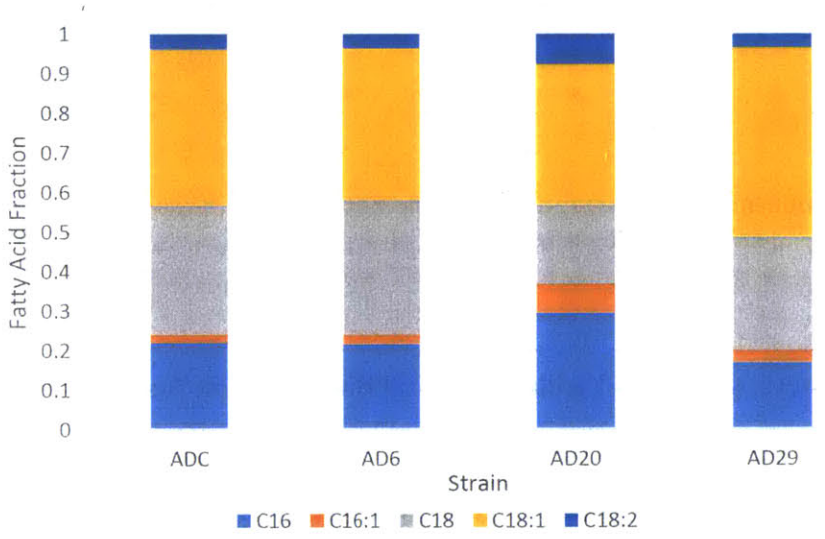

Figure 4.3: Fatty acid profiles of strains grown on glucose. Percentages shown here are the average of six measurements.

4.3.2 Fermentation performance of multi-gene overexpression strains on acetate

In addition to testing the fermentation performance of our multi-gene overexpression strains on glucose, we also assessed their fermentation performance on acetate. For this experiment, strains were grown in shake flask cultures of 50 mL YSMa for 178 hours with two biological replicates for each strain. The results of this fermentation are shown in **Figure 4.4** and **Table 4.5**.

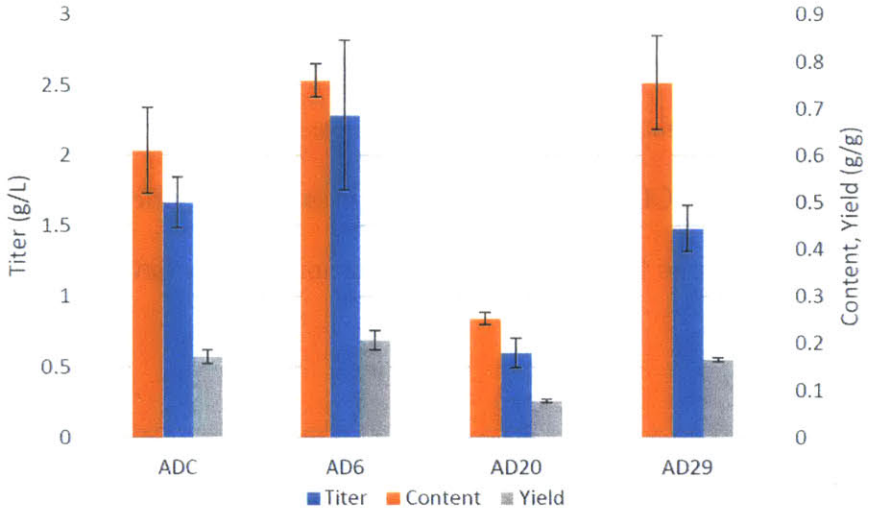

Figure 4.4: Fermentation performance of strains on acetate. Error bars represent the standard deviation of two biological replicates.

Table 4.5: Lipid Titers, Contents, Yields, and Total Biomass from Acetate Fermentations

Strain	Lip. Titer (g/L)	Lip. Content (g/g)	Lip. Yield (g/g)	DCW (g/L)
ADC	1.67	0.611	0.172	2.78
AD6	2.29	0.760	0.207	3.03
AD20	0.60	0.253	0.078	2.38
AD29	1.48	0.755	0.165	2.00

The control ADC strain produced a lipid titer, content, and yield of 1.67 g/L, 0.611 gram lipids per gram DCW, and 0.172 gram lipids per gram of acetate consumed, respectively. SLC1 overexpression led to increase in the lipid titer, content, and yield of the AD6 strain by 37%, 24%, and 20%, respectively. As opposed to the case on glucose, the lipid content increase in the AD6 strain along with the comparatively modest increase in biomass productivity suggests that a significant fraction of the increased lipid synthesis flux in this strain on acetate becomes TAGs as opposed to phospholipids. The performance of strain AD20 on acetate was similarly poor to its performance on glucose, with decreases in titer, content, and yield of 64%, 59%, and 55%, respectively compared to strain ADC; it is likely that the same NADH/energy consumption issue that hinders this strain's performance on glucose is similarly present here on acetate. Lastly, the overexpression of SOL3 in strain AD29 led to an increase in lipid content of 24%, and only slight, insignificant decreases in titer of 11% and in yield of 4%. The performance of this strain on acetate resembles slightly that of the SOL3 individual overexpression strain described previously [25], which was characterized by slow growth rate, low lipid titer and yield, but high lipid content. The speculated increase in NADPH availability in this strain directs higher lipid flux through the fatty acid synthesis pathway, which is immediately accessible for the imported acetate after conversion to acetyl-CoA, and away from any other biosynthetic pathways. In addition to being an inhibitor of Acc1 [32], palmitoyl-CoA has also been shown to be an inhibitor of Zwf1, catalyzing the first step of the oxidative pentose phosphate pathway, in a different yeast, *T. utilis* [33]; the much improved performance of the AD29 strain on acetate in contrast to the slowly growing SOL3-only overexpression strain (see Chapter 3) may be due to the overexpression of DGA2 decreasing the palmitoyl-CoA pool through the reaction with DAGs to form TAGs.

The fatty acid profiles of all four strains on acetate are illustrated in **Figure 4.5**. Strain AD6 showed a strong preference for oleate and against saturated fatty acids. Strain AD20 showed significant decreases in palmitoleate and oleate, with the oleate fraction being smaller than the stearate fraction; this fatty acid profile is typical of very low-oil or unhealthy cell cultures or those that are in the early stages of growth. Strain AD29 showed a preference for unsaturated fatty acids, which would be overproduced in the presence of a larger available NADPH pool.

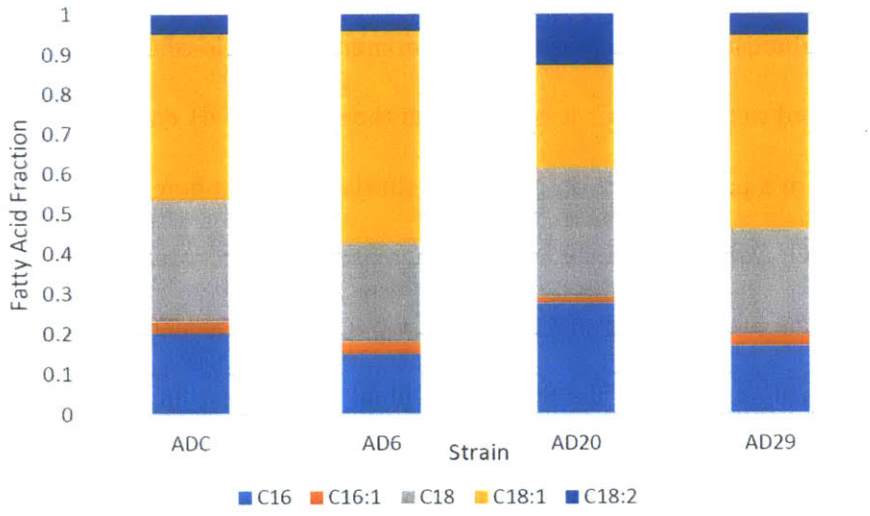

Figure 4.5: Fatty acid profiles of strains grown on acetate. Percentages shown here are the average of four measurements.

Unfortunately, as was the case on glucose, all of the increases in lipid titer, content, and yield over ADC by the other strains observed on acetate were not found to be statistically significant at the $\alpha = 0.05$ significance level, although the increases in lipid content and yield of strain AD6 over ADC were significant at the $\alpha = 0.1$ level (with p-values of 0.0829 and 0.0961, respectively). An experiment with more biological replicates would likely in this case increase certainty beyond that needed to achieve a higher level of statistical significance; in this case, the

lipid yield of strain AD6 we achieved on acetate will be the highest lipid yield on acetate ever reported in *Yarrowia* or any other oleaginous yeast.

4.4 Conclusions

In this study, we explored the potential of a few gene targets previously validated to increase lipogenesis when overexpressed in *Yarrowia lipolytica* as to their effectiveness in a high-oil background that already overexpresses the native ACC1 and DGA2 genes. Our results demonstrated that SLC1 overexpression in this background was more effective than overexpression of either GPD or SOL3, with the resulting strain demonstrating an increase in lipid titer/productivity of 40% on glucose and increase in titer, content, and yield of 37%, 24%, and 24%, respectively, over the background control. This finding, along with the severely detrimental effect of GPD overexpression and the substrate-dependent, mixed effects of SOL3 overexpression in this background, demonstrate the vast range of potential phenotypic effects that may arise from combinations of individually positive genetic modifications to a host strain, underscoring the degree of connectivity of a cell's metabolic network and the need for balance among its constituent pathways in order to obtain an optimal output from it. This observed difficulty in predicting synergy among genetic modifications can make even marginal improvements in strain performance valuable when assessing process economics. While it seems that overexpressing the native ACC1, DGA2, and SLC1 genes simultaneously leads to a *Yarrowia* strain superior to one that only overexpresses ACC1 and DGA2, especially for acetate fermentations, in which the lipid yields are the highest ever reported for this substrate, experiments with more replicates would need to be carried out to achieve more conclusive statistical proof of this.

References:

1) 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. *Progress in Lipid Research*, *48*(6), 375–87. doi:10.1016/j.plipres.2009.08.005
2) 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. *European journal of lipid science and technology*, *111*(12), 1221-1232.
3) 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 Technology*, *114*, 443–9. doi:10.1016/j.biortech.2012.02.091
4) 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. *Applied Microbiology and Biotechnology*, *58*(3), 308–12. doi:10.1007/s00253-001-0897-0
5) Fickers, P., Benetti, P.-H., Waché, Y., Marty, a, Mauersberger, S., Smit, M. S., & Nicaud, J.-M. (2005). Hydrophobic substrate utilisation by the yeast Yarrowia lipolytica, and its potential applications. *FEMS Yeast Research*, *5*(6-7), 527–43. doi:10.1016/j.femsyr.2004.09.004
6) Rude, M. a., & Schirmer, A. (2009). New microbial fuels: a biotech perspective. *Current Opinion in Microbiology*, *12*(3), 274–281. doi:10.1016/j.mib.2009.04.004
7) Xue, Z., Sharpe, P. L., Hong, S.-P., Yadav, N. S., Xie, D., Short, D. R., ... Zhu, Q. (2013). Production of omega-3 eicosapentaenoic acid by metabolic engineering of Yarrowia lipolytica. *Nature Biotechnology*, *31*(8), 734–40. doi:10.1038/nbt.2622
8) 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 (Reading, England)*, *153*(Pt 7), 2013–25. doi:10.1099/mic.0.2006/002683-0
9) Van Erp, H., Kelly, A. A., Menard, G., & Eastmond, P. J. (2014). Multigene Engineering of Triacylglycerol Metabolism Boosts Seed Oil Content in Arabidopsis. *Plant Physiology*, *165*(1), 30–36. doi:10.1104/pp.114.236430
10) Vanhercke, T., El Tahchy, A., Shrestha, P., Zhou, X. R., Singh, S. P., & Petrie, J. R. (2013). Synergistic effect of WRI1 and DGAT1 coexpression on triacylglycerol biosynthesis in plants. *FEBS Letters*, *587*(4), 364–369. doi:10.1016/j.febslet.2012.12.018
11) Tai, M., & Stephanopoulos, G. (2013). Engineering the push and pull of lipid biosynthesis in oleaginous yeast Yarrowia lipolytica for biofuel production. *Metabolic Engineering*, *15*, 1–9. doi:10.1016/j.ymben.2012.08.007
12) Yan, Y., & Liao, J. C. (2009). Engineering metabolic systems for production of advanced fuels. *Journal of Industrial Microbiology and Biotechnology*, *36*(4), 471–479. doi:10.1007/s10295-009-0532-0
13) Warnecke, T., & Gill, R. T. (2005). Organic acid toxicity, tolerance, and production in Escherichia coli biorefining applications. *Microbial Cell Factories*, *4*, 25. doi:10.1186/1475-2859-4-25
14) Chávez-Béjar, M. I., Balderas-Hernandez, V. E., Gutiérrez-Alejandre, A., Martinez, A., Bolívar, F., & Gosset, G. (2013). Metabolic engineering of Escherichia coli to optimize

melanin synthesis from glucose. *Microbial Cell Factories*, *12*, 108. doi:10.1186/1475-2859-12-108
15) Görgens, J. F., Van Zyl, W. H., Knoetze, J. H., & Hahn-Hägerdal, B. (2001). The metabolic burden of the PGK1 and ADH2 promoter systems for heterologous xylanase production by saccharomyces cerevisiae in defined medium. *Biotechnology and Bioengineering*, *73*(3), 238–245. doi:10.1002/bit.1056
16) Tai, M. (2012). Metabolic Engineering of Oleaginous Yeast for the Production of Biofuels. Ph.D. Thesis. Massachusetts Institute of Technology.
17) Thelen, J. J., & Ohlrogge, J. B. (2002). Metabolic engineering of fatty acid biosynthesis in plants. *Metabolic Engineering*, *4*, 12–21. doi:10.1006/mben.2001.0204
18) 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 Biotechnology Journal, 5: 431–441. doi: 10.1111/j.1467-7652.2007.00252.x
19) 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. *Metabolic Engineering*, *30*, 27–39. doi:10.1016/j.ymben.2015.02.007
20) 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. *Journal of Biological Chemistry*, *276*(37), 34840–34846. doi:10.1074/jbc.M105174200
21) Sambrook, J., & Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. CSHL press, Cold Spring Harbor, NY, USA.
22) 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. *Nature Methods*, *6*(5), 343–345. doi:10.1038/nmeth.1318
23) Chen, D. C., Beckerich, J. M., & Gaillardin, C. (1997). One-step transformation of the dimorphic yeast Yarrowia lipolytica. *Applied Microbiology and Biotechnology*, *48*, 232–235. doi:10.1007/s002530051043
24) 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.
25) Silverman, A. M.*, Qiao, K.*, Xu, P., and Stephanopoulos, G. Functional overexpression and characterization of lipogenesis-related genes in the oleaginous yeast *Yarrowia lipolytica*. Submitted.
26) Carman, G. M., & Han, G.-S. (2011). Regulation of phospholipid synthesis in the yeast Saccharomyces cerevisiae. *Annual Review of Biochemistry*, *80*, 859–883. doi:10.1146/annurev-biochem-060409-092229
27) Daum, G., Wagner, a., Czabany, T., & Athenstaedt, K. (2007). Dynamics of neutral lipid storage and mobilization in yeast. *Biochimie*, *89*(2), 243–248. doi:10.1016/j.biochi.2006.06.018
28) Ansell, R., Granath, K., Hohmann, S., Thevelein, J. M., & Adler, L. (1997). The two isoenzymes for yeast NAD+-dependent glycerol 3-phosphate dehydrogenase encoded by GPD1 and GPD2 have distinct roles in osmoadaptation and redox regulation. *EMBO J*, *16*(9), 2179–2187.

29) Ratledge, C. (2002). Regulation of lipid accumulation in oleaginous micro-organisms. *Biochemical Society Transactions*, *30*(6), 47–50.
30) Bruinenberg, P. M., Van Dijken, J. P., & Scheffers, W. a. (1983). A Theoretical Analysis of NADPH Production and Consumption in Yeasts. *Microbiology*, *129*(4), 953–964. doi:10.1099/00221287-129-4-953
31) Nelson, D. L. & Cox, M. M. (2005). *Lehninger principles of biochemistry* (4th ed.). New York: W.H. Freeman. Pg. 799.
32) Tehlivets, O., Scheuringer, K., & Kohlwein, S. D. (2007). Fatty acid synthesis and elongation in yeast. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1771*(3), 255–270. doi:10.1016/j.bbalip.2006.07.004
33) Natter, K., & Kohlwein, S. D. (2013). Yeast and cancer cells - Common principles in lipid metabolism. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1831*(2), 314–326. doi:10.1016/j.bbalip.2012.09.003

Chapter 5

Knockout of SNF1 Protein Kinase and its Effect on High-Oil Strains

5.1 Introduction

Yarrowia lipolytica is an oleaginous yeast that has recently received significant attention as a model host for single cell oil (SCO) processes for the production of biodiesel. This organism can naturally accumulate lipids as up to 36% of its biomass, with over 90% of these lipids in the form of triacylglyerols (TAGs) [1]. While other oleaginous fungi may naturally accumulate higher lipid contents [2], *Yarrowia* has a distinct advantage in that its genome has been sequenced [3] and lightly annotated, and genetic techniques and tools such as transformation [4] and promoters [5-7] for use in native gene overexpression or heterologous gene expression have been developed, allowing for the metabolic engineering of strains for improved lipid productivity and content or for production of fatty acids not natively produced, such as eicosapentaenoic acid (EPA) [8].

Previous engineering efforts for the purpose of improving lipid production in *Yarrowia* have largely focused on overexpressing native genes coding for enzymes directly involved in lipid synthesis, such as acetyl-CoA carboxylase (ACC1) and diacylglycerol acyltransferase (DGA2) [6]. While overproducing an enzyme can be an effective method of increasing the flux through a metabolic pathway of interest, in some cases, the enzyme activity at a particular node can be limited due to post-translational modifications made by enzyme regulators. Snf1 is a protein kinase, capable of phosphorylating Acc1 [9] at a particular serine residue (most likely Ser-1178 in *Yarrowia*, from comparison with sequence near the *Saccharomyces cerevisiae* Acc1 phosphorylation site, Ser-1157 [10]) thus inactivating the first committed step of fatty acid synthesis. Snf1 is the catalytic α subunit in the three subunit AMPK (AMP-activated protein kinase) complex along with one of three possible β subunits Sip1, Sip2, or Gal83, which are involved in localization of the complex to various cellular compartments, and the γ regulatory

subunit Snf4 [9]. Snf1 becomes activated by upstream regulatory kinases in response to decreased energy charge in the cell (ATP:AMP ratio) and is generally activated for adapting the cell to use carbon substrates other than glucose, although it may not be required for adaptation in *Yarrowia lipolytica* due to its normally incomplete glucose repression [11].

Gene knockouts are typically achieved by transformation of a linearized DNA molecule (or cassette) containing a marker gene (such as LEU2, URA3, or HYG) flanked by sequences homologous to regions just upstream and downstream of the ends of the target gene's coding sequence. When this cassette enters the cell, integration into the genomic DNA can be mediated by two different protein complexes, each producing different results regarding the location of the cassette in the genomic DNA in relation to that of the target gene locus; successful knockout of a gene of interest requires targeting the specific locus containing that gene. After exonucleases resect DNA double strand breaks (DSBs) in genomic DNA that occur during the transformation process, the homologous recombination (HR) protein complex directs the transformed cassette to a site in the gDNA homologous to sequences in the cassette; if the target locus is the site of a resected DSB, then the HR complex may "shuffle" DNA strands of the gDNA and cassette, potentially leading to strand crossovers, in which the target gene is replaced by the marker gene [12]. In the process of non-homologous end joining, DSBs are repaired by a different protein complex that ligates the ends of two DNA strands with no consideration for sequence homology or strand overhang complementarity [13]; during cassette transformation, the NHEJ complex can join the ends of a cassette to any pair of DNA ends formed from a DSB in a genome, therefore, cassette integration mediated by the NHEJ complex can occur anywhere in the genome, and targeting to the locus of the gene of interest is statistically extremely rare.

While others have reported on their ability to knock out genes in *Yarrowia* with relative frequency, our lab and some others have encountered significant difficulties in observing locus targeting in the resultant colonies after knockout cassette transformation. *Yarrowia lipolytica* has a relatively active NHEJ complex [14], which outcompetes HR complex proteins for binding the ends of the transformed cassette, leading to most transformants containing the marker gene in indiscriminate locations throughout their genomes. In our experience, the *Yarrowia lipolytica* strain po1g and its derivatives seem to have an especially overactive NHEJ complex, as screening several hundred transformant colonies using several different variations of traditional homology cassette transformation protocols (in terms of cassette homology site length, heat shock temperature and time, split marker cassettes, partial exonuclease digestion for single stranded ends on cassettes, etc.) for knockouts of various genes proved fruitless; similar transformations in strains with more reported gene targeting success, such as po1d, produced similar results.

To facilitate generating a reliable platform for gene knockouts in *Yarrowia*, we attempted to first construct a strain with reduced NHEJ activity by knocking out the gene coding for a principal component of the complex, KU70 [14]. As the traditional homology cassette transformations did not work, we looked into other transformation methods that may favor the HR complex over NHEJ or circumvent the NHEJ complex altogether. Eventually, our collaborators at the Pacific Northwest National Laboratory (PNNL) achieved a knockout of KU70 using *Agrobacterium tumefaciens*-mediated transformation. *A. tumefaciens* is a bacterium found in soil that can infect plants and cause tumors by injecting its Ti plasmid (or T-DNA) into their cells in response to sensing the acetosyringone that is secreted from the plants' wounds [15]. By inserting genes into the T-DNA, plants and fungi can acquire these genes by incubating

their cells with the modified plasmid-carrying *A. tumefaciens* in the presence of acetosyringone. Including marker genes flanked by sequences homologous to those flanking a target gene in the gDNA of fungi allows for HR to occur at that locus; it has been demonstrated previously in *A. awamori* [16] and *K. lactis* [17] that transformation mediated by *A. tumefaciens* significantly improves the relative frequency of HR-derived transformants versus NHEJ-derived transformants in comparison to traditional cassette transformation methods. This phenomenon has been attributed to the T-DNA being inserted as single stranded DNA (which in many organisms undergoes HR at higher frequencies than double stranded DNA) and in a complex with the VirD2 and VirE2 proteins [16].

As Snf1 is an inhibitor to a key enzyme in fatty acid biosynthesis, in this study, we used the ΔKu70 background strain generously provided to us by our collaborators at PNNL to construct an SNF1-null mutant of *Yarrowia lipolytica* and determined whether this modification leads to increased ability to produce lipids on both glucose and acetate. SNF1 has already been reported to have been knocked out in *Yarrowia* [18], so we attempted here to verify the previously observed phenotype and then test whether the ΔSNF1 mutation can act synergistically with overexpressions of native genes also previously verified to improve lipogenesis, namely, ACC1 and DGA2. Our results showed that we were able to successfully verify construction of our ΔSNF1 strain in the ΔKu70 background with relative ease; we were also able to verify that knockout of SNF1 does lead to increased lipid content of cells on glucose as reported previously, and this effect is significantly more pronounced on acetate. The additional overexpression modifications we have chosen to combine with ΔSNF1 did not lead to improved lipid production on glucose, but DGA2 overexpression did lead to a strain characterized by very high lipid contents and yields on acetate in the ΔSNF1 background, but only when not combined with

ACC1 overexpression. In addition to producing another strain (see Chapter 4) with improved lipid yield over our controls, and achieving what is the highest reported lipid content for a *Yarrowia* strain grown on acetate, in this study we again demonstrate the concept of how sensitive metabolic flux through a pathway can be in response to increased activity of that pathway's enzymes [19], illustrated here with an example of how flux versus pathway enzyme activity is not a monotonically increasing function, but rather can reach an optimal level and sharply decrease in response to further attempts at improvement.

5.2 Materials and Methods

5.2.1 Plasmid and *Yarrowia lipolytica* strain construction and verification

All *Escherichia coli* strains developed in this study are plasmid carrying transformants of the cloning strain DH10β (New England Biolabs, Ipswich, MA). Standard growth conditions for *E. coli* for the purposes of plasmid construction and generation have been previously described [20].

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). A list of all plasmids used in this study is shown in **Table 5.1**. Plasmids pMT038, pMT053, and pMT065 are described previously [6, 19]: all plasmids contain the LEU2 marker and all other genes use the native *Yarrowia* XPR2 terminator, plasmid pMT038 contains the *E. coli* LacZ gene under control of the Transcription Elongation Factor-1α (sans associated intron, and henceforth will be referred to as "TEF1αsi") promoter, which for the purposes of our experiments is meant to provide the LEU2 marker while causing the metabolic burden of expressing a gene (on a similar scale to other strains overexpressing native genes) that should not be relevant to the phenotype measured in our fermentations (see [6], [19], and Chapter 3), plasmid pMT053 contains the native DGA2 gene under control of the full TEF1α promoter, and plasmid

pMT065 contains a copy of the native ACC1 under the control of the hp4d promoter and a copy of DGA2 under the control of the full TEF1α promoter.

Plasmid pASAe contains the ACC1 coding sequence under control of the full TEF1α promoter. The *Yarrowia* ACC1 coding sequence is 7270 bases long, containing two introns. We therefore inserted this gene into our base plasmid using a piecewise approach in which we removed the introns from the cloned gene (which henceforth will be referred to as ACC1e). A list of all primers used in this study is given in **Table 5.2.** Plasmid pMT015, containing the full TEF1α promoter and described previously [6], was first PCR amplified using primers rAS316 and rAS317, and the second exon of ACC1 was PCR amplified from *Yarrowia* gDNA with rAS345 and rAS404. These amplicons were mixed together along with primers rAS343 and rAS344, which are complementary strands containing the first ACC1 exon and with one end complementary to the amplified pMT015 and the other complementary to the overhang of primer rAS345; these DNA strands were joined together using the Gibson Assembly Method [21] and Kit (NEB). The resulting plasmid was then amplified with rAS316 and rAS404, and the first half (approximately) of the third exon of ACC1 was amplified with rAS347 and rAS405; these fragments were joined by Gibson Assembly. Finally, this plasmid was amplified with rAS316 and rAS408 and the rest of the third exon from ACC1 was amplified with rAS407 and rAS348; these fragments were joined by Gibson Assembly to produce plasmid pASAe. Assembly products were confirmed after each step by PCR verification and sequencing.

Plasmid pSNF1KO contains a URA3 marker flanked upstream by a 1639 bp site (the "5'" site) homologous to the region in *Yarrowia* gDNA just upstream of the SNF1 coding sequence (NCBI Locus Tag: YALI0D02101g) and including the first 438 bases of the coding sequence, and downstream by a 1765 bp site (the "3'" site) homologous to a region in *Yarrowia* gDNA just

downstream of the SNF1 coding sequence and including the last 659 bases of the coding sequence. This plasmid was constructed by Gibson Assembly of the 3 fragments (the 5' site, marker, and 3' site) into a PCR amplified pUC19 plasmid. The pUC19 plasmid was amplified using primers rAS254 and rAS255, the 5' site was amplified from *Yarrowia* gDNA using primers rAS252 and rAS139, the URA3 marker was amplified from plasmid pMT092 (described in Chapter 4) with primers rAS142 and rAS143, and the 3' site was amplified from *Yarrowia* gDNA using primers rAS140 and rAS253.

Table 5.1: Plasmids and *Yarrowia* strains used in this study.

Plasmids			
Name	**Description**	**Auxotrophic Marker**	**Source**
pMT015	contains TEF1α promoter	LEU2	[6]
pMT038	contains TEF1αsi-LacZ	LEU2	[6]
pMT053	contains TEF1α-DGA2	LEU2	[6]
pMT065	contains hp4d-ACC1, TEF1α-DGA2	LEU2	[6]
pASAe	contains TEF1α-ACC1e	LEU2	this work
pUC19	base plasmid for knockout cassette propagation	None	NEB
pSNF1KO	contains SNF1 knockout cassette	URA3	this work

***Yarrowia* Strains**		
Name	**Genotype**	**Source**
Po1g	MATa, leu2-270, ura3-302::URA3, xpr2-332, axp-2	Yeastern Biotech
Po1z	As Po1g, with ΔURA3	[19]
YLKuKO	As po1g, with ΔKu70	PNNL
YLKuKO-U	As YLKuKO, with ΔURA3	this work
YLKSKO	As YLKuKO, with SNF1::URA3	this work
YL38K	As YLKuKO-U, with URA3+, LEU2+, and TEF1αsi-LacZ	this work
YLAeK	As YLKuKO-U, with URA3+, LEU2+, and TEF1α-ACC1e	this work
YL53K	As YLKuKO-U, with URA3+, LEU2+, and TEF1α-DGA2	this work
YL65K	As YLKuKO-U, with URA3+, LEU2+, and hp4d-ACC1, TEF1α-DGA2	this work
YL38S	As YLKSKO, with URA3+, LEU2+, and TEF1αsi-LacZ	this work
YLAeS	As YLKSKO, with URA3+, LEU2+, and TEF1α-ACC1e	this work
YL53S	As YLKSKO, with URA3+, LEU2+, and TEF1α-DGA2	this work
YL65S	As YLKSKO, with URA3+, LEU2+, and hp4d-ACC1, TEF1α-DGA2	this work

Yarrowia transformations in our lab were performed using the lithium acetate method [22]. The URA3 gene in strain YLKuKO was disrupted by transformation with a cassette containing the first 234 bp of the gene and the last 300 bp of the gene followed by recovery on 1 g/L 5'-fluoroorotic acid (5-FOA) plates, yielding strain YLKuKO-U. Strains YL38K, YLAeK, YL53K, and YL65K were constructed from YLKuKO-U first by complementation with the URA3 marker amplified from plasmid pMT092, followed by transformation of linearized plasmid pMT038, pASAe, pMT053, or pMT065, respectively.

Strain YLKSKO was achieved with the successful knockout of SNF1 by transformation with the cassette from pSNF1KO amplified with primers rAS252 and rAS253, followed by recovery of the transformation mixture on uracil dropout agar plates with acetate as a sole carbon source. The knockout strain was confirmed first by attempting to PCR amplify the entire coding sequence of the gene using primers rAS394 and rAS395 from the gDNA of transformant colonies and performing gel electrophoresis on the finished reaction mixture in the presence of SYBR Safe DNA Staining Dye (Life Technologies, Grand Island, NY) alongside a 2-Log DNA Ladder (NEB). The knockout cassettes were designed to contain the first 438 bp of the 5' region of the coding sequence in the 5' site region, and the last 659 bp of the 3' region of the coding sequence in the 3' site region; the combined length of these fragments along with the URA3 marker between them was significantly larger (2312 bp) than the native coding sequence (1740 bp). Visualizing the gel on a Proteinsimple AlphaImager HP (Bio-Techne, Minneapolis, MN) gel imager was used for the first confirmation of a knockout, as the larger amplicon from the knockout cassette was easy to differentiate from the smaller band that would be obtained from the native coding sequence. Results from a colony that showed the larger band only were considered successful for this first step. In the second step, RNA was extracted from colonies that "passed" the first step and a q-RT-

PCR analysis was performed on them and their SNF1 expression values were compared to those of strain YLKuKO (see **Section 5.2.2**).

Table 5.2: Primers used in this study

Name	Sequence
rAS316	CATCTTGAGGTACCTCCATG
rAS317	ctgcggttaGTACTGCAAAAAGTGCTGGTC
rAS343	agcacttttgcagtactaaccgcagCGACTGCAATTGAGGACACTAACACGTCGGTTTTTCAG**TATGGCTTCAGGATCTTCAACGC**
rAS344	**GCGTTGAAGATCCTGAAGCCATA**CTGAAAAACCGACGTGTTAGTGTCCTCAATTGCAGTCGctgcggttagtactgcaaaaagtgct
rAS345	CTAACACGTCGGTTTTTCAGTATGGCTTCAGGATCTTCAACGC
rAS404	gaggtacctcaagatgCTTGTTGATAACTGTATGACCTCCG
rAS347	GTCATACAGTTATCAACAAGGTCCTCATCGCTAACAACGGTATTG
rAS405	gaggtacctcaagatgAATCAGCTCGAGACACCGAGG
rAS407	CCTCGGTGTCTCGAGCTGATTCCGTCTCCGACTTTTCGTACAC
rAS348	gaggtacctcaagatgTCACAACCCCTTGAGCAGCTC
rAS408	GTGTACGAAAAGTCGGAGACGGAATCAGCTCGAGACACCGAGG
rAS254	GAGAGAGAGAAAGGCTGCCAGGGGATCCTCTAGAGTCGACCTGC
rAS255	CCAACAGTTCCAGGGTGTAATCAGGGGTACCGAGCTCGAATTCAC
rAS252	TGAATTCGAGCTCGGTACCCCTGATTACACCCTGGAACTGTTGGG
rAS139	GGGCATGTTGTTGTGTTTCTCG**TTA**TCTATGGCAGTACTCGACTGC
rAS140	GTCAATGGCTCTCTGGGCGAATCCGGCAAGCCTAAGCAC
rAS253	GCAGGTCGACTCTAGAGGATCCCCTGGCAGCCTTTCTCTCTCTC
rAS142	GCAGTCGAGTACTGCCATAGA**TAA**CGAGAAACACAACAACATGCCC
rAS143	GTGCTTAGGCTTGCCGGATTCGCCCAGAGAGCCATTGAC
rAS394	ATGGCGACCGAACACGTGGAAC
rAS395	TTACTTCTCACTCTCCTTCTGAGAAC
rAS106	CCTGAACCGAATCACCGTCC
rAS107	TCACGTGGAACAAGGTACTCG
rAS434X	CCGAGGTTATCAGTGGCAAGC
rAS435X	ACAGCATGACGTATAGAATGACACC
MTR001	TCCAGGCCGTCCTCTCCC
MTR002	GGCCAGCCATATCGAGTCGCA

5.2.2 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 [6], with some modification. Samples from 50 mL shake flask cultures of the YLKuKO strain and each YLKSKO candidate transformant growing on YSMg media (see

Section 5.2.3) were harvested at 36 hours after inoculation and pelleted. The pellets were resuspended in 1 mL Trizol reagent (Life Technologies, Grand Island, NY), 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 of the knockout of SNF1 in YLKSKO candidate transformant colonies 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. The *Yarrowia* actin gene (YALI0D08272g) was used as a reference, and a segment of its cDNA was amplified with primers MTR001 and MTR002. Fragments of the SNF1 cDNA were amplified using two different primer pairs (Pair 1: rAS106 and rAS107, Pair 2: rAS434X and rAS435X) to ensure precision in transcript quantification; both pairs bind to the section of the cDNA that was meant to be replaced with the URA3 marker during homologous recombination. Raw fluorescence data was then input into the Real Time PCR Miner program [23] 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 the relative expression levels with confidence intervals for SNF1 in the candidate YLKSKO

transformants relative to the YLKuKO control strain. Three replicate wells were analyzed for each gene (primer pair) and strain combination.

5.2.3 Media formulations and shake flask fermentations

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 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), 0.79 g/L Complete Supplement Mixture (no dropout) (Sunrise Science Products, San Diego, CA), 50 g/L glucose, and 2.21 g/L ammonium sulfate (Sigma) for an atomic C:N ratio of 50. Yeast synthetic medium containing acetate (YSMa) used for fermentations contains 1.7 g/L Yeast Nitrogen Base (YNB, without amino acids or ammonium sulfate), 0.79 g/L Complete Supplement Mixture (no dropout), 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 27 hours (for YPD cultures) or 30 hours (for YPA cultures) 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 24 hours (for glucose cultures) or 28 hours (for acetate cultures) 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 250 mL shake flasks containing 50 mL of YSMg or YSMa to an

initial OD_{600} of 0.05 (for glucose cultures) or 0.025 (for acetate cultures). Flasks were closed with 34mm-45mm foam plugs (Jaece Industries Inc., North Tonawanda, NY) and placed in an Innova 43 Incubator Shaker (New Brunswick Scientific, Edison, NJ) and shaken for 120 hours (for glucose fermentations) or 145 hours (for acetate fermentations) at 250 rpm and 30°C. 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 flasks were used for each strain in both fermentations.

5.2.4 Lipid derivatization and quantification

Derivatization of fatty acids from lipids into fatty acid methyl esters (FAMEs) was achieved with a modified form of the direct transesterification method [24]. Samples of 200 µL each were taken at the fermentation end points (120 hours for glucose, 145 hours for acetate) and 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.

5.3 Results and Discussion

5.3.1 Verification of SNF1 knockout in YLKSKO transformant colonies

Laboratory strains of model yeasts such as *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* tend to be highly genetically tractable, and gene targeting in these organisms can be achieved using DNA cassettes with short homology regions (~50 bp) at a high per colony efficiency [16] because of the strong relative activity of their HR complexes when repairing DSBs. In many other lower eukaryotes, gene targeting at acceptable rates is likely only with long (1 kb or more) homology regions; in the strains of *Yarrowia lipolytica* we have used, the NHEJ complex is so active, that even with very long (1.6 kb or more) homology regions, we were unable to confirm any successfully locus-targeted products after screening several hundred transformants. **Figure 5.1A** shows the results of PCR confirmations in which the entire SNF1 coding sequence (1740 bp) could still be detected after transformation of strain Po1z with the

pSNF1KO cassette. In order to make knockouts easier to acquire in *Yarrowia*, we attempted to disrupt the KU70 gene (NCBI Locus Tag: YALI0C08701g), which codes for a key component of the NHEJ complex, and sought an alternative method to the traditional cassette transformation. Our collaborators at PNNL were able to successfully knock out KU70 using *Agrobacterium tumefaciens*-mediated transformation, a method typically used for transformation of plant cells [25], and sent us the resulting ΔKU70 strain (referred to as YLKuKO). After the relatively trivial disruption of the URA3 gene in this strain (to generate YLKuKO-U), we attempted transformation of this new background with the pSNF1KO cassette as usual. We inoculated 12 colonies in YSMg (with Ura dropout) for gDNA screening; only five grew sufficiently for replating and gDNA extraction. The results of the PCR confirmation assay for the SNF1 knockout in the YLKuKO-U background are shown in **Figure 5.1B**.

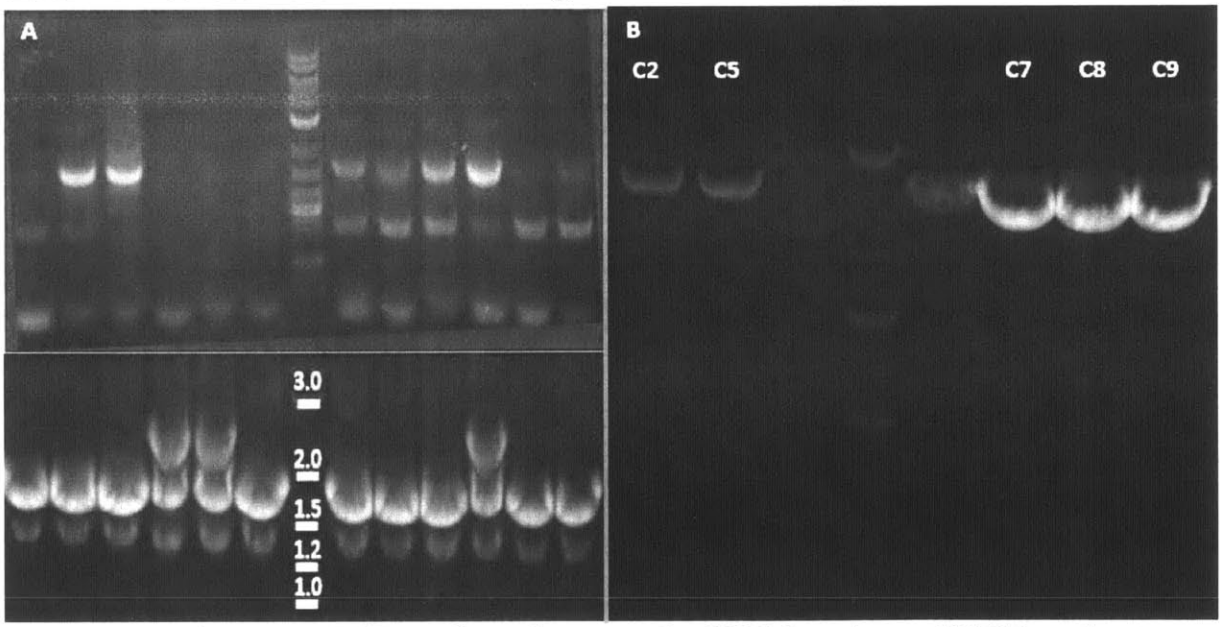

Figure 5.1: PCR confirmation assay for SNF1 knockout. SNF1 knockout transformant colonies were confirmed by PCR of their gDNA with primer rAS394 and rAS395. A band of 2312 bp represents the knockout cassette consisting of 5' and 3' end regions of the SNF1 coding sequence flanking the URA3 marker gene. A band of 1740 bp represents the unmodified SNF1 coding sequence. (A) Results of 24 representative colonies (out of hundreds) from transformations of strain Po1z. (B) Results of five colonies from transformation of strain YLKuKO-U. C2, C5, etc. represents the colony number (out of 12 original transformants).

While none of the po1z background colonies shown in **Figure 5.1A** were shown to be knockouts due to the presence of the 1740 bp SNF1 coding sequence band or absence of the 2312 bp knockout cassette, all of the YLKuKO-U colonies in **Figure 5.1B** produced the knockout cassette band without a detectable intact SNF1 coding sequence band. As the knockout cassette band is larger, it should be harder to amplify, so the absence of the smaller coding sequence band from our assay suggests (but does not prove) its absence from the genome of the colonies altogether. Although ostensibly rare, the knockout cassette could have integrated by NHEJ multiple times into the same cell, in which case, the primers used in our PCR confirmation assay could preferentially bind to and amplify the knockout cassette over the SNF1 coding sequence. We therefore needed a second method to ensure that the results from our first assay could be trusted. We extracted RNA from cultures of these colonies and measured the expression level of SNF1 in each using Quantitative Reverse Transcription PCR (qRT-PCR) with two sets of primers, and compared these results to the YLKuKO background strain. The results of transcriptional analysis are shown in **Figure 5.2**.

For a true knockout, we should expect that expression of the gene of interest should not just be significantly smaller than the control, but be undetectable, only allowing for the normal small amount of off-target binding and amplification beginning at the later cycles of the qPCR. **Figure 5.2** shows that despite all five colonies "passing" the first confirmation assay, only three (C5, C7, and C8) show nearly undetectable increases in fluorescence (the result for Pair 1 for C8 was exactly zero, as no off-target amplification was detected). The results from both primer pairs are also in strong agreement with each other, showing the high precision of our method and thusly demonstrating the abolishment of expression of SNF1 in those colonies. It has previously been reported [14] that the knockout of Ku70 in *Yarrowia* and other organisms significantly increases

the frequency of HR-derived product in future transformations. Although our baseline levels of HR frequency with the Po1z strain (and others) are very divergent from those found in much of the literature, with the ΔKU70 background, we have demonstrated that we can now achieve gene targeting in *Yarrowia* at acceptable rates on the order of those reported for other ΔKU70 strains.

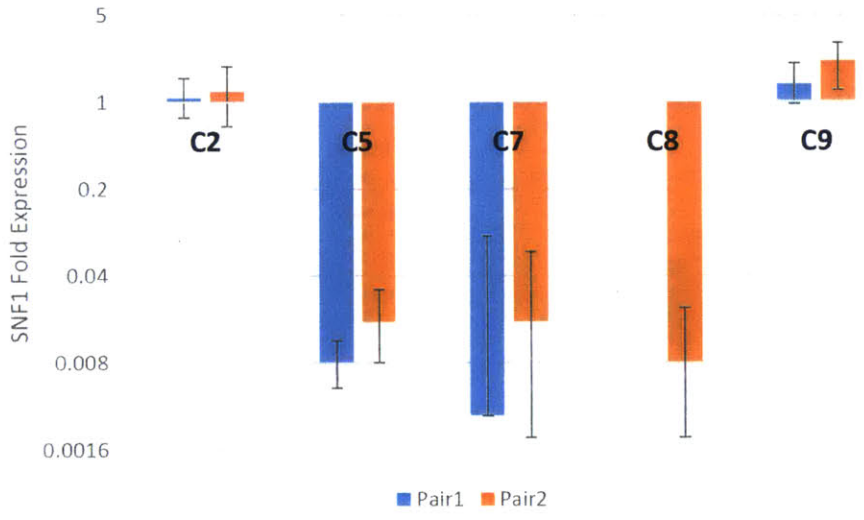

Figure 5.2: SNF1 expression of potential YLKSKOcolonies relative to control strain YLKuKO as determined by qRT-PCR analysis. Primer Pair 1: rAS106/rAS107, Primer Pair 2: rAS434X, rAS435X. Actin was used as a reference gene. Three wells per primer pair and strain combination were used.

5.3.2 Assessment of lipid production phenotype in ΔSNF1 strain on glucose and acetate and comparison to previously reported results

After qRT-PCR analysis on transcripts of SNF1 in our candidate YLKSKO colonies, we selected colony C8 for further transformations, as Primer Pair 1 showed no detectable fluorescence increase over 50 cycles of amplification, and Primer Pair 2 showed the smallest off-target effects (giving a readout of SNF1 expression of at most 0.8% of the control strain). We transformed this strain with pMT038 to approximate a "wild-type" background (meaning effectively no additional modifications besides ΔKU70 and ΔSNF1) and named this new strain

YL38S. We constructed control strain YL38K by complementing YLKuKO-U with the URA3 marker and also transforming it with pMT038, and compared these two strains in shake flask fermentations of glucose and acetate. As a ΔSNF1 mutant *Yarrowia* strain has already been reported on in literature [18] with respect to lipid production phenotype on glucose, we wanted to further validate our knockout by comparing our glucose fermentation results to the literature. The results of our glucose fermentation are shown in **Figure 5.3**.

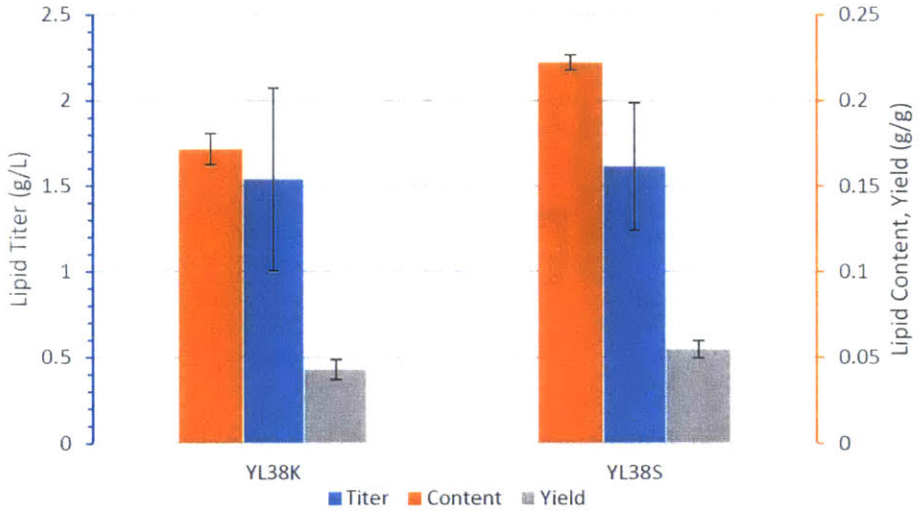

Figure 5.3: Fermentation performance of control and ΔSNF1 strains on glucose. Fermentations were performed in supplemented (with CSM) glucose media at an atomic C:N of 50 for 120 hours. Lipid titer, content, and yield were measured; error bars represent the standard deviations from three biological replicates.

Figure 5.3 shows that the lipid titer, content, and yield of our ΔSNF1 mutant was 1.61 g/l, 0.222 g/g, and 0.055 g/g, respectively, while our control strain produced a lipid titer, content, and yield of 1.54 g/L, 0.171 g/g, and 0.043 g/g, respectively. The 29.5% increase in lipid content and 27.7% increase in lipid yield for our mutant were statistically significant (with p-values of 5.22×10^{-4} and 0.0284, respectively), however our increase in lipid titer was not significant. These results are in line with the previously published data [18] comparing wild-type *Yarrowia* to ΔSNF1 mutants in high C:N "oleaginous" glucose media, in which the lipid titer of the mutant

was slightly but not significantly higher, and the lipid content of the mutant showed a significant increase of 46.8%, which is on the order of the increase shown in our results. Therefore, our results from the glucose fermentation confirm that we have validated the phenotype of our mutant in addition to the genotype. We went on to determine if our now overwhelmingly validated ΔSNF1 mutant also improved lipid production on acetate as well.

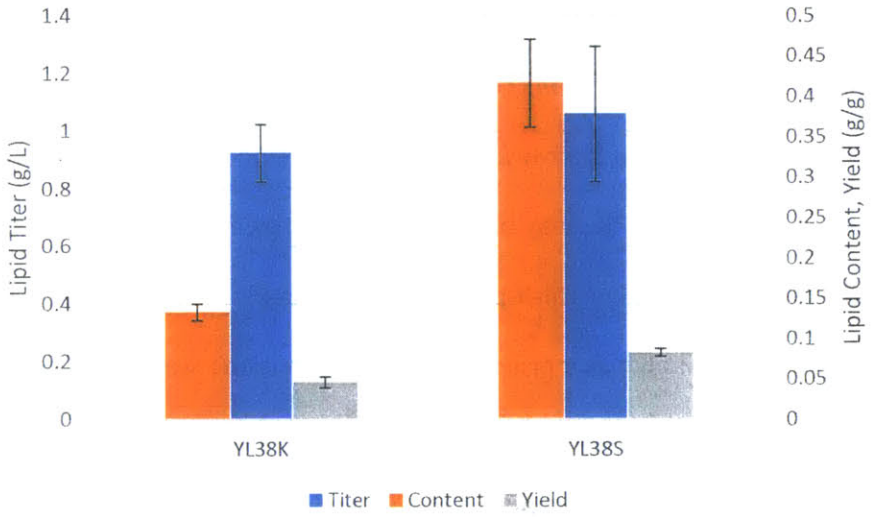

Figure 5.4: Fermentation performance of control and ΔSNF1 strains on acetate. Fermentations were performed in supplemented (with CSM) acetate media at an atomic C:N of 40 for 145 hours. Lipid titer, content, and yield were measured; error bars represent the standard deviations from three biological replicates.

The first step in acetate metabolism is the formation of acetyl-CoA, the substrate of ACC1, of which SNF1 is a negative regulator. We hypothesized that the increased activity of ACC1 combined with the high-flux, short-length metabolic path (compared to glucose) between carbon source and enzyme substrate would mean one of two things was likely to happen on acetate as a result of knocking out SNF1: either the increases in yields in the ΔSNF1 strain will be significantly higher than those on glucose, or the amount of ATP used by the cell in activating the acetate and being unable to regulate further energy consumption by ACC1 will result in a

severe energy deficit leading to growth defects and an inability to complete biosynthesis of fatty acids. **Figure 5.4** shows that the former hypothesis was fortunately correct, as the lipid titer, content, and yield of our YL38S strain on acetate was 1.06 g/L, 0.417 g/g, and 0.083 g/g, respectively, compared to the YL38K control strain's titer, content, and yield of 0.925 g/L, 0.133 g/g, and 0.046 g/g, respectively. Like the case for glucose, the increase in lipid titer from our YL38S strain was not statistically significant, but the 214% increase in lipid content and the 80% increase is yield are, with p-values of 4.54×10^{-4} and 7.98×10^{-4}, respectively.

Aside from inactivating Acc1 in *Yarrowia*, there has been evidence to suggest that the mammalian homolog to Snf1, AMPK, may phosphorylate and thereby inactivate enzymes of glycerolipid synthesis such as sn-glycerol-3-phosphate acyltransferase (GPAT) in rat hepatocytes [26]. In addition to its involvement as a protein-level regulator, Snf1 was also shown previously to influence transcription levels of many *Yarrowia* genes [18], including those involved in fatty acid synthesis; deletion of SNF1 led to small but significant (~50%) increases in FAS1, FAS2, and GPD expression, and seemed to reduce the activity of the fatty acid β-oxidation pathway. Inactivation of Acc1 by Snf1 in *Saccharomyces cerevisiae* has been shown to decrease expression of the INO1 gene, coding for inositol-1-phosphate synthase, which is in yeast the rate-limiting step in the biosynthesis of inositol-containing phospholipids in the absence of extracellular inositol in the medium [27]; as phospholipids are required for cell growth and division, this may explain the significant decreases in growth rate in the YL38S strain as measured by total biomass accumulated by the ends of the fermentations. In sum, our results, combined with other information from literature, show that Snf1, as a global regulator of cellular energy usage, is involved in the regulation of fatty acid and glycerolipid synthesis at many different points along these pathways; deletion of this gene therefore leads to a significantly

more lipogenic phenotype in *Yarrowia*, with the changes in phenotype more dramatic on acetate than on glucose due to the larger acetyl-CoA accessibility in cells growing on the former substrate.

5.3.3 Effect of native gene overexpressions on fermentation performance of the ΔSNF1 mutant

Having validated the more lipogenic phenotype of our ΔSNF1 mutant on two different substrates, our next goal was to determine whether we could observe any synergistic effects between the ΔSNF1 modification and one or both overexpressions of native genes ACC1 and DGA2 when combined in the same strain. Using plasmids pASAe, pMT053, and pMT065, we constructed strains overexpressing ACC1, DGA2, and both ACC1 and DGA2, respectively, in the control ΔKU70 background (yielding strains YLAeK, YL53K, and YL65K) and in the ΔKU70 ΔSNF1 background (yielding strains YLAeS, YL53S, and YL65S). Like the strains without overexpression modifications, we tested fermentation performance of these six strains in shake flasks on glucose and acetate. The results of the glucose fermentation are shown in **Figure 5.5** and **Table 5.3**.

Table 5.3: Lipid titer, content, and yield in native gene overexpressing strains with and without ΔSNF1 modification in glucose fermentation.

Strain	Lipid Titer (± SD)	Lipid Content (± SD)	Lipid Yield (± SD)
YLAeK	1.20 ± 0.11	0.174 ± 0.010	0.044 ± 0.003
YL53K	2.82 ± 0.76	0.397 ± 0.031	0.108 ± 0.004
YL65K	6.38 ± 1.50	0.599 ± 0.045	0.158 ± 0.010
YLAeS	1.61 ± 0.13	0.225 ± 0.045	0.059 ± 0.004
YL53S	2.62 ± 0.17	0.402 ± 0.062	0.093 ± 0.007
YL65S	0.98 ± 0.12	0.171 ± 0.006	0.052 ± 0.001

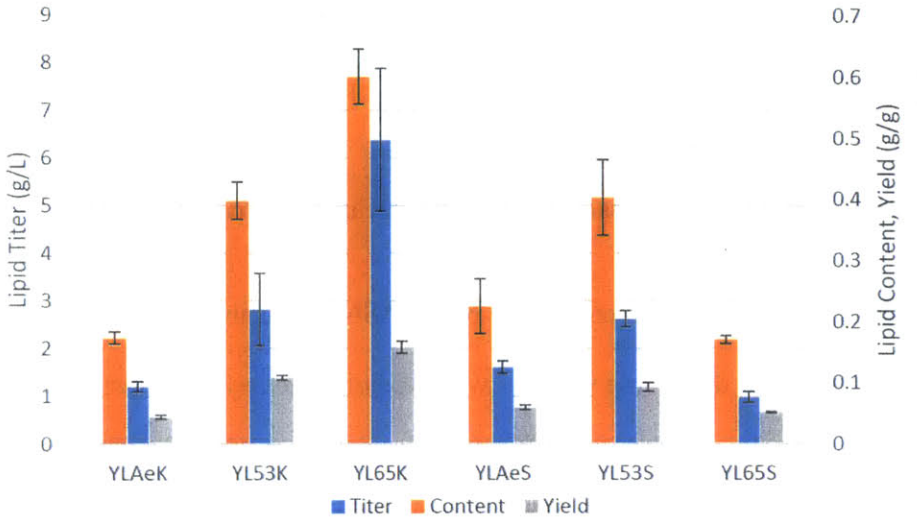

Figure 5.5: Fermentation performance of strains YLAeK, YL53K, YL65K, YLAeS, YL53S, and YL65S on glucose. Fermentations were performed in supplemented (with CSM) glucose media at an atomic C:N of 50 for 120 hours. Lipid titer, content, and yield were measured; error bars represent the standard deviations from three biological replicates.

It has been shown previously that overexpression of ACC1 lead to a doubling of lipid content with a minimal effect on lipid titer on glucose [6], however, this was for low (~20) C:N atomic ratios in the medium used; for higher atomic ratios, the performance of wild-type and ACC1 overexpressing strains are usually similar in our experience (data not shown). It is therefore unsurprising that strains YLAeK and YLAeS do not behave on glucose in a manner that is statistically significantly different from their non-ACC1 overexpressing counterparts in terms of any of our three metrics (lipid titer, content, and yield). In contrast to the improved performance of our ΔSNF1 mutant in the case of ACC1 overexpression or no overexpression, we observed no difference between strains YL53K and YL53S, both overexpressing only DGA2. Dga2 catalyzes the final step in producing TAGs, the "storage molecule" for fatty acids in *Yarrowia*; it is believed that overproduction of this enzyme has a large effect on neutral lipid production at least in part due to the quick removal of fatty acyl-CoA molecules produced from the fatty acid

synthase (FAS) complex, which are known feedback inhibitors of Acc1 [28, 29]. It is possible that on glucose, the acetyl-CoA pool in the cytosol, produced by ATP:citrate lyase (see Chapter 2), is already entirely used for lipids due to the increased activity of Acc1 caused by the lessening of this feedback inhibition. In the case of our ACC1 and DGA2 overexpression strains, we observed the previously reported [6] highly synergistic behavior of these two overexpressions in driving lipid synthesis on glucose; however, the deletion of SNF1 to this configuration had a largely adverse effect, with strain YL65S producing a lipid titer, content, and yield that was only 15%, 28%, and 33% of that produced by the control strain YL65K.

Table 5.4: Lipid titer, content, and yield in native gene overexpressing strains with and without ΔSNF1 modification in acetate fermentation.

Strain	Lipid Titer (± SD)	Lipid Content (± SD)	Lipid Yield (± SD)
YLAeK	0.72 ± 0.15	0.130 ± 0.016	0.045 ± 0.005
YL53K	1.66 ± 0.12	0.405 ± 0.020	0.134 ± 0.006
YL65K	2.22 ± 0.10	0.469 ± 0.012	0.152 ± 0.003
YLAeS	1.04 ± 0.12	0.376 ± 0.018	0.089 ± 0.003
YL53S	1.56 ± 0.38	0.767 ± 0.043	0.200 ± 0.020
YL65S	1.04 ± 0.16	0.445 ± 0.014	0.079 ± 0.040

Figure 5.6 and **Table 5.4** show the results of our ΔSNF1 and control strains in acetate fermentations. ACC1 overexpression in our YLAeK strain resulted in a similar phenotype to the YL38K strain, only growing slightly slower, resulting in a slight decrease in lipid titer produced. The YLAeS strain even more closely resembled its non-ACC1 overexpressing counterpart. The combination of DGA2 overexpression and SNF1 knockout in the YL53S strain led to the large synergistic effect of a highly lipogenic phenotype, with a lipid content of 76.7% and an overall lipid yield of 20.0%. Compared to transcriptional upregulation of ACC1 (as in strain YL65K), the elimination of post-translational inhibition of Acc1 in a DGA2-overexpressing background caused increases in the lipid content and yield by 63.3% and 31.6%, respectively, both of which were highly statistically significant (p-values of 1.57×10^{-4} and 7.19×10^{-3}, respectively). The

availability of cytosolic acetyl-CoA as a fraction of the total carbon entering the cell is extremely high on acetate, so as before with the strains that do not overexpress any native genes, we expect that increases in lipid content and yield would be more dramatic on acetate than on glucose. The reduction in growth rate [18] that was common to all ΔSNF1 strains tested and is also observed in Saccharomyces cerevisiae [30] could be responsible for the smaller lipid titer in YL53S compared to YL65K, although the difference in lipid titer between YL53S and YL53K was not significant. The improved performance of YL53S over YL65K suggests that Snf1 is relatively active on acetate; because activating acetate is energetically expensive (Acs is an AMP-forming enzyme in *Yarrowia*, effectively costing 2 ATP molecules per acetate molecule), the energy charge in the cell may be low enough for Snf1 to inactivate Acc1 [9, 28], however, our results demonstrate that the apparent threshold for how large an energy deficit needs to be to activate Snf1 is significantly smaller than what would cause a major defect in overall cellular health. A survey of literature turns up few results for lipid accumulation of oleaginous yeasts on acetate [19, 31-33], but none of the reports we collected have claimed a lipid content as high as what we achieved in our YL53S strain. Like the case for glucose, the ACC1 and DGA2 overexpressing ΔSNF1 strain YL65S accumulated lipids relatively poorly, although it had a similar lipid content and yield to the other ΔSNF1 mutants aside from strain YL53S.

In strain YL65S, we have an example of how flux through a pathway can be overengineered, leading to less than optimal (in our case, much less than optimal) production of the desired chemical. For the case of this strain's poor lipid production performance, we can offer two possible explanations, which are not mutually exclusive: 1) overengineering the fatty acid synthesis flux leads to a significant energy deficit in the cell, preventing the completion of fatty acid synthesis, or 2) competition of the fatty acid synthesis pathway with other cytosolic acetyl-

CoA utilizing pathways leads to an inability of the cells to store the neutral lipids produced by DGA2.

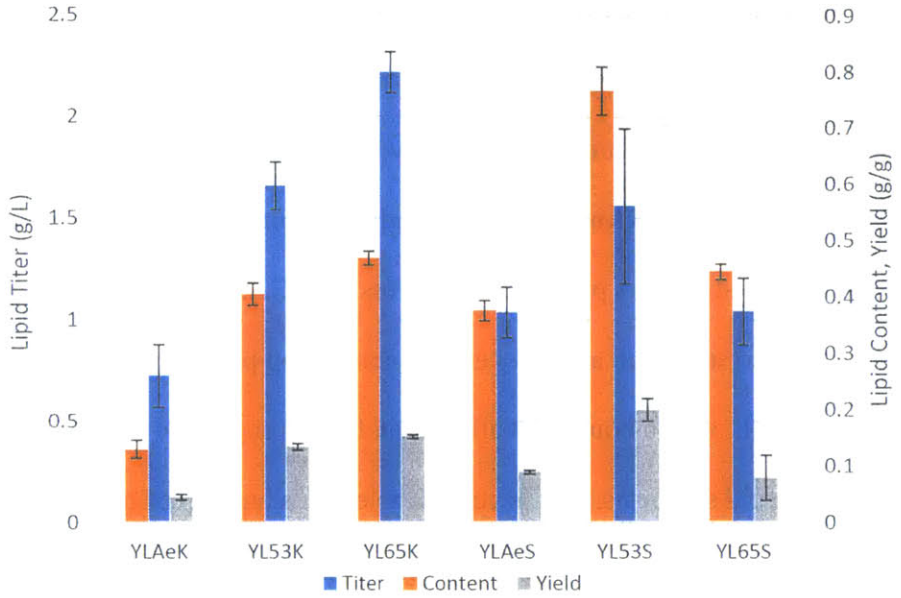

Figure 5.6: Fermentation performance of strains YLAeK, YL53K, YL65K, YLAeS, YL53S, and YL65S on acetate. Fermentations were performed in supplemented (with CSM) acetate media at an atomic C:N of 40 for 145 hours. Lipid titer, content, and yield were measured; error bars represent the standard deviations from three biological replicates.

Fatty acid synthesis is a highly energy intensive process. There is evidence to suggest ACC1 is transcriptionally regulated in yeast positively by the Ino2/Ino4 complex, which is most likely the Yas1/Yas2 complex in *Yarrowia* (see **Chapter 3**), and negatively by the Opi1 regulator, in response to the absence or presence, respectively, of inositol or choline [34], despite not having a consensus inositol-choline responsive element (ICRE) site in its promoter. However, there are other stretches in the yeast ACC1 promoter containing hexamer or pentamer repeats of unknown function [28], and ACC1 has been shown to be upregulated transcriptionally in G1 growth phase [28] in yeast, and in response to nitrogen depletion in the oleaginous yeast *R. toruloides* [35] and the oleaginous microalgae *C. pyrenoidosa* [36]. Previously, we discussed that Snf1 phosphorylates and inactivates Acc1 in response to cellular energy charge [9], and that acyl-CoA

products of the fatty acid synthesis pathway inhibit Acc1 activity [28, 29]. Therefore, as the first committed step of fatty acid biosynthesis, acetyl-CoA carboxylase activity is regulated in many different ways: at the levels of transcription, post-translational modification, and allosteric (feedback) inhibition. In strain YL65S, we prevent transcriptional regulation by overexpressing ACC1 using the constitutive hp4d promoter, we remove a major (if not the major) source of post-translational regulation in knocking out SNF1, and we significantly alleviate any feedback regulation in shrinking the acyl-CoA pool by overexpressing DGA2. The cell therefore is left with few, if any, options to stop such a largely energy consuming process, and this may lead to an energy deficit in the cell, which could globally lessen flux through many pathways needed by the cell, including fatty acid synthesis, as it is conceivable that a buildup of intermediate-length acyl-ACPs in the FAS complex may occur without the required cofactors for their completion.

As for our other explanation, yeast cells must provide a significant amount of cytosolic acetyl-CoA during the growth phase for the biosynthesis many different compounds required for proliferation, including intracellular and extracellular membrane constituents such as phospholipids and sterols [37]. These two types of molecules are synthesized by different pathways: phospholipids are synthesized using fatty acids, whereas sterols are derived from isoprenoids made in the mevalonate pathway [38]. In the event of Acc1 excessively outcompeting the acetoacetyl-CoA thiolase enzyme for acetyl-CoA, sterol biosynthesis may decrease below an adequate level. This combined with the finding that phospholipid biosynthesis also decreases with the removal of Snf1 due to the aforementioned downregulation of INO1 [27], could prevent or decrease the biosynthesis of membrane lipids, which are needed to form, among other things, the lipid bodies that store TAGs. If lipid body membrane synthesis lags behind the rate required based on the rate of TAG biosynthesis, it could prevent the storage

of TAGs which could make them more accessible to lipases or otherwise contribute to a significant feedback inhibition, significantly reducing the net rate of fatty acid biosynthesis.

5.4 Conclusions

While the overexpression of genes related to a pathway of interest can be an effective way to increase flux through that pathway, regulation at levels below transcription, such as post-translational modification of enzymes or inhibition by small molecule effectors can help the cell to resist or limit any changes made to the flux distribution by this method. Removal of these lower-level mechanisms of regulation can therefore potentially cause greater changes with fewer modifications, but in many cases, these lower levels of regulation are relatively poorly understood. Compounding this problem is that once a gene coding for an enzyme inhibitor is properly identified for removal, knocking out said gene through the targeting and homologous recombination of a transformed DNA cassette into the correct locus is challenging in many organisms, including *Yarrowia lipolyitca.*

Snf1 is a protein kinase known to inactivate Acc1, the first committed step in fatty acid synthesis. We successfully knocked out this inhibitor in *Yarrowia* using a strain specifically developed by our collaborators at PNNL to more frequently repair DSBs through homologous recombination through the removal of the KU70 gene, which codes for a part of the non-homologous end joining complex, a normally highly active DSB repair pathway that competes with HR and produces off-target cassette integrations. Our knockout strain achieved a lipid content 29.5% higher and a lipid yield 27.7% higher than our control strain on glucose, and a 214% increase in content and an 80% increase in yield over our control on acetate. While our ΔSNF1 strains were characterized by a significantly more lipogenic phenotype, they also showed a reduced growth rate. Adding native gene overexpressions of ACC1 and/or DGA2 did not

improve performance relative to the SNF1-expressing strains, except for the case of DGA2 overexpression on acetate, which produced a lipid content of 76.7% and an overall lipid yield of 20.0% in shake flasks. To our knowledge, this is the highest reported lipid content achieved on acetate in any oleaginous yeast, and our lipid yield is very similar to that of our previously constructed ACC1, DGA2, and SLC1 overexpression strain (see **Chapter 4**). Interestingly, merely adding the overexpression of ACC1 to this highly lipogenic strain yielded a strain with a relatively poor lipid producing phenotype. As Acc1 is regulated at the level of transcription, post-translational modification, and feedback inhibition, reducing or eliminating these methods of control over Acc1 activity by overexpressing ACC1, knocking out SNF1, and overexpressing DGA2, respectively, likely caused a significant energy deficit or other inability to efficiently produce other compounds required for normal cellular function, demonstrating that over-engineering a pathway of interest can easily disrupt the relative balance between pathways that is necessary to maintain an optimal flux distribution for the synthesis of a desired compound.

References:

1) 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. *Progress in Lipid Research*, *48*(6), 375–87. doi:10.1016/j.plipres.2009.08.005
2) Meng, X., Yang, J., Xu, X., Zhang, L., Nie, Q., & Xian, M. (2009). Biodiesel production from oleaginous microorganisms. *Renewable Energy*, *34*(1), 1–5. doi:10.1016/j.renene.2008.04.014
3) Dujon, B., Sherman, D., Fischer, G., Durrens, P., Casaregola, S., Lafontaine, I., … Souciet, J.-L. (2004). Genome evolution in yeasts. *Nature*, *430*(6995), 35–44. doi:10.1038/nature02579
4) Chen, D. C., Beckerich, J. M., & Gaillardin, C. (1997). One-step transformation of the dimorphic yeast Yarrowia lipolytica. *Applied Microbiology and Biotechnology*, *48*, 232–235. doi:10.1007/s002530051043
5) Madzak, C., Tréton, B., & Blanchin-Roland, S. (2000). Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous proteins in the yeast Yarrowia lipolytica. *Journal of Molecular Microbiology and Biotechnology*, *2*(2), 207–216.
6) Tai, M., & Stephanopoulos, G. (2013). Engineering the push and pull of lipid biosynthesis in oleaginous yeast Yarrowia lipolytica for biofuel production. *Metabolic Engineering*, *15*, 1–9. doi:10.1016/j.ymben.2012.08.007
7) Blazeck, J., Liu, L., Redden, H., & Alper, H. (2011). Tuning gene expression in Yarrowia lipolytica by a hybrid promoter approach. *Applied and Environmental Microbiology*, *77*(22), 7905–14. doi:10.1128/AEM.05763-11
8) Xue, Z., Sharpe, P. L., Hong, S.-P., Yadav, N. S., Xie, D., Short, D. R., … Zhu, Q. (2013). Production of omega-3 eicosapentaenoic acid by metabolic engineering of Yarrowia lipolytica. *Nature Biotechnology*, *31*(8), 734–40. doi:10.1038/nbt.2622
9) Hedbacker, K., & Carlson, M. (2008). SNF1/AMPK pathways in yeast. *Frontiers in Bioscience : A Journal and Virtual Library*, *13*, 2408–2420. doi:10.2741/2854
10) Choi, J. W., & Da Silva, N. a. (2014). Improving polyketide and fatty acid synthesis by engineering of the yeast acetyl-CoA carboxylase. *Journal of Biotechnology*, *187*, 56–59. doi:10.1016/j.jbiotec.2014.07.430
11) Flores, C., & Gancedo, C. (2005). Yarrowia lipolytica Mutants Devoid of Pyruvate Carboxylase Activity Show an Unusual Growth Phenotype. *Eukaryotic Cell*, *4*(2), 356–364. doi:10.1128/EC.4.2.356
12) San Filippo, J., Sung, P., & Klein, H. (2008). Mechanism of eukaryotic homologous recombination. *Annual Review of Biochemistry*, *77*, 229–257. doi:10.1146/annurev.biochem.77.061306.125255
13) Daley, J. M., Palmbos, P. L., Wu, D., & Wilson, T. E. (2005). Nonhomologous end joining in yeast. *Annual Review of Genetics*, *39*, 431–451. doi:10.1146/annurev.genet.39.073003.113340
14) Verbeke, J., Beopoulos, A., & Nicaud, J.-M. (2013). Efficient homologous recombination with short length flanking fragments in Ku70 deficient Yarrowia lipolytica strains. *Biotechnology Letters*, *35*(4), 571–6. doi:10.1007/s10529-012-1107-0

15) De Groot, M. J., Bundock, P., Hooykaas, P. J., & Beijersbergen, a G. (1998). Agrobacterium tumefaciens-mediated transformation of filamentous fungi. *Nature Biotechnology*, *16*(9), 839–842. doi:10.1038/nbt0998-839
16) Michielse, C. B., Arentshorst, M., Ram, a. F. J., & Van Den Hondel, C. a M. J. J. (2005). Agrobacterium-mediated transformation leads to improved gene replacement efficiency in Aspergillus awamori. *Fungal Genetics and Biology*, *42*(1), 9–19. doi:10.1016/j.fgb.2004.06.009
17) Bundock, P., Mroczek, K., Winkler, A.A., Steensma, H.Y., Hooykaas, P.J., 1999. T-DNA from Agrobacterium tumefaciens as an efficient tool for gene targeting in Kluyveromyces lactis. Mol. Gen. Genet. 261, 115–121.
18) Seip, J., Jackson, R., He, H., Zhu, Q., & Hong, S. P. (2013). Snf1 is a regulator of lipid accumulation in Yarrowia lipolytica. *Applied and Environmental Microbiology*, *79*(23), 7360–7370. doi:10.1128/AEM.02079-13
19) Tai, M. (2012). Metabolic Engineering of Oleaginous Yeast for the Production of Biofuels. Ph.D. Thesis. Ch. 7. Massachusetts Institute of Technology.
20) Sambrook, J., & Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. CSHL press, Cold Spring Harbor, NY, USA.
21) 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. *Nature Methods*, *6*(5), 343–345. doi:10.1038/nmeth.1318
22) Chen, D. C., Beckerich, J. M., & Gaillardin, C. (1997). One-step transformation of the dimorphic yeast Yarrowia lipolytica. *Applied Microbiology and Biotechnology*, *48*, 232–235. doi:10.1007/s002530051043
23) Sheng Zhao, Russell D. Fernald. Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J. Comput. Biol.* 2005 Oct;12(8):1045-62.
24) 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.
25) Klee, H., Horsch, R., & Rogers, S. (1987). Agrobacterium-Mediated Plant Transformation and its Further Applications to Plant Biology. *Annual Review of Plant Physiology*. doi:10.1146/annurev.pp.38.060187.002343
26) Muoio, D. M., Seefeld, K., Witters, L. a, & Coleman, R. a. (1999). AMP-activated kinase reciprocally regulates triacylglycerol synthesis and fatty acid oxidation in liver and muscle: evidence that sn-glycerol-3-phosphate acyltransferase is a novel target. *The Biochemical Journal*, *338 (Pt 3)*, 783–791. doi:10.1042/0264-6021:3380783
27) Shirra, M. K., Patton-Vogt, J., Ulrich, a, Liuta-Tehlivets, O., Kohlwein, S. D., Henry, S. a, & Arndt, K. M. (2001). Inhibition of acetyl coenzyme A carboxylase activity restores expression of the INO1 gene in a snf1 mutant strain of Saccharomyces cerevisiae. *Molecular and Cellular Biology*, *21*(17), 5710–5722. doi:10.1128/MCB.21.17.5710-5722.2001
28) Tehlivets, O., Scheuringer, K., & Kohlwein, S. D. (2007). Fatty acid synthesis and elongation in yeast. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1771*(3), 255–270. doi:10.1016/j.bbalip.2006.07.004
29) Nielsen, J. (2009). Systems biology of lipid metabolism: From yeast to human. *FEBS Letters*, *583*(24), 3905–3913. doi:10.1016/j.febslet.2009.10.054

30) Natter, K., & Kohlwein, S. D. (2013). Yeast and cancer cells - Common principles in lipid metabolism. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1831*(2), 314–326. doi:10.1016/j.bbalip.2012.09.003
31) Fei, Q., Chang, H. N., Shang, L., Choi, J. D. R., Kim, N., & Kang, J. (2011). The effect of volatile fatty acids as a sole carbon source on lipid accumulation by Cryptococcus albidus for biodiesel production. *Bioresource Technology*, *102*(3), 2695–2701. doi:10.1016/j.biortech.2010.10.141
32) Christophe, G., Deo, J. L., Kumar, V., Nouaille, R., Fontanille, P., & Larroche, C. (2012). Production of oils from acetic acid by the oleaginous yeast Cryptococcus curvatus. *Applied Biochemistry and Biotechnology*, *167*(5), 1270–1279. doi:10.1007/s12010-011-9507-5
33) 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 Technology*, *114*, 443–9. doi:10.1016/j.biortech.2012.02.091
34) Hasslacher, M., Ivessa, a. S., Paltauf, F., & Kohlwein, S. D. (1993). Acetyl-CoA carboxylase from yeast is an essential enzyme and is regulated by factors that control phospholipid metabolism. *Journal of Biological Chemistry*, *268*(15), 10946–10952. doi:10.1074/jbc.M802685200
35) Zhu, Z., Zhang, S., Liu, H., Shen, H., Lin, X., Yang, F., ... Zhao, Z. K. (2012). A multi-omic map of the lipid-producing yeast Rhodosporidium toruloides. *Nature Communications*, *3*, 1112. doi:10.1038/ncomms2112
36) Fan, J., Cui, Y., Wan, M., Wang, W., & Li, Y. (2014). Lipid accumulation and biosynthesis genes response of the oleaginous Chlorella pyrenoidosa under three nutrition stressors. *Biotechnology for Biofuels*, *7*(1), 17. doi:10.1186/1754-6834-7-17
37) Zinser, E., Paltauf, F., & Daum, G. (1993). Sterol composition of yeast organelle membranes and subcellular distribution of enzymes involved in sterol metabolism. *Journal of Bacteriology*, *175*(10), 2853–2858. doi:0021-9193/93/102853-06
38) Nielsen, J. (2014). Synthetic Biology for Engineering Acetyl Coenzyme A Metabolism in Yeast. *mBio*, *5*(6), 14–16. doi:10.1128/mBio.02153-14.Copyright

Chapter 6

Determination of Lipases Potentially Responsible for Late-Stage Lipostatic Behavior in *Yarrowia lipolytica* and Characterization of Knockouts

6.1 Introduction

The likelihood of single cell oil (SCO)-derived biodiesel to supplant traditional fossil fuels as a widely used energy source depends on establishing and maintaining high productivity in the processes that make it. A large factor in determining reactor productivity is the choice of host strain (meaning the species of organism and any genetic modifications) used to catalyze the reactions from substrate to product. *Yarrowia lipolytica* is a fast-growing oleaginous yeast, with wild-type strains able to accumulate lipids to account for over 36% of its dry cell weight in some settings [1]. Over 90% of these lipids are stored in the form of triacylglycerols (TAGs); biodiesel can be made from TAGs through transesterification with short chain alcohols to form fatty acid alkyl esters [2]. *Yarrowia lipolytica* is capable of consuming many different carbon sources, with acetate being one particularly cost-efficient example as a major component of many agro-industrial or municipal waste streams [3].

The biochemical process of TAG synthesis in *Yarrowia* is illustrated in **Figure 6.1**. The first committed step towards product formation in cells is the carboxylation of acetyl-CoA to form malonyl-CoA by the Acetyl-CoA Carboxylase 1 (ACC1) enzyme [4]. In the process of fatty acid elongation, the malonyl-CoA first attaches to the acyl carrier protein (ACP) of the fatty acid synthase (FAS) enzyme complex, while an acetyl-CoA molecule attaches to the β-keto-acyl-ACP synthase domain of the complex. Malonyl-CoA is then simultaneously decarboxylated and condensed with acetyl-CoA to form an acyl chain with a β-keto group. The β-keto group of this acyl chain in then reduced twice with electrons from NADPH for form a saturated acyl chain [5]. Seven rounds of elongation of the acetyl or acyl carbon chain by two carbon addition from malonyl-CoA followed by reductions at the β position yield a sixteen carbon acyl chain, which is cleaved from the FAS to form palmitoyl-CoA.

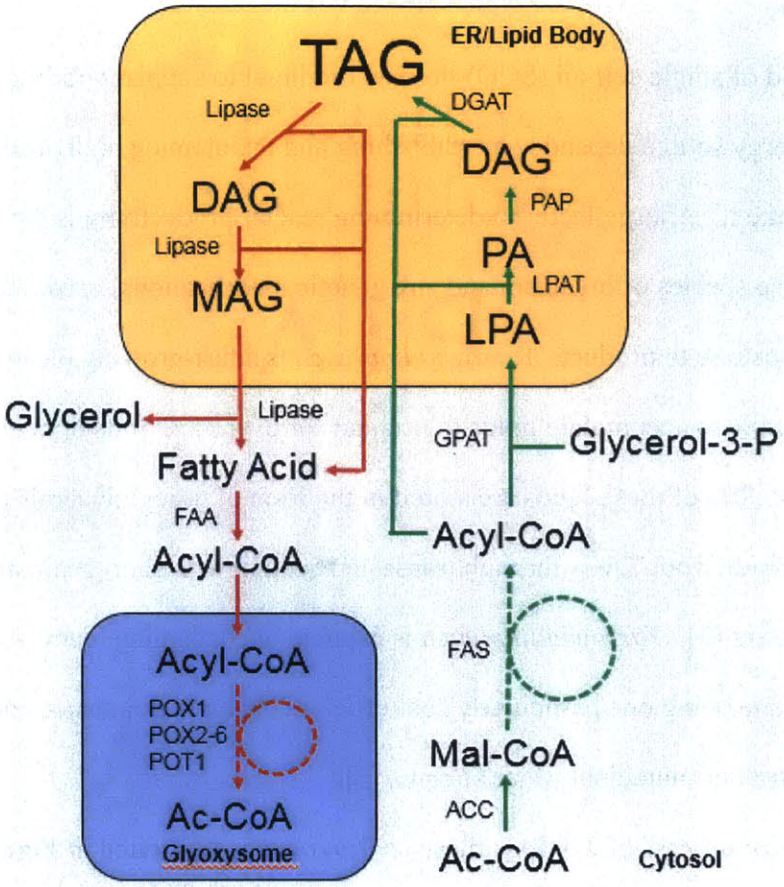

Figure 6.1: TAG synthesis and degradation pathways in *Yarrowia lipolytica*. Abbreviations: Ac-CoA: acetyl-CoA, Mal-CoA: malonyl-CoA, Glycerol-3-P: glycerol-3-phosphate, LPA: lysophosphatidic acid, PA: phosphatidic acid, DAG: diacylglycerol, TAG: triacylglycerol, MAG: monoacylglycerol, ACC: acetyl-CoA carboxylase 1, FAS: fatty acid synthase complex, GPAT: glyercol-3-phosphate acyltransferase, LPAT: lysophosphatidic acid acyltransferase, PAP: phosphatidic acid phosphatase, DGAT: diacylglycerol acyltransferase, FAA: fatty acyl-CoA ligase, POX1: acyl-CoA oxidase, POX2-6: (3*R*)-hydroxyacyl-CoA dehydrogenases, POT1: β-ketoacyl thiolase, ER: endoplasmic reticulum.

The fatty acyl-CoA pool in *Yarrowia* consists of palmitoyl-CoA, (C16:0), palmitoleoyl-CoA (C16:1), steroyl-CoA (C18:0), oleoyl-CoA (C18:1), and linoleoyl-CoA (C18:2). Palmitoyl-CoA, the product of the FAS complex, can be converted to the other four through one or more reactions involving fatty acyl-CoA elongases or one or two fatty acyl-CoA desaturases [6]. Fatty acyl-CoAs then enter the ER if they are to become incorporated into TAG molecules. In the Kennedy pathway for TAG biosynthesis, a fatty acyl-CoA molecule is transesterified with

glycerol-3-phosphate (Glycerol-3-P) on the first carbon by a glycerol-3-phosphate acyltransferase (GPAT) enzyme to form lysophosphatidic acid (LPA) [7]. A second fatty acyl chain is to the second carbon of the glycerol phosphate head group by a lysophosphatidic acid acyltransferase (LPAT) to form phosphatidic acid (PA). Next, the committed step in forming TAGs (as opposed to phospholipids) involves removal of the phosphate from the glycerol head group by a phosphatidic acid phosphatase (PAP) to form a diacylglycerol (DAG). Lastly, the TAG molecule is finished with the addition of a third fatty acyl chain onto the third carbon of the glycerol head group. This third chain can come from a fatty acyl-CoA molecule, and in this case the reaction is catalyzed by diacylglycerol acyltransferase (DGAT), or it can be donated by a phospholipid molecule, in which case the reaction is catalyzed by phospholipid:diacylglycerol acyltransferase (LRO) [8]. Finished neutral lipids like TAGs are then stored in phospholipid monolayer-bound lipid particles or lipid bodies, which contain the enzymes needed to later catabolize them if the cell has need for the carbon or energy stored within [8].

One of the most attractive features of microbial biocatalysis for producing a chemical of interest as opposed to conventional organic or inorganic catalysis is the high selectivity of its reactions in terms of position, functional group, and stereochemistry [9]. However, for these processes, the reactions taking place to convert one chemical to another are not the only ones occurring within the cells, and these other reactions can result in loss of product yield due to the formation of by-products or biomass, or even due to consumption of the product before it could be removed from the reactor. It has been observed that as a bioreactor culture of *Yarrowia lipolytica* reaches stationary phase, the cells begin to consume their own intracellular lipids, even when there is still significant amounts of carbon in the medium; this phenomenon has been documented in the case where the carbon substrate in the medium is glucose [6] and in the case

of fatty acid mixtures [10]. It is speculated that the reason for intracellular lipid turnover could involve the rate of carbon incorporation from extracellular sources being unable to keep up with the metabolic demand of the cells in stationary phase [11]. Because the TCA cycle in oleaginous organisms tends to become inactive during stationary phase (more specifically, upon nitrogen depletion from the medium [12]), it is possible that the difficulty in generating sufficient NADH for ATP production changes the cells' preference to intracellular lipid consumption in these cases due to the higher amount of NADH obtained per molecule of fatty acid converted to acetyl-CoA than would be obtained from glucose, despite both molecules requiring two ATP spent to activate them. For the case of intracellular vs. extracellular fatty acids, extracellular fatty acids would likely cost additional ATP to import into the cell (for pumping out protons to maintain intracellular pH) [13].

Figure 6.1 also illustrates the TAG degradation process in *Yarrowia*. TAG degradation involves the sequential removal of free fatty acids from the glycerol head group by enzymes called lipases [7]. The free fatty acids are released from the lipid bodies into the cytosol, and are then activated by fatty acyl-CoA ligase (Faa) to form fatty acyl-CoA [14]. Fatty acyl-CoA is then transported into peroxisomes or glyoxysomes, and undergoes a process called β-oxidation, the first step of which involves oxidizing the fatty acyl-CoA to a trans-2-enoyl-CoA by acyl-CoA oxidase (Pox1). Water is then added across the carbon-carbon double bond, and the resulting hydroxyl group is oxidized with NAD+ by one of the other Pox enzymes (Pox2-6, depending on the chain length [15]) to form a 3-ketoacyl-CoA. Lastly, a peroxisomal oxoacyl thiolase (Pot1) enzyme catalyzes the replacement of a free CoA with the first two carbons in the acyl chain, producing acetyl-CoA. Successive rounds of this β-oxidation process for are needed for each two carbon unit to be freed from the acyl-CoA chain.

As an organism highly adept at producing and consuming lipids and other hydrophobic substrates, *Yarrowia lipolytica* possesses a large number of genes with various unique or overlapping functions involved in these processes, especially compared to non-oleaginous yeasts such as *Saccharomyces cerevisiae*. A recent paper described 16 genes in *Yarrowia* that code for putative lipases and 4 genes that code for putative esterases [16], enzymes can also potentially release fatty acids from glycerolipid molecules. We performed a BLAST search on the National Center for Biotechnology Information's (NCBI: http://www.ncbi.nlm.nih.gov/) *Yarrowia lipolytica* protein sequence database for enzymes with significant homology to intracellular lipases Tgl3, Tgl4, and Tgl5 from *Saccharomyces cerevisiae* and similar enzymes from *Debaromyces hansenii*, and found three additional genes that code for similar enzymes in *Yarrowia* that were not listed in the aforementioned paper. As one or more of these lipases or esterases may be responsible for intracellular lipid consumption in stationary phase, the aim of this study is to improve productivity in *de novo* lipid synthesis in *Yarrowia* on acetate as a carbon source by knocking out the most likely culprit genes.

Some similar attempts to improve the lipid content of *Yarrowia* by knocking out genes involved in lipid degradation were previously successful. Knocking out the six POX genes and the mitochondrial glycerol-3-phosphate dehydrogenase (GUT2, which converts Glycerol-3-P into dihydroxyacetone phosphate, or DHAP) led to a more than twofold increase in lipid content after 24 hours of growth on glucose [15]. Another study [17] determined that knocking out the TGL3 and TGL4 genes in *Yarrowia* (in this chapter referred to as the "D17" gene and "F10" gene, respectively) increases the TAG accumulation in 0.5% glucose and 3.0% oleic acid medium by almost twofold, and that disruption of the TGL4 gene abolished TAG degradation under complete starvation conditions. While the latter paper purports to conclusively prove that

TGL3 and TGL4 are the only intracellular lipases present in *Yarrowia*, the experiments performed therein tested only conditions of extracellular fatty acid accumulation and complete starvation; it is possible that one or more of the other 21 genes described previously could be unexpressed or inactive under those conditions, but expressed and/or active in different contexts such as the one tested here: stationary phase bioreactors using acetate as a sole carbon source.

We begin our study by observing the lipid synthesis behavior of a previously constructed high-oil *Yarrowia* strain MTYL065, which overexpresses the native ACC1 and DGA2 genes [18], in a moderately aerated bioreactor growing on acetate. We took RNA samples at time points at the beginning and end of the run, and performed a differential expression analysis on the 23 putative lipase/esterase genes in *Yarrowia*. We identified three genes with significantly higher expression at the late stage than at the early stage, and determined that these genes may potentially be responsible for the lipostatic behavior observed towards the end of the bioreactor run. Using the previously discussed ΔKu70 strain (see **Chapter 5**) made for us by our collaborators at Pacific Northwest National Laboratory (PNNL), we created three strains that have each of the three genes of interest disrupted, all also overexpressing ACC1 and DGA2, and compared their ability to produce lipids on acetate to strain MTYL065. Unfortunately, our results do not show improvement in lipid content from any of the strains newly constructed for this study.

6.2 Materials and Methods

6.2.1 Fermentation conditions

YPA medium used in this study contains 10 g/L Yeast Extract (BD Biosciences, New Jersey), 20 g/L Bacto™ Peptone (BD), and 25 g/L anhydrous sodium acetate (Avantor Performance Materials, Center Valley, PA). YSMa medium used for shake flask fermentations

contains 1.7 g/L Yeast Nitrogen Base (YNB, without amino acids or ammonium sulfate) (Amresco, Solon, OH), 0.79 g/L Complete Supplement Mixture (no dropout) (Sunrise Science Products, San Diego, CA), 50 g/L anhydrous sodium acetate, and 1.61 g/L ammonium sulfate (Sigma-Aldrich, St. Louis, MO) for an atomic C:N ratio of 50.

Bioreactor fermentation was carried out in a 2 liter baffled, stirred-tank bioreactor. Two preculture stages were used to generate the inoculum for the reactor. In the first preculture, *Yarrowia* strain MTYL065 is inoculated into 3 mL rich YPA medium in three 14 mL polypropylene tubes with vented caps (VWR, Radnor, PA) and shaken for 16 hours at 200 rpm and 30°C. For the second preculture, 50 mL YPA in a 250 mL shake flask is inoculated to OD = 0.1 from the fastest-growing tube from the first preculture and shaken at 30°C at 200 rpm for 48 hours. The entire flask is poured into the bioreactor as an inoculum. The bioreactor initially contains 1.6 L of 4.25 g/L YNB, 2.5 g/L yeast extract, 30 g/L sodium acetate, and 2.4 g/L ammonium sulfate, with NaOH added to achieve a pH of 7.3. The bioreactor is run initially at 28°C, and with air sparging at 5 liters per minute. The impeller maintains the DO (dissolved oxygen) level at 20%, calibrated between pure nitrogen gas sparging (0%) and air sparging with the impeller rotating at 800 rpm (100%). After 72 hours, the air sparging flow rate was reduced to 4 LPM. As acetate is consumed in the reactor, the pH of the medium will rise. To maintain pH at 7.3 and add carbon and nitrogen feed simultaneously, a feed bottle is attached to the acid pump of the reactor controller. The feed bottle is first filled with 1 L of 30% v/v acetic acid (Sigma), 1.7 g/L YNB, and 15 g/L ammonium sulfate. At 99 hours, this bottle was refilled with 1 L of 30% v/v acetic acid. The bioreactor was run for 242 hours, with OD_{600}, dry cell weight, and lipid samples taken daily along with MicroGC measurements of CO_2 in the bioreactor headspace. Dry cell weight samples are taken gravimetrically by passing a predetermined

amount of sample from the reactor (2 mL at 0 hours, 1 mL at one day, 0.5 mL at 2 days, 0.3 mL at 3 and 4 days, and 0.2 mL every day afterwards) through a pre-weighed cellulose acetate filter paper (Whatman, GE Healthcare Biosciences, Piscataway, NJ). Samples for RNA extraction were taken at 54 hours and at 242 hours to represent growth phase gene expression and late stationary phase gene expression, respectively.

For shake flask fermentations, strains are inoculated into 3 mL YPA medium in a 14 mL polypropylene tube with a vented cap and shaken for 27.5 hours at 200 rpm and 30°C for the first pre-culture. Yeast from the first pre-culture was then transferred into 250 mL shake flasks containing 20 mL of YSMa to achieve an initial optical density (OD_{600}) of 0.1 and shaken for 26 hours at 250 rpm and 30°C for the second pre-culture. Yeast from the second pre-culture was then centrifuged and washed once with YSMa medium and inoculated into 250 mL shake flasks containing 50 mL of YSMa to an initial OD_{600} of 0.1. Flasks were closed with 34mm-45mm foam plugs (Jaece Industries Inc., North Tonawanda, NY) and placed in an Innova 43 Incubator Shaker (New Brunswick Scientific, Edison, NJ) and shaken for 170 hours at 250 rpm and 30°C. Two dry cell weight and lipid samples were taken from each flask at 116 hours and at 170 hours. 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. Samples taken for lipid analysis were 400 µL. Three biological replicate flasks were used for each strain.

6.2.2 RNA extraction and differential expression analysis

The methods for RNA extraction and gene expression analysis used in this study have been previously described [18], with some modification. A 1 mL sample was collected from the bioreactor at 54 hours, and a 0.2 mL sample was collected at 242 hours. The cells from both samples are first pelleted, and then 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.

Differential expression analysis of each of the 23 putative lipase/esterase genes 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. **Table 6.1** lists the names of the genes, their NCBI Locus Tags, and the primers used during the q-RT-PCR to amplify 78-87 bp fragments corresponding to each gene's mRNA sequence. Raw fluorescence data was then input into the Real Time PCR Miner program [19] 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 relative expression levels of each gene compared to a reference gene (with confidence intervals) for the sample taken at 241.5 hours compared to the sample taken at 49.5 hours. The *Yarrowia* actin gene (YALI0D08272g) was used as a reference. Two replicate wells were analyzed for each gene and sample combination.

Table 6.1: Primers Used for Differential Expression Analysis

Name	Locus Tag	F-Primer	R-Primer	Amplicon Length (bp)
Putative Lipases from [16]				
LIP2	YALI0A20350g	CAC CCA CCG AGG AGA TAT CGT C	GGC CAG TCA ATA AAG ACC TCA CCA	84
LIP4	YALI0E08492g	AGT GCC CGA CTG GCT CAA TTA C	CTT GTC CAG CTT GGG GTA AAC G	82
LIP5	YALI0E02640g	GAT GTT GTC CCT CGA GTG CCC	GGG TTG ATC TCA GCC ACA CTA ATG	83
LIP7	YALI0D19184g	TGG AAA TCA ACC AGC AGC GAA G	ATC CGT CCC AGT TGG GCA GG	81
LIP8	YALI0B09361g	GAC AGC CCC GAG TCG GAA AC	GTG ATT TCC AGA CCG TTG CCT TC	82
LIP9	YALI0E34507g	GAC GGT TAC ACT CAG CAG TCT G	CCC TGG CAG AAG ACC ACG TTC	86
LIP10	YALI0F11429g	GGT CCT AAC CAC GCC AAC TGC	TCA AAC TGA GCG TAT GCC AAA TTG	83
LIP11	YALI0D09064g	CGG AGG CGA GGT GTA CAT TGA C	AGG ATT CTG CTG GCC GTC AC	82
LIP12	YALI0D15906g	ACT AAC CCT CCC CTA CGG ACC C	GGC TGC TGT TAT AAG AGA GGT GG	87
LIP13	YALI0E00286g	GCC AAC CGC AAT ACG CGA AC	TCG ATG ACC TCC TTG TCT GTC TG	82
LIP14	YALI0B11858g	GGT CTT CTC TTG TGC CGG TCC	ACA ATG TCT AAC CGA GCA AGA GC	78
LIP15	YALI0E11561g	GGC TCA CCC ACT GGA ACG ATA TC	TCA ATG TAC ACT TCT CCA ATG GCG	82
LIP16	YALI0D18480g	ATA CGT CAA TGC GTG CGC TG	GAG AAT GTT GAC TCG AGC CAG C	82
LIP17	YALI0F32131g	CTT ACC ATC AAC CCC GAA AGA CG	TGT GAA TAA CCA GCG GGC CAT C	83
LIP18	YALI0B20350g	CCA TTC ACT TGG AGG CAC TGC	CCA GCC GAA AAA ACT AGG GGA G	84
LIP19	YALI0A10439g	CGC TCA AGG TCA AGG AGA GAG	CCA GAA GGG AAA CCG AGG CAC	80
Putative Esterases from [16]				
LIP1	YALI0E10659g	ACT TCA CGC TCT CAT GGT GC	ATG GTC CCT GAC GAG AGG ATA AC	84
LIP3	YALI0B08030g	AGC TCA ATG TGG CTG TGT GC	AGT CAG GAG TCA GAT GCT GC	81
LIP6	YALI0C00231g	ATG CGG TTC CCA AGG TCG AG	TCG TCA GGA TGC ACT TTC TTC G	85
LIP20	YALI0E05995g	CAG AGC TAC TCG GCC AAC TG	GAA GGG TCA AAG GGA TCA ACC TGG	82
BLAST Search Genes				
D16	YALI0D16379g	GTT GCA GCG GTT TTC GGG GTC	CAC GGG TCG CCC AAA ATA TGC	81
D17	YALI0D17534g	GAC ATA GAG GCT ACC ACG GCT	TAG TCG AAC TGG GCG AGT CTG	85
F10	YALI0F10010g	ATC ACA ATC CTG CCC CAG GTC C	CGG TTG GTG TCC TTC ATG TAT GC	86
Reference Gene				
Actin	YALI0D08272g	TCC AGG CCG TCC TCT CCC	GGC CAG CCA TAT CGA GTC GCA	141

6.2.3 Plasmid and Strain construction and confirmation

All *Escherichia coli* strains developed in this study are plasmid carrying transformants of the cloning strain DH10β (New England Biolabs, Ipswich, MA). Standard growth conditions for *E. coli* for the purposes of plasmid construction and generation have been previously described [20].

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). Strain MTYL065 is described previously [18] (see **Chapter 4**) and was

constructed by transformation of Po1g with plasmid pMT065, containing a copy of ACC1 under the control of the hp4d promoter and a copy of DGA2 under the control of the full TEF1α promoter.

Knockout cassettes for lipase/esterase genes were constructed in pUC19 plasmids and transformed into *E. coli* for maintenance. DNA fragments of homologous regions near and including part of the gene to be disrupted (LIP10, LIP4, or LIP1) of at least 1.6 kb in size were amplified along with the native URA3 gene, and were cloned into a PCR-amplified linear pUC19 using the Gibson Assembly Method [21], with the URA3 gene being flanked by the upstream and downstream homology regions. **Table 6.2** lists the primers used in the construction of these three plasmids.

Table 6.2: Primers Used for Plasmid Construction in this Study

Amplicon	F-Primer [Underlined = Binding Site, **Bold** = Stop Codon]	R-Primer [Underlined = Binding Site, **Bold** = Stop Codon]	Homologous Region Length
pUC19	GGGGATCCTCTAGAGTCGACC	GGGTACCGAGCTCGAATTCAC	
LIP10-Up	GTGAATTCGAGCTCGGTACCCCTGATCATTGTTCACAATTGCACAACC	GGGCATGTTGTTGTGTTTCTCG**TTA**GGAAACCAACACACGAGGCAAG	2052
LIP10-M	CTTGCCTCGTGTGTTGGTTTCC**TAA**CGAGAAACACAACAACATGCCC	GGGTTGCTGAAAAGGAGACCAGCGCCCAGAGAGCCATTGAC	
LIP10-Dn	GTCAATGGCTCTCTGGGCGCTGGTCTCCTTTTCAGCAACCC	GGTCGACTCTAGAGGATCCCCGTAAACATGTCCAGTCAAGTTGATGCC	2208
LIP4-Up	GTGAATTCGAGCTCGGTACCCTCCAGCTTGACATGGACTGCC	GGGCATGTTGTTGTGTTTCTCG**TTA**GGTCGTATGCAATGTAGCCGG	1888
LIP4-M	CCGGCTACATTGCATACGACC**TAA**CGAGAAACACAACAACATGCCC	CGCTCTCCTCGTCTATGAACACGCCCAGAGAGCCATTGAC	
LIP4-Dn	GTCAATGGCTCTCTGGGCGTGTTCATAGACGAGGAGAGCG	GGTCGACTCTAGAGGATCCCCTCAGTCTCTCTGACCATCACACCG	1693
LIP1-Up	GTGAATTCGAGCTCGGTACCCGGTGGCTGCTCTCTCTTACTGGTG	GGGCATGTTGTTGTGTTTCTCG**TTA**CCGTGGATCCATACAAACACAGG	1881
LIP1-M	CCTGTGTTTGTATGGATCCACGG**TAA**CGAGAAACACAACAACATGCCC	GCTCAAACAGAAGTTCCAGCTGCGCCCAGAGAGCCATTGAC	
LIP1-Dn	GTCAATGGCTCTCTGGGCGCAGCTGGAACTTCTGTTTGAGC	GGTCGACTCTAGAGGATCCCCGTGGACGAGAAGCAGAAGACG	2066

To construct our lipase/esterase knockout strains, we used strain YLKuKO, developed by our collaborators at Pacific Northwest National Laboratory (PNNL) with the removal of the Ku70 gene from strain Po1g using *Agrobacterium tumefaciens*-mediated gene transfer (described in more detail in **Chapter 5**). The URA3 gene in strain YLKuKO was disrupted by transformation with a cassette containing the first 234 bp of the gene and the last 300 bp of the gene followed by recovery on 1 g/L 5'-fluoroorotic acid (5-FOA) plates, yielding strain YLKuKO-U. Strain YLKuKO-U was

then used in three transformations, each with a different lipase/esterase gene knockout cassette amplified from their respective plasmid using the "Up" F-Primer and "Dn" R-Primer (see **Table 6.2**) and recovered on uracil dropout plates to yield URA3-complemented knockouts of each (LIP10, LIP4, LIP1) gene. Confirmed knockouts were then complemented with plasmid pMT065 to create high-oil versions of the knockout strains with roughly the same background as the strain MTYL065 used in the bioreactor. A control strain for shake flask fermentations was created by transforming YLKuKO-U with the URA3 marker first and then with pMT065, yielding strain YL65K. *Yarrowia* strains were transformed using the lithium acetate method [22]. The *Yarrowia* strains (and their genotypes) used in shake flask fermentations to assess the effect of the gene knockouts are listed in **Table 6.3**.

Table 6.3: *Yarrowia lipolytica* Strains Used in this Study

Name	Genotype	Source
po1g	MATa, leu2-270, ura3-302::URA3, xpr2-332, axp-2	Yeastern Biotech
MTYL065	As po1g, with hp4d-ACC1, TEF1α-DGA2, LEU2+	[18]
YLKuKO	As po1g, with ΔKu70	PNNL
YLKuKO-U	As YLKuKO, with ΔURA3	This work
YL65K	As YLKuKO-U, with URA3+, LEU2+, and hp4d-ACC1, TEF1α-DGA2	This work
YL65KL10	As YL65K, with LIP10::URA3	This work
YL65KL4	As YL65K, with LIP4::URA3	This work
YL65KL1	As YL65K, with LIP1::URA3	This work

Knockouts were confirmed first by attempting to PCR amplify the entire coding sequence of the gene from the genomic DNA of transformant colonies and performing gel electrophoresis on the finished reaction mixture in the presence of SYBR Safe DNA Staining Dye (Life Technologies, Grand Island, NY) alongside a 2-Log DNA Ladder (NEB). The knockout cassettes were designed to contain the first 200-400 bp of the 5' region of the coding sequence in the "Up" homology region, and the last 200-400 bp of the 3' region of the coding sequence in the "Dn" homology region; the combined length of these fragments along with the URA3 marker between them was

in all cases significantly larger than the native coding sequence. Visualizing the gel on a Proteinsimple AlphaImager HP (Bio-Techne, Minneapolis, MN) gel imager was used for the first confirmation of a knockout, as the larger amplicon from the knockout cassette was easy to differentiate from the smaller band obtained from the native coding sequence. Results from a colony that showed the larger band only were considered successful for this first step. A second step to confirm the knockout was performed to prove locus targeting of the transformed cassette by amplifying a region of the gDNA around the lipase/esterase locus using primers that bind at least 50 bp outside of the homology regions used in the knockout cassette. A digestion assay on this amplicon was performed using the SacI-HF (NEB) enzyme for LIP10 and the EcoNI enzyme (NEB) for LIP4 and LIP1 knockouts. Visualization of the digested DNA fragments confirmed that locus replacement had taken place, as the number and/or location of restriction sites in the locus are different if part of the lipase/esterase gene is replaced with URA3. Proof of locus targeting and an inability to detect the intact coding sequence in the gDNA of the same colony was considered confirmation of a gene knockout.

6.2.4 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 [2]. Samples of volume V {uL} = ($4000/OD_{600}$) from the bioreactor and 400 uL from shake flask fermentations 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 the bioreactor and each biological replicate in the fermentations, two lipid samples were analyzed per time point.

6.3 Results

6.3.1 High-oil strain MTYL065 shows evidence of intracellular lipid consumption in the late stage of growth on acetate in a bioreactor

In order to determine which genes may be responsible for lipid catabolism in *Yarrowia* used in SCO processes using acetate as a carbon source, it was important to first recreate this behavior in an industrially relevant context so that cell samples could be drawn from it for further analysis. This necessitated the use of a bioreactor scale fermentation with a high-oil engineered strain to

approximate industrial scale with a suitable host organism as closely as possible. The 2 L scale acetate bioreactor was run for approximately 10 days using the previously developed MTYL065 strain with the operating conditions described in **Section 6.2.1**. We tracked the OD and lipid content (grams lipids/gram dry cell weight) as well as the carbon dioxide output from the cells in the bioreactor headspace, which serves as a proxy measurement for their metabolic activity [23]. The results from the bioreactor fermentation are shown in **Figure 6.2** and **Figure 6.3**.

Figure 6.2 shows that our bioreactor run achieved a high maximum lipid content of 61.2% at 194 hours, which also coincided with the maximum OD_{600} of 270. The lipid content in the reactor increases sharply in the early exponential growth phase up to 75.5 hours, and afterwards grows more steadily until reaching its maximum. After the 194 time point, the lipid content declines slightly, suggesting that lipid catabolism is taking place. **Figure 6.3** shows the carbon dioxide output of the cells throughout the bioreactor run. While there is a sharp decline in the specific CO_2 output in the early stage of the run, there is a noticeable stabilization starting at 145.75 hours. At the end of the bioreactor run, the specific CO_2 output in the stationary phase cells (55.1 μL per min

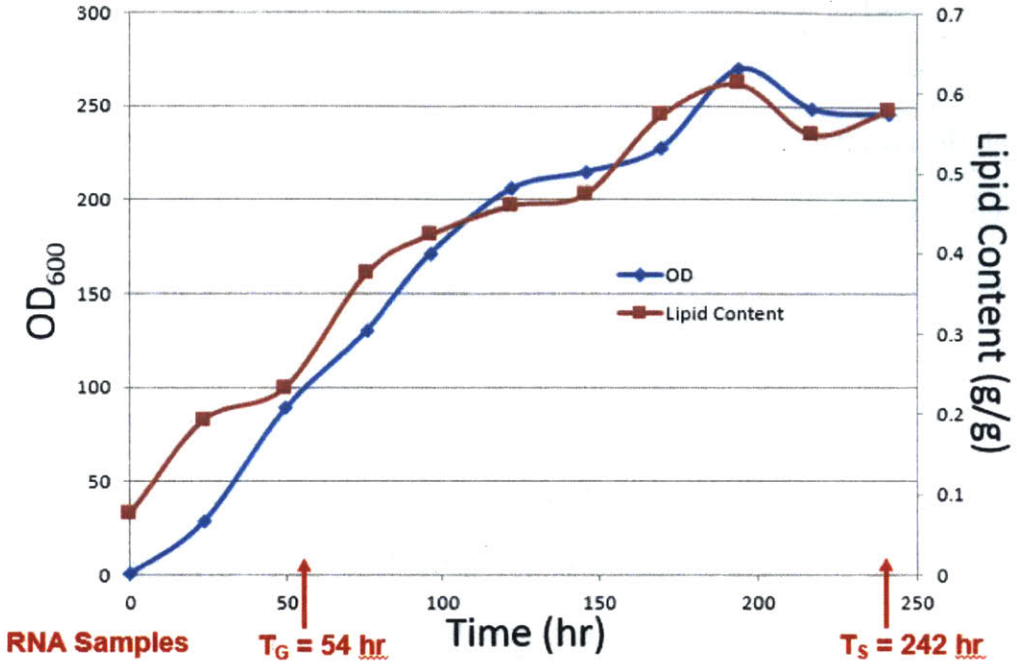

Figure 6.2: Performance of MTYL065 in 2 L scale acetate bioreactor. OD_{600} and lipid content were measured once per day for ten days. Lipid and dry cell weight samples taken for lipid content determination were taken in duplicate. Two RNA samples were taken, one at 54 hr (T_G) to represent the growth phase, and one at 242 hr (T_S) to represent the stationary phase.

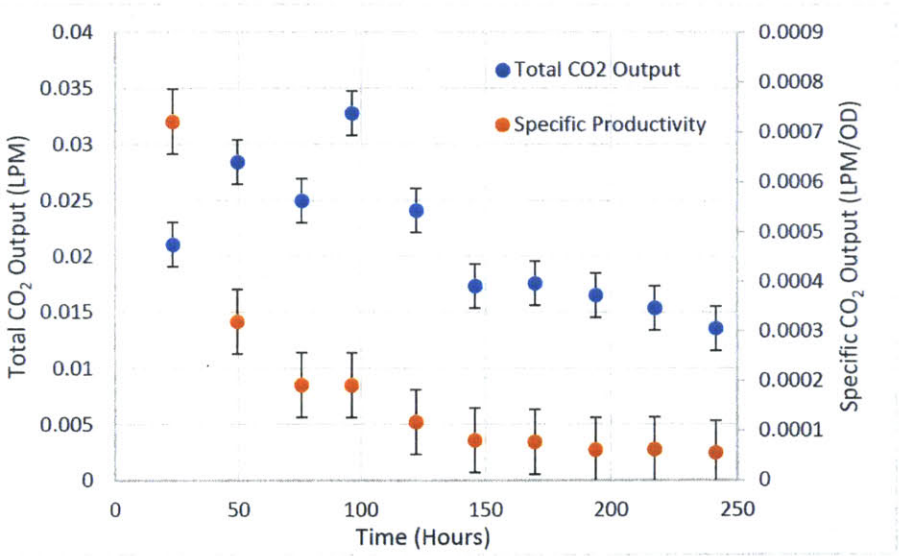

Figure 6.3: Carbon dioxide output in liters per minute (LPM) of MTYL065 in 2 L scale acetate bioreactor. Measurements were taken by attaching the output hose from the bioreactor condenser to a MicroGC. Error bars represent the standard deviations of three measurements once the readout of the instrument was stabilized.

per OD) is approximately 7.6% of the maximum specific productivity at the early growth phase, suggesting a small amount of metabolic activity is ongoing within cells in the late stationary phase.

6.3.2 Differential expression analysis reveals several putative lipase/esterase genes with significantly increased expression during stationary phase

A recent review [16] describes 16 putative lipase and four putative esterase genes discovered after the complete sequencing of the haploid *Yarrowia lipolytica* genome. In addition to these genes, we performed a BLAST search on all *Yarrowia* protein sequences for proteins that had homology to known intracellular lipases TGL3, TGL4, and TGL5 from *Saccharomyces cerevisiae* and similar proteins from *Debaromyces hansenii*. This search turned up three additional *Yarrowia* genes that we named "D16," "D17," and "F10," after the first three unique characters in their NCBI locus tag. We hypothesized that one or more of these genes, listed in **Table 6.1**, were responsible for the lipid catabolism behavior observed in stationary phase cultures by ourselves and others [6][10]. Because it would be a significant undertaking (and with no guarantee of the reward of a higher productivity strain) to knock out every gene of interest, and because very little is known about most of these genes [16], we needed some way to prioritize which genes we should focus on that were most likely to have significant activity in stationary phase cells.

Cell samples from our bioreactor were taken at 54 and 242 hours for the purpose of mRNA analysis. We hypothesized that the genes that were expressed highly in stationary phase compared to the growth phase were most likely to be responsible for the lipid catabolism observed in our bioreactor. The results of our differential expression analysis are given in **Figure 6.4** and **Table 6.4**, with **Figure 6.4** showing the range of expression ratio samples

generated by the REST 2009 bootstrapping program (see **Section 3.2.3**) of the gene fold expressions in stationary phase relative to growth phase and **Table 6.4** providing more absolute gene expression data from both phases in terms of relative abundance of transcripts from each gene relative to the actin reference gene.

The largest change from growth phase to stationary phase was seen in the LIP10 gene, where expression increased by a mean factor of approximately 107. Large changes to LIP4 (18.7) and LIP1 (12.1) were also observed. Also significant were changes to LIP17 (5.65) and LIP15 (4.28). Most other genes either showed insignificant changes or were less abundantly expressed in stationary phase.

Three genes that have been more extensively studied and are known to code for extracellular lipases are LIP2 [16], LIP7, and LIP8 [24]; the fold changes in these genes were 0.724, 1.623, and 0.829, suggesting that the cell does not significantly increase overexpression of these genes during intracellular lipid consumption. Two genes previously confirmed to be active intracellular lipases during conditions of starvation [17], TGL3 ("D17") and TGL4 ("F10") did not show significant fold-changes in expression (1.013 and 1.237, respectively), although these genes are expressed relatively highly throughout both phases of the fermentation, as shown in the "Growth LOG(Gene/Actin)" and "Stationary LOG(Gene/Actin)" lines in **Table 6.4**. Genes such as LIP8, LIP12, LIP13, LIP14, LIP16, LIP18, LIP3, LIP6, LIP20, and D17 are also relatively highly expressed in both phases, but only LIP3 and LIP6 show statistically significant small increases in expression from growth phase to stationary phase; all others had either similar expression levels or a decrease in expression between phases.

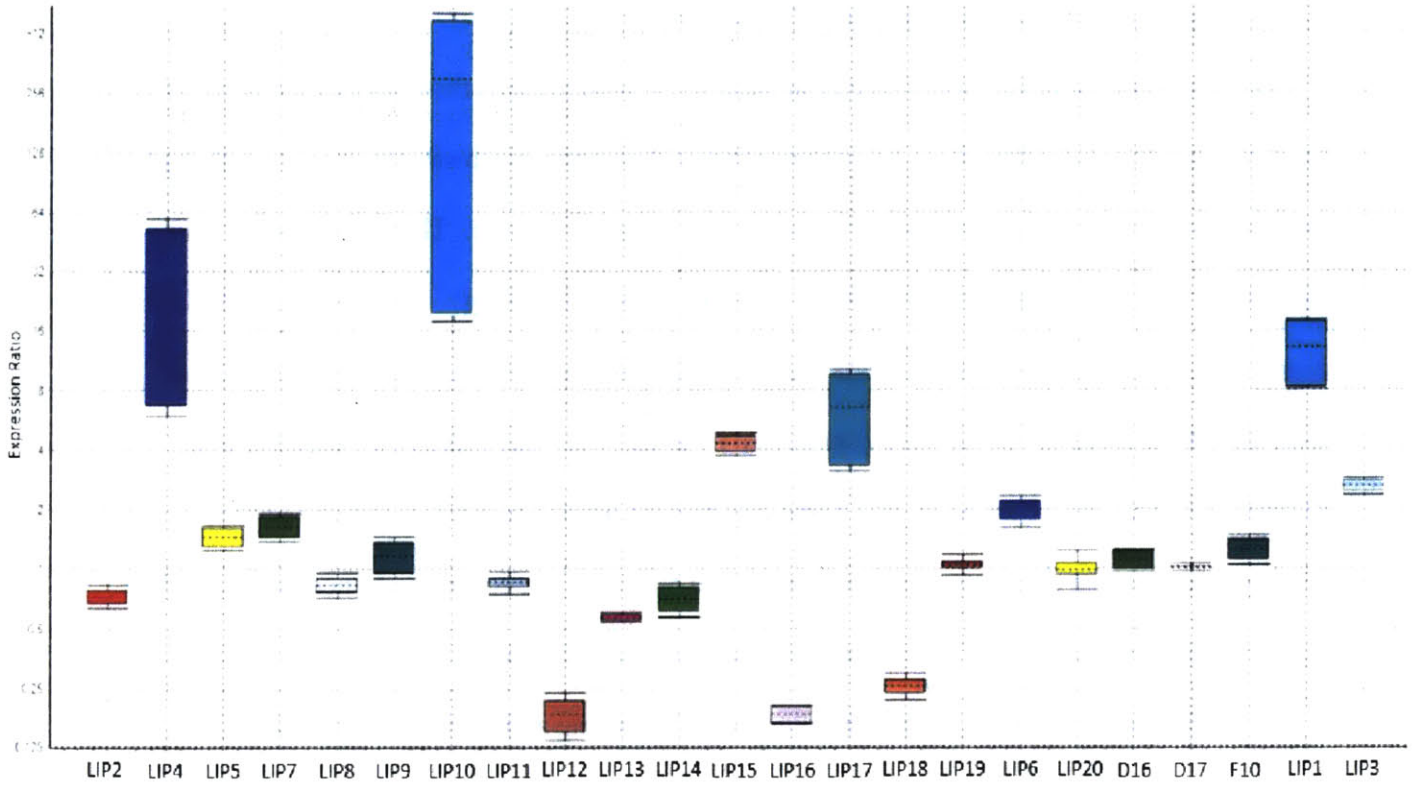

Figure 6.4: Expression ratios of putative lipase/esterase genes of stationary phase cells compared to growth phase cells. The dotted line in the box represents the median fold change of the samples generated by the REST 2009 statistical bootstrapping program, while the edges of the box represent second and third quartiles of generated samples (equivalent to a 50% confidence interval) and the ends of the whiskers represent the entire range of generated samples.

Table 6.4: Expression of putative lipase/esterase genes relative to actin, and mean normalized expression change from growth to stationary phase with 95% confidence intervals.

	Putative Lipases																Putative Esterases				BLAST Results		
	LIP2	LIP4	LIP5	LIP7	LIP8	LIP9	LIP10	LIP11	LIP12	LIP13	LIP14	LIP15	LIP16	LIP17	LIP18	LIP19	LIP1	LIP3	LIP6	LIP20	D16	D17	F10
Growth LOG(Gene/Actin)	-2.150	-4.100	-2.522	-2.765	-0.582	-2.103	-4.568	-2.137	0.375	-1.568	-1.504	-2.427	0.957	-3.064	-0.916	-2.257	-3.671	-1.639	-1.709	-0.723	-0.824	-0.313	-1.594
Stationary LOG(Gene/Actin)	-2.290	-2.829	-2.367	-2.555	-0.664	-2.048	-2.539	-2.208	-0.370	-1.813	-1.662	-1.796	0.222	-2.312	-1.512	-2.239	-2.587	-1.225	-1.422	-0.730	-0.784	-0.307	-1.502
LOG-Ratio	-0.140	1.272	0.156	0.210	-0.082	0.055	2.029	-0.070	-0.744	-0.245	-0.158	0.631	-0.735	0.752	-0.596	0.019	1.084	0.414	0.287	-0.006	0.040	0.006	0.092
Mean Relative Expression	0.724	18.685	1.431	1.623	0.829	1.134	106.954	0.851	0.180	0.569	0.694	4.279	0.184	5.650	0.253	1.044	12.141	2.596	1.936	0.986	1.096	1.013	1.237
95% CI LB	0.639	5.945	1.243	1.39	0.722	0.895	17.769	0.752	0.138	0.536	0.575	3.768	0.165	3.151	0.217	0.932	8.078	2.353	1.618	0.8	0.96	0.969	1.047
95% CI UB	0.821	58.815	1.648	1.896	0.951	1.437	644.25	0.964	0.236	0.604	0.839	4.861	0.205	10.134	0.296	1.17	18.247	2.865	2.318	1.218	1.251	1.059	1.462

6.3.3 Construction and confirmation of knockout strains

The results of the differential expression analysis pointed towards three genes that seemed to be strong candidates for knockout given their relative dormancy in growth phase along with high fold change in expression levels when the cells are in stationary phase. These genes are LIP10 (YALI0F11429g), LIP4 (YALI0E08492g), and LIP1 (YALI0E10659g). We constructed DNA cassettes to target each locus and disrupt their coding sequences by replacing a portion of them with the URA3 marker. These were constructed by assembling three fragments together in a plasmid: a URA3 gene flanked by two homology regions upstream and downstream of the coding sequence that also include the beginning and ending of the coding sequence, respectively. After transformation, we confirmed the knockout using two PCR-based assays.

In the first assay, we used primers binding to the beginning and ending of the coding sequence in order to attempt to amplify the entire coding sequence of the lipase/esterase from the gDNA of each transformant colony screened. Because the knockout cassettes also contained the primer binding sites, both the native coding sequence and the region of the knockout cassette containing the included coding sequence fragments around the URA3 marker could be amplified at the same time. The knockout cassette was purposely made to be the longer amplicon, making it more difficult to amplify than the native coding sequence; therefore, we can be more confident in the event that we can detect the knockout cassette band but not the native coding sequence band that this observation can be attributed to the lack of the native coding sequence in the gDNA. **Table 6.5** shows the expected bands for the native coding sequences and the knockout cassette regions amplified in this assay for all three genes.

In the second PCR assay, we used primers that bind just outside (50-100 bp away) the homology regions in the knockout cassette in order to amplify the entire locus and confirm that

homologous recombination had taken place there. After confirming amplification of the locus, the amplicon was digested by a restriction enzyme that would produce a different banding pattern from the native locus if coding sequence was disrupted with URA3: for LIP10, SacI-HF was used, and for LIP4 and LIP1, EcoNI was used. The expected banding patterns of the locus amplicon digestion assay are also listed in **Table 6.5**.

Table 6.5: Expected bands for lipase/esterase knockout confirmation assays

Amplicon/Banding Pattern	LIP10	LIP4	LIP1
Coding Sequence Detection - Native Gene	1023	1221	1461
Coding Sequence Detection - KO Cassette	1595	1804	1905
Locus Targeting Digest - Native Locus	3138 1889	4430	4129 690 21
Locus Targeting Digest - Successful Targeting	2587 1889 1174	2735 2281	2572 2004 690 21

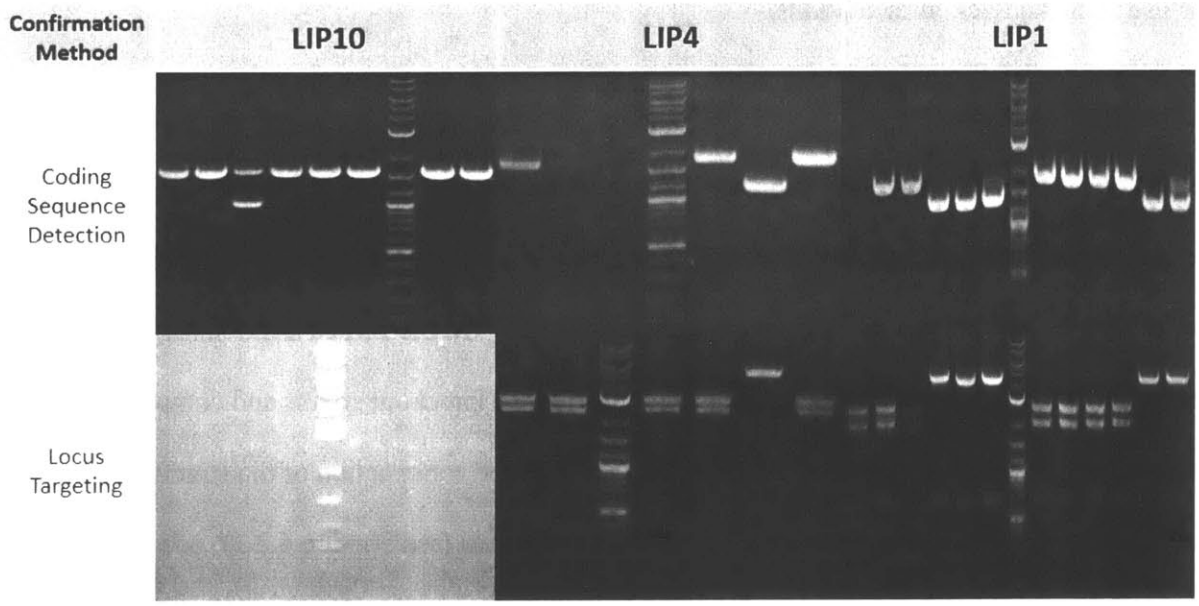

Figure 6.5: Gel image results of knockout confirmation PCR assays. Expected bands for each gene and assay are given in Table 6.5. In each image, the NEB 2-Log Ladder is the reference. For each gene, the results shown in both assays correspond to the same colonies in the same order.

Figure 6.5 shows the results of both knockout confirmation assays. The lanes from left to right in order correspond to the same colonies in both assay images for each gene. Colonies were considered "confirmed" for knockout of the gene of interest if both assays were conclusively passed; this means only the "KO Cassette" band was present in the coding sequence detection assay and the banding pattern for the locus digestion assay exactly matched that of the "Successful Targeting" list from **Table 6.5**. Of the colonies shown in the figure, seven out of eight were confirmed as knockouts in the case of LIP10, with only the third colony failing the coding sequence detection. Four of six LIP4 colonies were confirmed knockouts, as the third colony did not a conclusive band for the first assay, and the fifth colony clearly failed both. Six out of twelve LIP1 colonies passed both assays; the first colony did not give a conclusive band for the first assay, and the fourth, fifth, sixth, eleventh, and twelfth colonies clearly failed both assays. A confirmed colony for each knockout was then complemented with plasmid pMT065 for use in shake flask fermentations.

6.3.4 Shake flask fermentations with ACC1 and DGA2 overexpressing lipase/esterase knockout strains

After confirming the construction of all knockout strains and our remade control strain (using the ΔKu70 background), we tested whether disrupting any of the three genes of interest led to a strain with increased lipid productivity or cellular content. We performed shake flask fermentations in YSMa media in triplicate with each of the knockout strains and compared their performance to the remade control. Due to the timing of the construction of the strains, we performed two separate experiments using similar procedures (see **Section 6.2.1**): one testing the YL65KL4 and YL65KL1 against the YL65K control, and one testing strain YL65KL10 against the control. Fermentations were run for 170 hours, with time points taken at 116 hours and 170 hours in anticipation of observing lipostatic or lipid catabolic behavior in the strains between the

two time points. The lipid titers and contents observed from our fermentations are given in **Figure 6.6**.

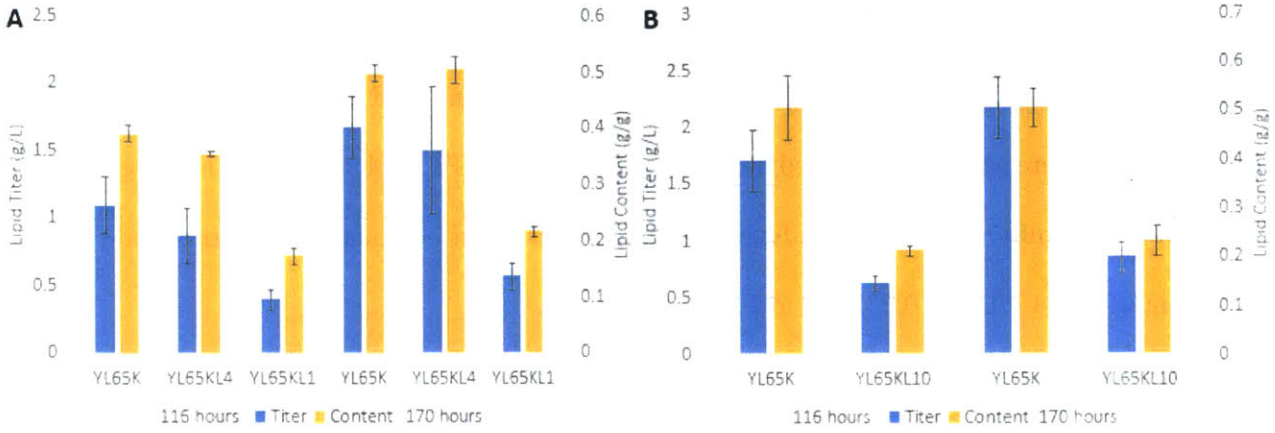

Figure 6.6: Lipid titer and content at 116 hours and 170 hours for fermentations with (A) the LIP4 and LIP1 knockout strains and (B) the LIP10 knockout strain. Error bars represent the standard deviations from three biological replicates.

The data in **Figure 6.6** clearly shows that none of the knockout strains we have created outperform our control strain. In our first experiment, our YL65K control strain had a lipid titer and content of 1.09 g/L and 0.390 g/g, respectively, at 116 hours, and 1.67 g/L and 0.496 g/g, respectively, at 170 hours. With intracellular lipids still being accumulated past 116 hours, we find that our LIP4 knockout strain performed very similarly at both time points to our control. The lipid titer and content at 170 hours were 1.50 g/L and 0.503 g/g, respectively, in this strain. The LIP1 knockout strain performed very poorly, with a slightly slower growth rate, but much lower capacity to store lipids compared to the control. The lipid titer and content in this strain at 170 hours were 0.566 g/L and 0.216 g/g, respectively.

In our second experiment, the lipid contents of both strains did not change significantly between time points, but the titers increased, suggesting that lipids and other non-lipid biomass were being produced at similar rates during this time. Therefore, it is possible that these cells did

not reach a stationary phase during the fermentation. However, the results anyway conclusively show that the effect of LIP10 knockout was detrimental to overall lipid production, not unlike that of the LIP1 knockout. The control strain in this fermentation produced a lipid titer and content of 1.71 g/L and 0.508 g/g at 116 hours, and 2.18 g/L and .508 g/g at 170 hours. The LIP10 knockout strain demonstrated a similar phenotype to LIP1, with a slightly slower growth rate but poor ability to produce lipids; this strain only managed to produce a lipid titer and content of 0.854 g/L and 0.232 g/g in 170 hours.

6.4 Discussion

Typically, oleaginous yeasts begin to increase the rate of neutral lipid production for energy and carbon storage in response to the depletion of a nutrient needed for growth, typically nitrogen [25]. After nitrogen depletion, the AMP deaminase enzyme converts AMP into IMP (inosine monophosphate), lowering the cellular concentration of AMP needed to activate the mitochondrial isocitrate dehydrogenase enzyme in the TCA cycle. With no nitrogen to produce biomass with, the cells enter a stationary phase, with the incoming carbon being directed towards lipid production. Despite no longer having the burden of preparing biomass for cell division and growth, a significant amount of energy is still required for cell maintenance (as well as lipid production). However, as the TCA cycle in this stationary phase has little activity, other catabolic processes are needed to obtain that energy. While it may seem logical to assume that with excess carbon in the media, the extracellular carbon would be imported and catabolized to produce the energy for cell maintenance and that excess energy produced in this way could be used to store more of the extracellular carbon in the form of neutral lipids, it has been observed previously [6][10] that after a certain point in stationary phase, the net flux of fatty acid and lipid production in *Yarrowia lipolytica* turns negative, even in the presence of significant amounts of

extracellular carbon. This is likely due to either the rate of intracellular lipid consumption being faster than that of substrate import and consumption [11], or because importing a substrate may be more energetically expensive in terms of activation (phosphorylation or thioesterification with CoA) and maintaining homeostasis than using intracellular lipids for energy. Although elimination of the POX genes prevents fatty acid degradation [15], free fatty acids and their CoA thioesters may be inhibitory to enzymes involved in fatty acid synthesis, such as ACC1 [18, 26]. It is therefore important to sequester fatty acids in lipid bodies in the form of TAGs, and prevent their mobilization by active intracellular lipases.

In this study, we ran a bioreactor with a high-oil (ACC1- and DGA2- overexpressing) strain on acetate and observed this catabolic behavior ourselves. The lipid content of our bioreactor reached a maximum of 0.612 g/g, which is similar to previous reports of acetate bioreactors with this strain [27], suggesting that this behavior is usually present in this strain. The differential expression analysis performed on growth phase and late stationary phase mRNA samples from the bioreactor turned up three genes that met the criteria we hypothesized would make them most likely to be responsible for the lipid mobilization in the bioreactor: LIP10, LIP4, and LIP1 all were not highly expressed during growth phase, but had very significantly increased expression in the late stationary phase. Unfortunately, the confirmed knockout strains we constructed and complemented with the ACC1 and DGA2 overexpression plasmid performed poorly in preliminary shake flask fermentations, with the LIP4 knockout performing similarly to the high-oil control and the LIP10 and LIP1 knockouts showing greatly diminished capacity for lipid production despite only a slight increase in growth rate.

While it seems reasonable to assume that Lip4 does not have an important role in mobilizing TAGs under the conditions tested here, and may not even be localized inside the cell at all,

despite the large increase in expression seen only in high oil content cells, the reasons for the behavior of LIP10 and LIP1 knockouts remain unclear. It was recently demonstrated in a paper [17] that was released after we performed our bioreactor run, BLAST search of intracellular lipases, and differential expression analysis, that the "D17" and "F10" genes in this study are indeed intracellular lipases in *Yarrowia lipolytica* (called TGL3 and TGL4, respectively), or are at least involved in TAG catabolism, with significant activity during growth on oleic acid and in starvation conditions. However, like the Tgl4 enzyme from *Saccharomyces cerevisiae* also exhibits LPAT activity [28], we recently found that overexpression of the native TGL4 gene in *Yarrowia* leads to a small but significant increase in lipid production [29] (see **Chapter 3**), meaning it likely has a similar multifunctionality. Tgl3 has also been speculated to be a positive regulator of Tgl4 instead of an actual lipase. It is possible that LIP1 and LIP10 have some other function aside from lipase activity that makes them important to lipid productivity, perhaps with more of a regulatory role that would not require them to have high expression levels to be effective in the cell.

One detail we noticed when looking at our differential expression analysis results was that the three confirmed extracellular lipases [16], LIP2, LIP7, and LIP8, which we would not expect to show expression levels that correlate with lipid content, indeed were not differentially expressed to a large extent. This may give us some indication that our expression results are reliable. However, expression of the TGL3 and TGL4 genes, which are known to have lipase-related activity, also did not show a correlation between expression and lipid content. For the case of these two genes, it seems that the way they are regulated is not based on transcription, but on protein localization [17]. With our three candidate lipases/esterases demonstrated to be ineffective at increasing lipid productivity when knocked out, it is possible that choosing

different criteria upon which to pick gene targets for future work may be more fruitful in discovering more active intracellular lipases in *Yarrowia*, if they exist. It is possible that genes that are more highly expressed throughout the fermentation but still have increased expression at high lipid contents, such as LIP15 and LIP17, may be worth considering; LIP15 has also been previously observed to be upregulated after nitrogen depletion on glucose media, prompting speculation that it may be involved in lipid turnover [30]. Of course, a study with a TGL4 knockout in a high-oil strain such as our ACC1- and DGA2-overexpressing background strain is also a logical next step that could lead to a potentially very synergistic effect in engineering host organisms for improved productivity in SCO processes.

6.5 Conclusions

In this study, we attempted to discover genes that code for intracellular lipases in *Yarrowia lipolytica* that were active when the lipid content is high and are responsible for loss of product in bioreactor fermentations. Although a differential expression analysis on growth phase and late stationary phase cells from an acetate bioreactor with a high-oil strain implicated three genes that had expression levels that correlated positively with lipid content, the removal of all three genes had either no effect (for LIP4) or a detrimental effect (for LIP10 and LIP1) on overall lipid production in shake flask fermentations. The reasons for the poor performance of the LIP10 and LIP1 knockouts may be due to multifunctionality of these genes. While it may be important to characterize every lipase/esterase gene in *Yarrowia* to eventually engineer a strain with no capacity for intracellular lipid consumption, the TGL3 (YALI0D17534g) and TGL4 (YALI0F10010g) genes have already been shown to have such activity; a logical next step for SCO host improvement would involve knocking out these two genes in a high-oil background strain.

References:

1) 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. *Progress in Lipid Research*, *48*(6), 375–87. doi:10.1016/j.plipres.2009.08.005
2) 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.
3) 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 Technology*, *114*, 443–9. doi:10.1016/j.biortech.2012.02.091
4) Tehlivets, O., Scheuringer, K., & Kohlwein, S. D. (2007). Fatty acid synthesis and elongation in yeast. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1771*(3), 255–270. doi:10.1016/j.bbalip.2006.07.004
5) Nelson, D. L. & Cox, M. M. (2005). *Lehninger principles of biochemistry* (4th ed.). New York: W.H. Freeman. Pg. 789.
6) 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. *European Journal of Lipid Science and Technology*, *111*(12), 1221–1232. doi:10.1002/ejlt.200900055
7) Czabany, T., Athenstaedt, K., & Daum, G. (2007). Synthesis, storage and degradation of neutral lipids in yeast. *Biochimica et Biophysica Acta*, *1771*(3), 299–309. doi:10.1016/j.bbalip.2006.07.001
8) Daum, G., Wagner, a., Czabany, T., & Athenstaedt, K. (2007). Dynamics of neutral lipid storage and mobilization in yeast. *Biochimie*, *89*(2), 243–248. doi:10.1016/j.biochi.2006.06.018
9) Johannes, T. W., Simurdiak, M. R., & Zhao, H. (2006). Biocatalysis. *Encyclopedia of Chemical Processing*, 101–110. doi:10.1081/E-ECHP-120017565
10) 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. *Applied Microbiology and Biotechnology*, *58*(3), 308–12. doi:10.1007/s00253-001-0897-0
11) Papanikolaou, S., & Aggelis, G. (2010). Yarrowia lipolytica: A model microorganism used for the production of tailor-made lipids. *European Journal of Lipid Science and Technology*, *112*(6), 639–654. doi:10.1002/ejlt.200900197
12) Ratledge, C. (2002). Regulation of lipid accumulation in oleaginous micro-organisms. *Biochemical Society Transactions*, *30*(6), 47–50.
13) 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
14) Hiltunen, J. K., Mursula, A. M., Rottensteiner, H., Wierenga, R. K., Kastaniotis, A. J., & Gurvitz, A. (2003). The biochemistry of peroxisomal β-oxidation in the yeast Saccharomyces cerevisiae. *FEMS Microbiology Reviews*, *27*(1), 35–64. doi:10.1016/S0168-6445(03)00017-2

15) Beopoulos, A., Mrozova, Z., Thevenieau, F., Le Dall, M.-T., Hapala, I., Papanikolaou, S., ... Nicaud, J.-M. (2008). Control of lipid accumulation in the yeast Yarrowia lipolytica. *Applied and Environmental Microbiology*, *74*(24), 7779–7789. doi:10.1128/AEM.01412-08
16) Fickers, P., Marty, A., & Nicaud, J. M. (2011). The lipases from *Yarrowia lipolytica*: Genetics, production, regulation, biochemical characterization and biotechnological applications. *Biotechnology Advances*, *29*(6), 632–644. doi:10.1016/j.biotechadv.2011.04.005
17) Dulermo, T., Tréton, B., Beopoulos, A., Gnankon, A. P. K., Haddouche, R., & Nicaud, J. M. (2013). Characterization of the two intracellular lipases of Y. lipolytica encoded by TGL3 and TGL4 genes: New insights into the role of intracellular lipases and lipid body organisation. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1831*(9), 1486–1495. doi:10.1016/j.bbalip.2013.07.001
18) Tai, M., & Stephanopoulos, G. (2013). Engineering the push and pull of lipid biosynthesis in oleaginous yeast *Yarrowia lipolytica* for biofuel production. *Metabolic Engineering*, *15*, 1–9. doi:10.1016/j.ymben.2012.08.007
19) Sheng Zhao, Russell D. Fernald. Comprehensive algorithm for quantitative real-time polymerase chain reaction. *J. Comput. Biol.* 2005 Oct;12(8):1045-62.
20) Sambrook, J., & Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. CSHL press, Cold Spring Harbor, NY, USA.
21) 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. *Nature Methods*, *6*(5), 343–345. doi:10.1038/nmeth.1318
22) Chen, D. C., Beckerich, J. M., & Gaillardin, C. (1997). One-step transformation of the dimorphic yeast Yarrowia lipolytica. *Applied Microbiology and Biotechnology*, *48*, 232–235. doi:10.1007/s002530051043
23) Kroukamp, O., & Wolfaardt, G. M. (2009). CO2 production as an indicator of biofilm metabolism. *Applied and Environmental Microbiology*, *75*(13), 4391–4397. doi:10.1128/AEM.01567-08
24) Fickers P, Fudalej F, Le Dall MT, Casaregola S, Gaillardin C, Thonart P, et al. Identification and characterisation of LIP7 and LIP8 genes encoding two extracellular triacylglycerol lipases in the yeast Yarrowia lipolytica.Fungal Genet Biol 2005c;42:264–74.
25) 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
26) Thelen, J. J., & Ohlrogge, J. B. (2002). Metabolic engineering of fatty acid biosynthesis in plants. *Metabolic Engineering*, *4*, 12–21. doi:10.1006/mben.2001.0204
27) Tai, M. (2012). Metabolic Engineering of Oleaginous Yeast for the Production of Biofuels. Ph.D. Thesis. Massachusetts Institute of Technology.
28) Rajakumari, S., & Daum, G. (2010). Multiple functions as lipase, steryl ester hydrolase, phospholipase, and acyltransferase of Tgl4p from the yeast Saccharomyces cerevisiae. *Journal of Biological Chemistry*, *285*(21), 15769–15776. doi:10.1074/jbc.M109.076331
29) Silverman, A. M.*, Qiao, K.*, Xu, P., and Stephanopoulos, G. Functional overexpression and characterization of lipogenesis-related genes in the oleaginous yeast *Yarrowia lipolytica*. Submitted.

30) Morin, N., Cescut, J., Beopoulos, A., Lelandais, G., Le Berre, V., Uribelarrea, J.-L., ... 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.

Chapter 7

Attempts at Improving Lipid Yield and Productivity on Acetate Using Heterologous Gene Expression

7.1 Introduction

While most processes involving microbial biocatalysts to create commodity chemicals use simple sugars like glucose as a carbon source to demonstrate the proof-of concept, obtaining glucose for use in such processes represents a significant cost in assessing its economic viability [1]. The most typical feedstocks used in these processes include plant cellulose and hemicellulose, which requires treatment with acid and/or steam at high temperatures and pressures, followed by enzymatic hydrolysis to break down polymers in the biomass into a mixture of five- and six- carbon sugars [2]. This pretreatment of plant biomass is slow and expensive, representing a large fraction of the total processing time and operating costs, and preventing competition with low-cost products made from non-renewable sources [3]. This has prompted interest in other carbon sources aside from simple sugars.

Acetate and other volatile fatty acids (VFAs) are an attractive carbon source due to their abundance in agro-industrial and municipal wastewater streams [4], which can be obtained at a low or even negative cost. In particular, single cell oil (SCO) processes for producing biodiesel using such waste water streams have an advantage over many other commodity chemical products (including other biofuels) due to the oil staying inside the cells, allowing for easy separation and concentration of biodiesel precursors from any undesirable chemicals in the medium after fermentation, which reduces the amount of pretreatment required of the inputs to the process. *Yarrowia lipolytica*, as an oleaginous yeast that can naturally consume acetate, VFAs, long chain fatty acids, and even very hydrophobic substrates such as alkanes [5], has emerged as a model host organism for developing SCO processes, and the recent sequencing of the genome [6] has led to some annotation of genes and the development of genetic tools [7-9] by which the lipid productivity of this organism can be further enhanced.

To make lipids from acetate involves a relatively direct path compared to sugars. Acetate first is imported into the cell as acetic acid by passive diffusion at low pH [10] or through symport with a hydrogen ion at high pH [11]. Acetate is converted into acetyl-CoA by Acs, an (AMP-forming) acetate-CoA ligase [12]. Acetyl-CoA, as a key intermediate in many biosynthetic pathways [13], is a reactant in the first committed step of fatty acid synthesis, a carboxylation to malonyl-CoA by the enzyme acetyl-CoA carboxylase (Acc1) [9]. The reaction catalyzed by Acc1 has an energy cost of one ATP. Elongation of fatty acids in the fatty acid synthase (FAS) complex requires the reduction of the growing acyl chain by two NADPH molecules for each two carbon unit added. While the pathway from acetyl-CoA to lipids may be straightforward, acetyl-CoA must distributed three ways to provide the carbon for lipids (fatty acid synthesis pathway), the NADPH reducing equivalents required for fatty acyl-CoA elongation, and the energy needed for its activation from acetate and for the activity of Acc1.

The question of which pathway(s) provides the main source(s) of cytosolic NADPH for fatty acid synthesis in *Yarrowia lipolytica* has received a lot of interest recently. Generally, there are three NADPH-producing enzymes that are thought to provide cytosolic reducing equivalents for elongation and desaturation of acyl-CoAs, with even phylogenetically similar organisms sometimes having dramatically different flux distributions among the three: 1) malic enzyme [14], 2) NADP+-specific isocitrate dehydrogenase [14], or 3) glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase from the oxidative pentose phosphate pathway (oxPPP) [15]. Previously, it has been determined that the overexpression of the native malic enzyme (MAE1) [15] gene or the native putative cytosolic NADP+-specific isocitrate dehydrogenase (IDP2) [16] (see Chapter 3) gene does not lead to improved lipid synthesis; Mae1 in *Yarrowia* was determined to be mitochondrial with a much higher affinity for NAD+ [15],

while Idp2 has been speculated by some to be mitochondrial as well [17], despite both the ESLPred (http://www.imtech.res.in/raghava/eslpred/index.html) and MitoProt (https://ihg.gsf.de/ihg/mitoprot.html) protein localization algorithms predicting cytosolic localization. The oxPPP, however, has been shown to be the main NADPH producing pathway in *Yarrowia* for lipid synthesis on glucose [18], and overproduction of 6-gluconolactonase, one of the enzymes of the pathway, has a significant effect on lipid flux when growing on both glucose and acetate [16]. For acetyl-CoA to reach the oxPPP, it must first enter the peroxisome through the carnitine acetyltransferase shuttle [19]. From here, the glyoxylate cycle (see Chapter 2) produces one succinate molecule from two acetyl-CoA molecules. The succinate then enters the mitochondrion, where it is converted into malate in the TCA cycle. Malate then leaves the mitochondrion, is converted to oxaloacetate by cytosolic malate dehydrogenase, and then to phosphoenolpyruvate (PEP) by PEP carboxykinase. PEP is then converted in the gluconeogenesis pathway to form glucose-6-phosphate, which can then be converted to carbon dioxide in the PPP to produce two NADPH reducing equivalents for each carbon removed; to make the glycerol backbone needed to store fatty acids as finished neutral lipids in the form of triacylglycerols (TAGs), PEP is instead converted to glycerol-3-phosphate is in the glyceroneogenesis pathway.

In producing the necessary NADPH from acetyl-CoA, some energy is generated in the cell; for every two acetyl-CoA molecules that take the most direct path through gluconeogenesis to go through the PPP (which would provide six NADPH reducing equivalents if run through the oxidative, non-oxidative, and gluconeogenic cycle until the carbon was fully oxidized to CO_2), two ATP are consumed, but a cytosolic NADH and a mitochondrial $FADH_2$ are generated, which can be re-oxidized to produce approximately four ATP (see Chapter 2). However, activating

acetate to acetyl-CoA using Acs costs two ATP per molecule. Therefore, there is a large net energy cost in converting extracellular acetate to intracellular acetyl-CoA for use in fatty acid synthesis and to produce the requisite NADPH, and the cell must divert a fraction of the acetyl-CoA to a high net energy producing pathway, such as the TCA cycle. Acetyl-CoA is transported into the mitochondrion by the carnitine acetyltransferase shuttle [19], and the oxidation of two carbons to CO2 yields 10 ATP per acetyl-CoA, or roughly eight ATP per extracellular acetate, minus any potential ATP costs for maintaining intracellular pH from improting a proton along with each acetate (see **Section 2.2.6** and **Section 2.3.2**) .

While overexpression of native genes is a commonly used tool to increase the flux through a target pathway, heterologous expression of genes can be used to create new pathways in a host organism that did not previously exist. In SCO processes in *Yarrowia*, heterologous expression has been used successfully in constructing pathways for synthesis of eicosapentaenoic acid (EPA) [20] and for metabolism of xylose to form lipids [21]. In this study, we attempt to improve the yield and productivity of lipids in high-oil *Yarrowia* cells growing on acetate by expressing heterologous genes to create more efficient alternative pathways to generate the precursors necessary for fatty acid synthesis. These new pathways are highlighted in **Figure 7.1**. Our two areas of focus include: 1) reducing the ATP cost to convert acetate to acetyl-CoA, and 2) providing a shorter metabolic path to produce NADPH from acetyl-CoA. We attempt to achieve the first goal by expressing the ADP-forming acetyl-CoA synthetase (bpACS) gene from *Bordetella pertussis*, which could potentially compete with the AMP-forming native Acs enzyme. In an attempt to achieve the second goal, we express the known cytosolic, NADP+-specific isocitrate dehydrogenase gene (scIDP2) along with a mitochondrial citrate:oxoglutarate transporter gene (scYHM2), both from *Saccharomyces cerevisiae* [22]; for this

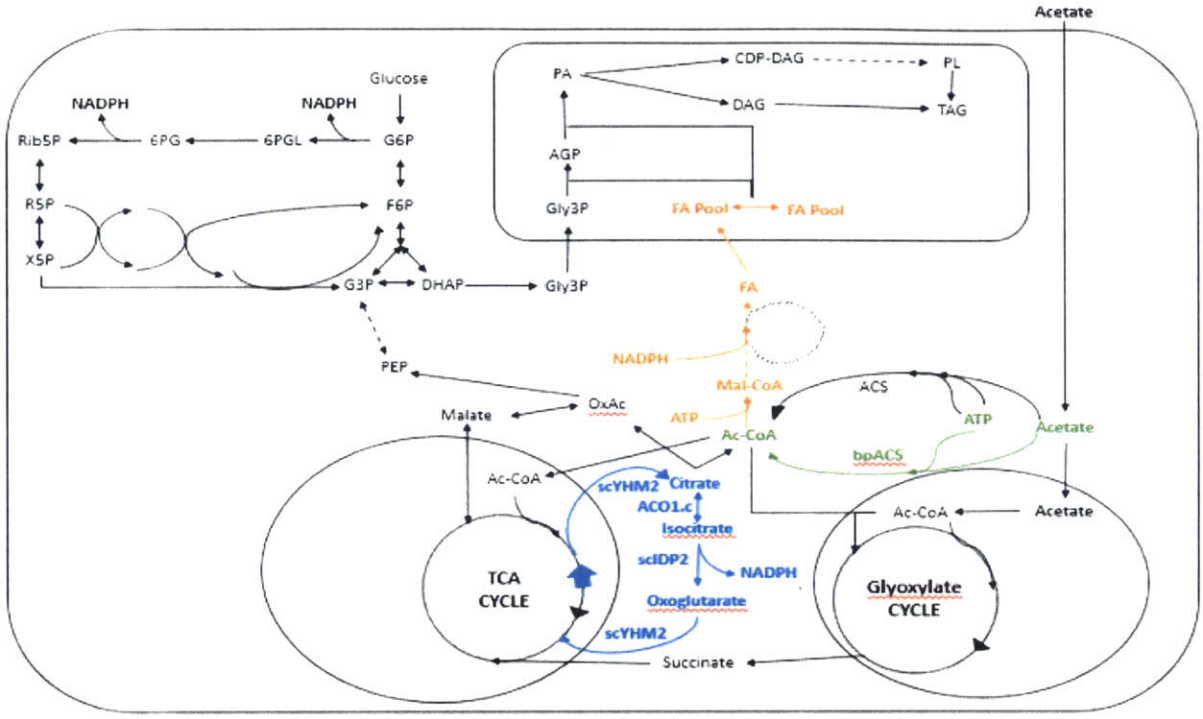

Figure 7.1: Major metabolic pathways for acetate during lipogenesis. For the purposes of lipid production, acetyl-CoA initially gets distributed among three major pathways: fatty acid synthesis in the cytosol (highlighted in orange), the glyoxylate cycle in the peroxisomes, and the TCA cycle in the mitochondria. The heterologous bpACS pathway is highlighted in green, and the heterologous scIDP2/scYHM2 pathway is highlighted in blue.

pathway, acetyl-CoA would be transported into the mitochondrion first, condensed with oxaloacetate to form citrate, and then the citrate would be exported out of the mitochondrion, aconitase will convert it to isocitrate, and IDP2 will produce a cytosolic NADPH (instead of mitochondrial NADH) and oxoglutarate, which will get transported back into the mitochondrion to continue with the TCA cycle. In addition to this pathway for acetyl-CoA, a cycle could be created in which mitochondrial citrate is exported into the cytosol by scYHM2, oxidized to form cytosolic NADPH, and then the resulting oxoglutarate is shuttled back into the mitochondrion by scYHM2 to be reductively carboxylated back into citrate at the cost of a mitochondrial NADH [22], and most likely, also an ATP, as reductive carboxylations generally require it [17]. We hypothesize that this pathway will improve lipid productivity by eliminating the requirement for

acetyl-CoA to traverse through several pathways in the cell's metabolic network to generate cytosolic NADPH.

We first demonstrate through pathway yield calculations that in conserving energy using the bpACS pathway, increases in the theoretical yield of lipids on acetate are significant; we then demonstrate that in using the scIDP2/scYHM2 pathway, the reduction in theoretical yield is minimal, and we hypothesize that the shorter path length will still allow for overall increases in productivity. Next, we express these two pathways in a *Yarrowia* strain previously engineered to overproduce lipids by overexpressing the native ACC1 and diacylglycerol acyltransferase (DGA2) genes, and grow them in shake flask fermentations on acetate to determine if expression of these genes is effective at increasing lipid yield or productivity. Our results indicate that the scIDP2/scYHM2 pathway may potentially increase productivity by increasing growth rate, but the NADPH generated does not directly contribute to lipogenesis.

7.2 Materials and Methods

7.2.1 Pathway yield calculations

In order to determine the maximum potential improvement or reduction to the yield of lipids on acetate for our heterologous pathways, we performed pathway yield calculations comparing the closed carbon, reducing equivalent, and energy balances in the newly constructed metabolic networks and compared their theoretical yields to those of the native metabolic network. For all native reactions and pathways, the carbon species stoichiometries and cofactor utilizations were obtained from [23], except for acetyl-CoA synthetase, which was determined from [24]. The condensed set of reactions for synthesis of 1 mol of lipids (in the form of glyceryl tripalmitate), the NADPH required (from the PPP), and the makeup energy from the TCA cycle from acetate is given in **Table 7.1**.

Table 7.1: System of equations for native production of palmitate, and required NADPH and ATP from acetate.

Eqn. No.	Rxn Product	G6P	Ac-CoA	ATP	NADH	NADPH	FADH2	Succ	OxAc	PEP	Gly3P	Palm
1	Overall TAG-Palmitate	0	-24	-21	0	-42	0	0	0	0	-1	3
2	Pentose Phosphate Pathway Cycle	-1	0	0	0	12	0	0	0	0	0	0
	Glyoxylate Cycle	0	-2	0	1	0	0	1	0	0	0	0
	Succinate to OxAc.c	0	0	0	1	0	1	-1	1	0	0	0
	PEPCK	0	0	-1	0	0	0	0	-1	1	0	0
3	Ac-CoA to PEP (Sum of Previous 3 lines)	0	-2	-1	2	0	1	0	0	1	0	0
4	Gluconeogenesis	1	0	-2	-2	0	0	0	0	-2	0	0
5	Glyceroneogenesis	0	0	-1	-2	0	0	0	0	-1	1	0
6	Acetate Activation	0	1	-2	0	0	0	0	0	0	0	0
7	TCA Cycle - Electron Transport	0	-1	10	0	0	0	0	0	0	0	0

The theoretical yield of palmitate on acetate is determined by first balancing the system of equations by adding Eqn. 1 (using the basis of 3 mols of palmitate per 1 mol glyceryl tripalmitate) to the correct linear combination of Eqns. 2-5 to cancel out everything except the "Ac-CoA," "ATP," "NADH," "FADH2," and "Palm" columns. Leftover NADH and FADH2 is then converted to ATP: 2.5 ATP per NADH, and 1.5 ATP per $FADH_2$. The cost to import and activate the acetate to Ac-CoA is 2 ATP per Ac-CoA needed (Eqn. 6), as described previously in **Section 2.3.2**. The ATP deficit run by the cell is then found by finding the ATP value of the NADH, FADH2, and Ac-CoA generated or required and adding these values to that left over in the ATP column. This energy deficit is made up using additional acetate at a rate of 8 ATP per Ac-CoA needed (from Eqns. 6 and 7). This additional Ac-CoA is added to the amount needed after the first step. Using a molecular weight of 59.04 g/mol acetate needed and a molecular weight of 255.42 g/mol palmitate and dividing the mass of palmitate by the mass of acetate gives the theoretical maximum yield.

For the case of using the bpACS heterologous pathway, Eqn. 6 is altered so that cost of acetate activation is 1 ATP per Ac-CoA needed. The case of using the scIDP2/scYHM2 enzymes to generate NADPH requires replacing the gluconeogenesis and PPP cycle equations with the heterologous pathway, and also requires keeping track of cytosolic and mitochondrial NADH separately. The condensed set of equations needed to find to theoretical maximum yield of palmitate on acetate using the scIDP2/scYHM2 pathway to make NADPH is given in **Table 7.2.**

Table 7.2: System of equations for production of palmitate from acetate, using scIDP2/scYHM2 pathway for NADPH generation.

Rxn No.	Rxn Product	AcCoA	ATP	NADH.c	NADH.m	NADPH	FADH2	Succ	OxAc	PEP	Gly3P	Palm
1	Overall TAG-Palmitate	-24	-21	0	0	-42	0	0	0	0	-1	3
2	Cit-Oxoglutarate IDP2/YHM2 Cycle	0	-1	0	-1	1	0	0	0	0	0	0
3	Ac-CoA-IDP2-YHM2/TCA	-1	1	0	2	1	1	0	0	0	0	0
	Glyoxylate cycle	-2	0	1	0	0	0	1	0	0	0	0
	Succinate to OxAc.c	0	0	1	0	0	1	-1	1	0	0	0
	PEPCK	0	-1	0	0	0	0	0	-1	1	0	0
4	Ac-CoA to PEP (Sum of Previous 3 lines)	-2	-1	2	0	0	1	0	0	1	0	0
5	Glyceroneogenesis	0	-1	-2	0	0	0	0	0	-1	1	0
6	Acetate Activation	1	-2	0	0	0	0	0	0	0	0	0
7	TCA Cycle - Electron Transport	-1	10	0	0	0	0	0	0	0	0	0

For this case, there are two reactions (Eqns. 2 and 3) that can provide NADPH for fatty acid synthesis: Eqn. 2 does not require Ac-CoA input, but causes a deficit in mitochondrial NADH, while Eqn. 3 does require Ac-CoA input, but can make up the deficit in mitochondrial NADH twice as efficiently as it can generate NADPH. Therefore, the correct linear combination of Eqns. 2-5 to add to Eqn.1 to minimize usage of Ac-CoA involves using a 2:1 ratio of Eqn. 2 to

Eqn. 3. The ATP conversions, ATP deficit, and additional acetate needed to make up the deficit are then calculated the same way as before.

7.2.2 Plasmid and *Yarrowia* strain construction

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). Standard growth conditions for *E. coli* for the purposes of plasmid construction and generation have been previously described [25].

Plasmid pMT015, described previously [9], 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, along with a gene coding for the native Leu2 enzyme as an auxotrophic marker. Plasmid pMT065, containing the native ACC1 gene under control of the hp4d promoter and the native DGA2 gene under the full TEF-1α promoter (with intron), along with the LEU2 marker, was described previously [9]. The plasmid p91T contains a URA3 marker and an insertion site for gene coding sequences flanked by the full TEF-1α promoter upstream and by the XPR2 terminator downstream, all between two sites homologous to regions in the *Yarrowia* genome around the extracellular lipase LIP2 gene; this plasmid was also described previously (see **Chapter 4**).

Table 7.3: Primers, plasmids, and *Yarrowia* strains used in this study.

Primers		
Primer Name	**Sequence [Underline = Binding Segment]**	**Amplicon**
rAS442	agcacttttgcagtactaacggcagGACAACGCCTCCGTTTCGC	bpACS coding sequence - gBlock 1
rAS443	GGGCTTGGTTGTAGGCCTGC	bpACS coding sequence - gBlock 1
rAS444	CTATAGTGCAGGCCTACAACCAAGCC	bpACS coding sequence - gBlock 2
rAS445	gaggtacctcaagatgTTAACCGGGTCCTAACATAGCGG	bpACS coding sequence - gBlock 2
rAS482	agcacttttgcagtactaaccgcagACCAAGATCAAGGTTGCAAACCCC	scIDP2 coding sequence
rAS483	gaggtacctcaagatgTTAGAGGGCGGCGGCTTCGAA	scIDP2 coding sequence
rAS484	agcacttttgcagtactaaccgcagCCTTCTACTACTAACACTGCTGCTG	scYHM2 coding sequence
rAS485	gaggtacctcaagatgTTAGTGTTTAGCAACGGGGGTCTC	scYHM2 coding sequence
rAS134	CTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGAT	Segment of p15K1
rAS135	CTACGTCTTGCTGGCGTTCGTTCGGACACGGGCATCTCAC	Segment of p15K1
rAS136	GTGAGATGCCCGTGTCCGAACGAACGCCAGCAAGACGTAG	Segment of p15K2
rAS137	ATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAG	Segment of p15K2
rAS378	agcacttttgcagtactaaccgcagGGCCTCAGCGTGCCCGCT	Native ACO1 coding sequence
rAS379	gaggtacctcaagatgTTACTTGTTGGCAGCCTTCTGGGC	Native ACO1 coding sequence
rAS452	cctgagtcatcAATTCCATGTGTAACACTCGCTCTGGAG	Segment of JMP62-HYG
rAS453	cccattatgCTAGGGATAACAGGGTAATTATCGCTTCGG	Segment of JMP62-HYG
rAS454	accctgttatccctagCATAATGGGGAAGGCCATCCAGC	Segment of p15I9
rAS455	agtgttacacatggaattGATGACTCAGGCGACGACGGAATTC	Segment of p15I9

Plasmids			
Name	**Gene(s) [All with native XPR2 Terminator]**	**Yeast Marker**	**Source**
pMT015	None	LEU2	[9]
pMT065	hp4d-ACC1, TEF1α-DGA2	LEU2	[9]
p91T	None	URA3	Chapter 4
p91J6	TEF1α-bpACS	URA3	This work
p15K1	TEF1α-scIDP2	LEU2	This work
p15K2	TEF1α-scYHM2	LEU2	This work
p15K12	TEF1α-scIDP2, TEF1α-scYHM2	LEU2	This work
p91K12	TEF1α-scIDP2, TEF1α-scYHM2	URA3	This work
JMP62-HYG	None	HYG	Nicaud Group, INRA
p15I9	TEF1α-ACO1.c	LEU2	This work
pJHI9	TEF1α-ACO1.c	HYG	This work
***Yarrowia* Strains**			
Name	**Genes (Over)expressed**	**Auxotrophy/Resistance**	**Source**
po1z	None	Leu-, Ura-	[21], Chapter 4
po165	(ACC1, DGA2)	Ura-	[21], Chapter 4
ADC	(ACC1, DGA2)	None	Chapter 4
YLASJ6	(ACC1, DGA2), bpACS	None	This work
YLASK12	(ACC1, DGA2), scIDP2, scYHM2	None	This work
YLASK12I9	(ACC1, DGA2), scIDP2, scYHM2, ACO1.c	Hyg+	This work

Plasmids, primers, and *Yarrowia* strains used in this study are listed in **Table 7.3**. Coding sequences for the bpACS, scIDP2, and scYHM2 genes were amplified by PCR from chemically synthesized gBlocks Gene Fragments (Integrated DNA Technologies, Coralville, IA). Sequences synthesized into gBlocks were codon optimized for *Yarrowia lipolytica* by inputting the protein sequence obtained from the National Center for Biotechnology Information website (http://www.ncbi.nlm.nih.gov) into the Codon Optimization online tool from Blue Heron Biotechnology (http://www.blueheronbio.com); for all sequence outputs, the "expr_opt" option was chosen for synthesis. For constructing the expression plasmid for bpACS, two amplified gBlocks (one of 1198 bp, the other 998 bp, with a 57 bp overlap between the two sequences to make the 2139 bp coding sequence) were joined with a PCR-amplified p91T using the Gibson Assembly method and Cloning Kit (NEB) [26].

For constructing the expression plasmid for scIDP2 and scYHM2, each gBlock (one for each coding sequence, 1233 bp and 939 bp, respectively) was amplified and individually joined with a PCR-amplified pMT015 vector by Gibson Assembly to create plasmid p15K1 and p15K2, respectively. A fragment of p15K1 containing part of the AmpR marker, the p15_ori origin of replication, and the scIDP2 gene was PCR amplified by with primers rAS134 and rAS135, and joined with a fragment of p15K2, amplified using primers rAS136 and rAS137 and containing the scYHM2 gene, the LEU2 marker, and the other part of the AmpR marker, to form plasmid p15K12. The fragment containing both coding sequences was then amplified using primers rAS482 and rAS485, and joined with an amplified p91T to form p91K12.

For constructing plasmid pJHI9, overexpressing the native aconitase ACO1 gene without the mitochondrial targeting tag (determined by MitoProt, and hence "targeting" the enzyme to the cytosol), the coding sequence of ACO1 (sans the N-terminal MLS tag) was amplified from

Yarrowia lipolytica genomic DNA using primers rAS378 and rAS379, and joined with a PCR-amplified pMT015 to form p15I9. The ACO1.c gene (with promoter and terminator) was then amplified using primers rAS454 and rAS455, and joined with a fragment from plasmid JMP62-HYG amplified with rAS452 and rAS453, to form plasmid pJH19. All plasmids constructed for this study were verified by PCR of the purified plasmid; p91J6, p15K1, p15K2, and p15I9 were also confirmed by plasmid sequencing.

To construct *Yarrowia lipolytica* strains YLASJ6 and YLASK12, po165 was transformed with p91J6 and p91K12, respectively, and recovered on Leu- and Ura- dropout YSM agar plates. To construct strain YLASK12I9, strain YLASK12 was transformed with pJHI9 and recovered on a Leu- and Ura- dropout YSM agar plate supplemented with 200 mg/L Hygromycin B (Sigma-Aldrich, St. Louis, MO). Strains were verified by PCR of transformant colony gDNA.

7.2.3 Fermentation conditions and lipid quantification analysis

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 acetate (YSMa) used for fermentations contains 1.7 g/L Yeast Nitrogen Base (YNB, without amino acids or ammonium sulfate) (Amresco, Solon, OH), 50 g/L anhydrous sodium acetate, and 2.02 g/L ammonium sulfate (Sigma) for an atomic C:N ratio of 40.

For fermentations, strains are inoculated into 3 mL YPA medium in a 14 mL polypropylene tube with a vented cap (VWR, Radnor, PA) and shaken for 28 hours at 200 rpm and 30°C for the first pre-culture. Yeast from the first pre-culture was then transferred into 250 mL shake flasks containing 20 mL YSMa to achieve an initial optical density (OD_{600}) of 0.1 and shaken for 39 hours at 250 rpm and 30°C for the second pre-culture. Yeast from the second pre-culture was

then centrifuged and washed once with YSMa medium and inoculated into 250 mL shake flasks containing 50 mL of YSMa to an initial OD_{600} of 0.05. Flasks were closed with 34mm-45mm foam plugs (Jaece Industries Inc., North Tonawanda, NY) and placed in an Innova 43 Incubator Shaker (New Brunswick Scientific, Edison, NJ) and shaken for 152 hours (for acetate fermentations) at 250 rpm and 30°C. 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 samples are taken gravimetrically by passing 2 mL through a pre-weighed cellulose acetate filter paper (Whatman, GE Healthcare Biosciences, Piscataway, NJ); the filter paper is then washed with MilliQ water and dried in a 60°C oven overnight. Two cell pellets (technical replicates) from 500 mL of culture per flask were used for lipid quantification. Three biological replicate flasks were used for each strain.

Derivatization of fatty acids from lipids into fatty acid methyl esters (FAMEs) was achieved with a modified form of the direct transesterification method [27]. Samples of 500 uL from acetate fermentations were taken at the end time point (152 hours) and 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.

7.3 Results and Discussion

7.3.1 Theoretical yield improvement for constructed strains

In order to demonstrate the potential utility of the heterologous pathways we built into *Yarrowia*, we performed pathway calculations for the theoretical maximum yield of palmitate on acetate. Using the equations and methods detailed in **Section 7.2.2**, we first calculated the theoretical yield for the native pathways from wild-type *Yarrowia*. The linear combination of Eqns. 2-5 and resulting energy deficit, makeup acetate, and theoretical yield for the native pathways are given in **Table 7.4**.

As shown in **Table 7.4**, the theoretical yield of palmitate on acetate in wild-type *Yarrowia lipolytica* is 0.255 g palmitate/g acetate. This calculation represents the most direct path from acetate to palmitate (stored in TAG form as glyceryl tripalmitate). The pentose phosphate

pathway in this model is the sole source of NADPH, as determined by process of elimination using evidence suggesting the lack of common alternatives for NADPH generation in other organisms. To produce the ATP needed for lipid synthesis and cover the energy cost of the acetate needed to make NADPH, the most efficient pathway, the TCA cycle, is used.

Table 7.4: Theoretical maximum yield of palmitate on acetate, calculated using native pathways in *Yarrowia lipolytica*.

Rxn Product	G6P	Ac-CoA	ATP	NADH	NADPH	FADH2	Succ	OxAc	PEP	Gly3P	Palm
Basis	0	-24	-21	0	-42	0	0	0	0	-1	3
Glyceroneogenesis	0	0	-1	-2	0	0	0	0	-1	1	0
Pentose Phosphate Pathway Cycle	-3.5	0	0	0	42	0	0	0	0	0	0
Gluconeogenesis	3.5	0	-7	-7	0	0	0	0	-7	0	0
Ac-CoA to PEP (Sum of Previous 3 lines)	0	-16	-8	16	0	8	0	0	8	0	0
Total:	0	-40	-37	7	0	8	0	0	0	0	3
ATP Value:		-80	-37	17.5		12					
		-87.5	ATP Deficit			50.9375	Ac for 3 Palmitate				
		10.9375	More Ac Needed				MW	Mass	Theoretical Yield		
						Ac	59.04	3007.35			
						Palm	255.42	766.26	0.255		

In an attempt to reduce the energy cost associated with the process of lipid production (as well as other biomass or byproduct formation requiring the use of the acetate carbon substrate as an input), we engineered a *Yarrowia* strain to express a gene coding for the ADP-forming acetyl-CoA synthetase gene from *Bordetella pertussis*. This enzyme can activate acetate to acetyl-CoA at a cost of one mol ATP per mol of acetate, whereas the native *Yarrowia* acetyl-CoA synthetase enzyme is AMP-forming, which due to the equilibrium of AMP with ATP and ADP, effectively costs two mol ATP per mol acetate. Because producing lipids is an energy intensive process, by saving energy on all acetate activation, we hypothesized significantly fewer moles of acetate

would be needed to be oxidized in the TCA cycle, leading to an improvement in the theoretical maximum yield of palmitate on acetate. Also by decreasing the cost of activating the sole carbon source in our fermentations, we hypothesized that the likely resulting higher cellular energy charge could potentially improve growth rate and productivity. The results of the energy deficit analysis and yield calculations are given in **Table 7.5**. **Table 7.5** shows that the theoretical maximum yield exclusively using the bpACS pathway to activate acetate is 0.287 g palmitate/g acetate, a 12.5% increase over the native metabolism.

Table 7.5: Theoretical maximum yield of palmitate on acetate, calculated using bpACS pathway for acetate activation.

Rxn Product	G6P	Ac-CoA	ATP	NADH	NADPH	FADH2	Succ	OxAc	PEP	Gly3P	Palm
Basis	0	-24	-21	0	-42	0	0	0	0	-1	3
Glyceroneogenesis	0	0	-1	-2	0	0	0	0	-1	1	0
Pentose Phosphate Pathway Cycle	-3.5	0	0	0	42	0	0	0	0	0	0
Gluconeogenesis	3.5	0	-7	-7	0	0	0	0	-7	0	0
Ac-CoA to PEP (Sum of Previous 3 lines)	0	-16	-8	16	0	8	0	0	8	0	0
Total:	0	-40	-37	7	0	8	0	0	0	0	3
ATP Value:		-40	-37	17.5		12					
		-47.5	ATP Deficit			45.278	Ac for 3 Palmitate				
		-5.278	More Ac Needed				MW	Mass	Theoretical Yield		
						Ac	59.04	2673			
						Palm	255.42	766.3	0.287		

As the metabolic pathway length from acetate to the NADPH-generating pentose phosphate pathway is large, we thought to engineer *Yarrowia* with an alternative pathway that includes an NADPH-generating cycle that does not require input of acetyl-CoA by heterologous expression of genes coding for the NADP+-dependent cytosolic isocitrate dehydrogase Idp2 and the citrate:oxoglutarate mitochondrial transporter Yhm2 from *Saccharomyces cerevisiae*. While we

hypothesized that the removal of the requirement for such a large path length should improve the productivity of NADPH in the cell, we also performed a theoretical maximum yield calculation to determine if yield could also be improved using this cycle, or if the theoretical yield is reduced, whether this may potentially outweigh the benefits of the hypothesized productivity increase. Although this cycle requires the input of mitochondrial NADH (and most likely ATP, and this scenario is reflected in our system of equations), this can be more efficiently produced from acetyl-CoA than NADPH. The theoretical maximum yield calculations are shown in **Table 7.6**. Using the scIDP2/scYHM2 as the sole source of cytosolic NADPH for lipid synthesis leads to a very slight theoretical yield decrease to 0.250 g palmitate/g acetate, a 2.0% decrease. We therefore have confirmed that the bpACS pathway is advantageous for conserving acetate during lipid production on the basis of theoretical yield, and that the scIDP2/scYHM2 cycle's negative effect on theoretical yield is miniscule and not likely to significantly impact the overall process when compared to its potential to increase lipid productivity.

Table 7.6: Theoretical maximum yield of palmitate on acetate, calculated using scIDP2/scYHM2 pathway for NADPH generation.

Rxn Product	Ac-CoA	ATP	NADH.c	NADH.m	NADPH	FADH2	Succ	OxAc	PEP	Gly3P	Palm
Basis	-24	-21	0	0	-42	0	0	0	0	-1	3
Glyceroneogenesis	0	-1	-2	0	0	0	0	0	-1	1	0
Cit-Oxog IDP2/YHM2	0	-28	0	-28	28	0	0	0	0	0	0
Ac-CoA-IDP2-YHM2/TCA	-14	14	0	28	14	14	0	0	0	0	0
Ac-CoA to PEP (Sum of Previous 3 lines)	-2	-1	2	0	0	1	0	0	1	0	0
Total:	-40	-37	0	0	0	15	0	0	0	0	3
ATP Value:	-80	-37	0			22.5					
		-94.5	ATP Deficit			51.81	Ac for 3 Palmitate				
		-11.81	More Ac Needed				MW	Mass	Theoretical Yield		
						Ac	59.04	3059			
						Palm	255.42	766	0.250		

7.3.2 Fermentation performance of constructed strains

The calculations from the previous section confirmed that our bpACS heterologous pathway could increase the overall yield of lipids of our acetate fermentations significantly by reducing the ATP cost of activating the acetate to acetyl-CoA, while our scIDP2/scYHM2 pathway, meant to boost productivity, brings with it only very minor reductions in theoretical yield. We also anticipated that, due to the likely increase in cellular energy charge using bpAcs and the shorter pathway length for generating NADPH from acetyl-CoA using scIdpP2/scYhm2, that both heterologous pathways could increase growth rate and lipid productivity in *Yarrowia*. We therefore constructed *Yarrowia* strains that expressed the genes coding for their respective enzymes, transforming each set of pathway genes in a high-oil producing background, a strain that overexpresses the native ACC1 and DGA2 genes [9].

We also anticipated that the scIdp2/scYhm2 pathway construct may not function as expected due to the cycle requiring the conversion of the citrate exported from the mitochondrion by scYhm2 into isocitrate in the cytosol so that it could be acted on by scIdp2. Due to yeast aconitase likely being localized in mitochondria and glyoxysomes [28] only, we attempted to pre-emptively troubleshoot this potential problem by constructing an additional strain in the same ACC1- and DGA2- overexpressing background that also expresses scIDP2, scYHM2, and a modified form of the native ACO1 gene in which the MTS (mitochondrial targeting sequence) from the N-terminus of the coding sequence as predicted by MitoProt was removed. The modified aconitase gene, here called ACO1.c, should be able to convert citrate to isocitrate in the cytosol, thus bridging the one node gap between the heterologous pathway enzymes.

To test for any increase in lipid yield and productivity from these pathways, we performed shake flask fermentations comparing the lipid productivity, content (grams lipids per gram dry

cell weight), and yield for four *Yarrowia* strains on minimal media using acetate as a sole carbon source: 1) strain ADC, our control strain for this experiment, overexpresses the native ACC1 and DGA2 genes, 2) strain YLASJ6 overexpresses ACC1 and DGA2 and expresses the bpACS gene under the full TEF1α promoter [21] (see Chapter 3), 3) strain YLASK12 overexpresses ACC1 and DGA2 and expresses the scIDP2 and scYHM2 genes under the TEF1α promoter, and lastly, 4) strain YLASK12I9, builds upon strain YLASK12 by also including the aforementioned ACO1.c gene under the TEF1α promoter. The lipid titer/productivity, content, and yield on acetate results of the shake flask fermentation are shown in **Figure 7.2**, and dry cell weights accumulated by these strains in this fermentation are shown in **Table 7.7**.

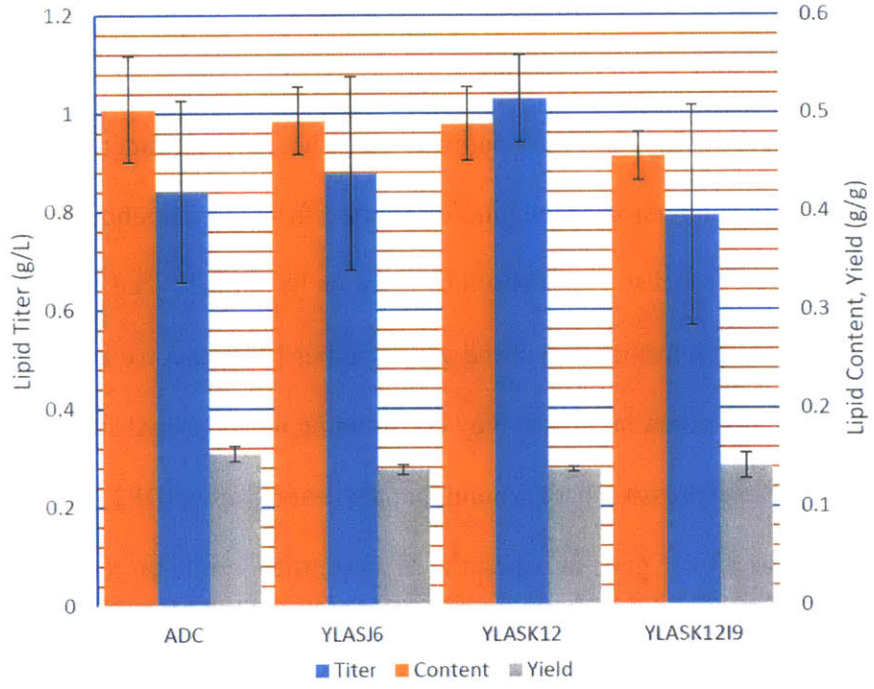

Figure 7.2: Fermentation performance of heterologous pathway expressing strains on minimal media containing 50 g/L sodium acetate and a C:N atomic ratio of 40 at 152 hours. Error bars represent the standard deviations from three biological replicates.

The results of the fermentation experiment show a lack of improvement in all areas from the newly constructed strains except for a very small, statistically insignificant increase in lipid titer/productivity from strain YLASJ6 (bpACS pathway) and a larger increase in titer/productivity of 22.5% from strain YLASK12 compared to the ADC control strain. The titer increase seen in YLASK12 is statistically significant to $\alpha = 0.1$, but not to $\alpha = 0.05$ ($p = 0.0934$). It is possible that an experiment using more biological replicates could give a result that is more significant to confirm the higher lipid productivity demonstrated by this strain thus far. The slight decrease in lipid content and yield in this strain, along with an increase in biomass (and hence, growth rate), suggests that any real increase in cellular NADPH availability by this strain was not directed specifically towards lipid biosynthesis, but was more favorably distributed towards other NADPH-consuming pathways (such as amino acid biosynthesis). It is also clear from our results that adding the ACO1.c gene to our scIDP2/scYHM2 pathway construct did nothing to facilitate its mitochondrial NADH to cytosolic NADPH cycle.

Table 7.7: Dry cell weight (DCW) measurements from heterologous pathway strains in acetate fermentation.

Strain	Avg. DCW (g/L)	St. Dev. DCW (g/L)
ADC	1.698	0.492
YLASJ6	1.792	0.453
YLASK12	2.103	0.054
YLASK12I9	1.720	0.413

While the pathways we tested would theoretically increase the maximum yield of lipids on acetate and/or the productivity of lipids on acetate, the fraction of these increases actually realized in the process depends strongly on the fraction of carbon flux through the heterologous pathway as opposed to the native pathway. Knocking out competing pathways has been

previously shown to be effective in increasing the flux of a common substrate through a desired heterologous pathway [29, 30], however, in the case of our specific pathways, this does not seem feasible. The native Acs enzyme in Yarrowia, like the AMP-forming Acs isoforms from *Saccharomyces cerevisiae* [31], may not just be involved in activation of acetate for metabolism, but also for providing acetyl-CoA in the nucleus for use in histone acetylation, allowing for the regulation of gene expression and cell cycle progression. It is also possible that the secondary reaction of hydrolyzing the pyrophosphate produced from AMP formation by the native Acs is necessary to increase the negative value of the free energy from the coupled reaction system and prevent reversibility of this reaction, which may be dependent on the relative intracellular concentrations of acetate and acetyl-CoA. For the case of the scIDP2/scYHM2 pathway, the native NADPH producing pathway for lipid synthesis on acetate is almost certainly the oxidative pentose phosphate pathway [16, 18]; knocking out a part of this pathway is likely a lethal mutation, as this is the main pathway by which precursors for nucleotides and some enzyme cofactors are synthesized [23]. However, the amount of acetate consumed and total biomass generated by all strains in this fermentation was very low. The YLASK12 strain generated the most biomass at 2.103 g/L. Using the empirical formula for average yeast biomass [32] for the non-lipid biomass in this strain, and assuming all nitrogen introduced into the cell as ammonium remains in the cell and is converted to biomass, it may be that there was a significant amount of nitrogen left in the media at the end; therefore, it may be that more significant productivity increases could be realized in this strain if the nitrogen in the media is lowered or the fermentation is allowed to run longer due to the amino acid biosynthesis pathways no longer being used after nitrogen depletion, in this way removing the competing pathways for the product of our heterologous pathway (NADPH) rather than the substrate acetyl-CoA.

7.4 Conclusions

Acetate is an important carbon substrate to consider when designing microbial biocatalysis processes for synthesizing low profit-margin commodity chemicals, as a large component of some agricultural, industrial, and municipal waste streams obtainable at low or negative cost. However, the use of an alternative substrate to the cellulose-derived hexose sugars traditionally used for these processes presents new challenges in engineering host strains for optimized carbon yield and productivity. In this study, we attempt to improve the lipid yield and productivity on acetate in an SCO process by introducing heterologous metabolic pathways into a strain of *Yarrowia lipolytica* previously engineered to achieve much higher than normal lipid production. Although our ATP-saving acetyl-CoA synthetase, bpACS, was calculated to improve yields if used significantly compared to the native pathway it replaced, this theoretical improvement was not realized in our fermentation experiment. We do have some evidence that our scIDP2/scYHM2 pathway does work to generate significant amounts of additional NADPH, however future experiments will need to lower the nitrogen content in the fermentation media in order to prevent this NADPH from being utilized for non-lipid biomass generation and should also include more biological replicates in order to more conclusively prove a statistically significant increase in the overall lipid productivity conferred by this pathway.

References:

1) Stephanopoulos, G. (2007). Challenges in engineering microbes for biofuels production. *Science*, (February), 801–805.
2) Hahn-Hägerdal, B., Galbe, M., Gorwa-Grauslund, M. F., Lidén, G., & Zacchi, G. (2006). Bio-ethanol - the fuel of tomorrow from the residues of today. *Trends in Biotechnology*, *24*(12), 549–556. doi:10.1016/j.tibtech.2006.10.004
3) Angenent, L. T., Karim, K., Al-Dahhan, M. H., Wrenn, B. a., & Domíguez-Espinosa, R. (2004). Production of bioenergy and biochemicals from industrial and agricultural wastewater. *Trends in Biotechnology*, *22*(9), 477–485. doi:10.1016/j.tibtech.2004.07.001
4) 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 Technology*, *114*, 443–9. doi:10.1016/j.biortech.2012.02.091
5) Fickers, P., Benetti, P.-H., Waché, Y., Marty, a, Mauersberger, S., Smit, M. S., & Nicaud, J.-M. (2005). Hydrophobic substrate utilisation by the yeast Yarrowia lipolytica, and its potential applications. *FEMS Yeast Research*, *5*(6-7), 527–43. doi:10.1016/j.femsyr.2004.09.004
6) Dujon, B., Sherman, D., Fischer, G., Durrens, P., Casaregola, S., Lafontaine, I., ... Souciet, J.-L. (2004). Genome evolution in yeasts. *Nature*, *430*(6995), 35–44. doi:10.1038/nature02579
7) Madzak, C., Tréton, B., & Blanchin-Roland, S. (2000). Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous proteins in the yeast Yarrowia lipolytica. *Journal of Molecular Microbiology and Biotechnology*, *2*(2), 207–216.
8) Blazeck, J., Liu, L., Redden, H., & Alper, H. (2011). Tuning gene expression in yarrowia lipolytica by a hybrid promoter approach. *Applied and Environmental Microbiology*, *77*(22), 7905–7914. doi:10.1128/AEM.05763-11
9) Tai, M., & Stephanopoulos, G. (2013). Engineering the push and pull of lipid biosynthesis in oleaginous yeast Yarrowia lipolytica for biofuel production. *Metabolic Engineering*, *15*, 1–9. doi:10.1016/j.ymben.2012.08.007
10) 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
11) Casal, M., Cardoso, H., & Leao, C. (1996). Mechanisms regulating the transport of acetic acid in Saccharomyces cerevisiae. *Microbiology*, (142), 1385–1390.
12) Kujau, M., H. Weber, and G. Barth. 1992. Characterization of mutants of the yeast Yarrowia lipolytica defective in acetyl-coenzyme A synthetase. Yeast 8:193-203.
13) Nielsen, J. (2014). Synthetic Biology for Engineering Acetyl Coenzyme A Metabolism in Yeast. *mBio*, *5*(6), 14–16. doi:10.1128/mBio.02153-14.Copyright
14) 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
15) 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. *Biotechnology Letters*, *35*(12), 2091–8. doi:10.1007/s10529-013-1302-7

16) Silverman, A. M.*, Qiao, K.*, Xu, P., and Stephanopoulos, G. Functional overexpression and characterization of lipogenesis-related genes in the oleaginous yeast *Yarrowia lipolytica*. Submitted.
17) Ratledge, C. (2014). The role of malic enzyme as the provider of NADPH in oleaginous microorganisms: A reappraisal and unsolved problems. *Biotechnology Letters*, *36*(8), 1557–1568. doi:10.1007/s10529-014-1532-3
18) 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. *Metabolic Engineering*, *30*, 27–39. doi:10.1016/j.ymben.2015.02.007
19) Hynes, M. J., Murray, S. L., Andrianopoulos, A., & Davis, M. a. (2011). Role of carnitine acetyltransferases in acetyl coenzyme a metabolism in Aspergillus nidulans. *Eukaryotic Cell*, *10*(4), 547–555. doi:10.1128/EC.00295-10
20) Xue, Z., Sharpe, P. L., Hong, S.-P., Yadav, N. S., Xie, D., Short, D. R., ... Zhu, Q. (2013). Production of omega-3 eicosapentaenoic acid by metabolic engineering of Yarrowia lipolytica. *Nature Biotechnology*, *31*(8), 734–40. doi:10.1038/nbt.2622
21) Tai, M. (2012). Metabolic Engineering of Oleaginous Yeast for the Production of Biofuels. Chapter 8. Ph.D. Thesis. Massachusetts Institute of Technology.
22) Castegna, A., Scarcia, P., Agrimi, G., Palmieri, L., Rottensteiner, H., Spera, I., ... Palmieri, F. (2010). Identification and functional characterization of a novel mitochondrial carrier for citrate and oxoglutarate in Saccharomyces cerevisiae. *The Journal of Biological Chemistry*, *285*(23), 17359–70. doi:10.1074/jbc.M109.097188
23) Nelson, D. L. & Cox, M. M. (2005). *Lehninger principles of biochemistry* (4th ed.). New York: W.H. Freeman.
24) Devirgilio, C., Burckert, N., Barth, G., Neuhaus, J. M., Boller, T., & Wiemken, a. (1992). Cloning and disruption of a gene required for growth on acetate but not on ethanol - the Acetyl-Coenzyme-A synthetase gene of saccharomyces cerevisiae. *Yeast*, *8 (12)*, 1043–1051.
25) Sambrook, J., & Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual. CSHL press, Cold Spring Harbor, NY, USA.
26) 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. *Nature Methods*, *6*(5), 343–345. doi:10.1038/nmeth.1318
27) 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.
28) McCammon, M. T., Veenhuis, M., Trapp, S. B., & Goodman, J. M. (1990). Association of glyoxylate and beta-oxidation enzymes with peroxisomes of Saccharomyces cerevisiae. *Journal of Bacteriology*, *172*(10), 5816–5827.
29) Atsumi, S., Cann, A. F., Connor, M. R., Shen, C. R., Smith, K. M., Brynildsen, M. P., ... Liao, J. C. (2008). Metabolic engineering of Escherichia coli for 1-butanol production. *Metabolic Engineering*, *10*(6), 305–311. doi:10.1016/j.ymben.2007.08.003
30) Soma, Y., Tsuruno, K., Wada, M., Yokota, A., & Hanai, T. (2014). Metabolic flux redirection from a central metabolic pathway toward a synthetic pathway using a metabolic toggle switch. *Metabolic Engineering*, *23*, 175–184. doi:10.1016/j.ymben.2014.02.008
31) Takahashi, H., McCaffery, J. M., Irizarry, R. a., & Boeke, J. D. (2006). Nucleocytosolic Acetyl-Coenzyme A Synthetase Is Required for Histone Acetylation and Global Transcription. *Molecular Cell*, *23*(2), 207–217. doi:10.1016/j.molcel.2006.05.040

32) Von Stockar, U., & Liu, J. S. (1999). Does microbial life always feed on negative entropy? Thermodynamic analysis of microbial growth. *Biochimica et Biophysica Acta - Bioenergetics*, *1412*(3), 191–211. doi:10.1016/S0005-2728(99)00065-1

Chapter 8

Summary, Conclusions, and Recommendations

8.1 Summary of this work

The primary goal of this work was to improve *de novo* lipid biosynthesis in *Yarrowia lipolytica* for use in single cell oil processes to produce biodiesel. As lipid synthesis is a complex phenomenon effectively requiring the concerted activities of an incredibly large and diverse set of proteins in the cell, we explored the effects of manipulating the activities of a subset of these proteins using the tools of metabolic engineering.

The panel of genes overexpressed in **Chapter 3** was a way of quickly screening for genes whose expression levels (near the range of values normal for wild-type *Yarrowia* cells) correlated positively with lipid productivity, content, or yield on glucose and acetate. With the previous development of a strain that overexpresses the native acetyl-CoA carboxylase (ACC1) and diacylglycerol acyltransferase (DGA2) genes to achieve a much improved lipid producing phenotype, we wanted to build upon this prior success by including other modifications that could act synergistically with these two to further improve this strain. While we identified many genes for which overexpression caused an increase in lipid production or lipogenic phenotype compared to a wild-type strain, three genes in particular stood out due to the magnitude of the effect of their overexpression (the acylglycerolphosphate acyltransferase SLC1) or because the most likely mechanism by which they improve lipogenesis by overexpression would not be redundant when combined with ACC1 and DGA2 (increasing availability of glycerol head groups with the glycerol-3-phosphate dehydrogenase GPD or NADPH with 6-phosphogluconolactonase SOL3). We constructed these three-gene overexpressing strains in **Chapter 4** and compared them to a strain that only overexpresses ACC1 and DGA2 in glucose and acetate fermentations; our results show that combining these two overexpressions with SLC1 overexpression may lead to a strain with higher lipid productivity on glucose and higher lipid

productivity, content, and yield on acetate, but more tests would need to be done to achieve higher levels of statistical proof of these increases.

In **Chapter 5**, we knock out the SNF1 gene in *Yarrowia*, coding for an Acc1-inhibiting protein kinase, and characterize the resulting strain. We observed an increase in lipid content and yield on both glucose and acetate in this strain, with the increase on acetate being much larger. We also attempted to make further improvements on this strain by overexpressing ACC1 and DGA2 in the ΔSNF1 background. We found that on acetate, overexpressing DGA2 in this background led to a highly lipogenic phenotype, with the strain achieving a lipid content of 76.7% and a lipid yield of 0.200 g lipids/g acetate.

In **Chapter 6**, we used differential expression analysis to identify lipase genes that could be responsible for TAG degradation in *Yarrowia* by determining which genes are more highly expressed during lipostatic phases than during growth and lipid production phases. Although the LIP1, LIP4, and LIP10 genes were all appreciably expressed in high lipid content cells after lipid production in a bioreactor reached its maximum point, and were not appreciably expressed during growth, the disruption of these three genes did not cause a more lipogenic phenotype in the resulting cells.

The work in **Chapter 7** is an attempt to improve lipid yield and productivity of *Yarrowia* on acetate through the expression of non-native pathways. We found that the expression of the cytosolic NADP+-specific isocitrate dehydrogenase and the mitochondrial citrate:oxoglutarate transporter from *Saccharomyces cerevisiae* for producing NADPH for fatty acid synthesis instead of using the oxidative pentose phosphate pathway potentially allows for higher lipid productivity in *Yarrowia* cells that overexpress the native ACC1 and DGA2 genes; however,

more experiments on this would need to be performed in order to more conclusively prove this result.

8.2 Highlights in strain construction from this work

Before the start of this work, the strain of *Yarrowia lipolytica* with the highest reported lipid production on glucose and acetate was the ACC1 and DGA2 overexpressing strain MTYL065 previously produced in our lab [1, 2]. For the case of glucose, there is some evidence to suggest that the inclusion of SLC1 overexpression along with these two modifications leads to a strain with increased productivity while maintaining similar levels of lipid content and yield. In our shake flask experiment detailed in **Chapter 4**, our AD6 strain produced a lipid titer, content, and yield of 5.63 g/L, 0.687 g lipids/g DCW, and 0.207 g lipids/g glucose, compared to the control ADC strain's titer, content, and yield of 4.03 g/L, 0.684 g lipids/g DCW, and 0.195 g lipids/g glucose. This could possibly represent a success in SCO strain development on glucose, however, statistical tests performed on our fermentation data reveal that the increase in lipid titer by this strain was not statistically significant despite its seemingly large magnitude.

There currently has not been much research performed in the area of using acetate for SCO processes outside of our lab, with only a few papers discussing fermentations with acetate included either as the sole carbon source, in a mixture of different carbon sources [3], or in a two-stage fermentation setup with glucose used to first develop biomass and acetate used to produce only lipids [4, 5]. These other studies use wild-type strains of a few oleaginous yeasts as a proof-of-concept that these organisms can metabolize acetate to produce lipids, but metabolic engineering for the purpose of optimizing a strain for improved lipid production on acetate has not been extensively explored. The strain MTYL065 has therefore been the state-of-the-art for metabolic engineering of an SCO strain in producing lipids from acetate since the start

of this work [2, 6]. **Table 8.1** consists of what may be an exhaustive list of published studies involving the use of oleaginous yeasts to produce lipids from acetate, and compares two strains from this work to the results of these previous studies in terms of the lipid contents and titers achieved by the strains in their reported fermentation setup. This table shows that our strain AD6 (see **Chapter 4**) has the highest reported lipid yield of any oleaginous yeast strain on acetate (0.207 g lipids/g acetate), while our strain YL53S (see **Chapter 5**) has the highest reported lipid content of any oleaginous yeast on acetate (0.767 g lipids/g DCW). It must be noted, however, in the case of AD6, the lipid yield increase compared to our control strain (which was effectively MTYL065 reconstructed for this experiment) was only statistically significant to the $\alpha = 0.1$ level.

Table 8.1: Comparison of fermentation performance on acetate of select strains constructed for this work with previous literature.

Author(s)	Year	Species	Modifications	Setup [Acetate/acetic acid is sole carbon source unless specified]	Lipid Content	Lipid Yield
Tai, M.	2012	*Yarrowia lipolytica*	hp4d-ACC1, TEF1α-DGA2	Batch bioreactor	62%	15.2%
Chakraborty, S.	2014	*Yarrowia lipolytica*	hp4d-ACC1, TEF1α-DGA2	Bioreactor with pH control using dilute feed of acetic acid	60%	16%
Fontanille, et al.	2012	*Yarrowia lipolytica*	None	Two-stage fed batch bioreactor: 1) glucose, 2) acetate	40.7%	16%
Christophe, et al.	2012	*Cryptococcus curvatus*	None	Two-stage fed batch bioreactor: 1) glucose, 2) acetate	50.89%	15%
Christophe, et al.	2012	*Cryptococcus curvatus*	None	Two-stage shake flask: 1) glucose, 2) acetate	47.30%	19%
Fei, et al.	2011	*Cryptococcus albidus*	None	Shake flask with VFAs (8:1:1 acetate:propionate:butyrate)	27.8%	16.7%
This work	2015	*Yarrowia lipolytica*	hp4d-ACC1, TEF1α-DGA2	Shake flask, C:N = 40, no CSM	61.1%	17.2%
This work	**2015**	***Yarrowia lipolytica***	**hp4d-ACC1, TEF1α-DGA2, TEF1α-SLC1**	**Shake flask, C:N = 40, no CSM**	**76.0%**	**20.7%**
This work	**2015**	***Yarrowia lipolytica***	**TEF1α-DGA2, ΔSNF1**	**Shake flask, C:N = 40, with CSM**	**76.7%**	**20.0%**

8.3 A note on some recurring general themes throughout this work

Previous efforts in our lab towards the metabolic engineering of *Yarrowia lipolytica* for enhanced lipid production resulted in the development of a strain that overexpresses ACC1 and DGA2 (strain MTYL065). This strain is capable of producing high lipid titers, with those lipids routinely making up greater than 60% of the total biomass; in bioreactors, the yield of lipids on both glucose and acetate after nitrogen depletion in the medium were reported to approach theoretical yield. The goal of this work, the development of a strain with lipid production characteristics greater than this strain, either through additional modifications to it, or a different combination of modifications altogether, has been proven repeatedly to be a difficult achievement.

Efficient *de novo* triacylglycerol production from growing cells must involve an optimal balance between production of biomass and lipids; as the metrics of lipid productivity, content, and yield for a strain increase beyond a certain point, a tradeoff between producing biomass for increased productivity at the expense of yield or producing higher yields at the cost of decreased productivity becomes guaranteed in the absence of alternative process designs such as multi-stage reactors, media supplements, cell recycling, etc., which are beyond the focus of this work. Like the case for yield, as a product stored intracellularly, the lipid content of cells also hits a ceiling beyond which the productivity of the overall process will necessarily shrink.

In addition to the potential tradeoffs between different production metrics as these increase, *Yarrowia lipolytica* in general shows a confounding lack of consistency in terms of growth rates and end point fermentation characteristics in biological replicates both across and within experiments. This is likely due to the ability of this strain to form a true hyphal morphology characterized by slow growth rate and low lipid accumulation in response to stress, but it is

commonplace within our experiments to see a large variation in growth rates among replicate flasks made with the same stock of media and inoculated from the same stock of pre-cultured cells. The effect of this is the large standard deviations in the lipid production metrics when comparing newly developed strains in our experiments, which happen to be typically on the order of our expected improvements in those metrics between experimental and control strains. Our efforts to make marginal improvements to or compared with already highly engineered strains therefore frequently end in an inability to conclusively prove strain improvement to our preferred level of statistical significance ($\alpha = 0.05$). This running theme shows up in the improvement in lipid titer on glucose in strain AD29 ($p = 0.1289$, see **Chapter 4**), in the improvement to lipid content ($p = 0.0829$) and yield ($p = 0.0961$) of strain AD6 on acetate (see **Chapter 4**), and in the improvement in lipid titer on acetate of strain YLASK12 ($p = 0.0934$, see **Chapter 7**); fortunately, we were able to achieve our desired level of statistical significance with the lipid content and yield improvements of strain YL53S (see **Chapter 5**) when compared to YL65K (which is essentially MTYL065 in a ΔKU70 background).

A second running theme throughout this work is the idea that the overengineering of a particular pathway is relatively easy to achieve and frequently results in well-below optimal flux of that pathway. Strains AD20 (see **Chapter 4**) and YL65S (see **Chapter 5**) are instructive examples of this. Lipid synthesis is an energy intensive process that generally occurs concurrently with reduced TCA cycle activity (a high energy-yielding pathway). Strain AD20 combines overproduction of lipid synthesis enzymes with the overexpression of an enzyme that consumes cytosolic NADH, the main source of energy for the cell in the absence of TCA cycle activity, which likely leads to an energy deficit in the cell. Likewise, strain YL65S combines three modifications that remove all three levels (transcriptional, post-translational, effector) of

regulation on the first committed step of fatty acid synthesis, resulting in a strain either unable to conserve energy for cellular maintenance or use cytosolic acetyl-CoA for biosynthesis of other compounds needed by the cell.

8.4 Recommendations for future work

In many instances, engineering a part of a pathway of interest introduces bottlenecks elsewhere in the pathway that may be difficult to detect. As lipid biosynthesis involves a multitude of cellular inputs (carbon, ATP, NADPH, etc.) and enzymes, and in light of the difficulties in achieving significant strain improvement through rational engineering of already highly engineered strains detailed in **Section 8.3**, combinatorial and inverse metabolic engineering methods may prove more effective at generating strains with further improvements to lipid production. Transcription factor or basal transcription machinery engineering methods such as global transcription machinery engineering (gTME), use error prone PCR to create libraries of genes coding for proteins involved in the transcription of many other genes. In this way, the expression profiles of potentially thousands of genes can be altered simultaneously. Alternatively, a highly productive strain can be mutagenized, for example, using UV light or ethyl methanesulfonate to result in an even more productive strain, which can later have its genome sequenced to determine which specific changes led to the additional improvements. In both cases, development of a fast, reliable screening method for strains with improved production is critical to the success of these engineering methods; in the case of lipid synthesis, staining colonies or running fluorescence-activated cell sorting (FACS) with lipophilic Nile Red can be used as a screen for cells with high lipid contents, and passaging floating cells (as floating is an indicator of high lipid content) repeatedly could enrich for a strain with high growth rate and lipid content, as well.

As for rational approaches, there is still considerable work to be done regarding the efficient production of lipids from xylose, a major component of hemicellulose and a sugar not natively consumed by *Yarrowia.* Previous efforts related to this [2] did not produce a strain with a lipid content or yield near comparable to those achieved on glucose or acetate. It has been hypothesized that the consumption of NADPH in the Xyl12 pathway required to convert xylose to a glycolytic or pentose phosphate pathway intermediate is a major cause of poor lipid production from this carbon source. One possible solution could include the use of the xylose isomerase (XylA) pathway, which does not consume NADPH.

To summarize, the tools of metabolic engineering have allowed for significant progress towards the improvement of *Yarrowia lipolytica* strains for biodiesel production. Rational engineering methods can still be easily applied in contexts where production occurs significantly below theoretical maximum levels. Efforts towards further improvement of strains after successful rational engineering approaches should focus on screening the outputs of combinatorial or inverse engineering methods to achieve the incremental progress that is reflective of a nearly optimized strain; in developing a process for producing a high-volume, low-profit margin commodity, these marginal improvements represent tangible steps closer towards the economic viability of alternative fuels and could make the difference between maintaining the status quo or achieving significant reductions in the potentially devastating effects of anthropogenic climate change.

References:

1) Tai, M., & Stephanopoulos, G. (2013). Engineering the push and pull of lipid biosynthesis in oleaginous yeast Yarrowia lipolytica for biofuel production. *Metabolic Engineering*, *15*, 1–9. doi:10.1016/j.ymben.2012.08.007
2) Tai, M. (2012). Metabolic Engineering of Oleaginous Yeast for the Production of Biofuels. Ph.D. Thesis. Massachusetts Institute of Technology.
3) Fei, Q., Chang, H. N., Shang, L., Choi, J. D. R., Kim, N., & Kang, J. (2011). The effect of volatile fatty acids as a sole carbon source on lipid accumulation by Cryptococcus albidus for biodiesel production. *Bioresource Technology*, *102*(3), 2695–2701. doi:10.1016/j.biortech.2010.10.141
4) 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 Technology*, *114*, 443–9. doi:10.1016/j.biortech.2012.02.091
5) Christophe, G., Deo, J. L., Kumar, V., Nouaille, R., Fontanille, P., & Larroche, C. (2012). Production of oils from acetic acid by the oleaginous yeast Cryptococcus curvatus. *Applied Biochemistry and Biotechnology*, *167*(5), 1270–1279. doi:10.1007/s12010-011-9507-5
6) Chakraborty, S. (2014). Exploring Volatile Fatty Acids (VFAs) as a Novel Substrate for Microbial Oil Production. Ph.D. Thesis. Massachusetts Institute of Technology.