THE EXPERIMENTAL AND COMPUTATIONAL STUDY OF RECOMBINANT

Saccharomyces cerevisiae FOR BETA-CAROTENE PRODUCTION

A Dissertation

by

MELANIE ROSE DESESSA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	M. Nazmul Karim
Committee Members,	Katy Kao
	Arul Jayaraman
	Phanourios Tamamis
Head of Department,	M. Nazmul Karim

August 2018

Major Subject: Chemical Engineering

ABSTRACT

        Terpenes are an important and diverse group of chemicals used in many

industries, including nutraceuticals, pharmaceuticals, alternative fuels, and personal care

products. In yeast, the natural mevalonate pathway provides a starting point for

genetically engineering terpene-producing strains. One such example of the

advantageous exploitation of the mevalonate pathway is the work done in the Kao lab

with *Saccharomyces cerevisiae*. A strain of *S. cerevisiae*, named SM14, was engineered

to produce beta-carotene. Beta-carotene is a carotenoid that is used as an orange food

dye, in supplements due to its antioxidant properties, and in vitamins as a precursor to

vitamin A. This document describes the further study of strain SM14, both

experimentally and computationally, to understand and improve the production of beta-

carotene. SM14 was cultivated in a bench-scale bioreactor and key metabolites and off

gas were analyzed. In addition, this experimental data was further studied using a flux

balance analysis program in Matlab® (Mathworks®) known as Constraint-Based

Reconstruction and Analysis (COBRA). This program allowed the intracellular behavior

of SM14 to be extensively studied. The dynamic behavior of the strain was also

examined using the COBRA strategy in combination with other optimization strategies.

The intracellular behavior of the continuous production of beta-carotene was also studied

using COBRA models in conjunction with a continuous kinetic model. Further utilizing

these computational techniques, the Optknock tool within the COBRA toolbox was used

to target beneficial gene knockouts. The CRISPR-Cas9 system was examined as a

ii

method of efficiently performing these knockouts as well as more traditional knockout

methods. The results of these experiments are discussed.

DEDICATION

This thesis is dedicated to my family, and to my friends who are family.

I may never make it to Mars, so:

I claim this thesis for my Auntie Marian.

ACKNOWLEDGEMENTS

CONTRIBUTORS AND FUNDING SOURCES

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Motivation

The search for alternative methods for the production of fuels and chemicals is motivated by several different reasons. Alternative fuels have become increasingly attractive due to increasing costs and adverse environmental impact of traditional petroleum-based fuels. In addition, alternative fuels are the best option when striving towards sustainability and energy independence [1]. Pharmaceuticals are a similarly wide-reaching and important industry. The use of pharmaceuticals is only expected to rise, but the cost burden on the consumer is unsustainable [2, 3]. In the fuel and pharmaceutical industries, as well as many more important industries, biotechnology has the potential to lower costs and environmental impact [1, 4-8]. While research and development of biotechnological solutions is ongoing, further work can be done to streamline the development and further improvement of production processes. Computational tools, such as kinetic modeling, metabolic flux modeling, and computational genetic modification targeting, can expedite and improve the understanding of bioprocesses. In addition, the study and improvement of continuous bioprocesses can further reduce costs [9]. Understanding the behavior of the organism within a continuous production process as well as the process itself, is important for further improvement.

1.2 Goals of This Work and Outline of Thesis

The combination of computational and experimental techniques for the further understanding of terpenoid production in yeast, using beta-carotene as a case study, is the theme of this thesis.

The overall goal of this work involves the marriage of experimental and computational analysis work to develop a microbial engineering strategy for the beta-carotene production in *S. cerevisiae.* It is the hope that future work can employ this strategy, shown in Figure 1 for the production of many different bioproducts in different organisms and using different sources of carbon.
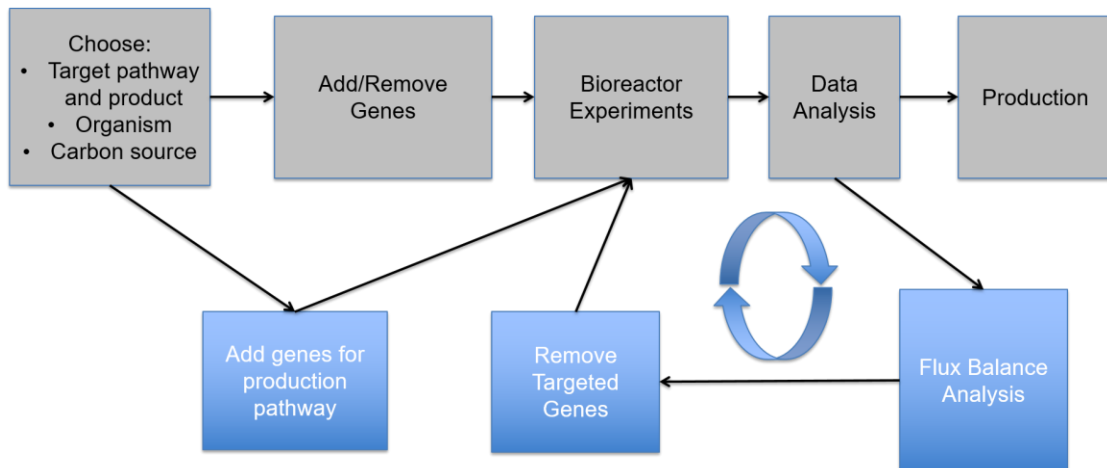
Figure 1: The microbial engineering strategy employed in this work for the production of beta-carotene in *S. cerevisiae*.

The specific goals accomplished in this work can be summarized as follows:

1. Analyze a batch bioreactor process for the production of beta-carotene, specifically the overflow metabolism behavior and multiple substrate utilization;

2. Perform steady-state flux balance analysis of the beta-carotene production process using a stoichiometric yeast cell model and the COBRA toolbox in Matlab;

3. Expanding on the tools used and knowledge gained in Goal 2 used in conjunction with a extracellular kinetic model, develop a process for the dynamic flux balance analysis of the batch beta-carotene production process;

4. Study the intracellular behavior of a continuous beta-carotene production process by combining an extracellular kinetic model of the continuous process with the tools used in Goals 2 and 3;

   and

5. Use additional flux balance analysis techniques to target beneficial genetic modifications of a beta-carotene-producing strain of yeast, and attempt the modifications experimentally to verify the computational methods.

In Chapter II, the group of chemicals known as terpenes, and specifically the subgroup of modified terpenes known as terpenoids, is introduced. The pathway within the yeast strain *Saccharomyces cerevisiae* that can be modified to produce terpenoids, known as the mevalonate pathway, is discussed. The terpenoid beta-carotene is introduced, and it's relevance as an important nutraceutical as well as its use as a case study, is considered. Background on *S. cerevisiae* is given, specifically the difference between anaerobic and aerobic growth as well as the Crabtree effect. Finally, this chapter summarizes the construction and directed evolution of a strain of *S. cerevisiae* that produces beta-carotene.

Chapter III discusses the development of a bench-scale batch bioreactor process for the production of beta-carotene. The operation of the bioreactor is summarized, and the methods for the quantification of all relevant products and byproducts are discussed. The results of the bioreactor experiments performed for this work are shown and analyzed.

Chapter IV of this thesis introduces the concept of flux balance analysis. This chapter also discusses a stoichiometric model of the intracellular reactions in *S. cerevisiae*. The alterations made to this model to represent the beta-carotene production strain are summarized. This chapter then discusses the steady-state flux balance analysis work performed. The results of the analysis performed with the model to calculate the theoretical beta-carotene yield of the strain, as well as to verify the model's ability to simulate overflow metabolism, are discussed.

In Chapter V, the kinetic modeling of a batch beta-carotene production process is discussed. A dynamic flux balance analysis procedure, using the previously described stoichiometric model and the results from the kinetic model, is proposed. The results of this dynamic flux balance analysis, including the fluxes through key intracellular pathways, are summarized and discussed.

Chapter VI of the thesis introduces the computational gene targeting tool known as Optknock. The use of this tool to target beneficial genetic deletions in the beta-carotene production strain of *S. cerevisiae* is discussed. The deletions targeted by Optknock are summarized, and the relevant reactions and possible effects are discussed.

In Chapter VII of this thesis, experimental methods for the deletion of genes in the beta-carotene production strain of *S. cerevisiae* are discussed. The history of the CRISPR-Cas9 system is summarized, and the method of using this tool for gene deletions is described. To verify a CRISPR-Cas9 plasmid, the *ADE2* gene was knocked out of a strain of *S. cerevisiae*. The results of this work are discussed. In addition, the more standard method of yeast deletions, homologous recombination, is described. Using the traditional knockout method, it was attempted to knock out a subset of the genes targeted by the Optknock tool discussed in the previous chapter. The results of these experiments are summarized.

Chapter VIII discusses the intracellular analysis of the beta-carotene-producing *S. cerevisiae* strain in a continuous production process. The expansion of the previously discussed batch kinetic model into a continuous kinetic model is illustrated. The concentrations of biomass, ethanol, acetate, glucose, and beta-carotene are shown for a range of dilution rates, and the corresponding flux values are also shown. The use of these flux values as input to the intracellular stoichiometric model is described, and the resulting intracellular fluxes are shown for key pathways within the cell.

Chapter IX summarizes the contents of the previous chapters. The relevance of the work is reiterated, and areas for future research directions are discussed.

1.3 Novelty of Work

The work described in this document contain several novel contributions. Specifically, Chapters 5 and 6 of this document describe novel work performed using an intracellular flux model and the COBRA optimization toolbox in Matlab®. In Chapter 5, the dynamic flux balance analysis dual-optimization procedure has not been performed in other works, especially with a full-genome model of recombinant *S. cerevisiae*. In Chapter 6, analysis of the intracellular behavior of a continuous beta-carotene production process was performed for the first time. While the methods and work discussed in Chapters 3, 4, 7, and 8 are not in themselves novel, they have led to new and further understanding of the beta-carotene production strain of *S. cerevisiae*, as well as the batch bioreactor process for beta-carotene production.

2. BACKGROUND: BETA-CAROTENE PRODUCTION IN YEAST

2.1 Terpenes and Terpenoids

Terpenes are a diverse class of chemicals used in many industries, as precursors or final products. These industries include nutraceuticals, pharmaceuticals, alternative fuels, and personal care products. The basic unit of terpenes is a hydrocarbon molecule known as isoprene. The isoprene unit consists of five carbons. A molecule consisting of one isoprene unit is known as a hemiterpene. Molecules containing two isoprenes bonded together is called a monoterpene, while those containing three units are known as sesquiterpenes. Terpenes that have modifications, such as added oxygen, are known as terpenoids. Isoprenoids, therefore, comprise many plant metabolites, such as hormones, sterols, and carotenoids [10]. Figure 2, adapted from the review by Chappell, shows the isoprenoid biosynthesic, or mevalonate pathway [11]. This pathway begins with the condensation of three acetyl-CoA units to produce 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA). HMG-CoA is then converted to mevalonate, a reaction catalyzed by HMG-CoA reductase. HMG-CoA is then phosphorylated and decarboxylated by mevalonate kinase, mevalonate-5-phosphate kinase, and mevalonate 5-diphosophate decarboxylase to form isopentenyl diphosphate (IPP). IPP is condensed with dimethylallyl diphosphate (DMAPP) to form geranyl diphosphate. A second IPP unit is added to generate farnesyl diphosphate (FPP), and another added IPP unit forms geranylgeranyl diphosphate. Many varied molecules are formed from IPP, FPP, GPP, and GGPP, as summarized in the figure. These products have a variety of functions from

growth hormones and anti-fungals to pharmaceuticals and perfume [12-18]. Examples of industrially important terpene or terpenoid products are bisabolene, a precursor to alternative fuel bisabolane, taxadiene, a precursor to cancer-fighting drug Paclitaxel, as well as natural rubbers prevalent in many common products [4, 19].These and many other terpene derivatives are naturally produced, mainly by plants, and are difficult to synthesize. The importance of these chemicals has brought about increased demand, but current production methods have resulted in high costs and low availability. It is desired, therefore, to develop alternative production methods to help alleviate the costs of and high demand for these products [7, 20, 21].
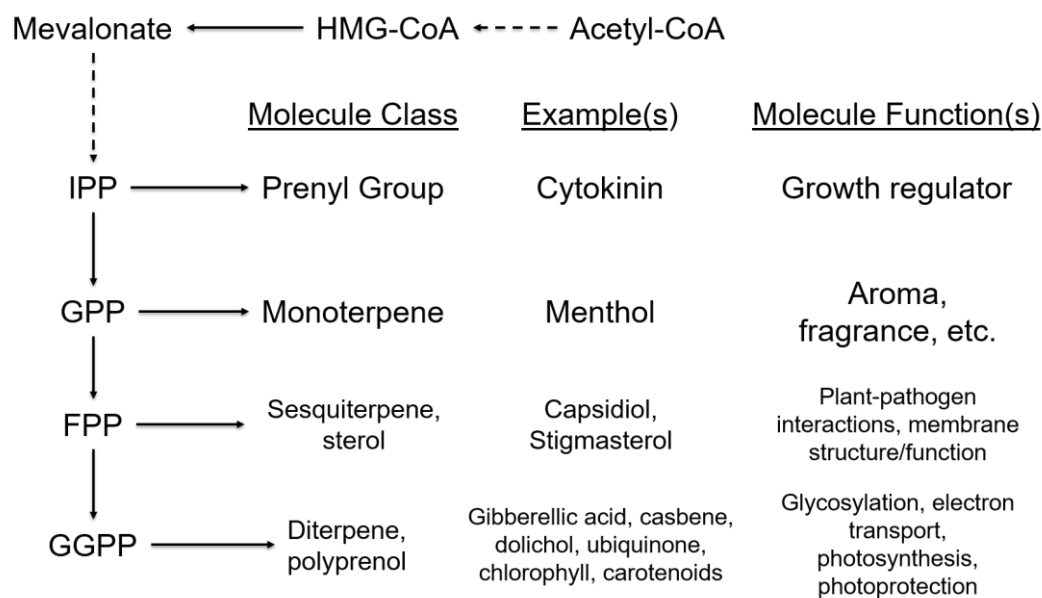
Figure 2: The isoprenoid biosynthetic, or mevalonate pathway. IPP: isopentenyl diphosphate, GPP: geranyl diphosphate, FPP: farnesyl diphosphate, GGPP: geranylgeranyl diphosphate. Examples of native plant terpenoids are shown at the point at which they're produced within the pathway.

2.2 Mevalonate Pathway in Yeast

A viable method of producing terpenes at higher levels is modifying the mevalonate pathway in an easy-manipulated organism, such as *Saccharomyces cerevisiae* [5, 22, 23]. The mevalonate pathway in yeast is shown below in Figure 3. As described in the previous section, a portion of acetyl-CoA in the cell is converted to HMG-CoA via acetoacetyl-CoA thiolase and 3-hydroxy-3-methylglutaryl coenzyme A synthase. HMG-CoA is then converted to mevalonate, which in turn is converted to mevalonate-5P and mevalonate-5-PP via mevalonate kinase and phosphomevalonate kinase, respectively. Mevalonate-5-PP is converted to isopentenyl diphosphate (IPP) via mevalonate pyrophosphate decarboxylase. IPP is converted to dimethylallyl-pyrophosphate via isopentenyl diphosphate dimethylallyl diphosphate isomerase [24]. IPP and dimethylallyl-pyrophosphate are intermediates along the way to eventual production of squalene and ergosterol, important components for cellular structure [25]. The native mevalonate pathway in yeast can be edited to branch off of FPP to produce diterpenes such as taxadiene, or off of GGPP to produce larger molecules such as bisabolene [26].
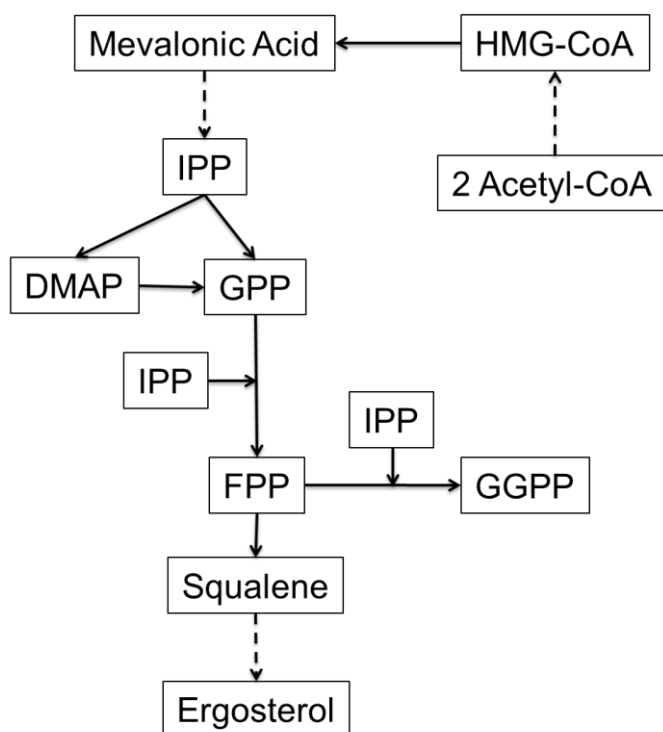
Figure 3: The mevalonate pathway in S. cerevisiae. The dotted lines indicate intermediate reactions not shown, while the solid lines indicate the shown reaction.

2.3 Case Study: Beta-carotene

An important terpene called beta-carotene is produced naturally in plants, fungi, and some bacteria. It is used naturally in photosynthetic plants and microbes for light absorption as well as a defense against photo-oxidation [27]. Beta-carotene is a bright orange-pigmented compound, and is used as a food dye due to this coloring. The molecule is a polyunsaturated hydrocarbon containing 40 carbon atoms, terminated by beta-rings on each end. This structure is shown in Figure 4. Beta-carotene has antioxidant properties, and has been shown to have positive vision effects [28]. One molecule of beta-carotene can be cleaved into two molecules of vitamin A, and therefore is used in vitamins as a natural precursor to vitamin A. It is produced naturally in low

levels, which requires a large amount of time and resources to meet the current demand. In order to meet rising demand, as well as lower costs for production, recombinant organisms may be used to produce beta-carotene in much higher levels [29-33]. In this paper, we have used the production of beta-carotene in *S. cerevisiae* as a case study in order to better understand the mevalonate pathway and its importance in the production of terpenes in this organism.
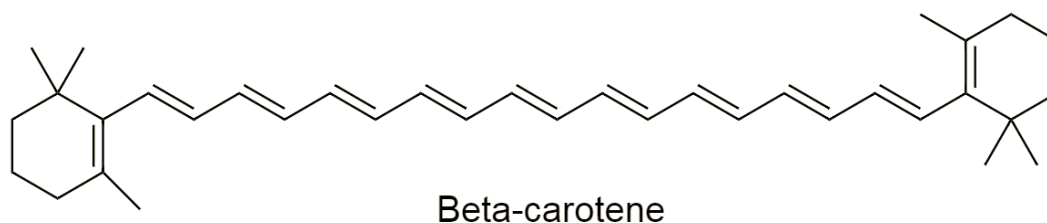
Beta-carotene

Figure 4: Chemical structure of beta-carotene.

2.4 Yeast Growth and Respiration

The growth of yeast is highly dependent on many environmental factors, including temperature, pH, and aeration. These conditions can be tuned to optimize the growth and overall health of the cells, as well as, in the case of aeration, direct metabolic flux towards certain pathways. Yeast can grow under anaerobic or aerobic conditions, with much different outcomes. For the scope of this work, only growth using glucose as a carbon source will be considered. In addition, while different strains of yeast can

produce different products under anaerobic conditions, only the relevant ethanol-

producing *S. cerevisiae* will be considered in this work.

2.4.1 Anaerobic vs. Aerobic Growth

In the presence of oxygen, *S. cerevisiae* undergoes aerobic respiration. Under

these conditions, *S. cerevisiae* consumes glucose and oxygen and produces carbon

dioxide and water, among other byproducts in lesser amounts [34, 35]. The pathways

that illustrate aerobic glucose consumption and respiration are the glycolysis pathway

and the TCA cycle. These pathways are illustrated in Figure 5. After glucose has been

taken up by the cell and phosphorylated, glucose-6-phosphate is converted to fructose-6-

phosphate via glucose-6-phosphate isomerase [36]. Fructose-6-phosphate is then

converted to fructose-1,6-bisphosphate by phosphofructokinase [37]. Fructose-1,6-

bisphophate is converted to glyceraldehyde-3-phosphate via the aldolase enzyme [38].

Glyceraldehyde-3-phosphate is then converted to 3-phosphoglycerate via 3-

phosphoglycerate kinase [39]. 3-phosphoglycerate is converted to 2-phosphoglycerate by

phosphoglycerate mutase [40]. 2-phosphoglycerate is then converted to

phosphoenolpyruvate via the enolase enzyme [41]. In the final step of the glycolysis

pathway, phosphoenolpyruvate is converted to pyruvate by the pyruvate kinase enzyme

[42]. The TCA cycle, important in aerobic respiration, begins, and produces carbon

dioxide and water in its overall reaction [43].
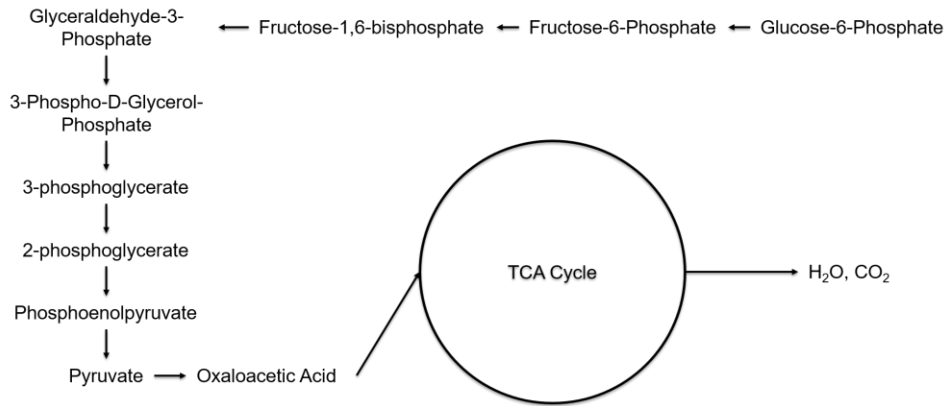
Figure 5: The glycolysis pathway and overall TCA cycle reaction in *S. cerevisiae*.

When grown without oxygen, *S. cerevisiae* undergoes aerobic respiration. Under these conditions, *S. cerevisiae* consumes glucose and produces ethanol, carbon dioxide, and a sufficient amount of energy to continue to grow, among other byproducts in lesser amounts. This process is known as fermentation. The glucose fermentation pathway is shown in Figure 6. The production of pyruvate in this pathway is similar to the glycolysis pathway previously described. At this point however, in the absence of oxygen, pyruvate is converted to ethanol, via alcohol dehydrogenase, and acetate, via aldehyde dehydrogenases [44, 45].

13

Fructose-1,6-bisphosphate ⟵ Fructose-6-Phosphate ⟵ Glucose-6-Phosphate

↓

Glyceraldehyde-3-Phosphate

↓

3-Phospho-D-Glycerol-Phosphate

↓

3-phosphoglycerate

↓

2-phosphoglycerate

↓

Phosphoenolpyruvate ⟶ Pyruvate ⟶ Acetaldehyde ⟶ Ethanol

Acetate ↑ (from Acetaldehyde)
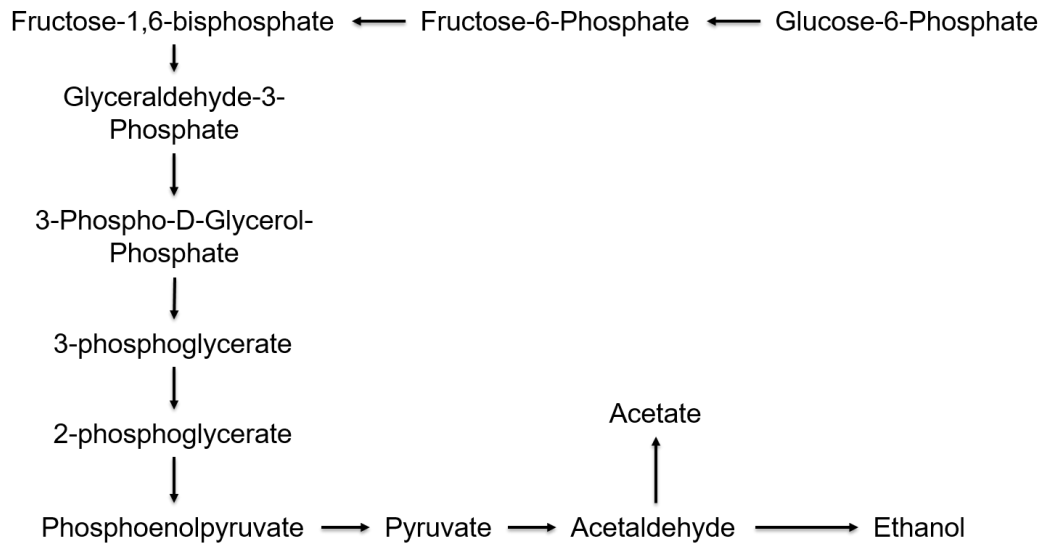
Figure 6: The glucose fermentation pathway in *S. cerevisiae*.

2.4.2. Overflow metabolism, or the Crabtree Effect

During aerobic respiration, if the amount of glucose is at sufficiently high levels, many organisms may undergo overflow metabolism [46, 47]. When this occurs in *S. cerevisiae*, in addition to normal production of carbon dioxide, a significant amount of ethanol is produced [35]. This phenomenon, also known as the Crabtree effect, is illustrated below in Figure 7. In many production processes, it is attempted to avoid overflow metabolism, as carbon is diverted from the desired product and the cells don't grow efficiently. This can be done by maintaining the glucose concentration at low levels [34].
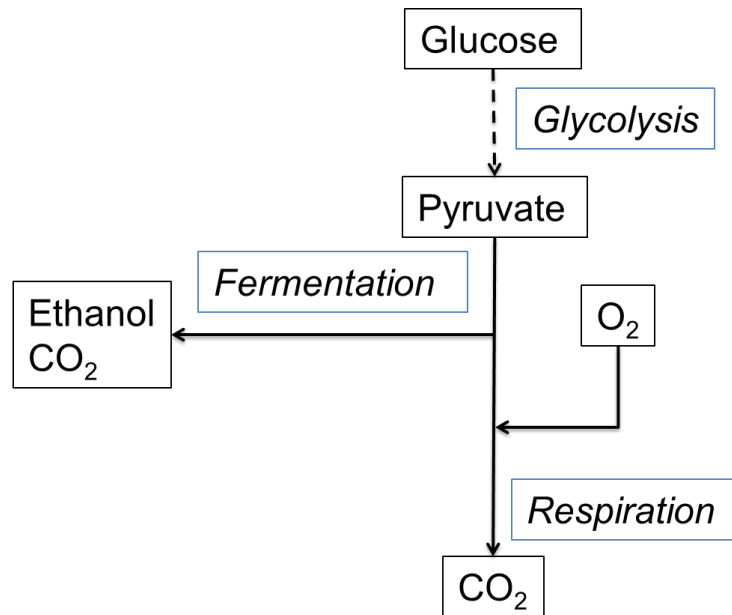
Figure 7: A simplified flowchart of yeast metabolism, illustrating the Crabtree effect.

2.5 Strain Construction

2.5.1 Chromosomal Modification and Added Pathway Description

A strain of *Saccharomyces cerevisiae* capable of producing beta-carotene was constructed by introducing genes from *Xanthophyllomyces dendrorhous* into the chromosome of *S. cerevisiae* [48]. Figure 8 below depicts the additional reactions and metabolites branching off the mevalonate pathway after the genes have been introduced. While the conversion of farnesyl-diphosphate (FPP) to geranylgeranyl-diphosphate (GGPP) is native to *S. cerevisiae*, the gene *crtE* was added in an attempt to increase flux through this reaction. The gene *crtI* facilitates both the conversion of phytoene to neurosporene, as well as the conversion of neurosporene to lycopene. The gene *crtYB* facilitates the conversion of GGPP to phytoene, neurosporene to 7,8-dihydro-beta-

15

carotene, and lycopene to beta-carotene. Information about the mechanisms of each of these reactions, as well as rates and reversibility, is available in the study done by Moise et al [49]. Because of its hydrophobic properties, when beta-carotene is produced in *S. cerevisiae*, it accumulates in the cell membrane and is not exported by the cell [50] .
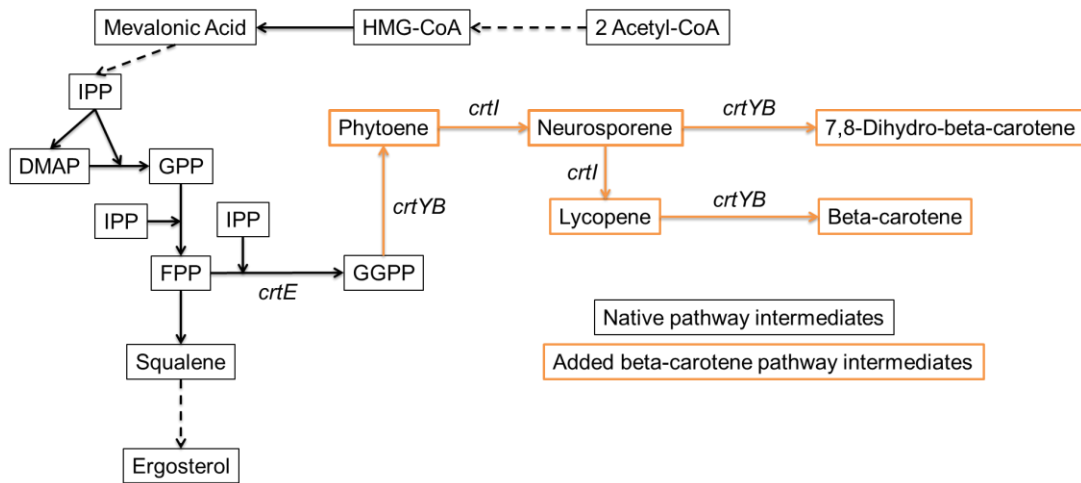
Figure 8: The mevalonate pathway in *S. cerevisiae* with the added beta-carotene production branch. Metabolites boxed in orange indicate non-native products.

2.5.2 Adaptive Evolution

After chromosomal modification, this strain was further refined by Reyes et al using adaptive evolution techniques (Reyes, Gomez, and Kao 2014). The strain was put under oxidative stress by culturing in the presence of hydrogen peroxide. In this way, the natural anti-oxidant properties of beta-carotene were exploited, and the strain adapted to produce more beta-carotene to combat the effects of the hydrogen peroxide. The strain used in this work is the highest level beta-carotene producer found after the evolution experiments were completed, known as strain SM14.

3. BIOREACTOR STUDIES

3.1 Introduction

The conditions and protocol for the bioreactor experiments in this work were based on previous work done in the Kao lab, as well as previous work in this lab [50-52]. In the work done by Olson, YPD and YNB media supplemented with glucose at a concentration of 20 g/L were used. The culture from an initial bioreactor run with YNB media was used to create 1 mL frozen stocks to be used in subsequent bioreactor runs. One of these 1 mL frozen stocks was used to start a 50 mL culture, grown for 48 hours at 30°C and 170 rpm, which was in turn used to inoculate the a 7L bioreactor with a 3L working volume. The bioreactor was run for 72 hours, the pH was controlled at set point using 2M NaOH and 2M HCl, and the temperature was controlled at 30°C. The agitation and air flow were maintained at a constant value. In Olson's studies, several pH set points were tested, namely 4, 4.5, and 5, and two agitation speeds, 400 and 800 rpm, were also tested. In those studies, it was found that the beta-carotene production rate and the ethanol consumption rate were higher at higher aeration rates, and so a constant agitation speed of 800 rpm was chosen for subsequent experiments. It was found that YNB medium, which is more defined than YPD, allowed for more consistency between runs, and was therefore used in subsequent experiments. It was also found that beta-carotene production decreased as the pH increased, and therefore the pH set point of 4 was used in subsequent experiments.

17

The pH set point of 4, the agitation speed of 800 rpm, and the medium composition of YNB + 20 g/L glucose continued to be used in the further work done by Jaladi and Ordonez. In these further studies, instead of starting from a separate frozen stock each time, a single frozen stock of SM14 was kept at -80°C, an agar YNB + 20 g/L glucose plate was struck out using this stock, and the 50 mL inoculum culture was started using a single colony from this plate.

3.2 Materials and Methods

3.2.1 Bioreactor Operation

The bioreactor experiments performed in this study were carried out in a 3L or 7L autoclavable, glass vessel (Applikon, Foster City) with a 1.3L or 3L working volume, respectively. YNB minimal medium containing 20 g/L glucose was used. The components of this medium are listed in Table 13 in the appendix. Directly prior to the start of the experiment, a fresh YNB + 20 g/L glucose agar plate was struck out with SM14 from a frozen stock. This plate was grown at 30°C for approximately 48 hours. A seed flask started by inoculating 50 mL of YNB minimal medium with a single SM14 colony from this plate was used as the inoculum for the bioreactor. For the larger bioreactor (3L working volume), the entire 50 mL volume was used, while for the smaller bioreactor (1.3L working volume), 22 mL of the inoculum was used to maintain the same ratio of inoculum to working volume. The environmental conditions were controlled in order to replicate a previous experiment done by Ordonez et al when developing an extracellular kinetic model [52]. These conditions were optimized for operation in batch mode in the Kao lab [50]. The temperature was maintained at 30°C

18

using a heating jacket. The pH was maintained at 4 by using an acid/base dual control with 2M NaOH and 0.5 HCl. For the 3L vessel, the agitation was set between 550 and 600 rpm, and the airflow was set to 2.6 L/min (2vvm) for the duration of the experiment. For the 7L vessel, the agitation was set to 800 rpm and the airflow was set to 6 L/min (2vvm). It was desired to collect as much data as possible from the bioreactor in order to later inform the model, but to also stay within the time and resources limitations. Therefore, samples were pulled from the bioreactor every two hours for 72 hours. Two samples were taken at each time point for contingency and statistics purposes. The absorbance of these samples was measured immediately to determine the cell density, and then the remainder of the sample volume was frozen at -20°C to be analyzed after the experiment's completion.

3.2.2 Quantification Methods

Online off gas measurements were taken using a BlueSens$^{©}$ off gas analyzer, and were recorded using the BlueVis software (BlueSens$^{©}$). The online pH, agitation rpm, temperature, dissolved oxygen (%DO), and base and acid addition were recorded every minute using BioXpert Lite software (Applikon $^{®}$).

Absorbance measurements were taken on a GENESYS^{TM} 10S UV-Vis spectrophotometer (Thermo Scientific^{TM}), at 600 nm wavelength. The absorbance was measured immediately after the sample was taken, and the samples were diluted with water as the experiment progressed to maintain a reading within linear range of the instrument.

The glucose, ethanol, and acetate in the bioreactor samples were measured using an LC 1260 HPLC (Agilent Technologies). Approximately 2 mL of each sample taken was spun down in microcentrifuge tubes at maximum speed for two minutes. The supernatant was then filtered through a 0.22 µm syringe filter to remove any solids or impurities. The supernatant was collected in an HPLC vial. The mobile phase used in the HPLC was 5 mM H_2SO_4, the flowrate was 0.6 mL/h, and the temperature was set to $50^{\circ}C$. The RID analyzer was used to quantify the metabolites.

The beta-carotene was measured using a colorimetric quantification method developed by Reyes et al [53]. In this method, 500 µL of cells were first separated from the supernatant via centrifugation at max speed (15,000 rpm). Then, 1000 µL of dodecane along with small glass beads were added to the spun-down cells. The cells were then broken open using a cell disruptor. This released the beta-carotene, which was extracted fully into the dodecane. When this method was completed for each sample taken from the bioreactor, the extracted beta-carotene/dodecane mixture was added to 96-well plate wells and the absorbance was read via plate-reader at 454 nm. The equation below was used to convert the absorbance measurements to beta-carotene concentration [48].

$$Beta-carotene\ Concentration\left(\frac{mg}{L}\right) = 2 * (9.1208 * Absorbance^2 +$$
$$9.7052 * Absorbance + 2.927) \tag{1}$$

3.3 Results and Discussion

Using the batch experiment method described above, fifteen experiments were run. Compared with historical experimental results, all of the bioreactor runs

20

underperformed in terms of biomass and beta-carotene produced. Figure 9 shows the comparison between the historical bioreactor data and a sample of the bioreactor data generated in this study. While the biomass concentration was slightly higher in the run performed for this study, the beta-carotene concentration was significantly lower. The trends and values for both glucose and ethanol concentration were fairly similar, but the acetic acid in the historical run's data was much lower. The behavior seen in the run shown in Figure 9 is representative of all fifteen experiments run while generating data for this study. Throughout the fifteen experiments, all environmental factors were considered and validated to ensure that the cause was not environmental. These factors include air flow rate, pH, temperature, agitation, and media component sources (including water). The mass flow controller for the air supplied to the bioreactor was determined to be accurate via external rotameter. The pH probes were replaced and re-calibrated using fresh pH buffer and an accurate external pH probe was used to verify pH readings. The temperature probes were tested using an accurate external temperature probe. The stir assemblies received preventative maintenance to insure proper function. All media components were replaced, and the filter for the ultra-pure water was replaced. After it was confirmed that the environmental factors were likely not the cause for the poor performance of the bioreactor, it was accepted that perhaps the SM14 strain had been somehow compromised in the course of running experiments.
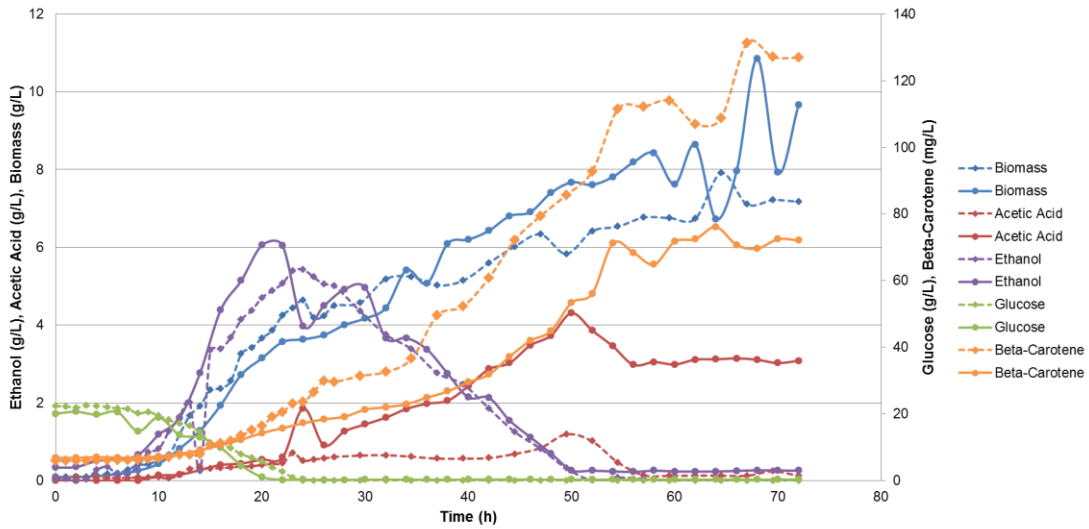
Figure 9: A comparison of historical bioreactor data and a sample bioreactor run performed in this study. The historical data is shown with dotted lines.

Because the data was primarily being used to inform and check the COBRA modeling work, it was decided to continue with three of the fifteen bioreactor runs that were run under the same conditions and whose results were well-replicated. These three runs are hereafter referred to as Trials A, B, and C. Trials A and C were run in a 3L bioreactor with an initial working volume of 1.3L. Trial B was run in a 7 L bioreactor with an initial working volume of 3L. Online data, including oxygen and carbon dioxide off gas data, were only collected for trials A and C. The off gas data in the form of carbon evolution rate (CER), oxygen uptake rate (OUR), and respiratory quotient (RQ) are shown in Figure 11 and Figure 12b. These values were calculated using equations 17-23 below. Table 1 contains additional information related to these equations. The average of the data collected from the three trials is shown in Figure 10 and Figure 12a. The comparison between these data and the historical data is shown in Table 2.

22

$$OUR\left(\frac{L}{min}\right) = \left[F_{AIR} * \left(\frac{\%O_{2,in}}{100}\right)\right] - \left[F_{AIR} * \left(\frac{\%O_{2,out}}{100}\right)\right] \qquad (2)$$

$$OUR\left(\frac{mmol}{min}\right) = \frac{P*\left[OUR\left(\frac{L}{min}\right)\right]}{R*T} * 1000 \qquad (3)$$

$$OUR\left(\frac{mmol}{gDCW*h}\right) = \frac{\left[OUR\left(\frac{mmol}{min}\right)\right]}{X} * 60 \qquad (4)$$

$$CER\left(\frac{L}{min}\right) = \left[F_{AIR} * \left(\frac{\%CO_{2,in}}{100}\right)\right] - \left[F_{AIR} * \left(\frac{\%CO_{2,out}}{100}\right)\right] \qquad (5)$$

$$CER\left(\frac{mmol}{min}\right) = \frac{P*\left[CER\left(\frac{L}{min}\right)\right]}{R*T} * 1000 \qquad (6)$$

$$CER\left(\frac{mmol}{gDCW*h}\right) = \frac{\left[CER\left(\frac{mmol}{min}\right)\right]}{X} * 60 \qquad (7)$$

$$RQ = \frac{CER\left(\frac{mmol}{gDCW*h}\right)}{OUR\left(\frac{mmol}{gDCW*h}\right)} \qquad (8)$$

Table 1: Description of parameters used in the calculations of OUR and CER.

Parameter	Description/Value
F_{AIR}	Flow rate of air into/out of the bioreactor $\left(\frac{L}{min}\right)$
$\%O_{2,in}$	Percentage of oxygen in the inlet air (21%)
$\%O_{2,out}$	Percentage of oxygen in the air exiting the bioreactor
P	Air pressure (101.592 kPa)
R	Gas constant (8.314 $\frac{L*kPa}{K*mol}$)
T	Temperature of the bioreactor (K)
X	Biomass (gDCW)
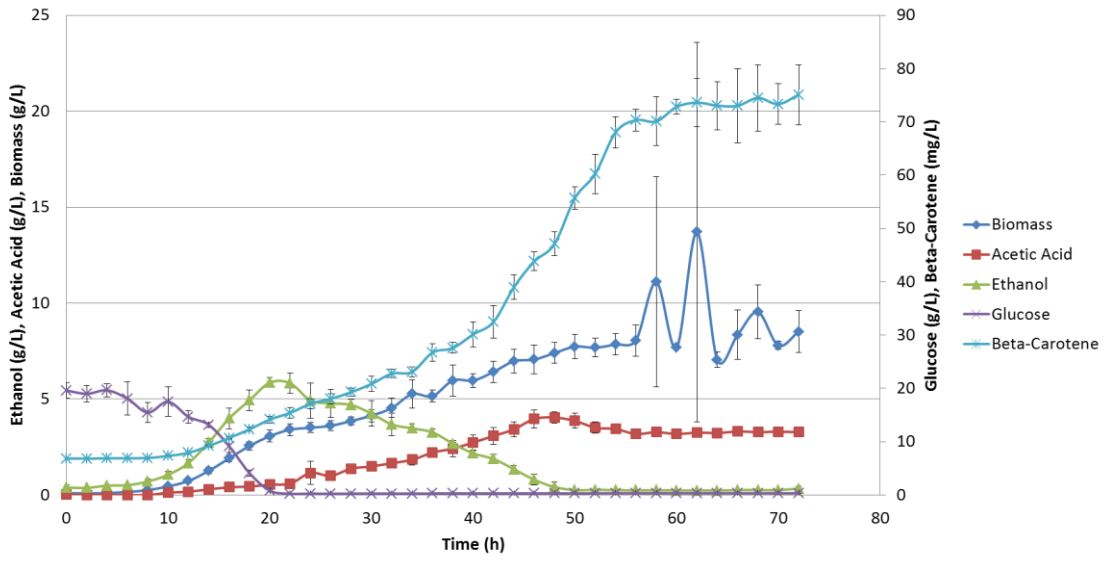
Figure 10: Average bioreactor metabolite data from three chosen runs. Error bars represent one standard deviation.
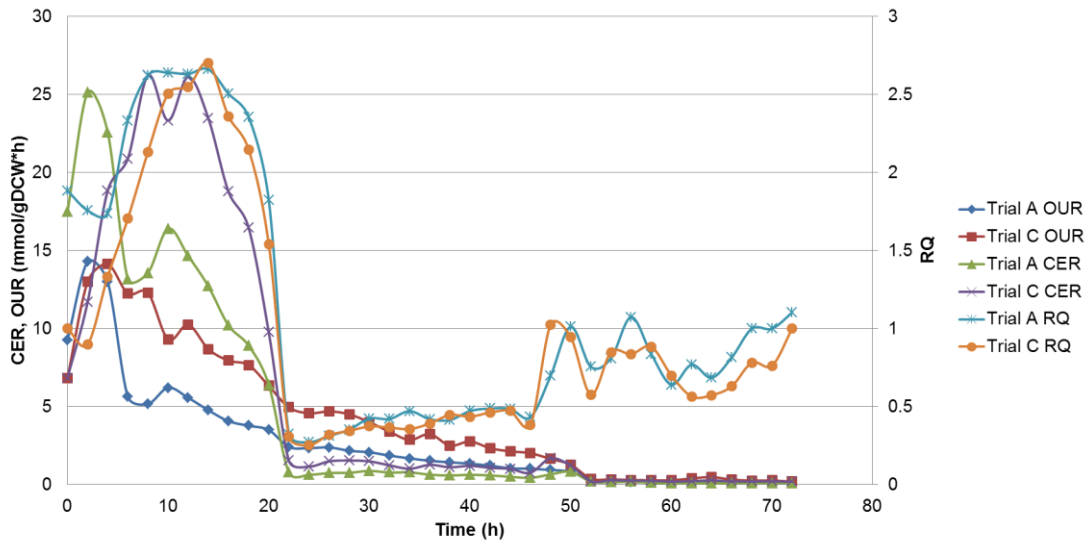
Figure 11: OUR, CER, and RQ values for Trials A and C.

24

Table 2: Summary of the metabolite concentrations from the three chosen trials for this study and the historical data.

	Maximum Beta-Carotene Titer (mg/L)	Maximum Biomass Concentration (g/L)	Maximum Ethanol Concentration (g/L)	Maximum Acetic Acid Concentration (g/L)
Average of Trials A, B, and C	75.10	13.70	5.86	4.04
Historical Data	131.30	7.91	5.43	1.19

As can be seen in Figure 10 and Figure 12a, the yeast first consume the available glucose and produce ethanol. This aerobic production of ethanol is due to overflow metabolism. After all of the glucose has been consumed, the cells begin to use the ethanol as a carbon source. When the ethanol has been consumed, the cells utilize a portion of the acetic acid to produce more beta-carotene. In addition to ethanol production, the Crabtree effect can be seen by examining the RQ values. As shown in Figure 12a, the RQ values during the first portion of the process (ethanol production) are greater than one, and then decrease to values of under one. Values of RQ that are greater than one indicate some fermentative metabolism (ethanol production) [34, 35]. In addition to the ethanol production and ethanol/acetic acid consumption phases, an important change in the metabolism of the cell occurs when the rate of beta-carotene production equals or outpaces the rate of biomass production. These key observations separate the beta-carotene production process into three distinct phases: the glucose consumption/ethanol production phase (phase 1), the ethanol consumption and low beta-carotene production phase (phase 2), and the ethanol/acetic acid and high beta-carotene

25

production phase (phase 3). These phases are represented by colored boxes in Figure 12a, and will be discussed further in section 4.2.
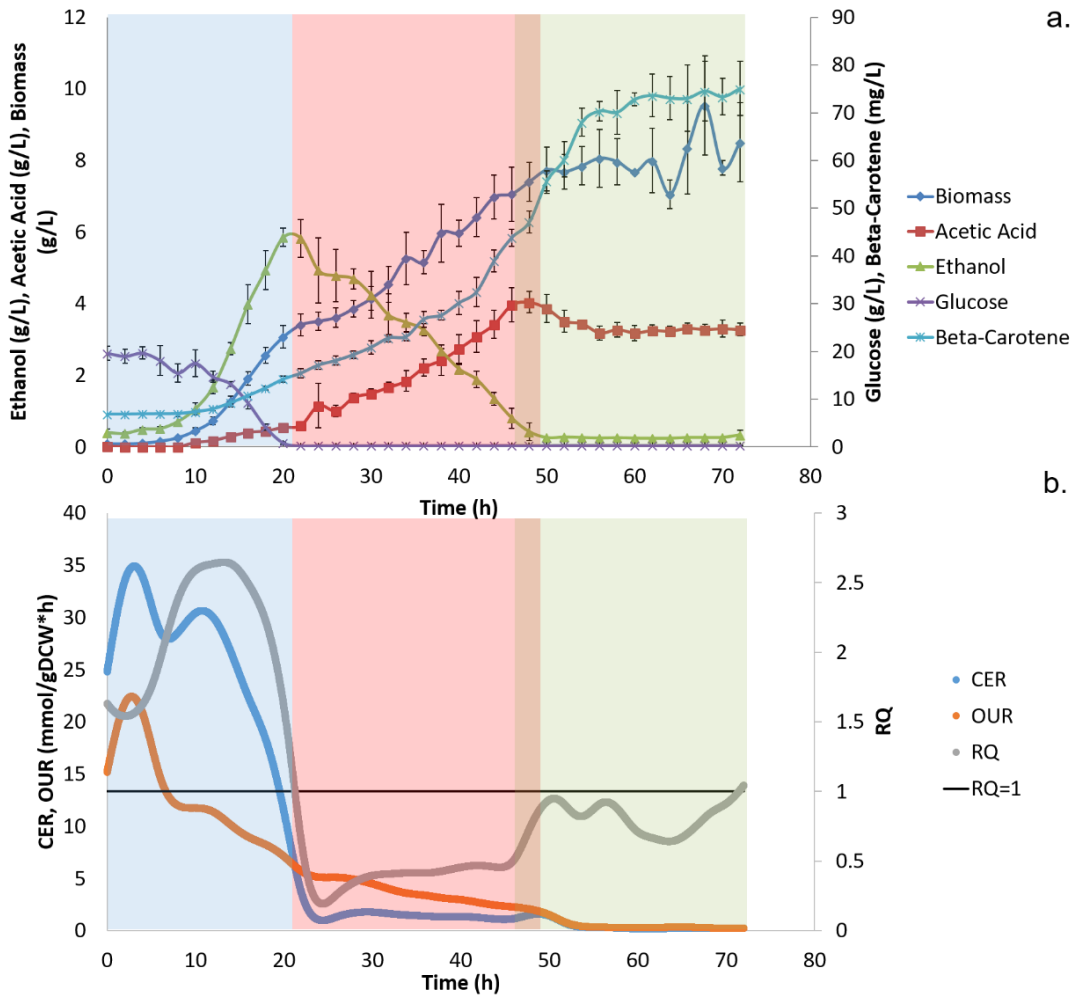
Figure 12: (a) Average growth and metabolite data for the three bioreactor runs performed in this study. The glucose consumption/ethanol production phase is shown in blue, the ethanol consumption phase is shown in red, and the acetic acid consumption phase is shown in green. (b) Average CER and OUR values for the three bioreactor runs performed in this study, as well as the calculated RQ values. The fermentative activity caused by the Crabtree effect occurs during the phase shown in blue, while the standard aerobic respiration phase is shown in red and green.

3.4 Conclusion

The analysis discussed in this study make it possible to better understand the behavior of strain SM14 in a batch bioreactor process. The growth of the strain, the glucose, biomass, ethanol, and acetate concentrations, and the carbon dioxide and oxygen levels in the off gas, were measured and analyzed. It was found through this analysis that there are three distinct phases of the batch bioreactor process, the glucose consumption/ethanol production phase, the ethanol consumption and low beta-carotene production phase, and the ethanol/acetic acid and high beta-carotene production phase. By examining the off gas data, it was confirmed that overflow metabolism is occurring. While the experiments performed in this work showed that the strain was underperforming compared with historical data, the current strain behavior was quite reproducible. In the future, it may be desired to perform more troubleshooting to determine how to improve the performance, including by following up on previous media optimization work performed in the lab [51].

4. STEADY-STATE FLUX BALANCE ANALYSIS

4.1 Introduction

Flux balance analysis is a powerful mathematical tool, used to analyze the flow of metabolites through a network of intracellular reactions [54]. This tool has been improved and developed over the years, and used hand in hand with other computational and experimental techniques to better understand the inner workings of native and recombinant organisms [55-57]. For instance, Plata et al studied the *Plasmodium falciparum* metabolic network to make predictions regarding malaria virulence [58]. In another study, Xu et al examined recombinant metabolic plant pathways in microbes using flux balance analysis [59].

In this work, it was desired to predict the intracellular metabolite fluxes of a beta-carotene-producing recombinant strain of *S. cerevisiae* using flux balance analysis (FBA). Flux is given in units of mmol/gDCW-h, and can be estimated by using the COnstraint-Based Reconstruction and Analysis (COBRA) toolbox [60-62]. This tool facilitates the solving of a linear optimization problem. In this work, the COBRA toolbox was used in Matlab® (Mathworks®) software. The system of linear equations solved using the COBRA toolbox is a group of all of the stoichiometric reactions present in a cell. A cellular objective (such as growth rate or productivity) is maximized, subject to substrate uptake, flux constraints, stoichiometry, and most importantly, a pseudo-steady state assumption. Equation 9 describes this optimization. Y, or the objective flux indicated by v_i, is maximized. The term v indicates the array of all the flux values within

the model, while S indicates a matrix containing the stoichiometric information for all the reactions within the model. The equation line $\sum S \cdot v = 0$ is the mathematical representation of the pseudo-steady state assumption, which indicates that the sum of all of the forward fluxes is equal to the sum of all of the reverse fluxes within the cell and entering or exiting the cell. Each individual reaction's flux, including substrate uptake fluxes, can be indicated by V, and the minimum and maximum values can be supplied by the user if they are known.

$$\max_{v} \quad Y = \sum_i w_i v_i \qquad\qquad (9)$$
$$subject\ to:$$
$$\sum S \cdot v = 0$$
$$v_{min} \leq V \leq v_{max}$$

The output of this optimization is a vector containing the flux values of all intracellular and exchange reactions of the cell. The growth rate is given in units of inverse hours (h^{-1}), and all other flux values are given as specific molar rates $\left(\frac{mmol}{gDCW*h}\right)$. This work uses a stoichiometric metabolic model of *S. cerevisiae* originally constructed by Herrgard et al with subsequent additions and improvements [63-68]. The model used in this work has been further edited to include beta-carotene production reactions.

4.2 Materials and Methods

The intracellular metabolic model of *S. cerevisiae* used in this study was adapted from a genome-scale network reconstruction developed by consensus. Versions 4 and 7 of this model (yeast_4.05 and yeast_7.6) were used in this work. These models were

edited to include the stoichiometric reactions along the beta-carotene production

pathway facilitated by the added genes from *X. dendrorhous*, as shown in equations 24-

27.

$$GGPP \xleftrightarrow{crtYB} 2\ Phytoene \tag{10}$$

$$Phytoene \xleftrightarrow{crtI} Neurosporene \tag{11}$$

$$Neurosporene \xleftrightarrow{crtI} Lycopene \tag{12}$$

$$Lycopene \xleftrightarrow{crtYB} Beta-carotene \tag{13}$$
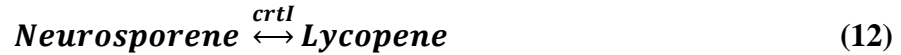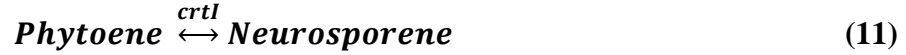
While details about the reaction mechanisms from Moise et al (see section 1.5.1)

were used to better understand the behavior of the genes involved in the beta-carotene

pathway, only basic stoichiometric information was used to inform the metabolic model.

In order to refrain from over-constraining the model, these reactions were defined as

reversible, and kept virtually unbounded (i.e. the upper and lower bounds of each

reaction were set to 1000 and -1000 $\frac{mmol}{gDCW \cdot h}$, respectively). As stated in section 1.5.1,

beta-carotene is not exported by the cell and accumulates in the cell membrane.

However, for modeling purposes, an additional exchange reaction was added to the

model, indicating a non-reversible transportation of produced beta-carotene out of the

cell, as shown in Equation 14.

$$Beta-carotene\ (intracellular) \rightarrow Beta-carotene\ (extracellular) \tag{14}$$

The glucose and oxygen uptake rates, as well as the ethanol, acetate, and beta-

carotene fluxes were set to various values in order to validate the model and study the

behavior of the cell. Growth and beta-carotene production were each set as objectives in respective studies in order to further understand the coupling of growth and production.

4.3 Results and Discussion

The intracellular yeast models described in section 4.1 were used in conjunction with the COBRA toolbox to study the qualitative behavior of the SM14 yeast strain, to determine theoretical yields of beta-carotene, and to confirm the function of the model in preparation for dynamic flux model optimization. While both the fourth and seventh versions of the intracellular yeast model were used in this study, the newer version of the model (version 7) was found to be more accurate and detailed, and therefore the results from that version are shown hereafter.

In order to study the theoretical yield of beta-carotene in SM14, the yeast model was provided with a non-limiting glucose uptake rate, and the biomass growth rate was set to 0.1 h^{-1}. The objective of the model was changed from growth to the beta-carotene exchange reaction (equation 28). Table 3 shows the inputs and results of the optimization of this yeast model. The flux through the yeast cell can be visualized using a COBRA Toolbox extension called Paint4Net [69]. Figure 13 shows the main glycolysis, TCA cycle, growth, and beta-carotene production pathways in the yeast model that was optimized for beta-carotene production. Many more reactions exist in the model, but any subset of reactions can be mapped using the extension. Clearly, this figure is a bit busy, but can be used to merely confirm flux and flux direction through the desired reactions (green lines denote forward flux through a reaction, blue lines denote reverse). Figure 14 is a zoomed-in portion of Figure 13 showing the flux of glucose, or

31

"s_0565" through "r_1714", or the glucose exchange reaction, as well as the flux

through the beginning of the glycolysis pathway. The green diamond indicates that the

glucose exchange reaction is an exchange reaction (i.e. flux is going into or out of the

cell). Figure 15 is a zoomed in portion of Figure 13 and shows a reaction called

"r_2111" which is the growth rate reaction in the yeast version 7 model. The oval named

"s_0450" denotes the biomass metabolite. The value seen within the green diamond

indicates that the growth rate is 0.1 h^{-1}. Figure 16 is a zoomed-in portion of Figure 13

showing the flux through the added beta-carotene production reactions. The reaction

called "Ex_betacarotene" is again denoted by a green diamond, indicating the transport

of beta-carotene in or out of the cell. The beta-carotene flux is shown to be

$0.5694\frac{mmol}{gDCW*h}$. This value of beta-carotene flux results in the theoretical yields to

glucose and biomass shown in Table 3.

Table 3: Input and output of a COBRA yeast model optimization to determine theoretical yield of beta-carotene to glucose.

Model Input	Value
Glucose Uptake Rate ($\frac{mmol}{gDCW*h}$)	-10
Maximum Growth Rate, $\mu_{max,G}$ (h^{-1})	0.1
Model Output	
Beta-carotene Exchange Rate ($\frac{mmol}{gDCW*h}$)	0.5694
Beta-carotene Yield to Glucose ($\frac{g\ Beta-carotene}{g\ Glucose}$)	0.1697
Beta-carotene Yield to Biomass ($\frac{g\ Beta-carotene}{gDCW}$)	3.0570
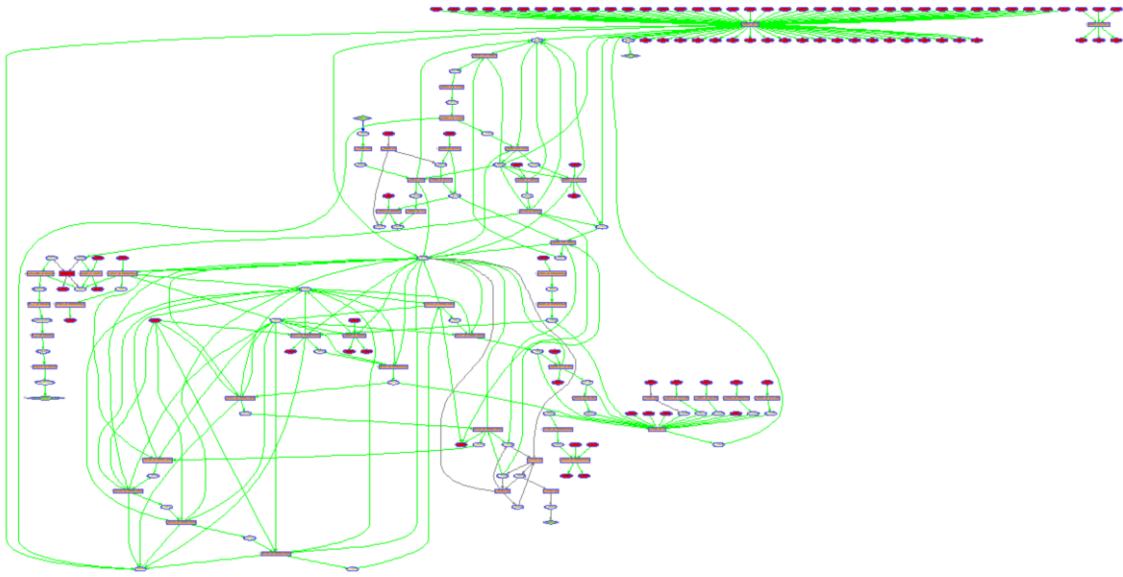
Figure 13: A flux map of the main glycolysis, TCA cycle, growth, mevalonate, and beta-carotene production pathways.
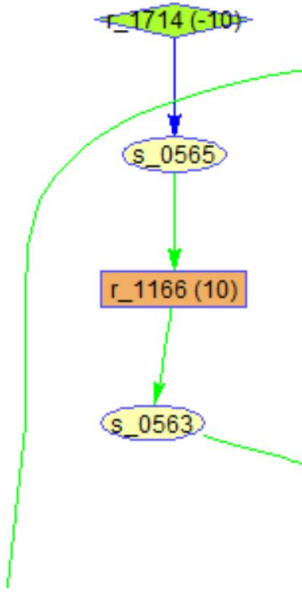
Figure 14: A zoomed-in portion of the flux map in Figure 13, showing the flux of the metabolite "s_0565", or glucose, through the reaction "r_1714", or glucose uptake rate, and continuing through the beginning of the glycolysis pathway.

33
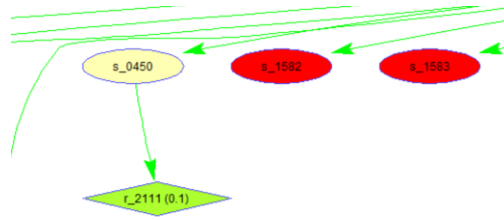
Figure 15: A zoomed-in portion of the flux map in Figure 13, showing the flux of the metabolite "s_0450", or biomass, through the reaction "r_2111", or growth rate.
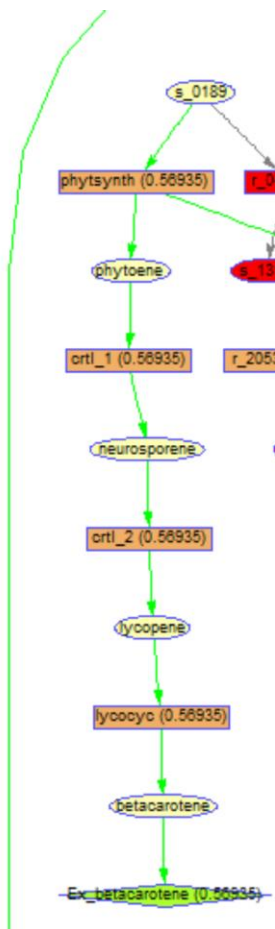
Figure 16: A zoomed-in portion of the flux map in Figure 13, showing the flux through the various reactions in the added beta-carotene production pathway.

In addition to examining the model function with respect to beta-carotene production, the model was tested at a high glucose uptake rate to confirm the function of the overflow metabolism pathways. Table 4 shows the results from this test. Figure 17 shows the main glycolysis, TCA cycle, growth, and beta-carotene production pathways in the yeast model that was optimized for growth when inputting a high glucose consumption rate. As explained above, many more reactions exist in the model, but only this particular subset was mapped. Figure 18 is a zoomed-in portion of Figure 17 showing the flux of glucose, or "s_0565" through "r_1714", or the glucose exchange reaction, as well as the flux through the beginning of the glycolysis pathway. Figure 19 is a zoomed in portion of Figure 13 and shows reaction "r_2111". The value seen within the green diamond indicates that the growth rate is 2.46 h^{-1}. This value is much higher than that of the optimization described above in Table 3. This is due to the much greater glucose uptake rate. This result indicates that the model may not be quantitatively accurate at high glucose uptake rates. Figure 20 is a zoomed-in portion of Figure 17 showing the flux through the ethanol fermentation pathway. The reaction called "r_1761" is denoted by a green diamond, indicating the transport of ethanol in or out of the cell. The ethanol flux is shown to be 190.02. While this value, like the growth rate, appears to not be quantitatively correct, the flux through this pathway indicates that overflow metabolism can be modeled using the COBRA optimization.

Table 4: Input and output of a COBRA yeast model optimization to confirm the function of the model and examine the behavior at high glucose uptake rates.

Model Input	Value
Glucose Uptake Rate ($\frac{mmol}{gDCW*h}$)	114.76
Oxygen Uptake Rate ($\frac{mmol}{gDCW*h}$)	12.81
Model Output	
Ethanol Flux ($\frac{mmol}{gDCW*h}$)	190.02
Growth Rate, $\mu_{max,G}$ (h^{-1})	2.46
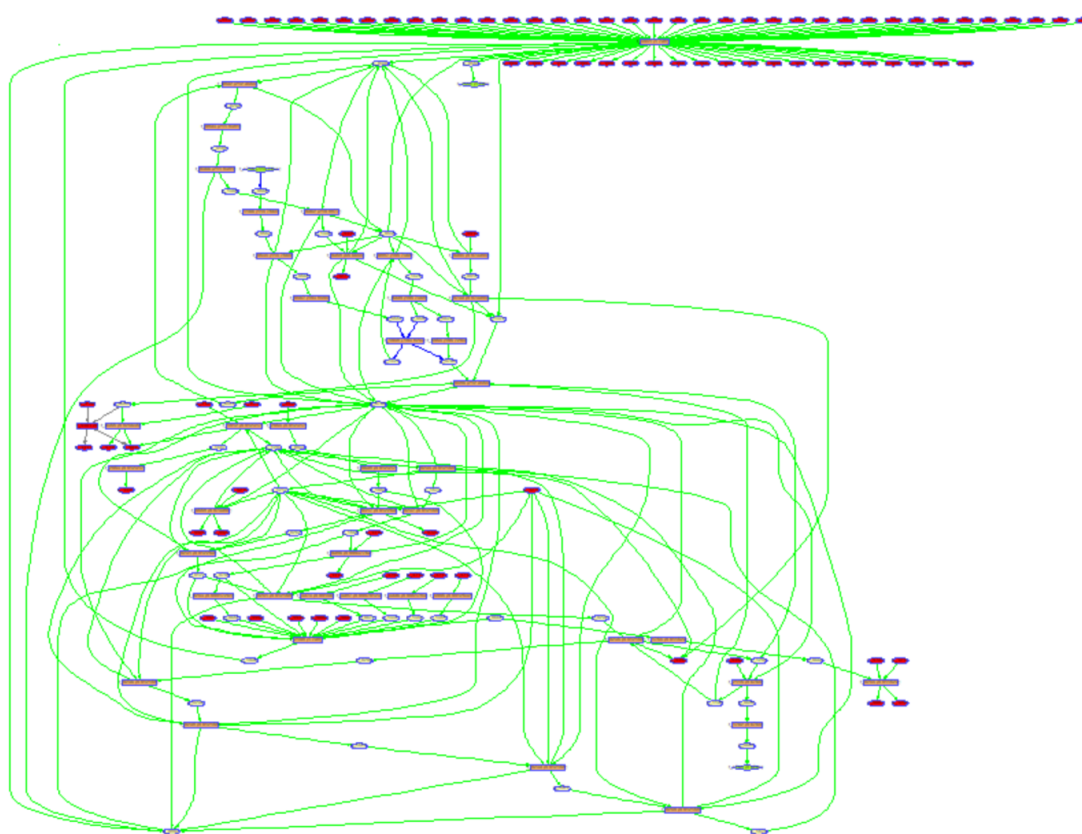
Figure 17: A flux map of the main glycolysis, TCA cycle, growth, and mevalonate pathways.
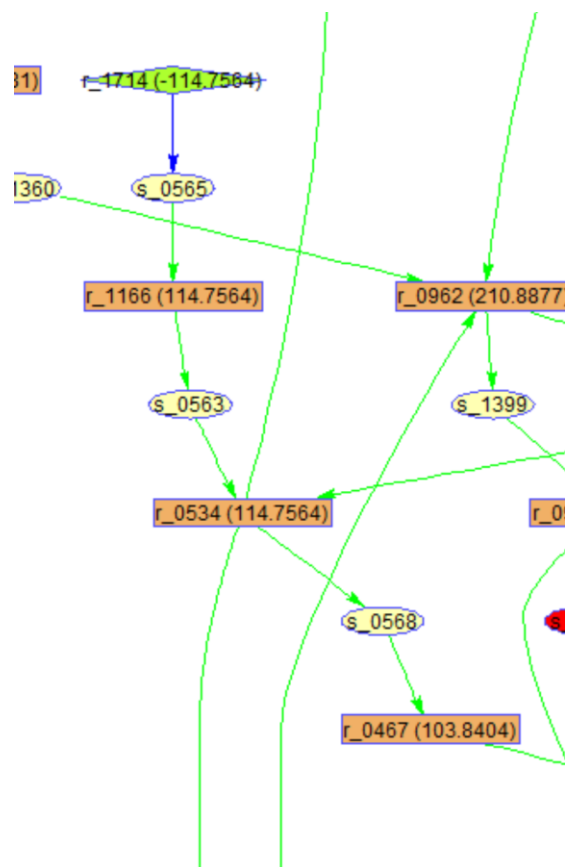
Figure 18: A zoomed-in portion of the flux map in Figure 17, showing the flux of the metabolite "s_0565", or glucose, through the reaction "r_1714", or glucose uptake rate, and continuing through the beginning of the glycolysis pathway.
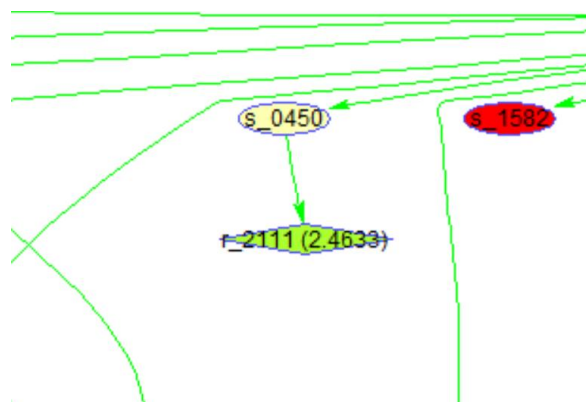
Figure 19: A zoomed-in portion of the flux map in Figure 17, showing the flux of the metabolite "s_0450", or biomass, through the reaction "r_2111", or growth rate.

37
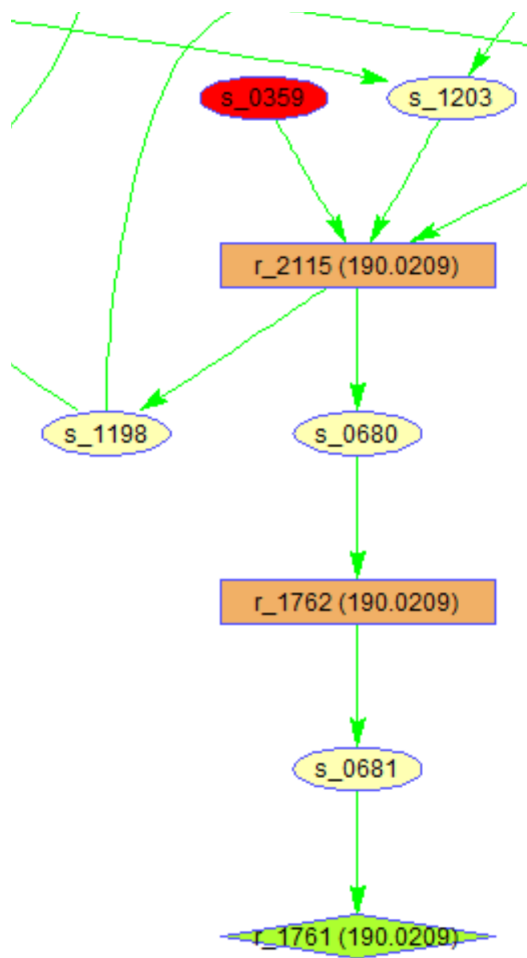
Figure 20: A zoomed-in portion of the flux map in Figure 17, showing the flux of the metabolite "s_0681", or ethanol, through the reaction "r_1761", or ethanol flux.

4.4 Conclusion

The methods developed in this study make it possible to better understand the qualitative intracellular behavior of a recombinant strain of *S. cerevisiae* that produces beta-carotene. Using the COBRA optimization toolbox with a metabolic model of *S. cerevisiae*, the beta-carotene yield to glucose was found to be 0.0194 g beta-carotene/g

glucose, and the beta-carotene yield to biomass was found to be 0.258 g beta-carotene/g biomass. It was also determined that the metabolic model could qualitatively demonstrate overflow metabolism. This preliminary work provides useful information, and also lays the groundwork for future flux balance analysis work, discussed in chapters 5, 6, and 7.

5. DYNAMIC FLUX BALANCE ANALYSIS

5.1 Introduction

In addition to the steady-state flux balance analysis previously performed, it was desired to predict the changes in intracellular fluxes over the course of the beta-carotene production bioreactor process. In this work, the COBRA model is optimized dynamically in order to predict changes in the flux of key pathway reactions over time. Dynamic flux balance analysis may be used to study many different organisms [70, 71]. These analyses have been used to study *E. coli* metabolism, such as the work done by Meadows et al and Mahadevan et al. Meadows et al developed a reactor-scale model of metabolism and growth consisting of 37 process variables and 123 intracellular reactions was developed to study industrial *E. coli* fermentation [72]. Mahadevan et al studied the diauxic growth in *E. coli* using dynamic flux balance analysis of an 85-reaction flux model. Dynamic flux balance analysis has also been used in conjunction with a kinetic model to examine the metabolism of *Shewanella oneidensis* MR-1 by Feng et al [73]. Dynamic flux balance analysis has been used by Hjerseted et al to study a fed-batch ethanol production process in *S. cerevisiae* [74, 75]. Here, dynamic flux balance analysis is used to model the multi-substrate, batch beta-carotene production process in *S. cerevisiae*. A full metabolic model containing 3493 reactions and 2220 metabolites was used in conjunction with a kinetic model of biomass, substrate, byproduct, and beta-carotene concentrations.

5.2 Materials and Methods

5.2.1 Kinetic Model Devlopment

The extracellular kinetic model used in conjunction with the newly-developed dynamic intracellular flux model was developed in our lab by Ordonez et al [52]. This model uses a modified version of the Monod equation to model the SM14 strain's use of glucose, ethanol, and acetic acid as carbon sources during beta-carotene production [76]. Equations for glucose, ethanol, and acetic acid consumption/production rates are also included in the model. The beta-carotene production is estimated using the Luedeking-Piret equation [77]. The model equations developed by Ordonez et al are shown in Equations 15-19. The biomass, glucose, ethanol, acetic acid, and product (beta-carotene) concentrations are represented by X, G, E, A, and P, respectively. The growth rates μ_G, μ_E, and μ_A represent growth on glucose, ethanol, and acetic acid, respectively, and they are shown in Equations 20-22. The symbol Y indicates yield on the respective carbon sources. The symbol $a_{i,j}$ represents the inhibition of the jth substrate on the organism's use of the ith substrate, and β is the coefficient for the non-growth-associated production of beta-carotene. The symbols χ_E and χ_A represent the inhibition of ethanol and acetic acid on the growth rate on glucose. The optimal values of the yields, along with the other parameters, can be re-estimated using experimental bioreactor data. The output of this model includes parameter estimations for all equations as well as growth rate and volumetric titer values for glucose, ethanol, acetic acid and beta-carotene.

$$\frac{dX}{dt} = r_X = (\mu_G + \mu_E + \mu_A) X \qquad (15)$$

$$\frac{dG}{dt} = r_G = -\frac{\mu_G X}{Y_{X/G}} \qquad (16)$$

$$\frac{dE}{dt} = r_E = k_1 \mu_G X - \frac{\mu_E X}{Y_{X/E}} \qquad (17)$$

$$\frac{dA}{dt} = r_A = (k_2 \mu_G + k_3 \mu_E) X - \frac{\mu_A X}{Y_{X/A}} \qquad (18)$$

$$\frac{dP}{dt} = r_P = (\alpha_1 \mu_G + \alpha_2 \mu_E + \alpha_3 \mu_A) X + \beta X \qquad (19)$$

$$\mu_G = \left(\frac{\mu_{max,G} \chi_E \chi_A G}{K_{SG} + G + a_{ge} E + a_{ga} A} \right) \qquad (20)$$

$$\mu_E = \left(\frac{\mu_{max,E} E}{K_{SE} + E + a_{eg} G + a_{ea} A} \right) \qquad (21)$$

$$\mu_A = \left(\frac{\mu_{max,A} A}{K_{SA} + A + a_{ag} G + a_{ae} E} \right) \qquad (22)$$

$$\chi_E = 1 - 4.1 \times 10^{-6} \cdot E^3 + 1.4 \times 10^{-4} \cdot E^2 - 9.1 \times 10^{-3} \cdot E \qquad (23)$$

$$\chi_A = 1 - 0.011 \cdot A^2 - 0.021 \cdot A \qquad (24)$$

5.2.2 Intracellular Model Development

The intracellular metabolic model of *S. cerevisiae* used in this study was adapted from a genome-scale network reconstruction developed by Dobson et al [65]. The seventh version of this model (yeast_7.6) was used as a starting point [64]. This model was modified to include the stoichiometric reactions along the beta-carotene production pathway facilitated by the added genes from *X. dendrorhous*, as shown in Equations 25-28.

$$GGPP \xLeftrightarrow{crtYB} 2\ Phytoene \qquad (25)$$

$$Phytoene \xleftrightarrow{crtI} Neurosporene \qquad (26)$$

$$Neurosporene \xleftrightarrow{crtI} Lycopene \qquad (27)$$

$$Lycopene \xLeftrightarrow{crtYB} Beta - carotene \qquad (28)$$

Information about the mechanisms of each of these reactions, as well as rates and reversibility, is available in the study done by Moise et al [49]. While these details were used to better understand the reactions, only basic stoichiometric information was used to inform the metabolic model. In order to refrain from over-constraining the model, these reactions were defined as reversible, and kept virtually unbounded (i.e. the upper and lower bounds of each reaction were set to 1000 and -1000 $\frac{mmol}{gDCW \cdot h}$, respectively). For modeling purposes, an additional exchange reaction was added to the model, indicating a non-reversible transportation of produced beta-carotene, as shown in Equation 29.

$$Beta - carotene\ (intracellular) \rightarrow Beta - carotene\ (extracellular) \qquad (29)$$

5.2.3 Dynamic Flux Balance Analysis Procedure

By combining the information gained from the previously-developed kinetic model with flux balance analysis techniques, it is possible to analyze intracellular fluxes and changes in the flux throughout the beta-carotene production process. This is accomplished by first converting the output of the kinetic model into the proper flux units and discretizing the data into time increments. The oxygen uptake rate is also needed as an input into the model, and carbon evolution rate is used for validation of the model. Because this information about the contents of the off gas of the bioreactor are

43

not included in the previously developed extracellular kinetic model, the experimental

data is taken directly from the database of collected bioreactor data, smoothed,

discretized, and converted to flux units. All of these data are then used as an input to the

model, in the optimization, and/or as confirmation of the model's results. An outer

optimization around the FBA optimization is created, in which the difference between

the specific growth rates from the kinetic model (v_x^M) and the inner FBA optimization

(v_x^C) is minimized. This outer optimization is shown below.

$$\min Z = \left(v_X^M - v_X^C\right)^2 \qquad (30)$$

This minimization is accomplished by optimizing the values of the weights (w)

in the objective function of the inner FBA optimization. The genetic algorithm solver in

Matlab® is used for this outer optimization. After this dual-level optimization is complete

at each time point, the final output is the values of all intracellular and exchange reaction

fluxes at each time point$\left(v_{i,j}\right)$. This procedure is summarized in Figure 21, and the

objective function is discussed in subsequent sections.
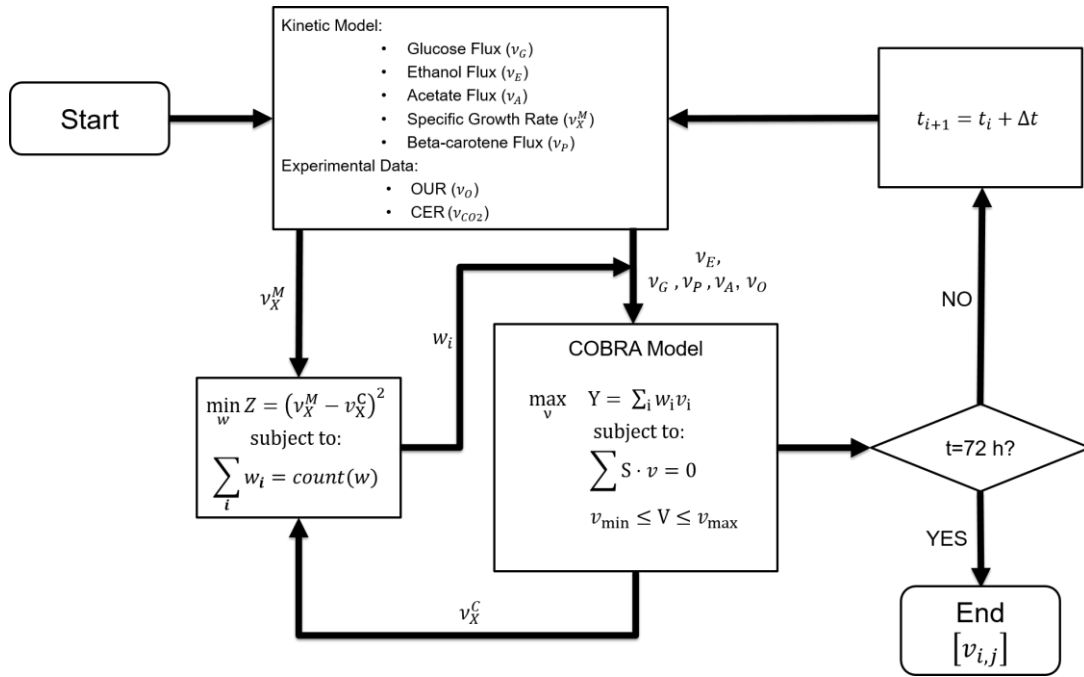
Figure 21: Flowchart depicting the dynamic FBA procedure. Superscript "M" indicates value calculated by extracellular kinetic model. Superscript "C" indicates value calculated by COBRA model optimization. Superscript "E" indicates value calculated from experimental data. "X" denotes biomass.

5.3 Results and Discussion

5.3.1 Bioreactor Data and Kinetic Modeling

As can be seen in Figure 22a, in the first phase of the process, the yeast consume the available glucose and produce ethanol. This aerobic production of ethanol is due to overflow metabolism, also known as the Crabtree effect [34, 35, 46, 47]. After the glucose has been consumed, the cells begin to use the ethanol as a carbon source during the second phase of the process. When the ethanol has been consumed and beta-carotene production has begun, the cells utilize a portion of the acetic acid to produce more beta-carotene in the third phase of the process. In addition to ethanol production, the Crabtree effect can be seen by examining the respiratory quotient (RQ). The RQ is

45

the ratio of the carbon evolution rate (CER) to the oxygen uptake rate (OUR), as shown in Equation 31.

$$RQ = \frac{CER\left(\frac{mmol}{gDCW*h}\right)}{OUR\left(\frac{mmol}{gDCW*h}\right)} \tag{31}$$

As shown in Figure 22b, the RQ values during the first phase of the process (ethanol production) are greater than one, and then decrease to values of less than one. Values of RQ that are greater than one indicate some fermentative metabolism (ethanol production) [34, 35]. Using the kinetic model and the bioreactor experiments described above, the parameters of the model were estimated, and the model was fitted to the experimental data. The model fit and the experimental data are shown in Figure 23. The estimated optimal parameters for the kinetic model are provided in the Appendix.
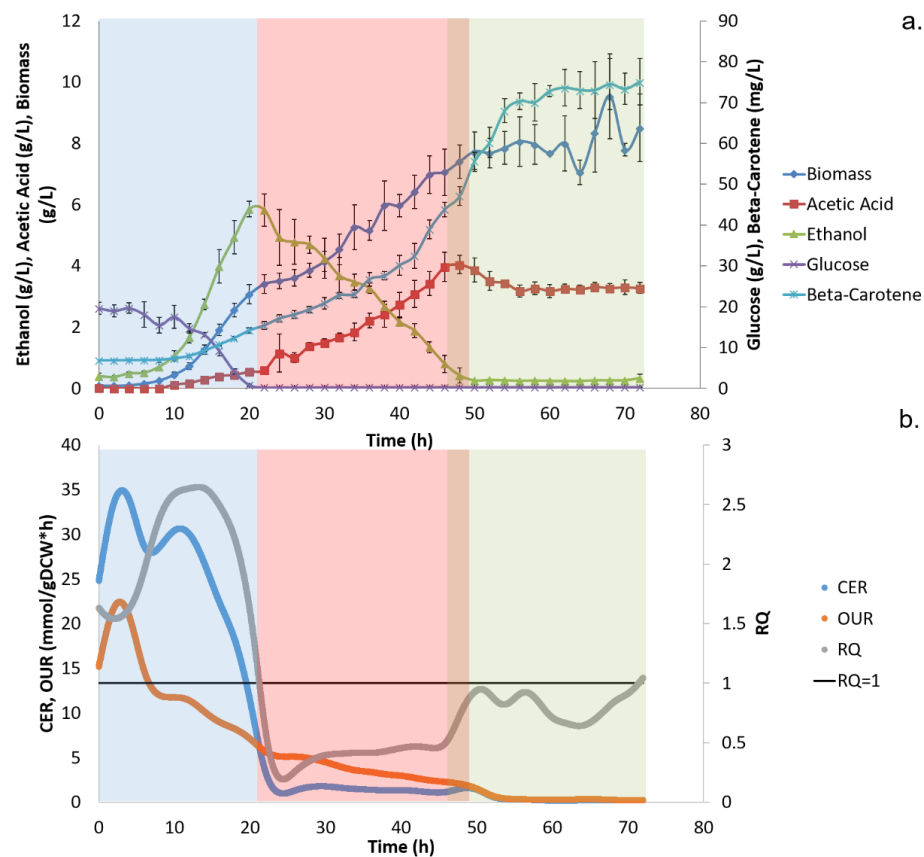
Figure 22: (a) Average growth and metabolite data for the three bioreactor runs performed in this study. The glucose consumption/ethanol production phase is shown in blue, the ethanol consumption phase is shown in red, and the acetic acid consumption phase is shown in green. The error bars signify one standard deviation from the mean. **(b)** Average CER and OUR values for the three bioreactor runs performed in this study, as well as the calculated RQ values. The fermentative activity caused by the Crabtree effect occurs during the phase shown in blue, while the standard aerobic respiration phase is shown in red and green.
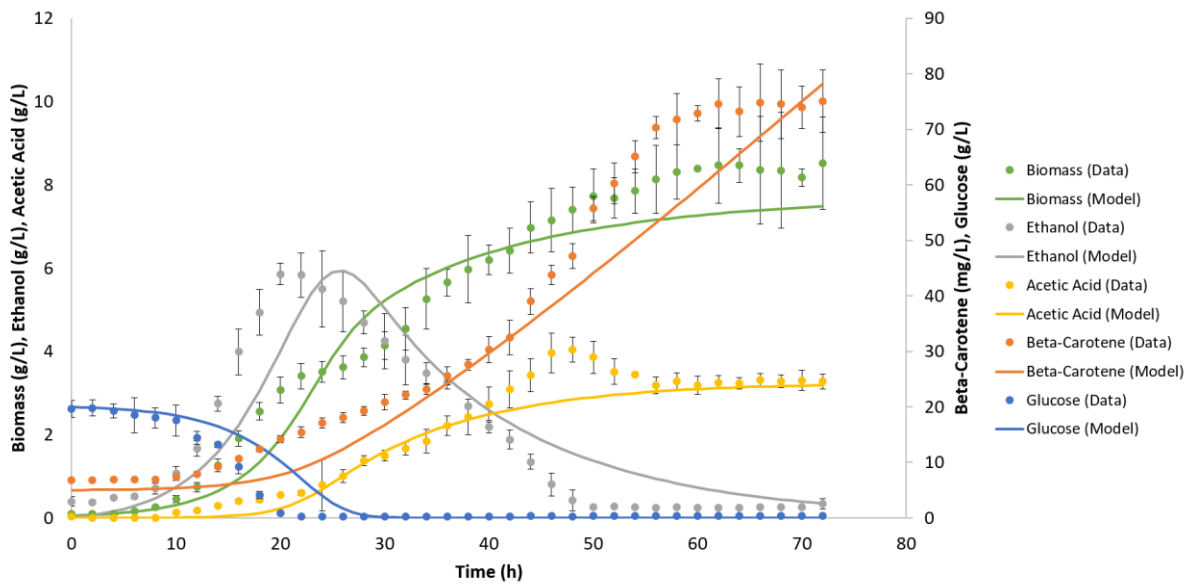
Figure 23: Experimental data (shown in dots) and the kinetic model fit to these data (solid lines). The error bars signify one standard deviation from the mean.

It is of note that when examining the results of the kinetic model, the biomass concentration and beta-carotene concentration appear to continue to increase at the end of the 72-hour bioreactor process. Figure 24 shows the same kinetic model fit, extended to 500 hours. As the figure shows, the beta-carotene concentration does indeed continue to increase exponentially, and the biomass concentration continues to increase as well. The kinetic model lacks any terms or parameters to describe cell death, which explains this continual increase. Because the biomass continues to increase and beta-carotene production is growth-coupled, the lack of any description of cell death likely also contributes to the increasing beta-carotene production. Work has been done, for instance the work done by Kono, to improve upon the Monod equation to describe all of the

48

phases of cell growth, including death [78]. The implementation of the cell growth

kinetics described in Kono's work could help improve the prediction of kinetic model.
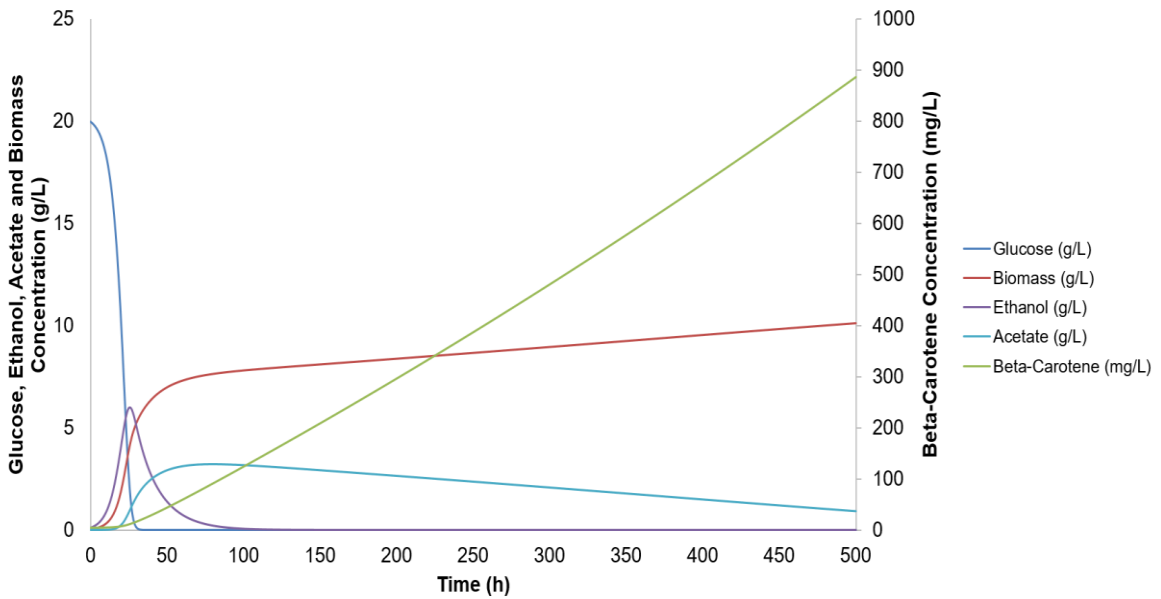
Figure 24: The kinetic model fit to the experimental data, solved to t=500 hr.

5.3.2 Volumetric and Specific Rates

To further analyze the data produced by the kinetic model fitting, the volumetric

rates of the three carbon sources (glucose, ethanol, and acetic acid), as well as the

volumetric rate of beta-carotene production, were calculated and discretized to hour

increments (g/L*h). These data are shown in Figure 25. When examining the volumetric

rates, three distinct phases can be seen. While these phases are slightly different in

length to those found by examining the key concentrations (Figure 12), they depict

similar behavior. During the first phase, the glucose consumption and ethanol production

49

rates increase to their maximum. During the second phase, the glucose consumption and ethanol production rates decrease, and then the rate becomes negative, as ethanol begins to be utilized as a carbon source. In addition, the production rate of beta-carotene continues to increase and an inflection point is seen at the point that the acetic acid production rate reaches a maximum. In the third and final phase of the process, the glucose consumption rate decreases to negligible values, the ethanol consumption rate and acetic acid decreases, and the beta-carotene production rate continues to increase.
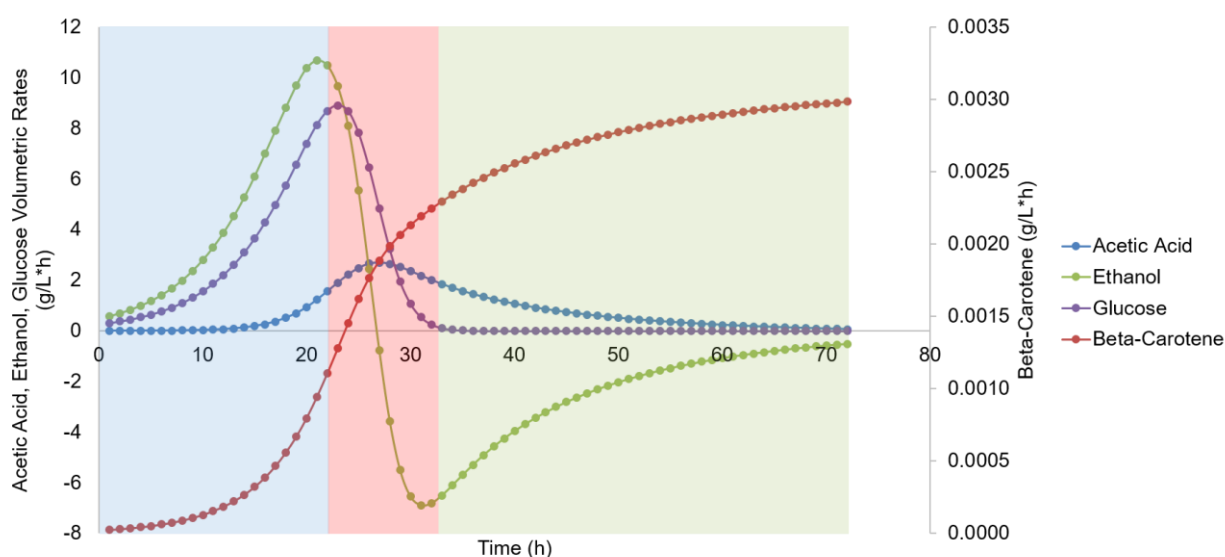
Figure 25: Volumetric rate data for the batch bioreactor process. The three distinct phases described in the text are indicated by the blue (first phase), red (second phase), and green (third phase) shading.

To prepare the data to be used in conjunction with the COBRA model, the results of the kinetic model fitting were used to calculate the specific fluxes (mmol/gDCW*h) of glucose, ethanol, acetic acid, and beta-carotene, as well as the specific growth rate (h^{-1}). These data were discretized to hour increments and are shown in Figure 26, along

with the experimental values of CER and OUR. As seen in Figure 26b, the first and second phases are identified by the shift from ethanol production to ethanol consumption. In Figure 26a, the second phase ends and the third phase begins when the specific acetic acid rate reaches a maximum value, and the specific beta-carotene rate becomes a non-negative value. In the case of the specific beta-carotene rate, the negative values during the beginning of the process do not indicate the uptake of beta-carotene, but rather the relatively negligible amount of beta-carotene produced relative to the amount of biomass produced. Figure 27 shows the beta-carotene flux in more detail, including the portion of the bioreactor run during which the flux is positive.
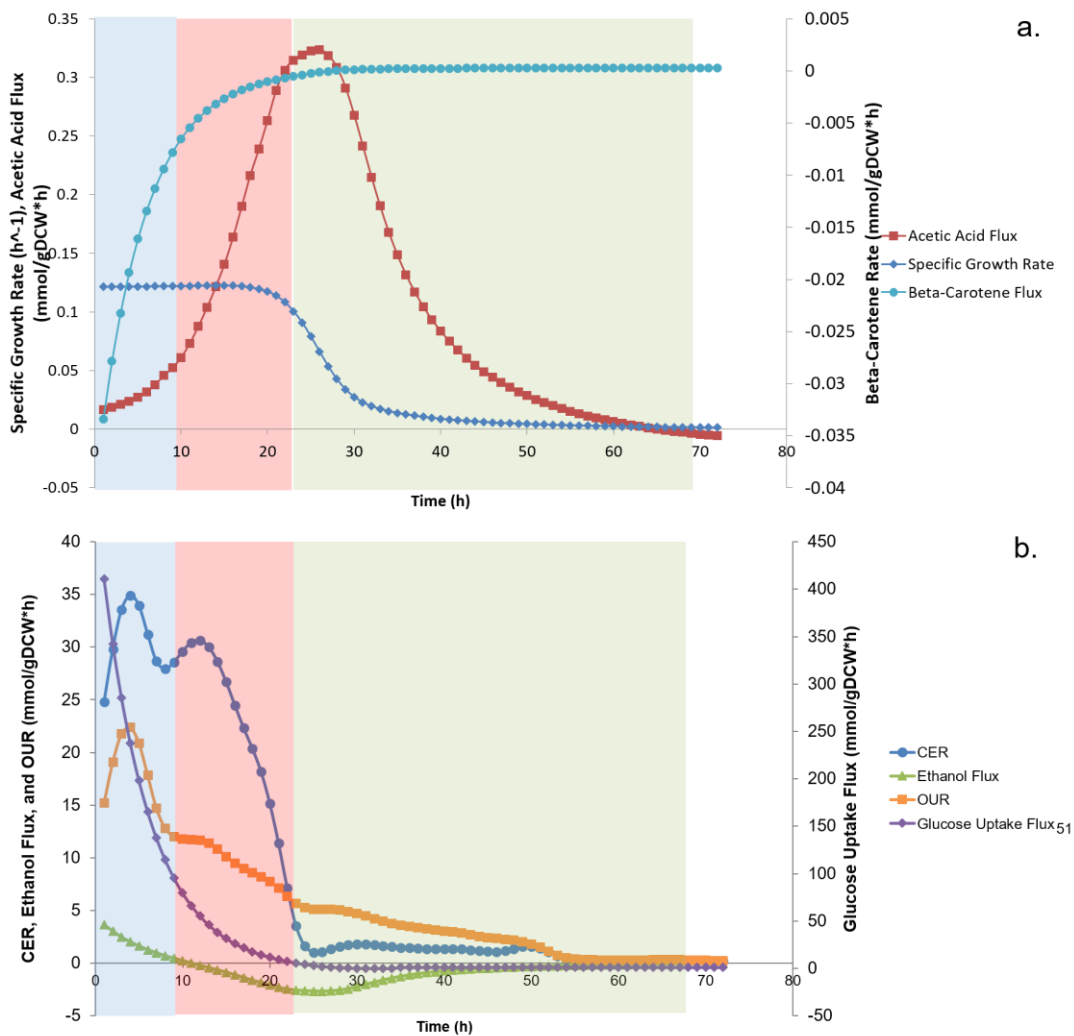
Figure 26: The data generated from the kinetic model, converted to standard flux units (mmol/gDCW*h). (a) Acetic acid, specific growth, and beta-carotene rates. (b) CER, ethanol, OUR, and glucose rates. The three distinct phases described in the text are shown in the blue, red, and green boxes.
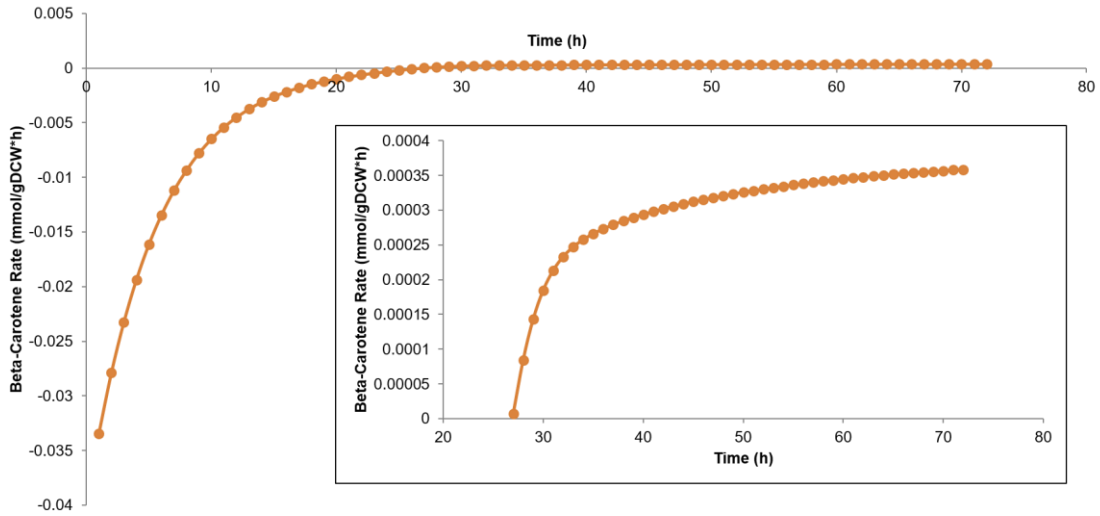
Figure 27: Beta-carotene flux calculated from the kinetic model data. The inset plot depicts the latter portion of the bioreactor run, when the flux of beta-carotene is positive.

5.3.3 Objective Function Development

The standard objective when optimizing the intracellular flux model is to maximize the specific growth rate of the cell [54]. In order to properly represent and study the competitive nature of cell growth and beta-carotene production in the SM14 strain, the objective function was expanded into a multi-objective function. The main metabolic pathways were examined, and important reactions, as well as reactions in branch points at which metabolite flux moves along both growth and production pathways, were identified. These key reactions were added to create a multi-objective function in the form:

$$Y = \sum_i w_i v_i \tag{32}$$

$$\sum_i w_i = count(w) \tag{33}$$

Each objective reaction flux (v_i) is weighted, with the sum of the weights (w_i) equal to the number of reactions in the objective function.

After examining the volumetric and specific flux data and understanding the importance of the beta-carotene production phases specifically, it was decided to focus the dynamic flux balance analysis on the latter portion of the process, including the higher beta-carotene production and the utilization of both ethanol and acetic acid. The metabolic pathways along which flux propagates during these phases are different, and therefore different objective functions were created. The objective functions are described in Equations 34 and 35.

$$Y_{phase\ 2} = w_1 v_{growth} + w_2 v_{acetyl-CoA\ synthetase} + w_3 v_{hexokinase} \qquad (34)$$

Equation 34 shows the objective function for the second phase, the ethanol consumption and low beta-carotene production phase. During the ethanol utilization phase of the process, the cell metabolism has shifted to utilizing ethanol as a carbon source, and the new terms in the objective function reflects this shift. The acetyl-CoA synthetase reaction, shown in Figure 28, converts one molecule each of acetate, ATP, and coenzyme A to one molecule each of acetyl-CoA, AMP, and pyrophosphate [79]. This reaction is a key step in the ethanol degradation pathway in yeast.
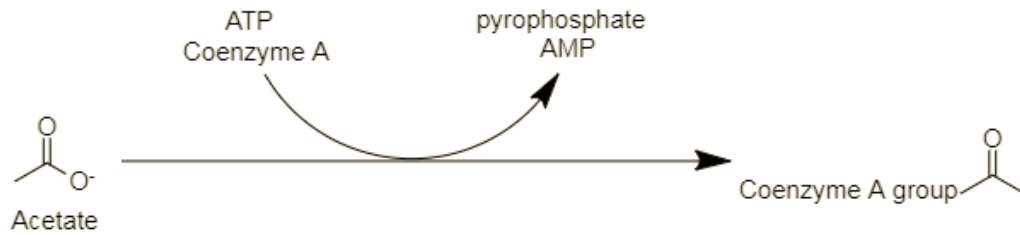
Figure 28: The reaction catalyzed by acetyl-CoA synthetase, as part of the ethanol degradation pathway.

The hexokinase reaction, shown in Figure 29, converts one molecule each of glucose and ATP to one molecule each of glucose-6-phosphate and ADP. This reaction is part of the glucose fermentation pathway, and is also shown to be important when utilizing ethanol as a carbon source [80].
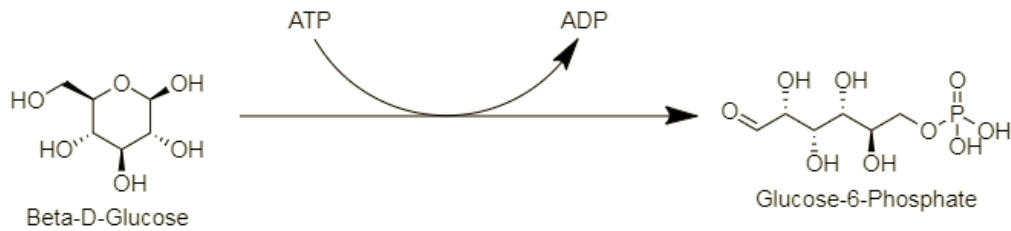
Figure 29: The reaction catalyzed by hexokinase, as part of the glucose fermentation superpathway.

$$Y_{phase\ 3} = w_1 v_{growth} + w_2 v_{beta-carotene} + w_3 v_{farnesyltranstransferase} +$$

$$w_4 v_{acetyl-CoA\ synthetase} + w_5 v_{hexokinase} \qquad (35)$$

Equation 35 shows the objective function for the high beta-carotene production phase. The second term is the weighted value of flux through the beta-carotene exchange reaction, or the rate at which beta-carotene is being produced. The third term in the

objective function is the weighted value of the flux through the farnesyltranstransferase

reaction. This reaction, shown in Figure 30, converts an isopentenyl diphosphate

molecule and a trans,trans-farnesyl diphosphate molecule to geranygeranyl diphosphate

and a pyrophosphate molecule. This reaction is a branch point at the end of the

mevalonate pathway and the larger ergosterol biosynthesis pathway, responsible for

maintaining a large portion of the plasma membrane [81]. The fourth and fifth terms of

the objective function are the acetyl-CoA and hexokinase reactions, described above.
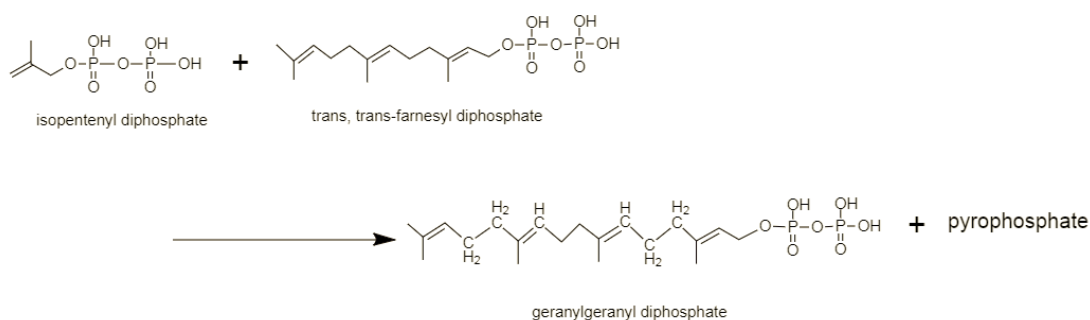
Figure 30: The reaction catalyzed by farnesyltranstransferase.

5.3.4 Dynamic Flux Balance Analysis Results and Discussion

The following results depict the final 50 hours of the production process, which

include the higher beta-carotene production and the utilization of both ethanol and acetic

acid. Figure 31 shows the initial results of the dynamic flux balance analysis. The

specific growth rate calculated by the kinetic model is shown in blue, and the output of

the dynamic flux balance analysis is shown in red. In the procedure depicted in Figure

21, the calculated CER was not supplied to the metabolic model. Instead, the CER in the

metabolic model was allowed to vary as a check of the accuracy of the dynamic flux balance analysis procedure. Figure 32 shows the comparison of the experimental CER values and the results of the COBRA model optimization.
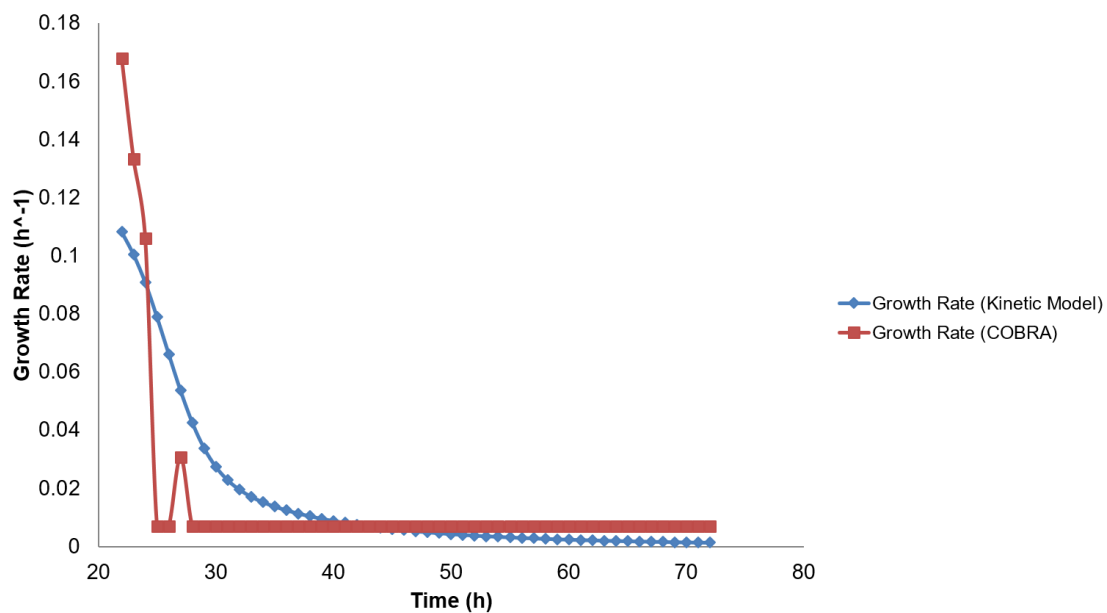
Figure 31: A comparison of specific growth rates calculated from the kinetic model fitting and from the COBRA model optimization, from hours 22 to 72 of the bioreactor run.
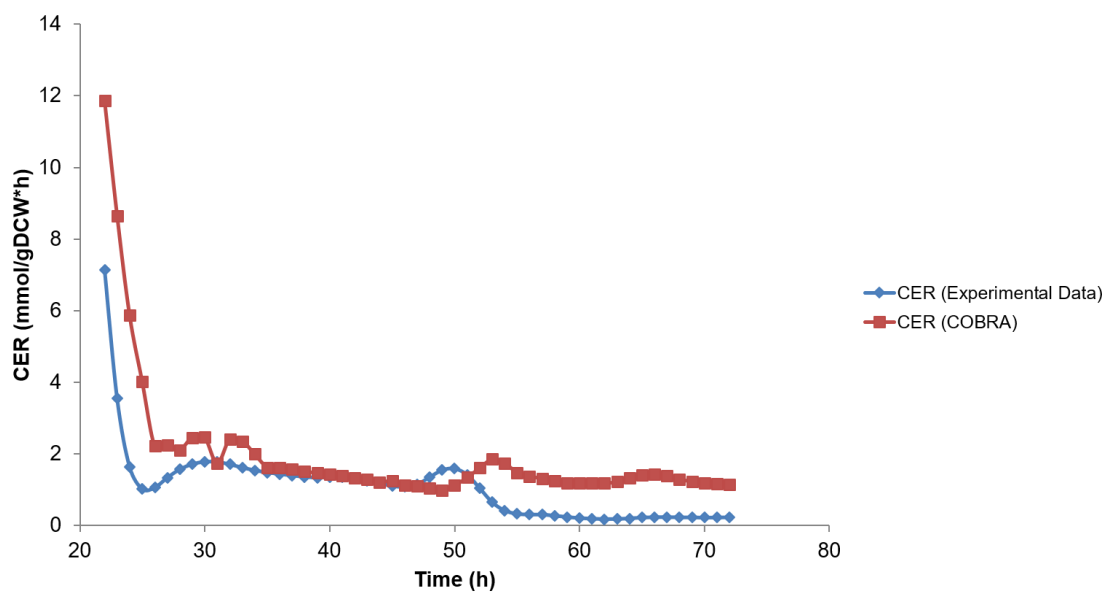
Figure 32: A comparison of carbon evolution rate (CER) calculated from the kinetic model fitting and from the COBRA model optimization, from hours 22 to 72 of the bioreactor run.

In addition to the dynamic flux balance analysis procedure described in Figure 21, a similar procedure is developed in an attempt to increase the accuracy of the model. In this procedure, shown in Figure 33, all of the kinetic model and experimental data are supplied to the metabolic model. The objective function is set to its default form, maximizing the specific growth rate. The outer optimization loop, minimizing the difference between the growth rate calculated from the kinetic model and the growth rate from the COBRA model optimization, remains. The growth rate comparison using this procedure is shown in Figure 34.
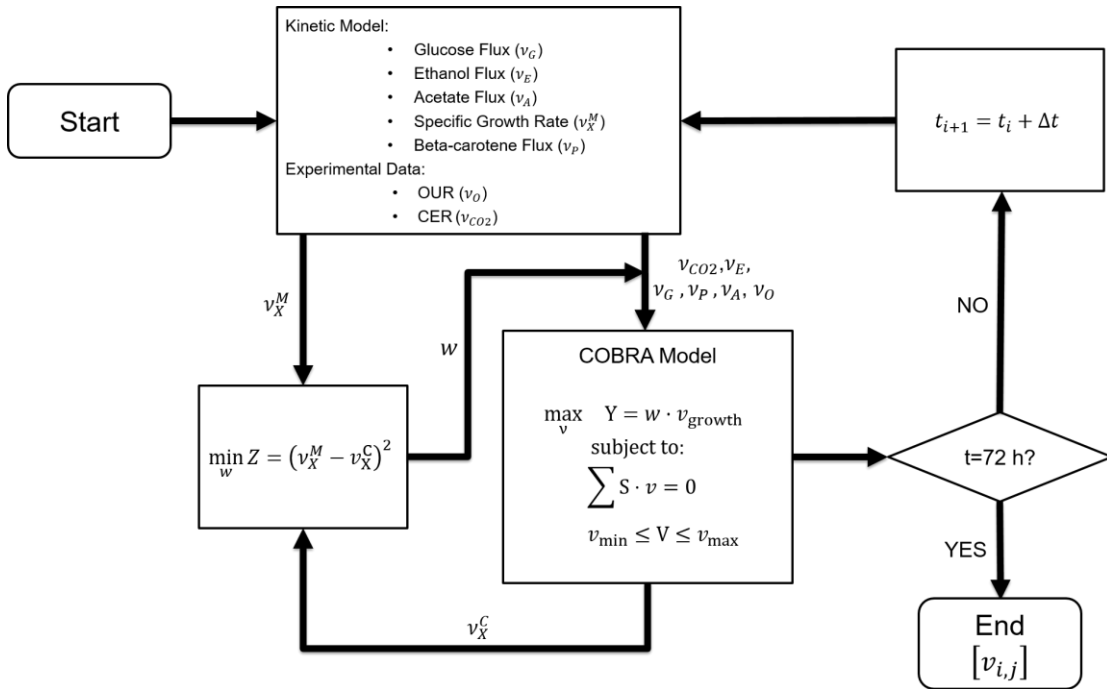
Figure 33: Flowchart depicting the alternate dynamic FBA procedure. Superscript "M" indicates value calculated by extracellular kinetic model. Superscript "C" indicates value calculated by COBRA model optimization. Superscript "E" indicates value calculated from experimental data. "X" denotes biomass.
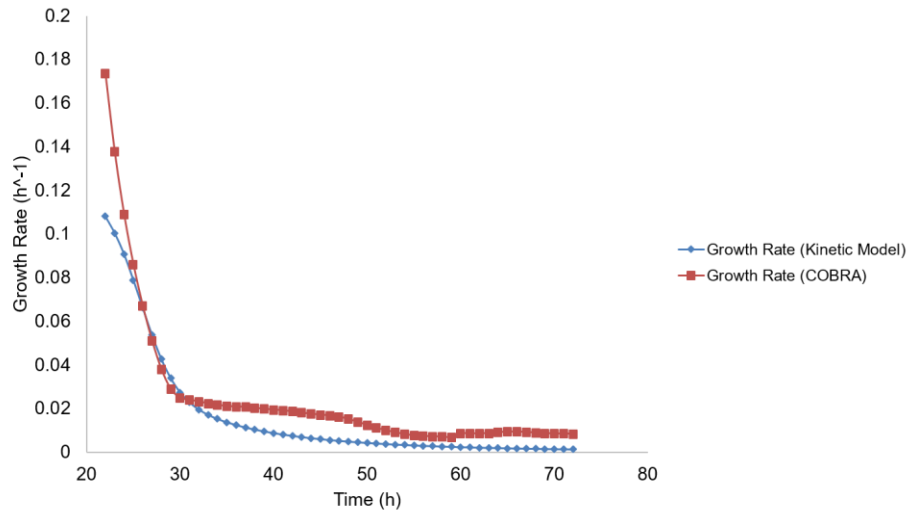
Figure 34: A comparison of specific growth rates calculated from the kinetic model fitting and from the alternate COBRA model optimization, from hours 22 to 72 of the bioreactor run.

The goal of both of the dynamic flux balance analysis methods used was to examine the behavior of several key pathways within the cell. The central metabolism of the cell was examined, including the glycolysis pathway, the beginning of the ethanol production pathway, and several reactions within the TCA cycle, shown in Figure 35. The key reactions within the glycolysis pathway are the phosphofructosekinase (A), phosphoglycerate kinase (B), and enolase (C) reactions [37, 39, 41]. The key reaction in ethanol production is the pyruvate decarboxylase (D) reaction [82]. The citrate synthase (E) and fumarate hydralase (F) reactions within the TCA cycle were also examined [83, 84]. The changes within the ergosterol biosynthesis pathway, which contains the mevalonate pathway, were also studied. The key reactions examined in this pathway are shown in Figure 36. These reactions include two reactions from the mevalonate pathway, hydroxymethylglutaryl-CoA reductase (G) and geranylgeranyltranstransferase (H) [85]. The farnesyltranstransferase (I) reaction at the beginning of the added beta-carotene pathway as well as the squalene synthase (J) reaction were also examined [24]. Finally, the dynamic changes of two important reactions involved in ethanol utilization were studied. The second reaction, acetyl-CoA synthetase (L), is also the main reaction involved in acetic acid utilization. These reactions are shown in Figure 37.

60

Fructose-1,6-Biphosphate ←—— A —— Fructose-6-Phosphate ←—————— Glucose-6-Phosphate

3-Phospho-D-Glycerol-Phosphate

| B

3-Phosphoglycerate

Phosphoenolpyruvate

| C

Pyruvate —→ Oxaloacetic Acid —— E —→ Citrate - - -→ Fumarate —— F —→ Malate

| D

Acetaldehyde

A: phosphofructokinase
B: phosphoglycerate kinase
C: enolase
D: pyruvate decarboxylase
E: citrate synthase
F: fumarate hydralase
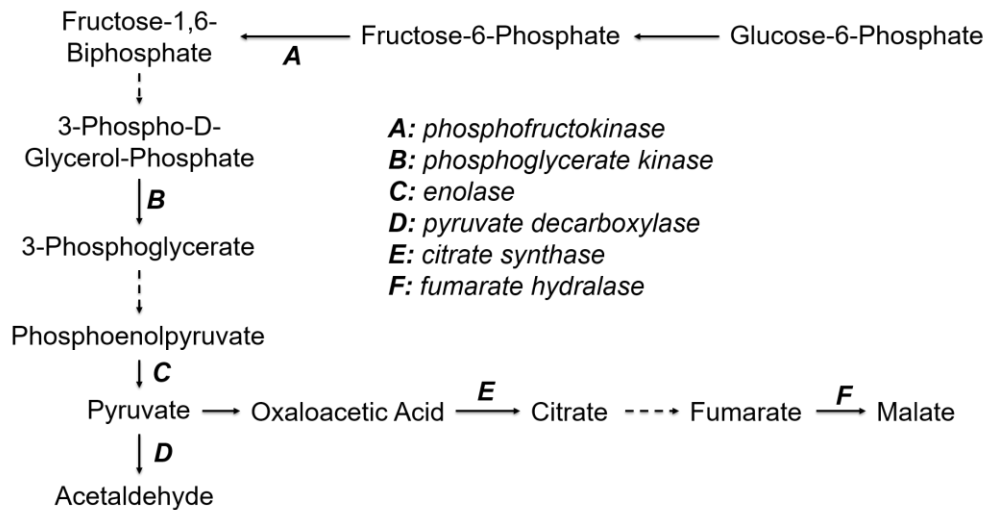
Figure 35: A portion of the central metabolic pathway in *S. cerevisiae*. The labeled reactions include two reactions from the glycolysis pathway (A, B, and C), a reaction from the glucose fermentation pathway (D), and two reactions from the TCA cycle (E and F).

Mevalonic Acid ←—— G —— HMG-CoA ←- - - - 2 Acetyl-CoA

GPP

IPP —→ | H

FPP ——————→ GGPP
IPP | I

J |

Squalene

G: hydroxymethylglutaryl-CoA reductase
H: geranyltranstransferase
I: farnesyltranstransferase
J: squalene synthase
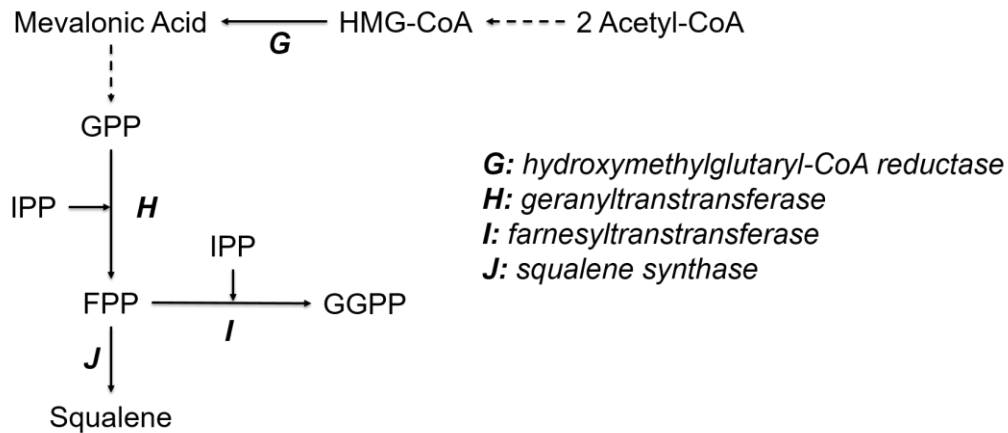
Figure 36: A portion of the ergosterol biosynthesis pathway in *S. cerevisiae*. The labeled reactions include two reactions from the mevalonate pathway (G and H), the key branch reaction at the beginning of the beta-carotene production pathway (I), and the squalene synthase reaction (J).

61

Acetaldehyde ◄─────── Ethanol

|
K
▼

Acetic Acid   **K:** *acetaldehyde dehydrogenase*
                     L: *acetyl-CoA synthetase*
|
L
▼

Acetyl-CoA
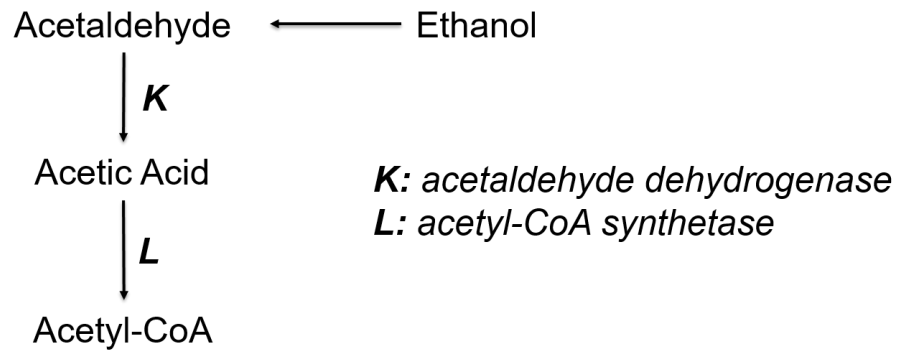
Figure 37: A portion of the ethanol and acetic acid utilization pathway in *S. cerevisiae*.

Figure 38a shows the flux of the reactions of the central metabolic pathway over time as calculated by the original dynamic flux balance analysis procedure. The glycolysis pathway reactions, phosphofructokinase (A), phosphoglycerate kinase (B), and enolase (C) follow the trend of the growth rate throughout the process. The trend of the pyruvate decarboxylase reaction (D) to zero flux corresponds to the shift from ethanol production to ethanol utilization, seen in Figure 25. The TCA cycle reactions, citrate synthase (E) and fumarate hydralase (F), are seen to spike during the high-growth and ethanol utilization phase. It has been seen that pyruvate limitation results in the remaining pyruvate being channeled into the TCA cycle, and this behavior agrees with this [86]. Figure 38b shows the flux of the reactions in the ergosterol biosynthesis pathway over time. As expected, the flux of the farnesyltranstransferase reaction (I) increases from zero when beta-carotene begins to be produced, while the flux of the squalene synthase reaction (J) decreases at the start of beta-carotene production, indicating a shift in flux through the added beta-carotene pathway. Figure 38c shows the flux of the acetaldehyde dehydrogenase and acetyl-CoA synthetase reactions. The

highest fluxes of these reactions are seen during the highest ethanol uptake portion of the

process, as well as the beginning of acetic acid uptake. Increased ethanol uptake results

in increased acetic acid production, and subsequently can increase the pool of acetyl-

CoA [87].
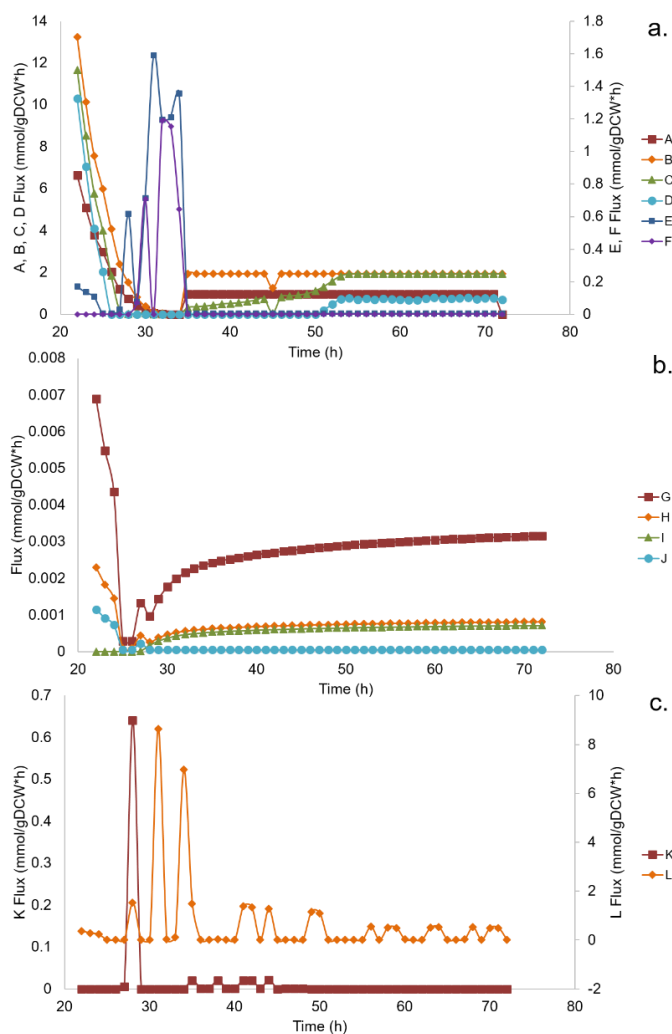
Figure 38: Results from the original (weight-varying) dynamic FBA method. Reactions A, B, C, D, E, and F correspond to the reactions in the central metabolic pathway (Figure 35). Reactions G, H, I and J correspond to the reactions in the ergosterol biosynthesis pathway (Figure 36). Reactions K and L correspond to the reactions in the ethanol utilization pathway (Figure 37).

Figure 39a shows the flux of the reactions of the central metabolic pathway over time as calculated by the alternate dynamic flux balance analysis procedure. Reactions A-E retain the same trends as discussed previously. The fumarate hydralase reaction (F) maintains a higher flux for longer, likely due to the increased time span of the formation of acetic acid seen in Figure39c. Figure39b shows the flux of the reactions in the ergosterol biosynthesis pathway over time. The trends of the flux of reactions G-J remain fairly similar to those seen and described previously. Figure 39c shows the flux of the acetaldehyde dehydrogenase and acetyl-CoA synthetase reactions. In this instance, because the ethanol and acetic acid fluxes are being provided to the model at all time points, the trends of these two reactions follow.
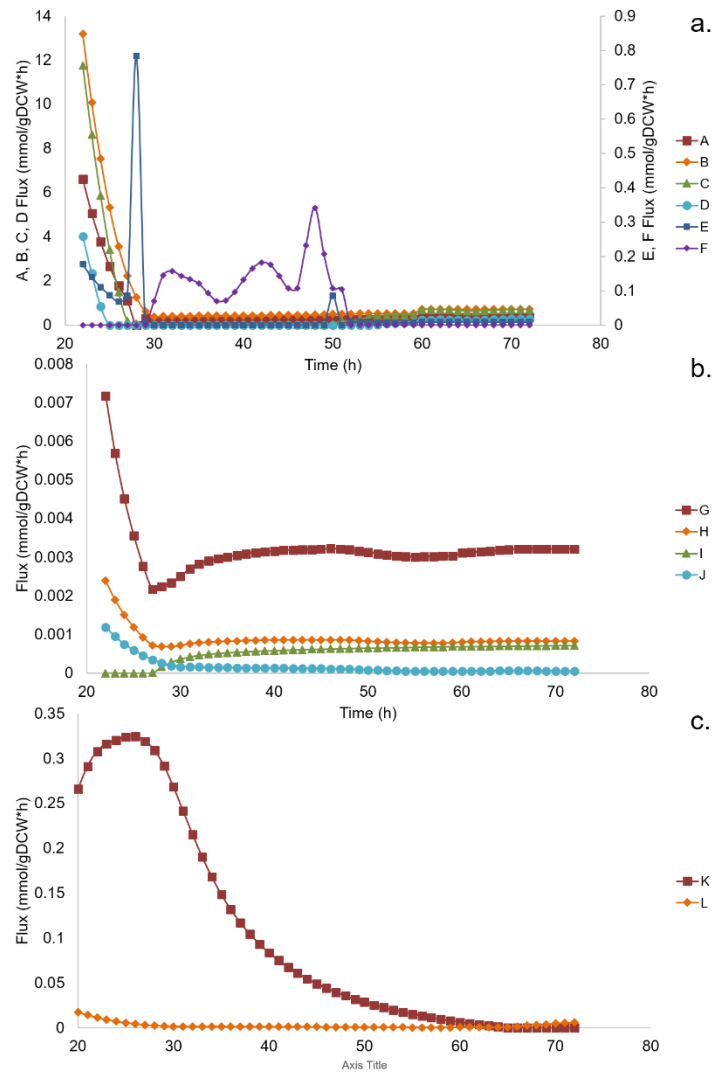
Figure 39: Results from the alternate dynamic FBA method. . Reactions A, B, C, D, E, and F correspond to the reactions in the central metabolic pathway (Figure 35). Reactions G, H, I and J correspond to the reactions in the ergosterol biosynthesis pathway (Figure 36). Reactions K and L correspond to the reactions in the ethanol utilization pathway (Figure 37).

The two dynamic FBA strategies presented make it clear that there is room for improvement in the predictive ability of the dynamic model. An additional improvement that could be made would be the inclusion of constraints associated with gene repression, toxicity, and/or inhibition. One can examine literature or perform

65

experiments to determine if certain metabolites or environmental conditions are repressive, such as done by Covert and Palsson [88]. Any genes found to be repressed can be removed from the model, or reaction fluxes can be constrained to a zero value. For the particular case of beta-carotene production, it may be of use to examine the effect of increased flux through the reactions in the objective functions and include additional information in the model constraints. Specifically, increasing flux through the mevalonate pathway by including farnesyltranstransferase in the objective function will increase isopentenyl pyrophosphate (IPP) concentrations and have a significant effect on the yeast cells, and this can be incorporated into the model [89, 90].

5.4 Conclusion

The methods developed in this study make it possible to better understand the behavior of a recombinant strain of *S. cerevisiae* that produces beta-carotene. The metabolic effects of the batch process were studied, and the intracellular behavior of the strain was examined. Growth, product, and by-product information from a kinetic model was used as input, and the dynamic behavior of intracellular reaction fluxes were estimated. By utilizing the method described, this work may be expanded to any terpene- or terpenoid-producing yeast strain. With minimal preliminary experimentation, kinetic models can be developed for other terpene- or terpenoid-producing yeast strains, intracellular flux models may be edited to add any non-native pathways, and similar dynamic analysis could be performed. The kinetic model could be improved by making further improvements to the description of cell growth such as adding parameters and

terms for cell death. The information gained from these studies would help streamline

and improve the process of producing many valuable products.

6. INTRACELLULAR ANALYSIS OF CONTINUOUS BIOPROCESSING STEADY-STATES

6.1 Introduction

A continuous beta-carotene production process has been proposed for strain SM14. An extracellular kinetic model of this continuous process was developed to describe the recombinant yeast cell growth, the consumption of glucose, the formation and depletion of acetic acid and ethanol, and beta-carotene production. It is desired to further understand the intracellular metabolic behavior of strain SM14 during continuous processing. The extracellular kinetic model is used in conjunction with an intracellular metabolic flux model to examine the behavior of intracellular reactions within the recombinant yeast during continuous beta-carotene production at a range of dilution rates.

6.2 Methods

6.2.1 Continuous Kinetic Model Development

The extracellular kinetic model used in conjunction with the newly-developed dynamic intracellular flux model was developed in our lab [52]. In further work, this batch kinetic model was then adapted to model a continuous process [91]. A dilution rate term was added to each batch equation, and the new system of equations, assuming a single substrate of glucose, is shown in equations 36-40. The equations for the specific growth rates and all the parameters remain the same as the batch system and are described above.

$$\frac{dX}{dt} = r_X = -\frac{F_{out}}{V}X + (\mu_G + \mu_E + \mu_A)X \tag{36}$$

$$\frac{dG}{dt} = r_G = -\frac{F_{out}}{V}(G - G_{in}) - \frac{\mu_G X}{Y_{X/G}} \tag{37}$$

$$\frac{dE}{dt} = r_E = -\frac{F_{out}}{V}E + k_1\mu_G X - \frac{\mu_E X}{Y_{X/E}} \tag{38}$$

$$\frac{dA}{dt} = r_A = -\frac{F_{out}}{V}A + (k_2\mu_G + k_3\mu_E)X - \frac{\mu_A X}{Y_{X/A}} \tag{39}$$

$$\frac{dP}{dt} = r_P = -\frac{F_{out}}{V}P + (\alpha_1\mu_G + \alpha_2\mu_E + \alpha_3\mu_A)X + \beta X \tag{40}$$

Using the continuous kinetic model described in section 1.8, the dilution rate was
set to a value, starting with 0.01 and increasing in increments of 0.005. The kinetic
model was solved using Matlab's ode45 solver and plotted to visualize steady states and
the dilution rate at which washout occurs. Each rate in Equation 36-40 was also set to
zero at each tested dilution rate in order to find steady state concentrations of glucose,
ethanol, acetate, beta-carotene, and biomass. The steady state concentrations of glucose,
ethanol, acetate and beta-carotene were converted to flux (mmol/gDCW*h), and the
values were used as input into the yeast metabolic model. The relevant equations to this
work are shown below.

$$D = -\frac{F_{out}}{V} \tag{41}$$

$$Glucose\ Flux\left(\frac{mmol}{gDCW*h}\right) = D * \frac{\left(1000 * \left(\frac{G}{180.16\frac{g}{mol}}\right)\right)}{X} \tag{42}$$

69

$$\text{Ethanol Flux}\left(\frac{mmol}{gDCW*h}\right) = D * \frac{\left(1000*\left(\frac{E}{46.07\frac{g}{mol}}\right)\right)}{X} \tag{43}$$

$$\text{Acetate Flux}\left(\frac{mmol}{gDCW*h}\right) = D * \frac{\left(1000*\left(\frac{A}{60.05\frac{g}{mol}}\right)\right)}{X} \tag{44}$$

$$\text{Beta} - \text{carotene Flux}\left(\frac{mmol}{gDCW*h}\right) = D * \frac{\left(\left(\frac{P}{536.87\frac{g}{mol}}\right)\right)}{X} \tag{45}$$

Equation 41 represents the dilution rate which is the outlet flow rate of the bioreactor divided by the volume of the reactor. The variables G, E, A, P, and X represent the glucose, ethanol, acetate, beta-carotene, and biomass concentrations, respectively.

6.2.2 Intracellular Model Development

The intracellular metabolic model of *S. cerevisiae* used in this study was adapted from a genome-scale network reconstruction developed by Dobson et al [65]. The seventh version of this model was used in our work [64]. This model was edited to include the stoichiometric reactions along the beta-carotene production pathway facilitated by the added genes from *X. dendrorhous*, as shown in equations 20-23. Information about the mechanisms of each of these reactions, as well as rates and reversibility, is available in the study done by Moise et al [49]. While these details were used to better understand the reactions, only basic stoichiometric information was used in our work to inform the metabolic model. In order to refrain from over-constraining the model, these reactions were defined as reversible, and kept "virtually unbounded" (i.e.

the upper and lower bounds of each reaction were set to 1000 and -1000 $\frac{mmol}{gDCW \cdot h}$,

respectively). For modeling purposes, an exchange reaction was added to the model,

indicating a non-reversible transportation of produced beta-carotene, as shown in

Equation 24. In addition, to allow for ethanol production, the ATP synthase reaction flux

was restricted by setting the upper and lower bounds of the reaction to zero [68].

$$GGPP \xrightarrow{crtYB} 2\ Phytoene \tag{46}$$

$$Phytoene \xrightarrow{crtI} Neurosporene \tag{47}$$

$$Neurosporene \xrightarrow{crtI} Lycopene \tag{48}$$

$$Lycopene \xrightarrow{crtYB} Beta-carotene \tag{49}$$

$$Beta-carotene\ (intracellular) \rightarrow Beta-carotene\ (extracellular)$$
$$\tag{50}$$
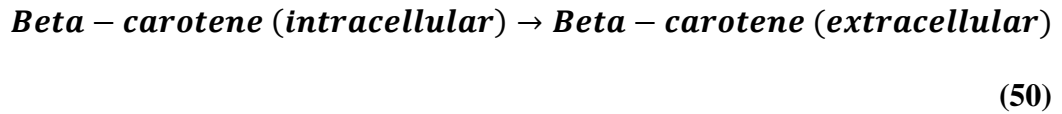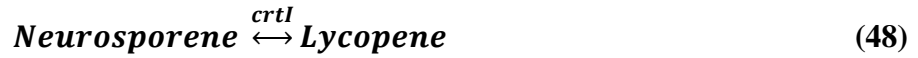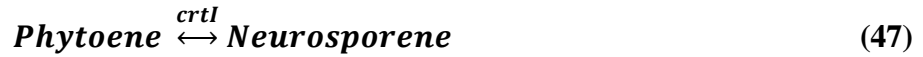
6.3 Results and Discussion

The continuous kinetic model discussed previously and shown in equations 10-

14 was solved at a range of dilution rates between 0.01 and 0.1 h^{-1}. The concentrations of

glucose, biomass, ethanol, acetate, and beta-carotene are plotted over time for the lower

and upper values in the range of dilution rates (Figure 40). The steady state values of the

various states (G, X, E, A, and P) for the range of dilution rates were found and are

plotted in Figure 41.a. Using equations 15-19, the steady state concentration values were

converted to steady state flux values. These flux values are plotted in Figure 41.b. After

examining the flux values, it is noted that while the acetate and ethanol fluxes and the

glucose uptake rate increase with increasing dilution rate, it appears that 0.07 h^{-1} is the

optimal dilution rate for beta-carotene production. Using these flux values as input to the intracellular metabolic flux model, the model was optimized, as described previously, using equation 9. The cellular objective used was the specific growth rate. Figure 42 shows the value of the objective at each dilution rate.
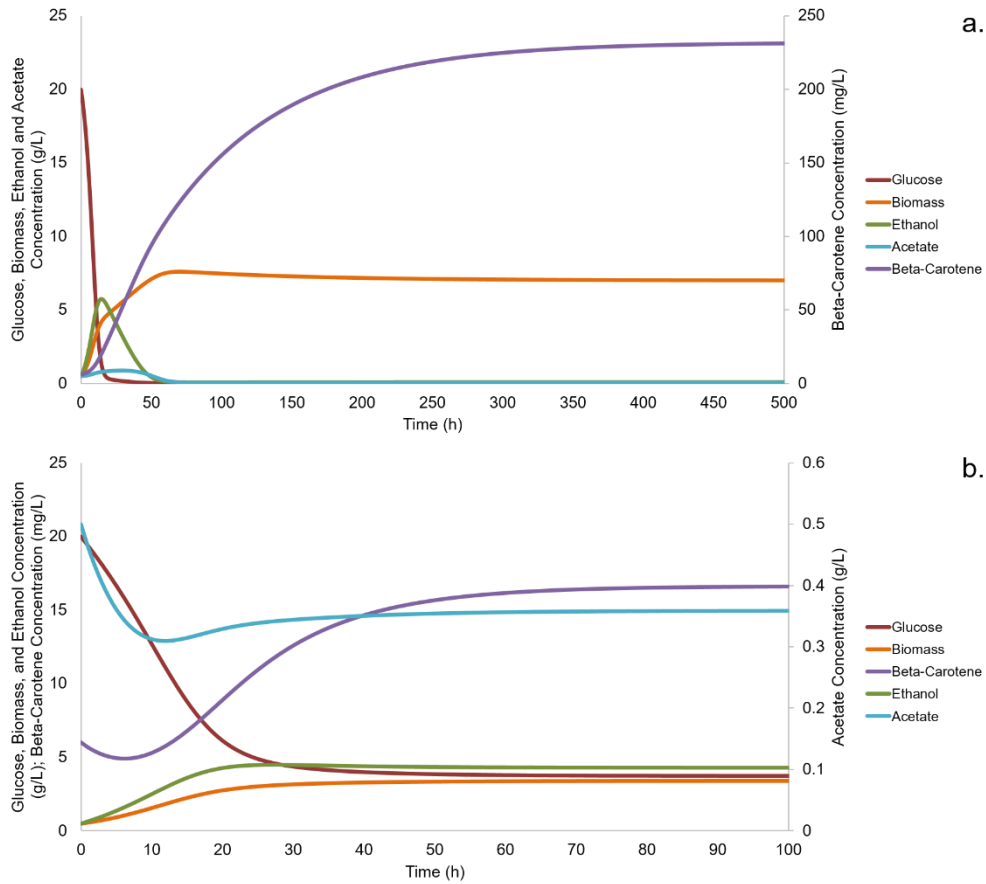
Figure 40: Concentration of glucose, biomass, ethanol, acetate, and beta-carotene over time calculated by the continuous kinetic model with a dilution rate of 0.01 h^{-1} (a) and 0.1 h^{-1} (b).
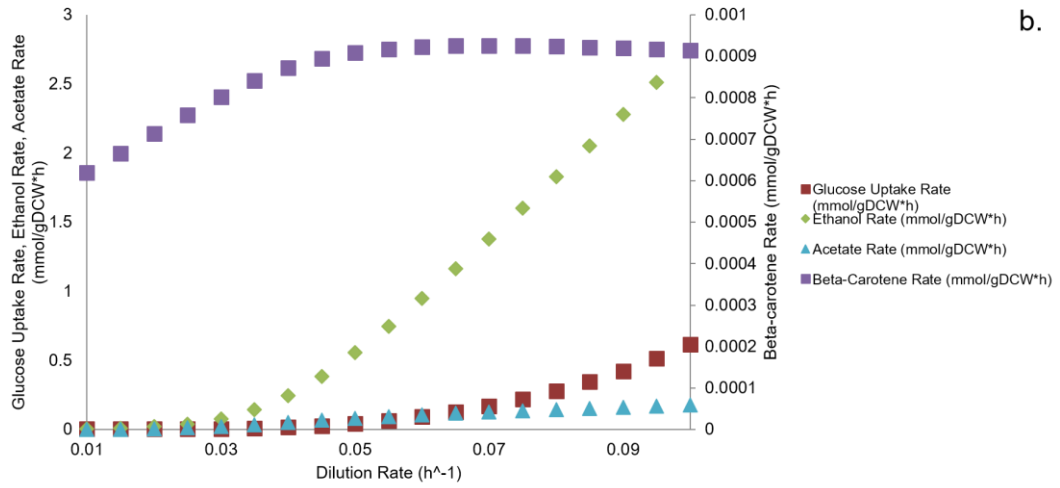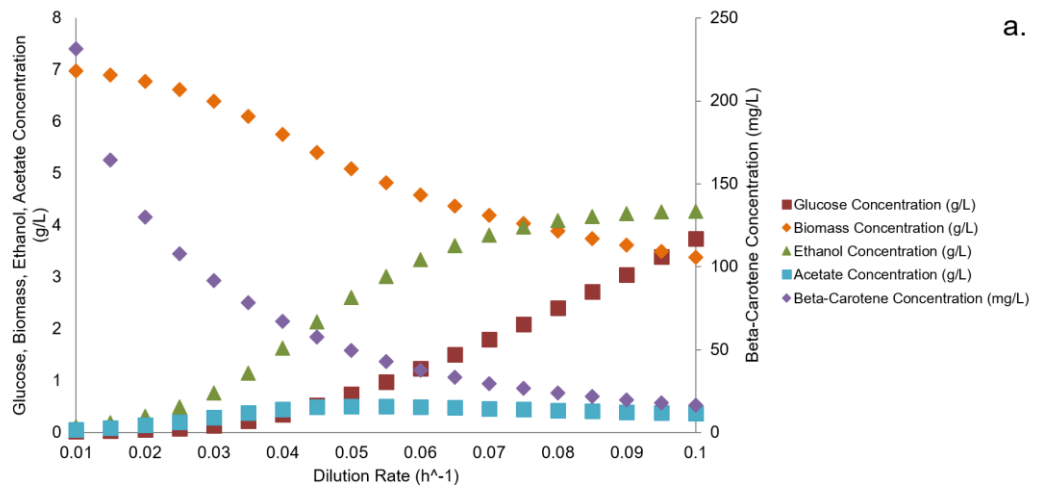
Figure 41: (a) Steady-state concentration values of glucose, biomass, ethanol, acetate, and beta-carotene at dilution rates between 0.01 and 0.1 h^{-1}. (b) Steady-state flux values of glucose, ethanol, acetate, and beta-carotene at dilution rates between 0.01 and 0.1 h^{-1}.
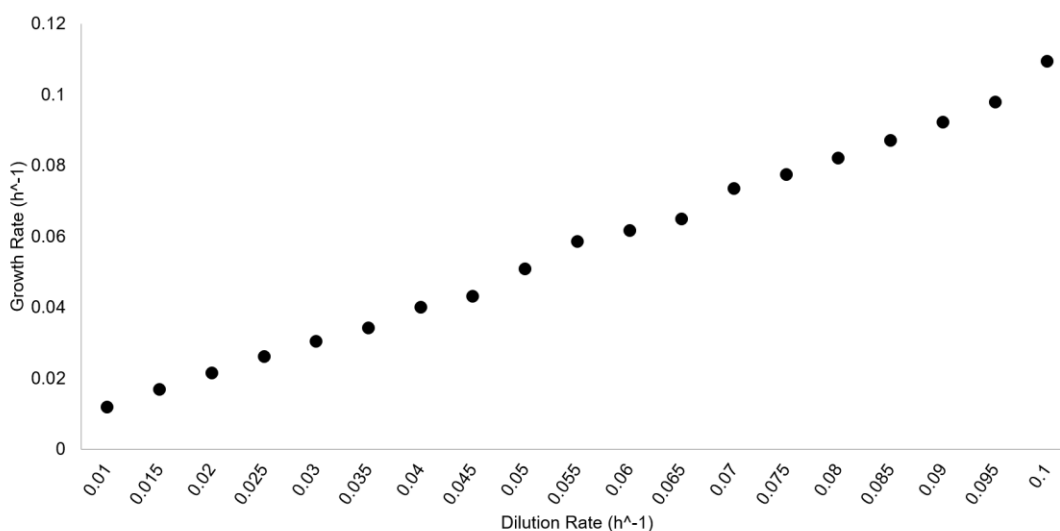
Figure 42: Specific growth rate values calculated by the COBRA model optimization at each dilution rate between 0.01 and 0.1 h⁻¹.

To better understand the intracellular behavior of the beta-carotene production strain, the intracellular fluxes of key reactions in important pathways may be examined after the intracellular flux model has been optimized for each dilution rate. Figure 43 shows a portion of the central metabolism of *S. cerevisiae*, including glycolysis and acetaldehyde formation. The fluxes of the four key reactions labeled in the figure are plotted in Figure 44 at each dilution rate. Between dilution rates of 0.01 and 0.075 h⁻¹, the fluxes of reactions A, B, and C reach an optimum value at 0.055 h⁻¹. Between dilution rates of 0.08 and 0.1 h⁻¹, the fluxes increase with increasing dilution rate. This is likely due to changes in glucose metabolism in response to increasing availability of glucose [35, 92]. The flux of reaction D increases proportionally with the glucose uptake rate (Figure 41). Figure 45 shows a portion of the mevalonate pathway, including the key branch point to the added beta-carotene pathway (G). This branch point and the

other labeled key reactions are plotted at each dilution rate in Figure 46. As expected, the

farnesyltranstransferase reaction flux reaches an optimum value at a dilution rate of 0.07

h^{-1}, the same optimal dilution rate observed when examining beta-carotene production.

The fluxes of E, F, and H increase with the dilution rate, due to the growth-coupled

nature of the mevalonate pathway and the greater ergosterol biosynthesis pathway [24].

Glyceraldehyde-3-Phosphate ← - - - - - - Fructose-6-Phosphate ← $\overset{A}{\rule{2cm}{0.4pt}}$ Glucose-6-Phosphate

↓ B

3-Phospho-D-Glycerol-Phosphate

A: *glucose-6-phosphate isomerase*
B: *glyceraldehyde-3-phosphate dehydrogenase*
C: *enolase*
D: *pyruvate decarboxylase*

⋮

2-phosphoglycerate

↓ C

Phosphoenolpyruvate

↓

Pyruvate $\overset{D}{\rule{1.5cm}{0.4pt}}$→ Acetaldehyde

Figure 43: A portion of the glycolysis pathway in *S. cerevisiae*. Key reactions are labeled A-D.

Figure 44: Steady-state flux values of the four key glycolysis pathway reactions labeled in Figure 43.

E: hydroxymethylglutaryl-CoA reductase
F: mevalonate kinase
G: farnesyltranstransferase
H: squalene synthase
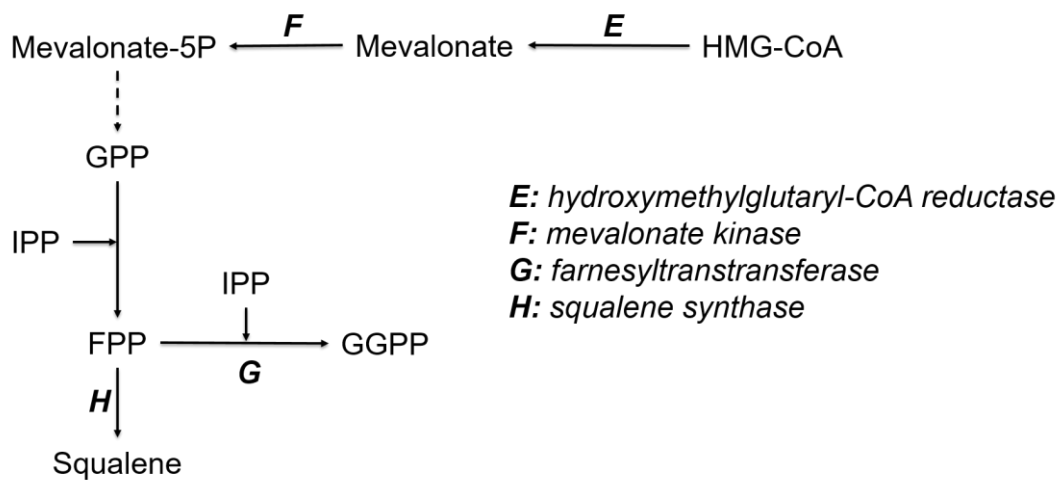
Figure 45: A portion of the mevalonate pathway in *S. cerevisiae*. Key reactions are labeled E-H.
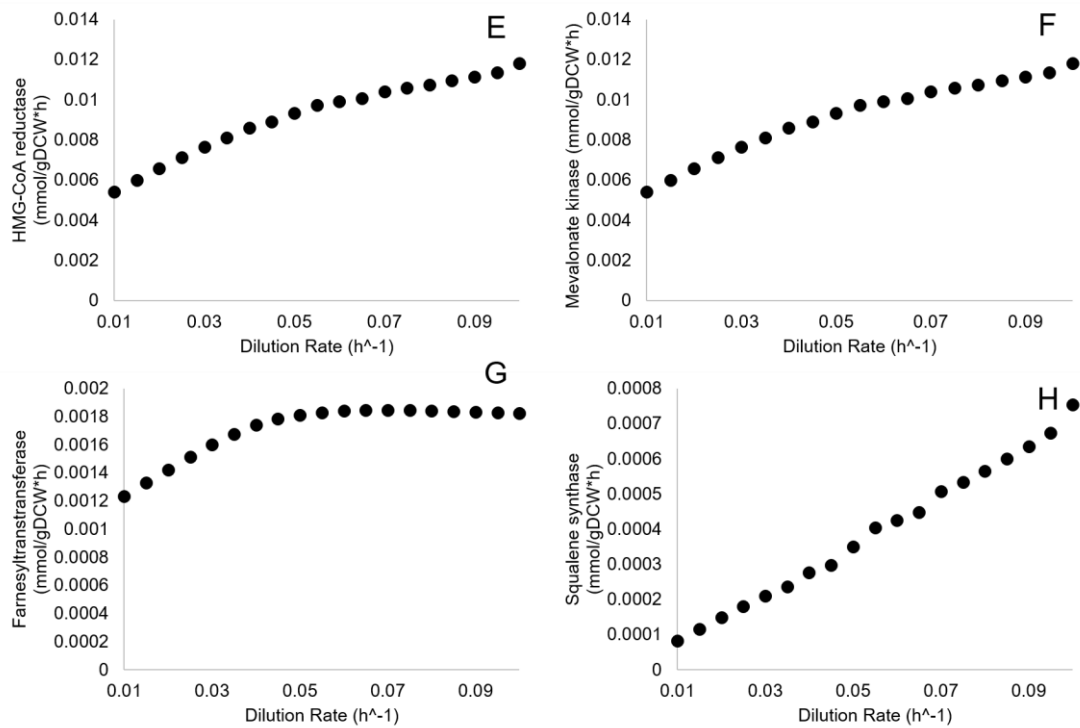
Figure 46: Steady-state flux values of the four key mevalonate pathway reactions labeled in Figure 45.

As discussed previously, Figure 41b shows that the maximum beta-carotene flux is seen at a dilution rate of 0.07 h-1. However, the glucose uptake rate and specific growth rate continue to increase. This indicates that there is room to improve the performance of this strain, and further metabolic engineering could be performed to carry out these improvements. Examining the full results of the metabolic flux model optimization could provide information about possible bottleneck reactions for beta-carotene production within the cell. In the previous work done by Raftery et al, it was shown that a multiple-feed process (glucose, ethanol, and/or media) increases the cellular and substrate yields of beta-carotene [91]. It is possible that integrating a similar

multiple-feed to the analysis performed in this work could provide further understanding of the optimal conditions for the continuous process.

6.4 Conclusion

In this study, a kinetic model of a continuous process for the production of beta-carotene by a recombinant strain of *S. cerevisiae* was used to study intracellular metabolic behavior. The kinetic model reactions were solved at various dilution rates, and the results were used as input to an intracellular stoichiometric model. The intracellular model was then solved using maximum specific growth rate as the cellular objective. The results of this model optimization make it possible to better understand the effect of dilution rate on the metabolic behavior of the yeast strain. The central metabolic pathway as well as the mevalonate pathway were examined in detail in this study. Effects of possible glucose repression and induction are seen when examining key glycolysis reactions. The optimal dilution rate for beta-carotene production when using a glucose feed is shown to be 0.07 h^{-1}. The procedure used in this study can be used in the future to explore the continuous production of many different terpenoid products. In addition, the kinetic model could be expanded to study different feeding schemes, including multiple substrates, which has been shown in previous work to have an effect on beta-carotene production.

7. PREDICTING BENEFICIAL GENETIC KNOCKOUTS FOR INCREASED BETA-CAROTENE PRODUCTION: OPTKNOCK

7.1 Introduction

An additional tool available within the COBRA toolbox is Optknock, a program that can identify beneficial gene knockouts [93]. Optknock identifies these beneficial gene knockouts by placing an additional optimization outside the objective maximization described above. The user can provide a desired percentage of the maximum growth rate, a maximum allowable number of gene knockouts, and the desired product or intermediate to be maximized. This outer optimization and its association with the inner COBRA optimization are shown in equation 48. Optknock and similar tools have been used in many studies to target beneficial reaction or gene deletions. Alper et al used these tools to identify improvements to a lycopene-producing *E. coli* strain [94]. Klein-Marcuschamer used similar targeting techniques to improve production of lycopene and other isoprenoid products [95]. In this work, it was desired to target genetic modifications to increase flux through the added beta-carotene pathway, as well as increase acetyl-CoA pools within the cell, and increase flux through the overall ergosterol biosynthesis pathway.

$$\underset{v}{max} \ bioengineering \ objective$$

$$subject \ to:$$

$$\left(\begin{array}{c} \underset{v}{max} \ biomass \ objective \\ subject \ to: \\ \sum Sv = 0 \\ v_{min} \leq V \leq v_{max} \end{array} \right) \tag{51}$$

$$v_{biomass} \geq v_{biomass,min}$$

$$\sum knocked \ out \ reactions \leq max. \ allowable$$

7.2 Methods

In order to determine beneficial gene knockouts, the Optknock tool was used. Tests using Optknock were performed systematically, using several different outer bioengineering target objectives. Yeast model version 7 was used for this work, and both the glucose uptake rate and oxygen uptake rate were set to 10 mmol/gDCW*h. The number of allowable knockouts was varied, as well as the minimum allowable growth flux. The list of possible reactions to be knocked out was narrowed to the main glycolysis, TCA cycle, and pyruvate-related reactions. The full list of these reactions can be found in the Appendix in Table 25. Table 5 shows the design for the Optknock optimizations performed.

Table 5: The design of Optknock optimizations performed, including target reactions, number of allowable knockouts, and the minimum allowable growth flux.

Target Reaction	Number of Reaction Deletions Specified	Minimum Allowable Growth Flux (h^{-1})
Ex_betacarotene	1, 2, 3, or ≤4	0.05
Pyruvate Dehydrogenase	1, 2, 3, or ≤4	0.05
Ergosterol Transport	1, 2, 3, or ≤4	0.05
Acetyl-coA Synthetase (cytoplasm and mitochondrion)	1, 2, 3, or ≤4	0.05
Farnesyltranstransferase	1, 2, 3, or ≤4	0.05
Squalene Synthase	1, 2, 3, or ≤4	0.05
HMG-CoA Reductase	1, 2, 3, or ≤4	0.05
Mevalonate Kinase (atp)	1, 2, 3, or ≤4	0.05
Growth	1, 2, 3, or ≤4	0.05
Pyruvate Decarboxylase	1, 2, 3, or ≤4	0.05
Hexokinase	1, 2, 3, or ≤4	0.05

As can be seen, the six key reactions used in the dynamic FBA multi-objective functions are included in this analysis, as well as several more important reactions. All of the reactions chosen as targets for the Optknock analysis were chosen, as described below, to increase the production of biomass, beta-carotene production, or both. The pyruvate dehydrogenase reaction, shown below in Figure 47, converts one molecule each of pyruvate, coenzyme A, and NAD to one molecule each of acetyl-CoA, carbon dioxide, and NADH [82]. The genes associated with the pyruvate dehydrogenase reaction are *PDB1, PDA1, LPD1, PDX1,* and *LAT1*. The *PDB1* gene encodes the E1 beta subunit of the pyruvate dehydrogenase (PDH) complex [96]. The *PDA1* gene encodes the E1 alpha subunit of the PDH complex [97]. The *LPD1* gene encodes the lipoamide dehydrogenase component of the PDH complex [98]. The *PDX1* gene encodes the E3 binding protein of the PDH complex [99]. The *LAT1* gene encodes the dihydrolipoamide

acetyltransferase component of the PDH complex [100]. Increasing flux through this reaction will increase acetyl-coA pools within the cell, which would theoretically in turn increase flux through several pathways, including the mevalonate pathway and beta-carotene production.
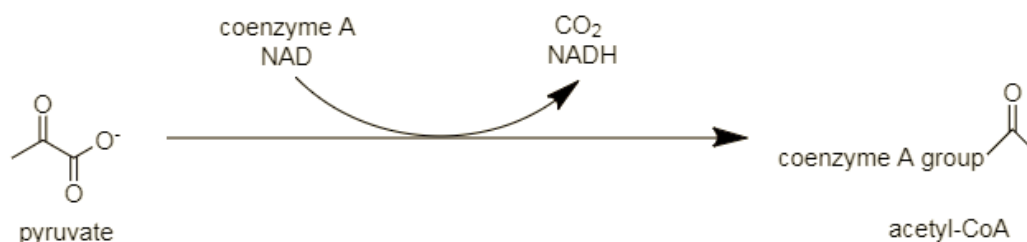
Figure 47: The reaction catalyzed by pyruvate dehydrogenase.

The ergosterol transport reaction is responsible for transporting ergosterol outside the cell, and is associated with the *PDR11* gene, which encodes an ABC transporter [101]. This reaction is important for overall growth and cell structure [24, 25]. The squalene synthase reaction, shown below in Figure 48, converts two trans, trans-farnesyl diphosphate molecules and one NADPH molecule to two pyrophosphate molecules and one molecule each of NADP+ and squalene. Squalene synthase is encoded by the *ERG9* gene, and catalyzes an important intermediate reaction involved in ergosterol synthesis [24, 25]. Increasing the flux through this reaction would theoretically increase the flux through the ergosterol biosynthesis pathway (including the mevalonate pathway), which could increase the production of beta-carotene. Increasing the flux through the squalene synthase reaction would also theoretically increase the formation of biomass, and

therefore due to the growth-coupled nature of beta-carotene production, could increase

the production of beta-carotene.

Figure 48: The reaction catalyzed by squalene synthase.

The HMG-CoA reductase reaction, shown in Figure 49, converts one molecule of

3-hydroxy-3-methyl-glutaryl-CoA (HMG-CoA) and two molecules of NADPH to two

molecules of NADP and one molecule each of coenzyme A and mevalonate. HMG-CoA

reductase is encoded by the *HMG2* gene, and this reaction is an important step in the

mevalonate pathway and the eventual production of beta-carotene in SM14 [102].

Figure 49: The reaction catalyzed by HMG-CoA reductase.

Following the HMG-CoA reductase reaction, the next step in the mevalonate pathway is

the mevalonate kinase reaction. Mevalonate kinase is encoded by the *ERG12* gene. This

reaction, shown in Figure 50, converts one molecule of mevalonate to one molecule of

mevalonate-5-phosphate, using one molecule of ATP for energy [103]. Increasing the

flux through these reactions within the mevalonate pathway would in turn increase the

flux through the entire pathway, which could theoretically increase the production of

both biomass and beta-carotene.
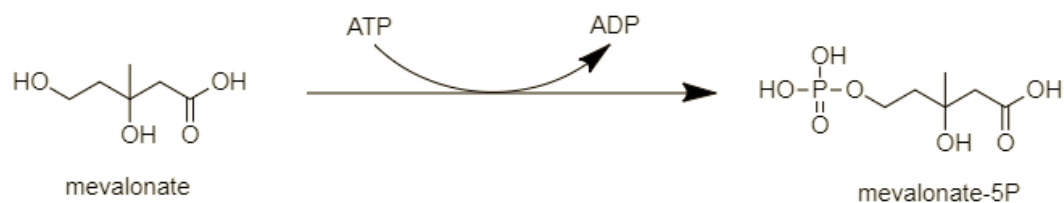
Figure 50: The reaction catalyzed by mevalonate kinase.

7.3 Results and Discussion

After performing Optknock optimizations systematically as laid out in Table 5,

several reaction knockouts were identified as beneficial to the increased flux of the

targeted reactions. These reactions, and the associated genes, are listed in Table 6 and

are described below. Some reactions are associated with more than one gene, and all

associated genes are included.

Table 6: The list of knocked out reactions identified by Optknock to be beneficial to specified target. The genes associated with each of the reactions are also listed.

Target Reaction	Number of Deletions Specified	Reaction(s) Identified	Associated Genes
Ex_betacarotene	1	glycine-cleavage complex (lipoamide), triose-phosphate isomerase	GCV3, GCV1, LPD1, GCV2, TPI1
Ex_betacarotene	2	malate synthase, pyruvate dehydrogenase	DAL7, PDB1, PDA1, LPD1, PDX1, LAT1
Ex_betacarotene	3	citrate synthase (peroxisomal), glycerol-3-phosphate dehydrogenase (NAD), glycine-cleavage complex (lipoylprotein)	CIT2, GPD1, GCV3, GCV1, LPD1, GCV2
pyruvate dehydrogenase	1	pyruvate carboxylase	PYC2, PYC1
pyruvate dehydrogenase	2	fumarase (cytoplasmic), pyruvate carboxylase	FUM1, PYC2
pyruvate dehydrogenase	3	glycine cleavage system, glycine-cleavage complex (lipoamide), glycine-cleavage complex (lipoylprotein)	GCV3, GCV1, LPD1, GCV2
pyruvate dehydrogenase	≤4	fumarase, glycine cleavage system, glycine-cleavage complex (lipoamide), glycine-cleavage complex (lipoylprotein)	FUM1, GCV3, GCV1, LPD1, GCV2
ergosterol transport	1	fumarase, oxoglutarate dehydrogenase	FUM1, KGD2, LPD1, KGD1
ergosterol transport	2	(R,R)-butanediol dehydrogenase, C-22 sterol desaturase (NADP)	BDH1, ERG5
ergosterol transport	3	2-methylcitrate synthase, glycine cleavage system, phenylpyruvate decarboxylase	CIT3, GCV3, GCV1, LPD1, GCV2, ARO10
acetyl-coA synthetase (cytoplasm)	1	pyruvate carboxylase	PYC2

Target Reaction	Number of Deletions Specified	Reaction(s) Identified	Associated Genes
acetyl-coA synthetase (cytoplasm)	2	pyruvate carboxylase, pyruvate dehydrogenase	PYC2, PDB1, LPD1, PDX1, LAT1
acetyl-coA synthetase (cytoplasm)	≤4	fumarase, malate dehydrogenase (peroxisomal), pyruvate carboxylase, pyruvate dehydrogenase	FUM1, DAL7, PYC2, PDB1, PDA1, LPD1, PDX1, LAT1
farnesyltranstransferase	1	succinate dehydrogenase (ubiquinone-6), citrate to cis-aconitate	SDH4, SDH3, SDH2, ACO1
farnesyltranstransferase	2	2-methylcitrate synthase, malate dehydrogenase (cytoplasmic)	CIT3, MDH2
farnesyltranstransferase	≤4	phenylpyruvate decarboxylase	ARO10
squalene synthase	1	2-methylcitrate synthase	CIT3
squalene synthase	2	2-methylcitrate synthase, pyruvate decarboxylase (acetoin-forming)	CIT3, PDC6
HMG-CoA reductase	1	2-methylcitrate synthase, diacylglycerol acyltransferase	CIT3, LRO1
HMG-CoA reductase	2	2-methylcitrate synthase, fumarate reductase	CIT3, FRD1
HMG-CoA reductase	3	glycine cleavage system, glycine-cleavage complex (lipoamide), glycine-cleavage complex (lipoylprotein)	GCV3, GCV1, LPD1, GCV2

Target Reaction	Number of Deletions Specified	Reaction(s) Identified	Associated Genes
HMG-CoA reductase	≤4	Fumarase (cytoplasmic), glycine cleavage system, glycine-cleavage complex (lipoamide), glycine-cleavage complex (lipoylprotein)	FUM1, GCV3, GCV1, LPD1, GCV2
mevalonate kinase (atp)	1	pyruvate dehydrogenase, cardiolipin synthase	PDB1, PDA1, LPD1, PDX1, LAT1, CRD1
mevalonate kinase (atp)	2	glycine-cleavage complex (lipoamide), phenylpyruvate decarboxylase	GCV3, GCV1, LPD1, GCV2, ARO10
mevalonate kinase (atp)	3	fumarate reductase, malate dehydrogenase (peroxisomal), pyruvate dehydrogenase	FRD1, MDH3, PDB1, PDA1, LPD1, PDX1, LAT1
mevalonate kinase (atp)	≤4	citrate synthase, glycerol-3-phosphatase, malate dehydrogenase, malate dehydrogenase (cytoplasmic)	CIT1, GPP2, MDH1, MDH2
acetyl-CoA synthetase (mitochondrion)	1	pyruvate dehydrogenase	PDB1, PDA1, LPD1, PDX1, LAT1
acetyl-CoA synthetase (mitochondrion)	2	pyruvate carboxylase, pyruvate dehydrogenase	PYC2, PDB1, PDA1, LPD1, PDX1, LAT1
acetyl-CoA synthetase (mitochondrion)	≤4	citrate synthase (peroxisomal), malate dehydrogenase, pyruvate carboxylase, pyruvate dehydrogenase	CIT2, MDH1, PYC2, PDB1, PDA1, LPD1, PDX1, LAT1
growth	1	(R,R)-butanediol dehydrogenase	BDH1

Target Reaction	Number of Deletions Specified	Reaction(s) Identified	Associated Genes
growth	3	C-5 sterol desaturase, isocitrate lyase, tetrahydrofolate aminomethyltransferase	ERG3, ICL1, GCV3, GCV1, LPD1, GCV2
pyruvate decarboxylase	1	pyruvate carboxylase	PYC2
pyruvate decarboxylase	2	malate dehydrogenase, pyruvate carboxylase	MDH1, PYC2
pyruvate decarboxylase	3	glycine cleavage system, glycine-cleavage complex (lipoamide), glycine-cleavage complex (lipoylprotein)	GCV3, GCV1, LPD1, GCV2
pyruvate decarboxylase	≤4	glycine cleavage system, glycine-cleavage complex (lipoamide), glycine-cleavage complex (lipoylprotein), pyruvate dehydrogenase	GCV3, GCV1, LPD1, GCV2, PDB1, PDA1, LPD1, PDX1, LAT1
hexokinase	1	2-methylcitrate synthase	CIT3
hexokinase	2	2-methylcitrate synthase, malate synthase	CIT3, MLS1
hexokinase	3	acetyl-CoA synthetase (mitochondrion), malate dehydrogenase, pyruvate carboxylase	ACS1, MDH1, PYC2
hexokinase	≤4	citrate synthase, malate synthase, pyruvate carboxylase	CIT1, DAL7, PYC2

The first reaction identified by Optknock is the glycine-cleavage complex. This reaction converts a dihydrolipoamide and NAD to a lipoamide and NADH. The genes *GCV3*, *GCV1*, *LPD1*, and *GCV2* are associated with this complex [98, 104-107]. The

next reaction identified is the malate synthase reaction. Malate synthase converts one molecule each of acetyl-CoA, glyoxylate, and water to one molecule each of malate and Coenzyme A, shown in Figure 51. This enzyme is encoded by the *DAL7* gene [108].
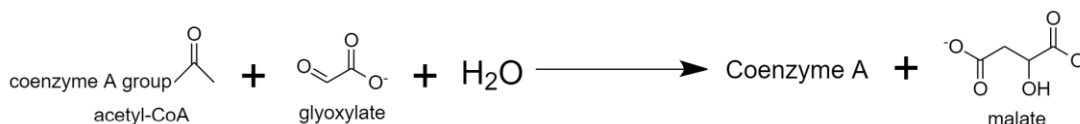
Figure 51: The reaction catalyzed by malate synthase.

The next reaction to be identified is the reaction catalyzed by pyruvate dehydrogenase, described in section 7.2 and in Figure 47. The reaction catalyzed by citrate synthase is also identified by Optknock. Citrate synthase converts one molecule each of acetyl-CoA, oxaloacetate, and water to one molecule each of citrate and coenzyme A, as shown in Figure 52. This enzyme is encoded by the *CIT1, CIT2,* and *CIT3* genes [109].
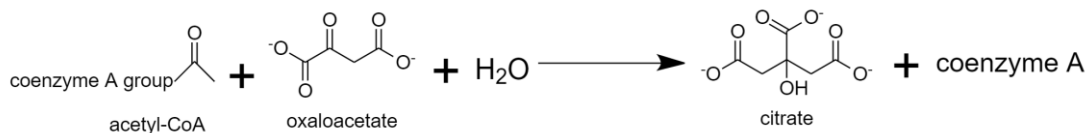
Figure 52: The reaction catalyzed by citrate synthase.

89

Optknock also identified the glycerol-3-phosphate dehydrogenase reaction. This

reaction, shown in Figure 53, converts one molecule each of dihydroxyacetone

phosphate and NADH to one molecule each of glycerol-3-phosphate and NAD.

Glycerol-3-phosphate dehydrogenase is encoded by the *GPD1* gene, and is a key

enzyme of glycerol synthesis [110].

Figure 53: The reaction catalyzed by glycerol-3-phosphate dehydrogenase.

The next reaction identified by Optknock is the pyruvate carboxylase reaction.

Pyruvate carboxylase as shown in Figure 54, converts one molecule each of ATP,

bicarbonate, and pyruvate to one molecule each of ADP, phosphate and oxaloacetic acid.

This enzyme is encoded by the *PYC2* gene [111].
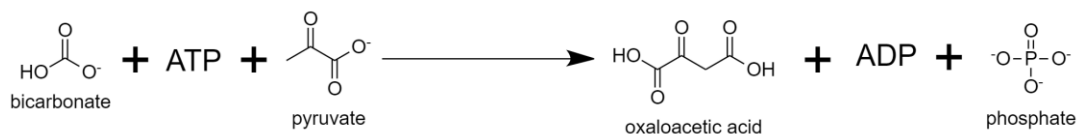
Figure 54: The reaction catalyzed by pyruvate carboxylase.

Optknock also identified the fumarase reaction as a beneficial knockout.

Fumarase, as shown in Figure 55, converts one molecule each of fumarate and water to

one molecule of malate. This enzyme is encoded by the *FUM1* gene, and is a part of the TCA cycle [84].
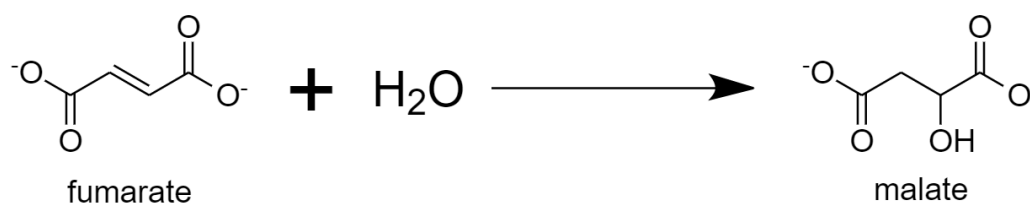
Figure 55: The reaction catalyzed by fumarase.

The next reaction that Optknock identified is the (R,R)-butanediol dehydrogenase reaction. (R,R)-butanediol dehydrogenase, as shown in Figure 56, converts one molecule each of (R,R)-2,3-butanediol dehydrogenase and NAD to one molecule each of (R)-acetoin and NADH. This enzyme is encoded by the *BDH1* gene [112].

Figure 56: The reaction catalyzed by (R,R)-butanediol dehydrogenase.

The C-22 sterol desaturase reaction was also identified by Optknock. In the yeast model, C-22 sterol desaturase converts one molecule each of ergosta-5.7,22,24(28)-trien-3beta-ol, hydrogen, NADPH, and oxygen, to one molecule each of ergosta-5,7,22,24(28)-tetraen-3beta-ol and NADP and two molecules of (1→3)-beta-D-glucan. C-22 sterol desaturase is encoded by the *ERG5* gene [24]. This reaction is shown in Figure 57.
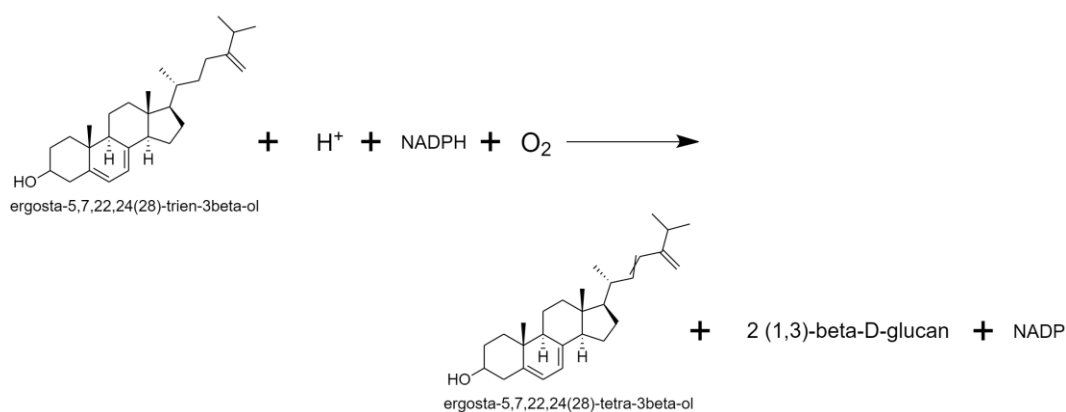
Figure 57: The reaction catalyzed by the C-22 sterol desaturase reaction.

The next reaction identified by Optknock is the 2-methylcitrate synthase reaction. 2-methylcitrate synthase, as shown in Figure 58, converts one molecule each of water, oxaloacetate, and propionyl-CoA to one molecule each of 2-methylcitrate, coenzyme A, and hydrogen. This enzyme is encoded by the *CIT3* gene [83].

92
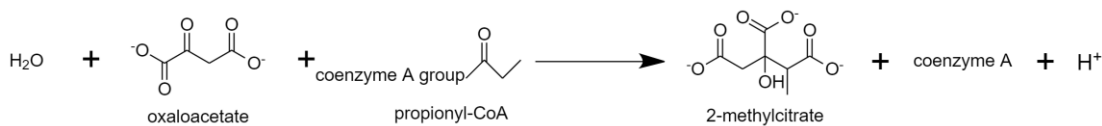
Figure 58: The reaction catalyzed by the 2-methylcitrate synthase reaction.

The phenylpyruvate decarboxylase reaction was also identified by Optknock. This enzyme, as shown in Figure 59, converts one molecule each of hydrogen and keto-phenylpyruvate to one molecule each of carbon dioxide and phenylaceetaldehyde. Phenylpyruvate decarboxylase is encoded by the *ARO10* gene [113].
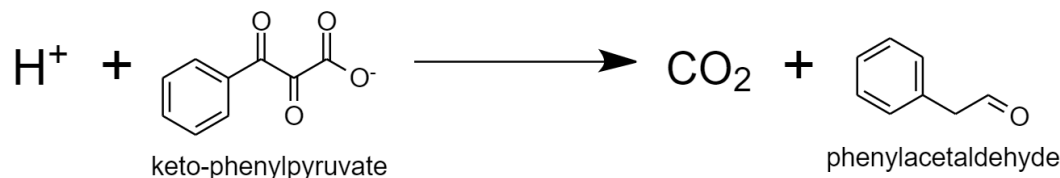
Figure 59: The reaction catalyzed by phenylpyruvate decarboxylase.

The malate dehydrogenase reaction was also identified by Optknock. As shown in Figure 60, malate dehydrogenase converts one molecule each of malate and NAD to one molecule each of oxaloacetate, NADH, and hydrogen. Malate dehydrogenase is encoded by the *MDH1* gene, and is a component of the TCA cycle [114].
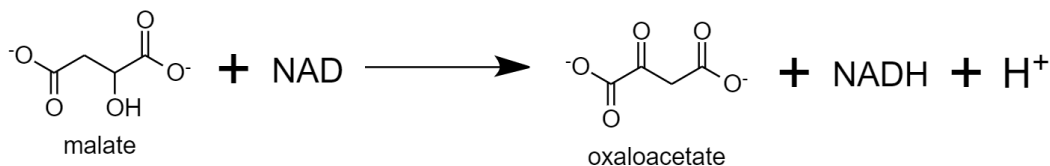
Figure 60: The reaction catalyzed by malate dehydrogenase.

The Optknock optimization also identified the succinate dehydrogenase reaction. Succinate dehydrogenase converts one molecule each of succinate and ubiquinone-6 to one molecule each of fumarate and ubiquinol-6, as shown in Figure 61. This enzyme's subunits are encoded by the *SDH4, SDH3,* and *SDH2* genes [115].

Figure 61: The reaction catalyzed by succinate dehydrogenase.

The fumarate reductase reaction was also identified by Optknock. Fumarate reductase, as shown in Figure 62, converts one molecule each of fumarate and FADH2 (a redox electron carrier) to one molecule each of succinate and FAD. This enzyme is encoded by the *FRD1* gene [116].
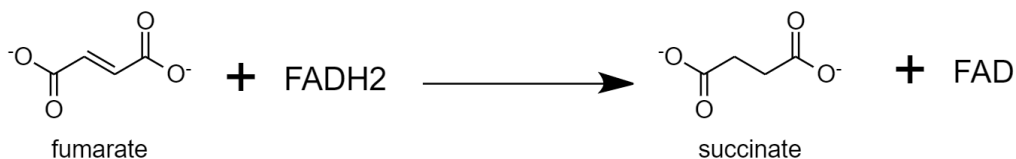
Figure 62: The reaction catalyzed by fumarate reductase.

The Optknock optimization also identified the C-5 sterol desaturase reaction. C-5 sterol desaturase as shown in Figure 63, converts one molecule each of episterol,

NADPH, hydrogen, and oxygen to one molecule each of ergosta-5,7,24(28)-trien-3beta-ol and NADP, and two molecules of water. This enzyme is encoded by the *ERG3* gene [117].
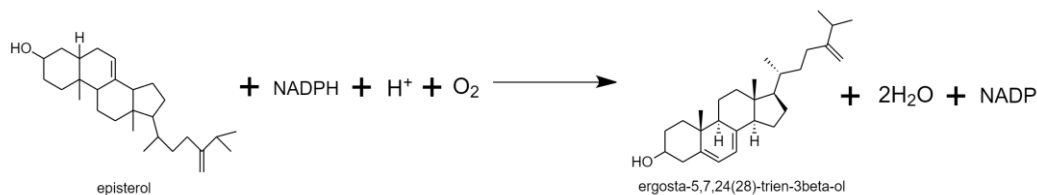
Figure 63: The reaction catalyzed by C-5 sterol desaturase.

The isocitrate lyase reaction was also identified by the Optknock optimization. As shown in Figure 64, isocitrate lyase converts one molecule of isocitrate to one molecule each of glyoxylate and succinate. This enzyme is encoded by the *ICL1* gene, and is a key component of the glyoxylate cycle [118].
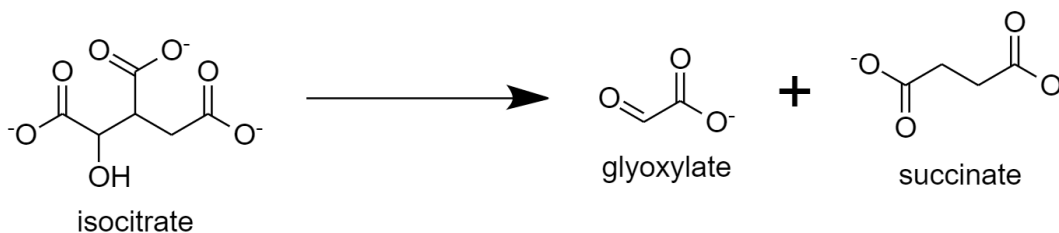
Figure 64: The reaction catalyzed by isocitrate lyase.

Finally, the tetrahydrofolate aminomethyltransferase reaction was identified by the Optknock optimization. This reaction is part of the glycine cleavage complex, and

converts one molecule each of 5,10-methenyl-THF and water to one molecule each of 5-formyltetrahydrofolic acid and hydrogen [98, 104-107].

7.4 Conclusion

The work performed in this study make it possible to quickly and efficiently target beneficial genetic modifications of the beta-carotene-producing SM14 strain of *S. cerevisiae*. Using the COBRA optimization tool known as Optknock in conjunction with a metabolic model of *S. cerevisiae*, several different key bioengineering objectives were targeted for optimization. These bioengineering objectives included beta-carotene production and other important intermediate reactions within the greater ergosterol biosynthesis pathway. The number of allowable reactions to be knocked out provided to the Optknock program was kept between one and three. Several reactions were identified as being beneficial to beta-carotene or intermediate production, and are listed above. Further experimental work may be performed to test the validity of these results. If Optknock is proven to be an accurate predictor of beneficial genetic modifications, this has the potential to greatly streamline strain improvement.

8. GENETIC MODIFICATION OF SM14 FOR INCREASED BETA-CAROTENE PRODUCTION

8.1 Introduction

8.1.1 Genetic Knockouts in Yeast: Traditional Method

Homologous recombination is the exchange of genetic information between different forms of a gene, or alleles. This phenomena serves several biological functions. During meiosis (cell division), homologous recombination serves to generate diversity among the progeny of the parent cells. Homologous recombination also ensures that the chromosome pairs are segregated correctly during meiosis [119]. Homologous recombination may be exploited to delete genes, and is widely used to target and delete genes in yeast [120]. This deletion procedure consists of (1) constructing a deletion module via PCR containing a selectable marker flanked by 5' and 3' homologous sequences, (2) transforming this deletion module into yeast, and (3) allowing homologous recombination to occur at the targeted site [121]. In this work, it was desired to delete genes that were targeted by Optknock to test the efficacy of the computational gene targeting tool. Three genes were chosen to be deleted individually: the citrate synthase (*CIT3)* gene, an acyltransferase gene (*LRO1*), and the dihydrolipoamide dehydrogenase (*LPD1*) gene. The specific procedure for the deletion of these genes is described below.

8.1.2 The Discovery of the CRISPR/Cas System

While genetic modification has been a large part of biochemical development for many years, a recent discovery has made genetic engineering much easier, faster, more accurate, and more efficient. While studying the native immune system of bacteria, sequences of DNA known as CRISPR (clustered regularly interspaced short palindromic repeats) were found, often in close proximity to cas (CRISPR-associated) genes. These CRISPR loci were also found to have sequences of homology with bacteriophage, and that this somehow aided in guiding the Cas proteins in the bacteria to the invading phage [122]. It was then discovered that a piece of this guiding mechanism involved sequences of homology in the CRISPR loci that were transcribed into RNA (CRISPR RNAs, or crRNAs) [123]. Scientists were then able to describe in more detail the CRISPR-Cas system, specifically that a specific Cas protein (Cas9) acts as a nuclease that cleaves target DNA [124]. Finally, it was found that a second RNA called trans-activating CRISPR RNA (tracrRNA) formed a duplex with crRNA and together this is what guides the Cas9 protein to its target [125]. All of these pieces and the mechanism described are depicted below in Figure 65. Much more work has been done and is still being done to learn even more about this process, but is beyond the scope of this work.

8.1.3 Genetic Modification using CRISPR-Cas9

After many years of studying the CRISPR-Cas9 system, scientists realized that it could be manipulated and used to target specific DNA sequences in an organism of choice.
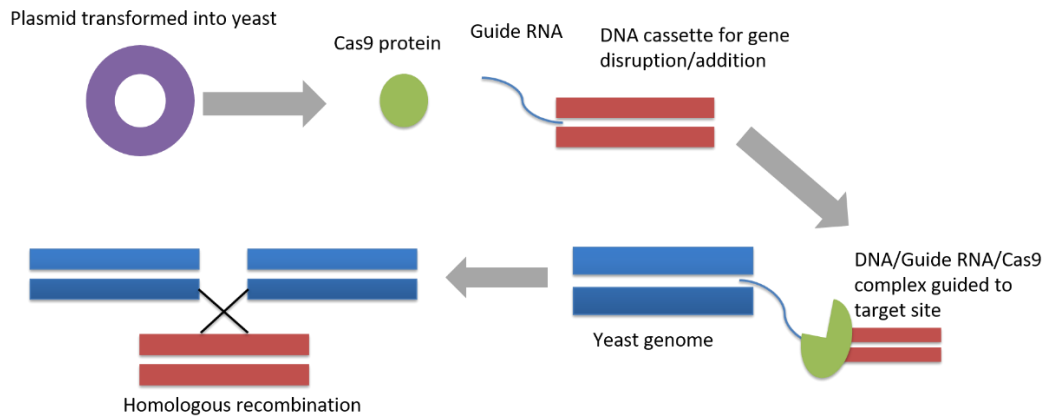
Figure 65: Visual depiction of CRISPR-Cas9 gene disruption process.

8.2 Methods

8.2.1 PCR and Overlap PCR Strategies for Traditional Gene Knockout Method

To perform the gene deletions in the SM14 strain, a hygromycin resistance gene was chosen as the selectable marker. Two different methods for constructing the deletion modules were attempted. The first method was to design primers to amplify the hygromycin resistance gene from plasmid pAG26, which was obtained from the Kao lab and was originally a gift from John McCusker [126]. The forward and reverse primers contained a 70 base pair sequence of homology with the gene in question (either *CIT3, LRO1,* or *LPD1*). These primers are shown in Table 17 in the Appendix. The second method used was to use overlap PCRs to create a deletion module with 500 base pairs of homology upstream and downstream of the hygromycin resistance gene. Three deletion modules were designed in this manner, and overlapping primers were designed to amplify the upstream homology region, the hygromycin gene, and the downstream

99

homology region separately. The forward primer for the upstream region and the reverse primer for the downstream region were then used to amplify the full deletion module. This is illustrated in Figure 66. The sequence of each deletion module, as well as the sequence of all the primers used, are shown in Table 17 and Table 18 in the Appendix. The PCRs performed for this work were completed using the Phusion® high-fidelity master mix kit or the Q5® high-fidelity master mix kit, both from New England BioLabs®. 50 µL reactions were run, and their contents are summarized in Table 20. The thermocycler program used for this work is shown in Table 22. The PCR products from both attempted methods were transformed into SM14, in a procedure laid out in section 8.2.3.
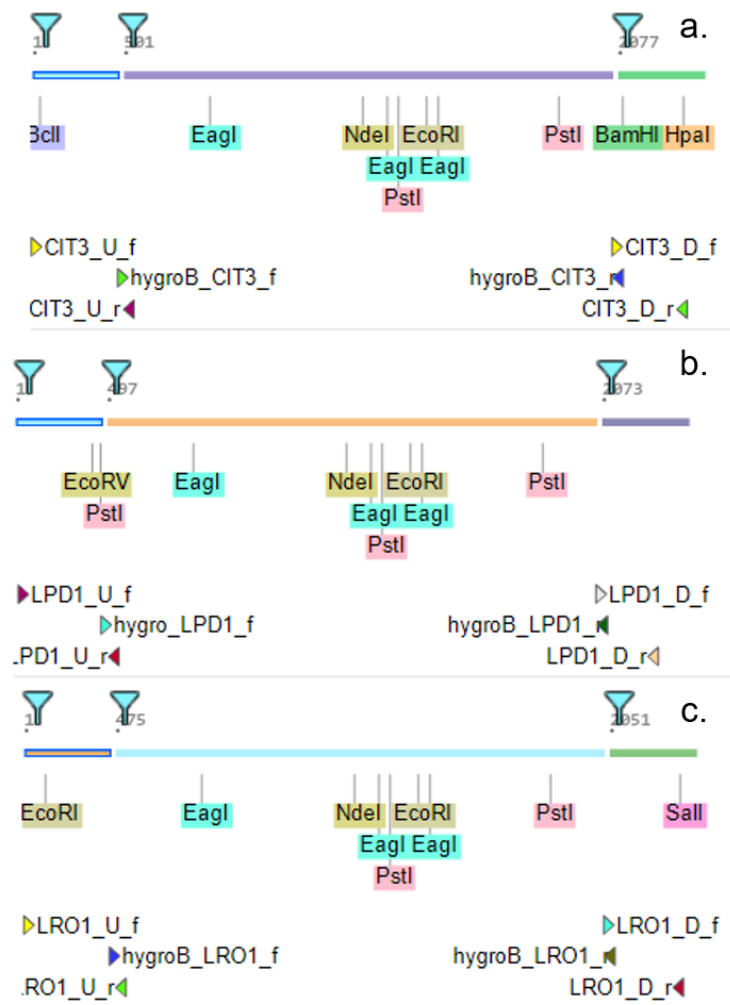
Figure 66: Deletion modules constructed to knock out genes from strain SM14. The smaller flanking portions are *CIT3* homology, and the center section is the hygromycin resistance gene. The primers and their amplification regions are shown. (a.) The module to knock out the *CIT3* gene. (b.) The module to knock out the *LPD1* gene. (c.) The module to knock out the *LRO1* gene.

8.2.2 Plasmid Construction

A CRISPR/Cas9 plasmid was obtained from Addgene. The plasmid, called

pCRCT, was a gift from Huimin Zhao (Addgene plasmid # 60621) [127]. Plasmid

pCRCT contains an ampicillin resistance gene for growth in bacteria. The plasmid also

contains a uracil selectable marker for growth in yeast. This plasmid encodes for all

necessary components for CRIPSR-Cas9 gene modification, namely Cas9, tracrRNA

and crRNAs. Also included in the design of the plasmid is a cloning site ideal for

inserting a specifically designed block of DNA to target any gene of choice. Figure 67

illustrates these components of pCRCT.
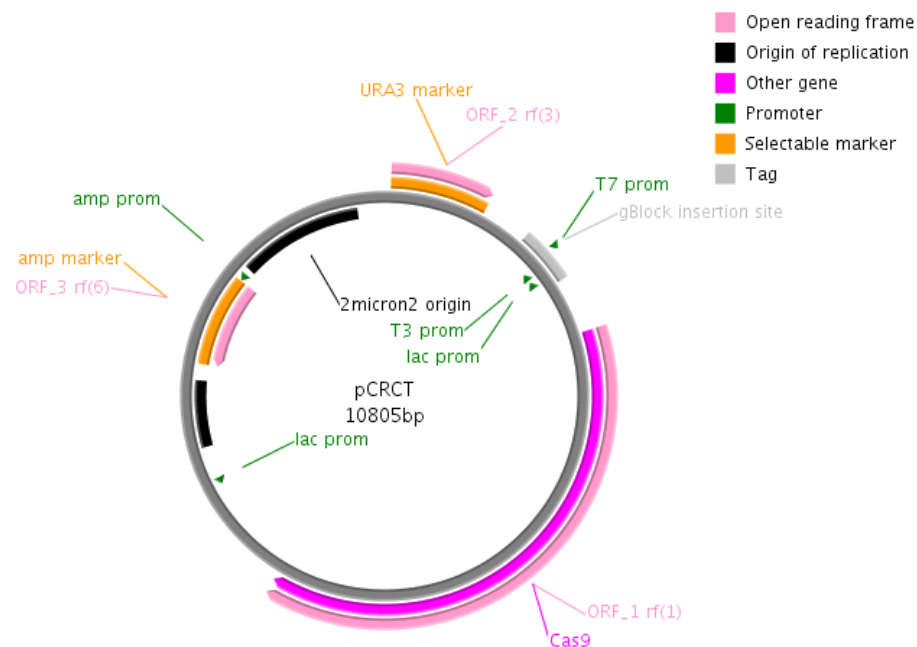
Figure 67: Diagram of the pCRCT plasmid. The three main components, the ampicillin resistance marker, the uracil marker, and the Cas9 protein and related RNAs, are shown. Also shown is the site at which the gBlock DNA may be inserted via restriction digest. [128]

This plasmid was received as bacteria in an agar stab, and a liquid culture of LB

medium was started upon arrival and grown overnight. Part of this culture was combined

with glycerol to create a stock for future use which was stored at -80°C. Plasmid pCRCT

was then purified from another portion of this culture using a plasmid miniprep kit

(Monarch®, New England Biolabs®)

For this work, a block of DNA called a gBlock®, constructed by Integrated DNA Technologies (IDT), was designed to disrupt the gene *ADE2* in *S. cerevisiae*. When this gene is not functional and external adenine is not provided, the yeast will turn a red color. This gene was chosen to be disrupted first because the red color would make it easy to confirm the function of the CRISPR/Cas9 plasmid. The gBlock contains two sequences, direct repeats, on each end. In between these direct repeats is a 100 base pair homologous donor sequence. This is a sequence that exists within the *ADE2* gene but contains an 8 base pair deletion in the middle. This deletion includes the 3 base pair protospacer adjacent motif (PAM) sequence, and the last 3 base pairs of a guide sequence. This guide sequence, a 20 base pair sequence contained in the *ADE2* gene, is also contained between the two direct repeats, directly following the homologous donor. The gBlock design is shown in Figure 68. The sequence of the gBlock can be found in Table 15.

Figure 68: Schematic of the gBlock constructed to disrupt the *ADE2* gene in *S. cerevisiae* strain BY4741.

To ensure a sufficient amount of the gBlock for cloning, the first step taken after receiving the gBlock from IDT is to amplify the DNA using PCR. First, a "master mix" is made which contains all components necessary for amplification. These components

103

are listed in Table 19. The primers used for the *ADE2* disruption gBlock are shown in Table 16. To create a 50 µL reaction, 47.5 µL of the master mix is added to a PCR tube. Then 2.5 µL of the gBlock DNA is added to the tube. The tube is then placed in a thermocycler, and the program shown in Table 21 is run. After the PCR is complete, the reaction was purified using a PCR cleanup kit (Monarch®, New England Biolabs®).

To clone the gBlock into pCRCT, a restriction digest is performed on both the insert and the plasmid. These digestions were done in separate tubes. For the gBlock digestion, 1 µg of the amplified and purified DNA is added to a tube, along with 1 µL of BsaI restriction enzyme, 3 µL of 10X buffer, and dH_2O as necessary to bring the final volume in the tube to 30 µL. For the plasmid digestions, 1 µg of the mini-prepped plasmid is added to a tube, along with 1 µL of BsaI restriction enzyme, 3 µL of 10X buffer, and dH_2O as necessary to bring the final volume in the tube to 30 µL. The tubes are incubated at 37°C for one hour, then the restriction enzyme is inactivated by incubating at 65°C for 20 minutes.

The digested plasmid and insert are then ligated together. All components necessary for the ligation, namely 25 µg of the digested plasmid, 75 µg of the digested insert, 1 µL of 10X buffer, 1 µL DNA ligase, and dH_2O as necessary to bring the final volume to 10 µL are added to a tube. The tube is incubated at room temperature for up to 2 hours. Another tube is run as a negative control, with all components except for the insert included. The result of this reaction will be the plasmid re-ligating with itself.

After the insert was ligated with the plasmid to create a new plasmid named pCRCT+ADE2dis, this new plasmid, as well as the negative control, is then transformed

into *E. coli* for plasmid replication and storage purposes. For this purpose, electrocompetent cells made by NEB® and stored at -80°C are used. An aliquot of these cells is removed from the freezer, thawed at room temperature and then placed in an ice bath. An electroporation cuvette is also placed in the ice bath. 10 pg to 25 ng of pCRCT+ADE2dis (or negative control) is added to the tube, and the tube is incubated on ice for 30 to 60 seconds. The electroporation apparatus is set to deliver a pulse of 25 μF capacitance, 2.5 kV, and 200 ohm resistance. The DNA/cell mixture is pipetted into the ice-cold electroporation cuvette and the outside of the cuvette is quickly dried. The electrical pulse is delivered, and 1 mL of SOC medium is added to the cuvette as quickly as possible. The entire contents of the cuvette are transferred to a test tube and mixed gently at 37°C for one hour. The contents of the test tube are then plated across several SOB+20 mM MgSO$_4$+AMP 100 agar plates. The plates are allowed to dry at room temperature and then incubated at 37°C for 12-16 hours. When colonies appear, a culture is started in LB medium from a single colony. The culture can then be used to create a frozen stock for later use of the plasmid, or used for immediate plasmid purification and yeast transformation.
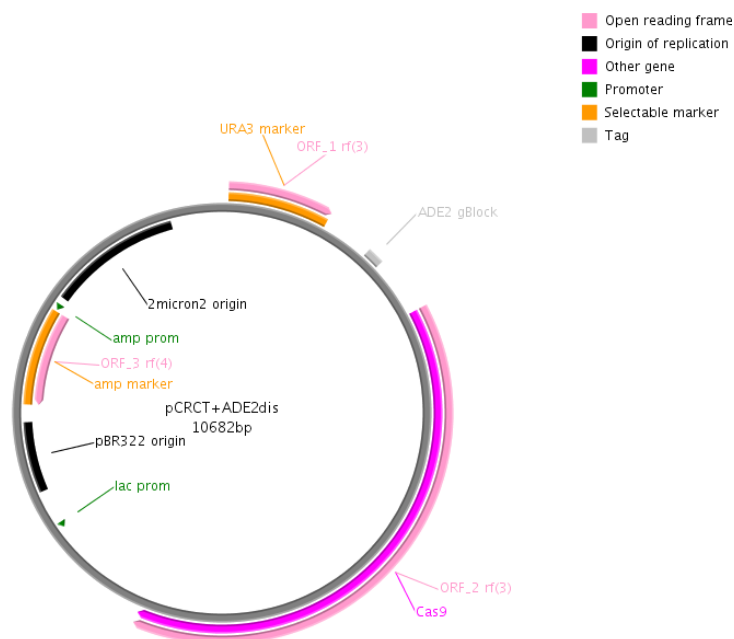
Figure 69: Diagram of the pCRCT+ADE2dis plasmid. The three main components, the ampicillin resistance marker, the uracil marker, and the Cas9 protein and related RNAs, are shown. Also shown is the site at which the ADE2 gBlock DNA was inserted [128].

When the SM14 yeast strain was made, the cassette containing the beta-carotene production genes was inserted into the chromosome at the URA3 locus. Therefore, the URA3 marker on the plasmids pCRCT, pCRCT+ADE2dis, or any subsequent disruption plasmids constructed would not be useful. Therefore, it was desired to construct a new plasmid that contained a hygromycin resistance gene in addition to the necessary CRISPR-Cas9 components. The hygromycin gene was taken from plasmid pAG26, which was obtained from the Kao lab and was originally a gift from John McCusker [126]. A map of this plasmid is shown in Figure 70.
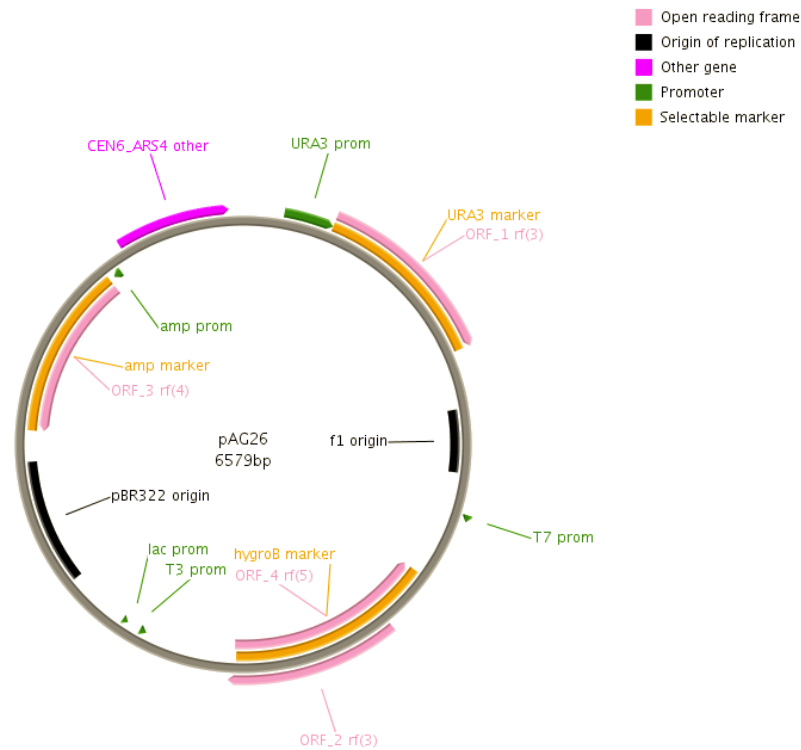
Figure 70: Diagram of the pAG26 plasmid. The ampicillin resistance marker, the uracil marker, and the hygromycin resistance marker are shown.

The hygromycin gene and its promoter was amplified from plasmid pAG26 using primers that also contained a BsaI restriction site, a junction to connect with a gene homology sequence, and a 5' DR sequence to complete the disruption gBlock. The sequences of the primers used can be found in the Appendix in Table 16. To test this method, the *CAN1* gene was chosen to be disrupted in SM14. If this gene was successfully disrupted, when grown in the presence of usually toxic canavanine, the cells will survive. The full gBlock design containing the CAN1 homology sequence, the guide sequence, and the hygromcin resistance gene, is shown in Figure 71. In the same manner that the adenine gene disruption gBlock was inserted into the pCRCT plasmid, so was

this newly-designed gBlock inserted into the same plasmid. In this case, three components were digested and ligated together rather than two as described above. This created a new plasmid called MDp001. Its design is identical to plasmid pCRCT+ADE2dis shown in Figure 69, except for the difference in inserted gBlock.
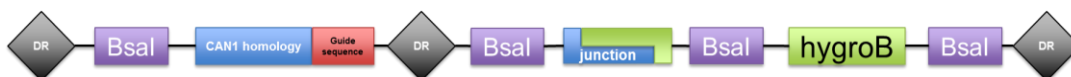
Figure 71: Schematic of the gBlock constructed to disrupt the *CAN1* gene in *S. cerevisiae* strain SM14, using hygromycin resistance as selective pressure for plasmid retention.

8.2.3 Yeast Transformation

For the CRISPR-Cas9 work, plasmid pCRCT+ADE2dis and the negative control were transformed into BY4741, an *S. cerevisiae* strain with a URA3 knockout. Plasmid pCRCT+CAN1dis and a negative control were transformed into strain SM14. A chemical transformation method was used. A lithium acetate mix, 50% PEG solution, and a PEG mix are made fresh. The recipes for each of these solutions can be found in the Appendix. First, a frozen stock of BY4741 or SM14 is plated on a YPD agar plate. A single colony from this plate is used to start an overnight culture in YPD liquid medium, grown at 30°C. A 50 mL culture from the overnight culture is then started and grown to an OD of approximately 3 OD. The cells are then spun at 3600 rpm for 5 minutes and resuspended in 10 mL of sterile H$_2$O. The cells are again spun down and resuspended in 10 mL of sterile H$_2$O. The cells are spun down again, and this time resuspended in 10 mL of the lithium acetate mix. The cells are spun down a final time, and resuspended in

108

100 µL of lithium acetate mix. The cells are then incubated at 30°C for at least 15 minutes. The cells are now competent and ready for transformation. In a microcentrifuge tube, purified pCRCT+ADE2dis or MDp001 is added, followed by 10 µL of salmon sperm DNA, 300 µL PEG mix, and finally 50 µL of competent cells (BY4741 or SM14, respectively). The tube is inverted gently and incubated at 30°C for at least 30 minutes. This tube is then heat shocked at 42°C for 20 minutes, spun down for 10-15 seconds, and the supernatant is decanted. The tube is then spun very briefly and the supernatant is decanted using vacuum. The pellet is then gently resuspended in 200 µL of YPAD medium and incubated at 30°C for at least one hour.

For the PCR-constructed deletion module work, the three deletion modules were transformed (separately) into SM14 using a similar method as described above. The differences came at the end of the protocol, where the SM14 cells were resuspended in 200 µL of YPD medium instead of YPAD, incubated at 30°C for at least one hour, and then plated on YPD+hygromycinB plates.

8.2.4 Gene Disruption Confirmation

For the BY4741 experiments, after the transformants have been recovered in YPAD medium for an hour, the cells are split, washed with sterile H_2O once, and then resuspended in synthetic complete liquid medium (SCM) without uracil, SCM without adenine, and SCM without both uracil and adenine. The cells were also plated on solid media of these three types. The cells were passed every 24 hours into fresh media.

For the SM14 experiments, after the transformants have been recovered, the cells were grown in YNB medium containing hygromycin. Several cell passes were

performed into fresh YNB+hygromycin medium to allow for gene disruption. The cells

were also plated on YNB plates containing canavanine to confirm gene disruption.

8.3 Results

8.3.1 Traditional Method

While the overlap PCRs proved difficult to carry out, amplifying the pAG26

plasmid with primers that contain 70 bp homology to either the *CIT3, LPD1,* or *LRO1*

gene was successful. Figure 72 shows the results of these PCRs. For each primer pair,

after confirming the optimal annealing temperature, 6-8 replicate reactions were run to

create enough DNA to transform into SM14. These replicate reactions were combined

and purified, and 700-1000 ng of DNA was transformed into SM14 for each construct.
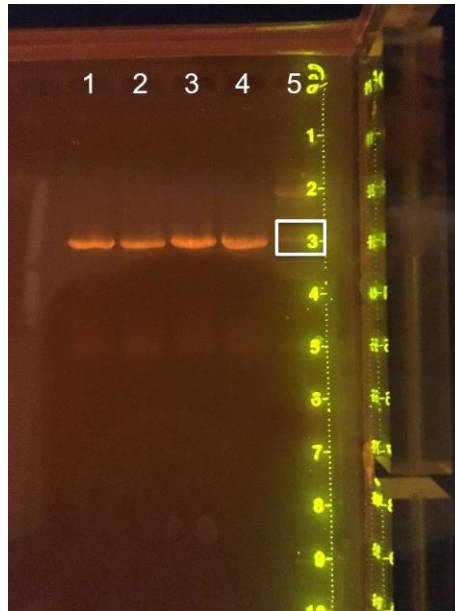
Figure 72: Blue light visualization of an agar gel containing DNA from the amplification of pAG26 with primers containing 70bp gene homology. 1) hygromycin resistance gene + 70bp upstream and downstream *CIT3* homology. 2) hygromycin resistance gene + 70bp upstream and downstream *LPD1* homology. 3 and 4) hygromycin resistance gene + 70bp upstream and downstream *LRO1* homology. 5) 1 kb DNA ladder. The boxed fragment is 2000 bp. The amplified products in lanes 1-4 are expected to be 1766 bp, and appear to be slightly below the indicated ladder fragment.

After transformation, colonies were seen on the agar plates containing the *LPD1* construct transformants as well as the *CIT3* construct transformants, seen in Figure 73 and Figure 74. However, additional testing confirmed that these colonies were false positives. Figure 75 shows the results of the confirmation tests. The contents of each lane of the gel are summarized and explained in Table 7.
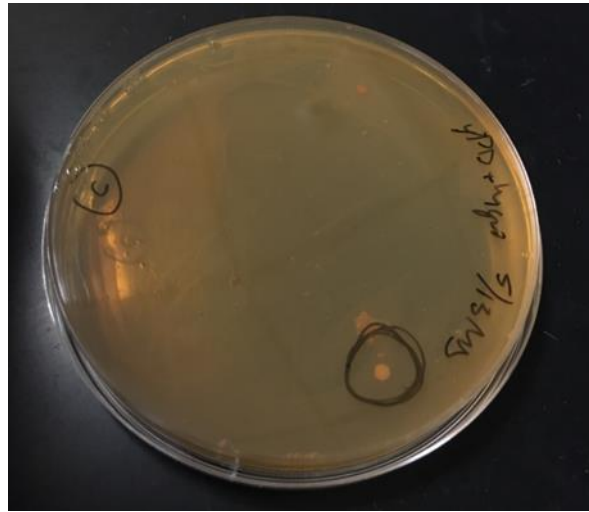
Figure 73: YPD+hygromycin agar plated with SM14 cells transformed with the *LPD1* construct. The colony is circled.

Figure 74: YPD+hygromycin agar plated with SM14 cells transformed with the *CIT3* construct. The colony is in the lower right portion of the plate.
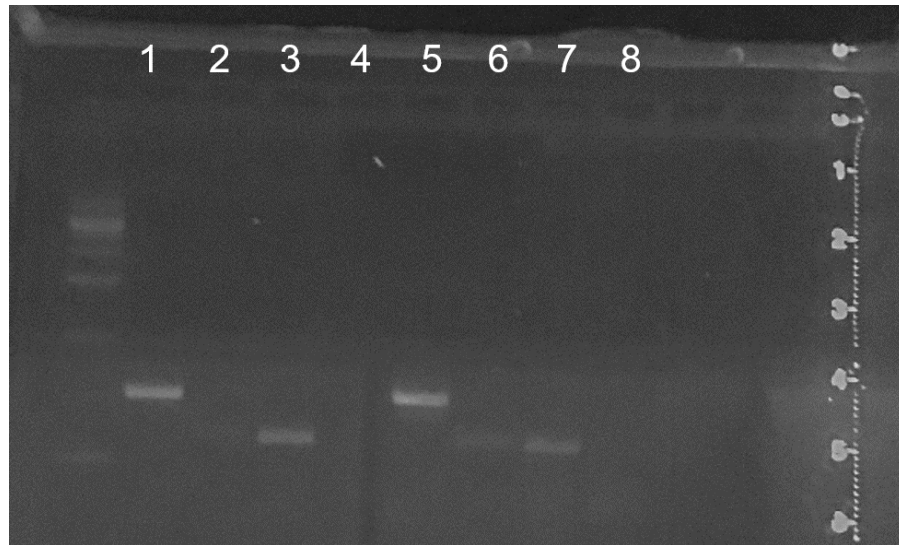
Figure 75: Visualization of an agar gel containing DNA from the amplification of SM14 and the colonies seen from the *CIT3* and *LPD1* deletion module transformations. The DNA ladder is shown to the left of the eight lanes.

Table 7: Description of results from PCRs run to confirm the knockouts in SM14.

Lane	Contents	Band Seen	Band Expected
1	SM14 genomic DNA w/ primers to amplify within CIT3 gene	~500 bp	517 bp
2	SM14 genomic DNA w/ primers to amplify CIT3 homology and hygromycin resistance gene	None	None
3	SM14 genomic DNA w/ primers to amplify within LPD1 gene	~300 bp	295 bp
4	SM14 genomic DNA w/ primers to amplify LPD1 homology and hygromycin resistance gene	None	None
5	Genomic DNA from CIT3 deletion module transformant w/ primers to amplify within CIT3 gene	~500 bp	None
6	Genomic DNA from CIT3 deletion module transformant w/ primers to amplify CIT3 homology and hygromycin resistance gene	None	302 bp
7	Genomic DNA from LPD1 deletion module transformant w/ primers to amplify within LPD1 gene	~300 bp	None
8	Genomic DNA from LDP1 deletion module transformant w/ primers to amplify LDP1 homology and hygromycin resistance gene	None	740 bp

8.3.2 CRISPR-Cas9 Method

114

After the pCRCT+ADE2dis plasmid was transformed into BY4741, the cells were plated on the three different media discussed in the methods section. Liquid cultures were also started in the three different media. These initial liquid cultures did not grow, and no further cell passes were performed. The plates took over the standard 48 hours to grow, but colonies eventually appeared. Figure 76 shows a SCM-URA agar plate containing the negative control cells as well as a SCM-URA agar plate containing the pCRCT+ADE2dis cells. The negative control cells were the usual white color, while the cells with the pCRCT+ADE2dis plasmid had a reddish color, indicating the successful disruption of the *ADE2* gene. One of these red colonies was used to restart liquid cultures in the three types of SCM media. These cultures were grown for 48 hours, and then new cultures were started with 5% inoculum from the initial flasks. A second cell pass was performed in this same manner, and these flasks are shown in Figure 77 after 48 hours of growth. As can be seen by the differences in color and density of the three flasks, selective pressure plays an important role in the effectiveness of the pCRCT+ADE2dis plasmid. The cells grew fairly robustly in media without uracil but with provided adenine, and were reddish in color, indicating the successful gene disruption. The cells grew fairly robustly in the media without adenine but with uracil, but were not reddish in color. This result indicates a loss of the plasmid due to no selective pressure on the host cells. The media containing neither adenine nor uracil did not provide a healthy environment for the cells, and the culture did not grow nearly as robustly as the other two.
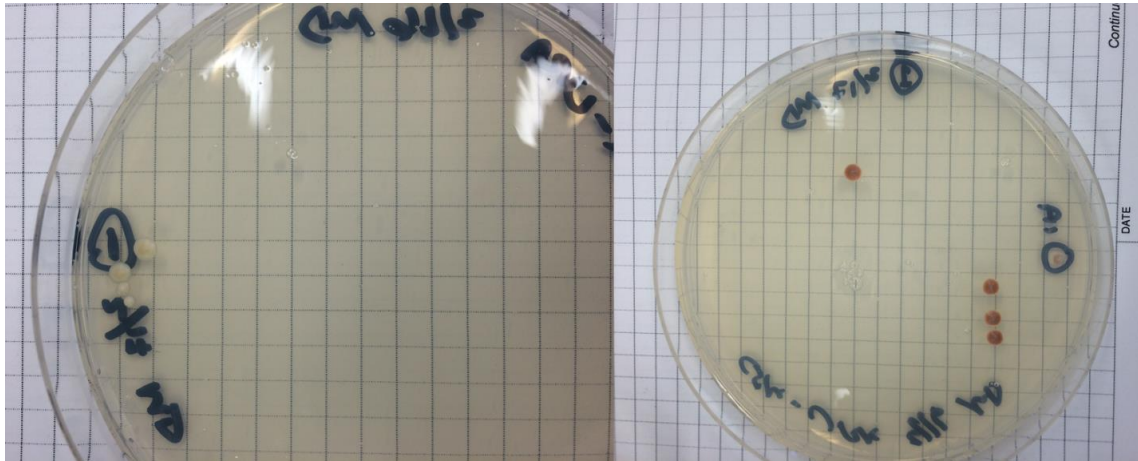
Figure 76: The transformed and plated negative control (left, white colonies) and pCRCT+ADE2dis (right, red colonies) on SCM-URA agar medium. The circled colony on the plate on the right was used to start additional liquid cultures.
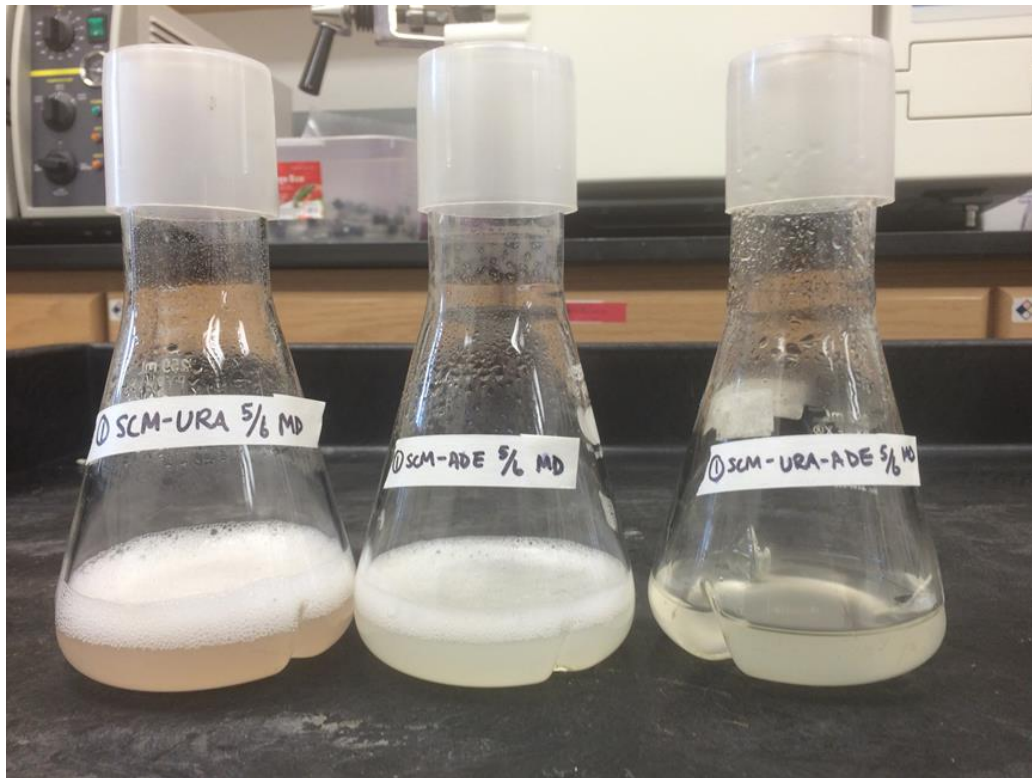
Figure 77: Second pass of BY4741 with pCRCT+ADE2dis plasmid after 48 hours, in three synthetic media formulae.

After plasmid MDp001 was transformed into SM14, the cells grew in hygromycin-containing medium, confirming the successful transformation of the plasmid, as well as the successful cloning of the hygromycin resistance gene into the pCRCT plasmid. When the cells were plated on medium containing canavanine, colonies grew, but took more than twice the normal amount of time to grow. These plates are shown in Figure 78.
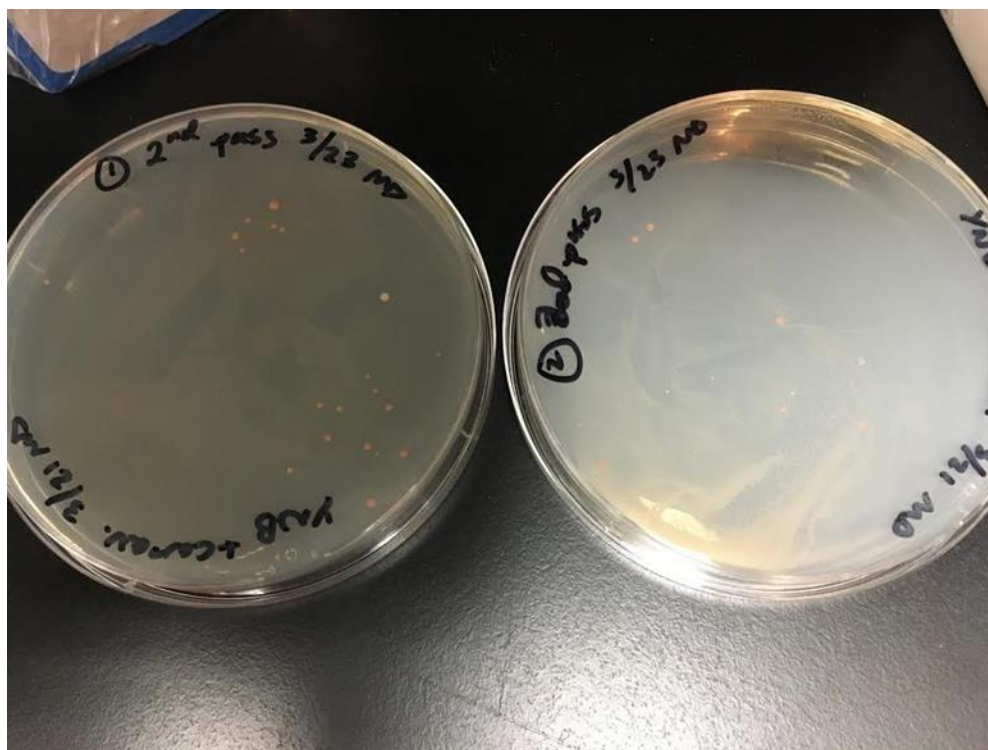
Figure 78: Yeast strain SM14 plated on YNB+canavanine medium after transformation with MDp001, and after two cell passes in hygromycin-containing liquid YNB medium.

Several colonies were sequenced, as well as the MDp001 plasmid. As can be seen in Figure 79, the canavanine gene disruption was not seen in the colonies that were

sequenced. The sequencing of MDp001 was inconclusive and was re-sequenced, but the

only thing that could be confirmed was that the hygromycin resistance gene was

successfully cloned into the plasmid. These results are shown in Figure 80.

Figure 79: Diagram showing the results of sequencing of the colonies grown after MDp001 transformation. The amplified canavanine gene from the colonies (top green block) is compared with the homology sequence contained within the inserted gBlock (bottom green/red block).

Figure 80: Diagram showing the results of sequencing purified MDp001 plasmid. The plasmid without inserts (pCRCT), the designed MDp001 plasmid, the hygromycin resistance gene, and the CAN1 disruption gBlock cassette are all compared.

118

8.4 Conclusion

Using the CRISPR/Cas9 system, the *ADE2* gene was successfully disrupted in yeast strain BY4741. The attempted disruption of the *CAN1* gene in SM14 was not successful due to difficulties in cloning attempts. In future work, an SM14 strain with a *URA3* knock out constructed in the Kao lab could be used for this work without any additional cloning of the pCRCT plasmid. Using a traditional gene knockout method, two genes identified by the Optknock tool were knocked out to create two separate single-knock out strains. Further work needs to be completed to confirm these strains and test their performance compared to the baseline SM14 strain. The genomic engineering work performed in this study using both genetic manipulation methods shows promise for future improvements of this strain of SM14. In conjunction with the results from the Optknock work performed, genes in SM14 can be targeted and new strains can be constructed that produce beta-carotene at higher levels and/or at an increased rate. It is also easy to envision the use of this versatile method of targeted strain engineering in different organisms for the production of many different chemicals.

9. CONCLUSIONS AND FUTURE WORK

In the bioreactor experiments performed in this study, the beta-carotene production was lower than historical runs, while the production of biomass, ethanol, and acetic acid was higher. As seen in section 3.3, the performance of strain SM14 seems to have declined over the course of its use, with no discernible cause. If this strain is to be used in future work, this decline in performance needs to be addressed and the cause pinpointed, if possible.

In this study, the *ADE2* gene was successfully disrupted in yeast strain BY4741 using CRISPR/Cas9. The attempted disruption of the *CAN1* in SM14 was not successful, although the work was abandoned before additional cloning troubleshooting could be completed. In future work, an SM14 strain with a *URA3* knock out could be used for this work without any additional cloning of the pCRCT gene. Pending further confirmation, the *CIT3* and *LPD1* genes were separately knocked out of SM14, and may be compared with baseline SM14 to determine if the production of beta-carotene increases. The engineering work performed in this study shows promise for future improvements of this strain of SM14. In conjunction with the results from the Optknock work performed, genes in SM14 can be targeted and new strains can be constructed that produce beta-carotene at higher levels and/or at an increased rate.

The steady-state flux balance analysis performed in this study showed that yeast intracellular models, in conjunction with the COBRA toolbox, can be used to predict behavior of certain key pathways. It may also be used to investigate the intracellular

behavior during the steady-state of a continuous bioreactor process, albeit only at low dilution rates. Furthermore, the dynamic flux balance analysis method developed in this work shows promise in predicting dynamic behavior. The intracellular model may not be able to be quantitatively validated with experimental results, but with more fine-tuning, such as the inclusion of oxygen uptake and cell death terms in the kinetic model and the inclusion of toxicity information, this model and method may be able to qualitatively predict intracellular flux trends. The intracellular behavior of SM14 during a continuous bioprocess was also examined. The optimal dilution rate value for beta-carotene production was found to 0.07 h^{-1}. In addition, the behavior of several main pathways and reactions within the cell over a range of dilution rates was studied. In future work, the flux values of all of the intracellular reactions may be inspected for bottlenecks, which could inform future genetic modifications.

Both the experimental and computational analysis work performed in this study comprise an overall metabolic engineering strategy that can be expanded and used for the production of many different bioproducts in different organisms and using different sources of carbon.

Several novel contributions were made in this work. Specifically this document describes novel work performed using an intracellular flux model and the COBRA optimization toolbox in Matlab®. The dynamic flux balance analysis dual-optimization procedure has not been performed in other works, especially in conjunction with a full-genome model of recombinant *S. cerevisiae*. The analysis of the intracellular behavior of a continuous beta-carotene production process was performed for the first time using the

COBRA toolbox and an intracellular flux model. The methods and work discussed in this document have led to new and further understanding of the beta-carotene production strain of *S. cerevisiae*, as well as the batch bioreactor process for beta-carotene production.

REFERENCES

1.   Salvi, B.L., K.A. Subramanian, and N.L. Panwar, *Alternative fuels for transportation vehicles: A technical review.* Renewable and Sustainable Energy Reviews, 2013. **25**: p. 404-419.

2.   PriceWaterhouseCoopers. *Pharma 2020: from vision to decision.* Available from: https://www.pwc.com/pharma2020.

3.   Cutler, D.M., *Rising Medical Costs Mean More Rough Times Ahead.* JAMA, 2017. **318**(6): p. 508-509.

4.   Guangyi Wang, W.T., and Robert R. Bidigare, *Terpenoids As Therapeutic Drugs and Pharmaceutical Agents*, in *Natural Products: Drug Discovery and Therapeutic Medicine*, L.Z.a.A.L. Demain, Editor. 2005, Humana Press, Inc.: Totowa, NJ. p. 197-227.

5.   Tippmann, S., et al., *From flavors and pharmaceuticals to advanced biofuels: production of isoprenoids in Saccharomyces cerevisiae.* Biotechnol J, 2013. **8**(12): p. 1435-44.

6.   Becker, J. and C. Wittmann, *Systems metabolic engineering of Escherichia coli for the heterologous production of high value molecules-a veteran at new shores.* Curr Opin Biotechnol, 2016. **42**: p. 178-188.

7.   Wang, J., et al., *Microbial production of value-added nutraceuticals.* Curr Opin Biotechnol, 2016. **37**: p. 97-104.

8.   Meadows, C.W., A. Kang, and T.S. Lee, *Metabolic Engineering for Advanced Biofuels Production and Recent Advances Toward Commercialization.* Biotechnol J, 2017.

9.   Konstantinov, K., *Continous bioprocessing: an interview with Konstantin Konstantinov from Genzyme. Interviewed by Prof. Alois Jungbauer and Dr. Judy Peng.* (1860-7314 (Electronic)).

10.   Zwenger, S.B., C., *"Plant Terpenoids: Applications and Future Potentials".* Biotechnology and Molecular Biology Reviews, 2008. **3**(1): p. 7.

11.   Chappell, J., *The Biochemistry and Molecular Biology of Isoprenoid Metabolism.* Plant Physiol, 1995. **107**(1): p. 1-6.

12.   Chojnacki, T. and G. Dallner, *The biological role of dolichol.* Biochemical Journal, 1988. **251**(1): p. 1-9.

13.   Gabay, O., et al., *Stigmasterol: a phytosterol with potential anti-osteoarthritic properties.* Osteoarthritis and Cartilage, 2010. **18**(1): p. 106-116.

14.   Kamatou, G.P., et al., *Menthol: a simple monoterpene with remarkable biological properties.* Phytochemistry, 2013. **96**: p. 15-25.

15.   Paleg, L.G., *Physiological Effects of Gibberellic Acid: I. On Carbohydrate Metabolism and Amylase Activity of Barley Endosperm.* Plant Physiology, 1960. **35**(3): p. 293-299.

16.   Perilli, S., L. Moubayidin, and S. Sabatini, *The molecular basis of cytokinin function.* Curr Opin Plant Biol, 2010. **13**(1): p. 21-6.

17.   Sitton, D. and C.A. West, *Casbene: An anti-fungal diterpene produced in cell-free extracts of Ricinus communis seedlings.* Phytochemistry, 1975. **14**(9): p. 1921-1925.

18.   Ward, E.W.B., C.H. Unwin, and A. Stoessl, *Postinfectional inhibitors from plants. XIII. Fungitoxicity of the phytoalexin, capsidiol, and related sesquiterpenes.* Canadian Journal of Botany, 1974. **52**(12): p. 2481-2488.

19.   Breitmaier, E., *Terpenes: Importance, General Structure, and Biosynthesis*, in *Terpenes*. 2006, Wiley-VCH Verlag GmbH & Co. KGaA. p. 1-9.

20.   Abdallah, I. and W.J. Quax, *A Glimpse into the Biosynthesis of Terpenoids*, in *International Conference on Natural Resources and Life Sciences*. 2017, KnE Life Sciences. p. 81-98.

21.   Zhang, L., et al., *Chassis and key enzymes engineering for monoterpenes production.* Biotechnol Adv, 2017.

22.   Chemler, J.A., Y.J. Yan, and M.A.G. Koffas, *Biosynthesis of isoprenoids, polyunsaturated fatty acids and flavonoids in Saccharomyces cerevisiae.* Microbial Cell Factories, 2006. **5**.

23.   Paramasivan, K. and S. Mutturi, *Progress in terpene synthesis strategies through engineering of Saccharomyces cerevisiae.* Crit Rev Biotechnol, 2017. **37**(8): p. 974-989.

24.   Paltauf, F.K., S; and Henry, SA, *Regulation and compartmentalization of lipid synthesis in yeast.*, in *The Molecular and Cellular Biology of the Yeast*

Saccharomyces: Gene Expression, E.P. Jones, JR; and Broach, JR, Editor. 1992, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY.

25. Parks, L.W. and W.M. Casey, *Physiological Implications of Sterol Biosynthesis in Yeast.* Annual Review of Microbiology, 1995. **49**(1): p. 95-116.

26. Liao, P., et al., *The potential of the mevalonate pathway for enhanced isoprenoid production.* Biotechnol Adv, 2016. **34**(5): p. 697-713.

27. Armstrong, G.A. and J.E. Hearst, *Carotenoids 2: Genetics and molecular biology of carotenoid pigment biosynthesis.* FASEB J, 1996. **10**(2): p. 228-37.

28. Grune, T., et al., *β-Carotene Is an Important Vitamin A Source for Humans.* The Journal of Nutrition, 2010. **140**(12): p. 2268S-2285S.

29. Ausich, R.L., *Commercial opportunities for carotenoid production by biotechnology.* Pure and Applied Chemistry, 1997. **69**(10): p. 2169-2173.

30. Frengova, G.I. and D.M. Beshkova, *Carotenoids from Rhodotorula and Phaffia: yeasts of biotechnological importance.* J Ind Microbiol Biotechnol, 2009. **36**(2): p. 163-80.

31. Johnson, E.A. and W.A. Schroeder, *Microbial carotenoids*, in *Downstream Processing Biosurfactants Carotenoids*. 1996, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 119-178.

32. Lee, P.C. and C. Schmidt-Dannert, *Metabolic engineering towards biotechnological production of carotenoids in microorganisms.* Appl Microbiol Biotechnol, 2002. **60**(1-2): p. 1-11.

33. Vachali, P., P. Bhosale, and P.S. Bernstein, *Microbial Carotenoids*, in *Microbial Carotenoids From Fungi: Methods and Protocols*, J.-L. Barredo, Editor. 2012, Humana Press: Totowa, NJ. p. 41-59.

34. Hagman, A. and J. Piškur, *A Study on the Fundamental Mechanism and the Evolutionary Driving Forces behind Aerobic Fermentation in Yeast.* PLoS ONE, 2015. **10**(1): p. e0116942.

35. Hagman, A., T. Säll, and J. Piškur, *Analysis of the yeast short-term Crabtree effect and its origin.* The Febs Journal, 2014. **281**(21): p. 4805-4814.

36. Aguilera, A. and F.K. Zimmermann, *Isolation and molecular analysis of the phosphoglucose isomerase structural gene of Saccharomyces cerevisiae.* Mol Gen Genet, 1986. **202**(1): p. 83-9.

37. Lobo, Z. and P.K. Maitra, *Phosphofructokinase mutants of yeast. Biochemistry and genetics.* J Biol Chem, 1983. **258**(3): p. 1444-9.

38. Schwelberger, H.G., S.D. Kohlwein, and F. Paltauf, *Molecular cloning, primary structure and disruption of the structural gene of aldolase from Saccharomyces cerevisiae.* Eur J Biochem, 1989. **180**(2): p. 301-8.

39. Blake, C.C. and D.W. Rice, *Phosphoglycerate kinase.* Philos Trans R Soc Lond B Biol Sci, 1981. **293**(1063): p. 93-104.

40. Lam, K.B. and J. Marmur, *Isolation and characterization of Saccharomyces cerevisiae glycolytic pathway mutants.* J Bacteriol, 1977. **130**(2): p. 746-9.

41. McAlister, L. and M.J. Holland, *Targeted deletion of a yeast enolase structural gene. Identification and isolation of yeast enolase isozymes.* J Biol Chem, 1982. **257**(12): p. 7181-8.

42. Boles, E., et al., *Characterization of a glucose-repressed pyruvate kinase (Pyk2p) in Saccharomyces cerevisiae that is catalytically insensitive to fructose-1,6-bisphosphate.* J Bacteriol, 1997. **179**(9): p. 2987-93.

43. Fernie, A.R., F. Carrari, and L.J. Sweetlove, *Respiratory metabolism: glycolysis, the TCA cycle and mitochondrial electron transport.* Curr Opin Plant Biol, 2004. **7**(3): p. 254-61.

44. Bennetzen, J.L. and B.D. Hall, *The primary structure of the Saccharomyces cerevisiae gene for alcohol dehydrogenase.* J Biol Chem, 1982. **257**(6): p. 3018-25.

45. Navarro-Avino, J.P., et al., *A proposal for nomenclature of aldehyde dehydrogenases in Saccharomyces cerevisiae and characterization of the stress-inducible ALD2 and ALD3 genes.* Yeast, 1999. **15**(10A): p. 829-42.

46. Crabtree, H.G., *The carbohydrate metabolism of certain pathological overgrowths.* Biochemical Journal, 1928. **22**(5): p. 1289-1298.

47. De Deken, R.H., *The Crabtree effect: a regulatory system in yeast.* J Gen Microbiol, 1966. **44**(2): p. 149-56.

48. Verwaal, R., et al., *High-level production of beta-carotene in Saccharomyces cerevisiae by successive transformation with carotenogenic genes from Xanthophyllomyces dendrorhous.* Appl Environ Microbiol, 2007. **73**(13): p. 4342-50.

49.   Moise, A.R., S. Al-Babili, and E.T. Wurtzel, *Mechanistic aspects of carotenoid biosynthesis.* Chem Rev, 2014. **114**(1): p. 164-93.

50.   Olson, M., *Metabolic Engineering of S. cerevisiae for Carotenoid Production Optimization*, in *Chemical Engineering*. 2014, Texas A&M University.

51.   Jaladi, T., *Development of a Fed-Batch Process to Produce Beta-Carotene Using Engineered Saccharomyces cerevisiae*, in *Chemical Engineering*. 2016, Texas A&M University.

52.   M. Carolina Ordonez, J.P.R., Tejasvi Jaladi, Xinhe Chen, Katy Kao, M. Nazmul Karim, *Modelling of batch kinetics of aerobic carotenoid production using Saccharomyces cerevisiae.* Biochemical Engineering Journal, 2016. **114**: p. 226-236.

53.   Reyes, L.H., J.M. Gomez, and K.C. Kao, *Improving carotenoids production in yeast via adaptive laboratory evolution.* Metabolic Engineering, 2014. **21**: p. 26-33.

54.   Orth, J.D., I. Thiele, and B.O. Palsson, *What is flux balance analysis?* Nature Biotechnology, 2010. **28**(3): p. 245-248.

55.   Boghigian, B.A., et al., *Metabolic flux analysis and pharmaceutical production.* Metabolic Engineering, 2010. **12**(2): p. 81-95.

56.   Keasling, J.D., *Synthetic biology and the development of tools for metabolic engineering.* Metabolic Engineering, 2012. **14**(3): p. 189-195.

57.   Stephanopoulos, G., H. Alper, and J. Moxley, *Exploiting biological complexity for strain improvement through systems biology.* Nature Biotechnology, 2004. **22**: p. 1261.

58.   Plata, G., et al., *Reconstruction and flux-balance analysis of the Plasmodium falciparum metabolic network.* Molecular Systems Biology, 2010. **6**(1).

59.   Xu, P., N. Bhan, and M.A.G. Koffas, *Engineering plant metabolism into microbes: from systems biology to synthetic biology.* Current Opinion in Biotechnology, 2013. **24**(2): p. 291-299.

60.   Becker, S.A., et al., *Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox.* Nat Protoc, 2007. **2**(3): p. 727-38.

61. Schellenberger, J., et al., *Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0.* Nat Protoc, 2011. **6**(9): p. 1290-307.

62. Vlassis, N., M.P. Pacheco, and T. Sauter, *Fast reconstruction of compact context-specific metabolic network models.* PLoS Comput Biol, 2014. **10**(1): p. e1003424.

63. Herrgard, M.J., et al., *A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology.* Nature Biotechnology, 2008. **26**(10): p. 1155-1160.

64. Aung, H.W., S.A. Henry, and L.P. Walker, *Revising the Representation of Fatty Acid, Glycerolipid, and Glycerophospholipid Metabolism in the Consensus Model of Yeast Metabolism.* Ind Biotechnol (New Rochelle N Y), 2013. **9**(4): p. 215-228.

65. Dobson, P.D., et al., *Further developments towards a genome-scale metabolic model of yeast.* BMC Syst Biol, 2010. **4**: p. 145.

66. Heavner, B.D., et al., *Yeast 5 - an expanded reconstruction of the Saccharomyces cerevisiae metabolic network.* BMC Syst Biol, 2012. **6**: p. 55.

67. King, Z.A., et al., *Next-generation genome-scale models for metabolic engineering.* Curr Opin Biotechnol, 2015. **35**: p. 23-9.

68. Heavner, B.D., et al., *Version 6 of the consensus yeast metabolic network refines biochemical coverage and improves model performance.* Database (Oxford), 2013. **2013**: p. bat059.

69. Kostromins, A. and E. Stalidzans, *Paint4Net: COBRA Toolbox extension for visualization of stoichiometric models of metabolism.* Biosystems, 2012. **109**(2): p. 233-9.

70. Antoniewicz, M.R., *Dynamic metabolic flux analysis--tools for probing transient states of metabolic networks.* Curr Opin Biotechnol, 2013. **24**(6): p. 973-8.

71. Jamshidi, N. and B.O. Palsson, *Mass action stoichiometric simulation models: incorporating kinetics and regulation into stoichiometric models.* Biophys J, 2010. **98**(2): p. 175-85.

72. Meadows, A.L., et al., *Application of dynamic flux balance analysis to an industrial Escherichia coli fermentation.* Metab Eng, 2010. **12**(2): p. 150-60.

73.   Feng, X., et al., *Integrating flux balance analysis into kinetic models to decipher the dynamic metabolism of Shewanella oneidensis MR-1.* PLoS Comput Biol, 2012. **8**(2): p. e1002376.

74.   Hjersted, J.L. and M.A. Henson, *Steady-state and dynamic flux balance analysis of ethanol production by Saccharomyces cerevisiae.* IET Syst Biol, 2009. **3**(3): p. 167-79.

75.   Hjersted, J.L., M.A. Henson, and R. Mahadevan, *Genome-scale analysis of Saccharomyces cerevisiae metabolism and ethanol production in fed-batch culture.* Biotechnol Bioeng, 2007. **97**(5): p. 1190-204.

76.   Monod, J., *Recherches sur la croissance des cultures bactériennes.* 1942.

77.   Luedeking, R. and E.L. Piret, *A kinetic study of the lactic acid fermentation. Batch process at controlled pH. Reprinted from Journal of Biochemical and Microbiological Technology Engineering Vol. I, No. 4. Pages 393-412 (1959).* Biotechnol Bioeng, 2000. **67**(6): p. 636-44.

78.   Kono, T., *Kinetics of Microbial Cell Growth.* Biotechnology and Bioengineering, 1968. **10**(2): p. 105-&.

79.   Kratzer, S. and H.J. Schuller, *Carbon source-dependent regulation of the acetyl-coenzyme A synthetase-encoding gene ACS1 from Saccharomyces cerevisiae.* Gene, 1995. **161**(1): p. 75-9.

80.   Herrero, P., et al., *Transcriptional regulation of the Saccharomyces cerevisiae HXK1, HXK2 and GLK1 genes.* Yeast, 1995. **11**(2): p. 137-44.

81.   Jiang, Y., et al., *BTS1 encodes a geranylgeranyl diphosphate synthase in Saccharomyces cerevisiae.* J Biol Chem, 1995. **270**(37): p. 21793-9.

82.   Pronk, J.T., H. Yde Steensma, and J.P. Van Dijken, *Pyruvate metabolism in Saccharomyces cerevisiae.* Yeast, 1996. **12**(16): p. 1607-33.

83.   Graybill, E.R., et al., *Functional comparison of citrate synthase isoforms from S. cerevisiae.* Arch Biochem Biophys, 2007. **465**(1): p. 26-37.

84.   Przybyla-Zawislak, B., et al., *Genetic and biochemical interactions involving tricarboxylic acid cycle (TCA) function using a collection of mutants defective in all TCA cycle genes.* Genetics, 1999. **152**(1): p. 153-66.

85.   Goldstein, J.L. and M.S. Brown, *Regulation of the mevalonate pathway.* Nature, 1990. **343**(6257): p. 425-30.

86. van Dijken, J.P., R.A. Weusthuis, and J.T. Pronk, *Kinetics of growth and sugar consumption in yeasts.* Antonie van Leeuwenhoek, 1993. **63**(3): p. 343-352.

87. Visser, D., et al., *Analysis of in vivo kinetics of glycolysis in aerobic Saccharomyces cerevisiae by application of glucose and ethanol pulses.* Biotechnol Bioeng, 2004. **88**(2): p. 157-67.

88. Covert, M.W. and B.O. Palsson, *Transcriptional regulation in constraints-based metabolic models of Escherichia coli.* J Biol Chem, 2002. **277**(31): p. 28058-64.

89. Chambon, C., et al., *Isolation and properties of yeast mutants affected in farnesyl diphosphate synthetase.* Curr Genet, 1990. **18**(1): p. 41-6.

90. George, K.W., et al., *Integrated analysis of isopentenyl pyrophosphate (IPP) toxicity in isoprenoid-producing Escherichia coli.* Metab Eng, 2018. **47**: p. 60-72.

91. Raftery, J.P., M.R. DeSessa, and M.N. Karim, *Economic improvement of continuous pharmaceutical production via the optimal control of a multifeed bioreactor.* Biotechnology Progress, 2017: p. n/a-n/a.

92. Johnston, M., *Feasting, fasting and fermenting. Glucose sensing in yeast and other cells.* Trends Genet, 1999. **15**(1): p. 29-33.

93. Burgard, A.P., P. Pharkya, and C.D. Maranas, *Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization.* Biotechnol Bioeng, 2003. **84**(6): p. 647-57.

94. Alper, H., et al., *Identifying gene targets for the metabolic engineering of lycopene biosynthesis in Escherichia coli.* Metabolic Engineering, 2005. **7**(3): p. 155-164.

95. Klein-Marcuschamer, D., P.K. Ajikumar, and G. Stephanopoulos, *Engineering microbial cell factories for biosynthesis of isoprenoid molecules: beyond lycopene.* Trends in Biotechnology, 2007. **25**(9): p. 417-424.

96. Miran, S.G., J.E. Lawson, and L.J. Reed, *Characterization of PDH beta 1, the structural gene for the pyruvate dehydrogenase beta subunit from Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 1993. **90**(4): p. 1252-6.

97. Steensma, H.Y., et al., *Molecular cloning of the gene for the E1 alpha subunit of the pyruvate dehydrogenase complex from Saccharomyces cerevisiae.* Eur J Biochem, 1990. **191**(3): p. 769-74.

98.   Ross, J., G.A. Reid, and I.W. Dawes, *The nucleotide sequence of the LPD1 gene encoding lipoamide dehydrogenase in Saccharomyces cerevisiae: comparison between eukaryotic and prokaryotic sequences for related enzymes and identification of potential upstream control sites.* J Gen Microbiol, 1988. **134**(5): p. 1131-9.

99.   Maeng, C.Y., et al., *Expression, purification, and characterization of the dihydrolipoamide dehydrogenase-binding protein of the pyruvate dehydrogenase complex from Saccharomyces cerevisiae.* Biochemistry, 1994. **33**(46): p. 13801-7.

100.   Stoops, J.K., et al., *On the unique structural organization of the Saccharomyces cerevisiae pyruvate dehydrogenase complex.* J Biol Chem, 1997. **272**(9): p. 5757-64.

101.   Wilcox, L.J., et al., *Transcriptional profiling identifies two members of the ATP-binding cassette transporter superfamily required for sterol uptake in yeast.* J Biol Chem, 2002. **277**(36): p. 32466-72.

102.   Basson, M.E., M. Thorsness, and J. Rine, *Saccharomyces cerevisiae contains two functional genes encoding 3-hydroxy-3-methylglutaryl-coenzyme A reductase.* Proc Natl Acad Sci U S A, 1986. **83**(15): p. 5563-7.

103.   Oulmouden, A. and F. Karst, *Nucleotide sequence of the ERG12 gene of Saccharomyces cerevisiae encoding mevalonate kinase.* Curr Genet, 1991. **19**(1): p. 9-14.

104.   McNeil, J.B., et al., *Cloning, and molecular characterization of the GCV1 gene encoding the glycine cleavage T-protein from Saccharomyces cerevisiae.* Gene, 1997. **186**(1): p. 13-20.

105.   Nagarajan, L. and R.K. Storms, *Molecular characterization of GCV3, the Saccharomyces cerevisiae gene coding for the glycine cleavage system hydrogen carrier protein.* J Biol Chem, 1997. **272**(7): p. 4444-50.

106.   Piper, M.D., et al., *Regulation of the balance of one-carbon metabolism in Saccharomyces cerevisiae.* J Biol Chem, 2000. **275**(40): p. 30987-95.

107.   Sinclair, D.A., S.P. Hong, and I.W. Dawes, *Specific induction by glycine of the gene for the P-subunit of glycine decarboxylase from Saccharomyces cerevisiae.* Mol Microbiol, 1996. **19**(3): p. 611-23.

108.	Fernandez, E., M. Fernandez, and R. Rodicio, *Two structural genes are encoding malate synthase isoenzymes in Saccharomyces cerevisiae.* FEBS Lett, 1993. **320**(3): p. 271-5.

109.	Kim, K.S., M.S. Rosenkrantz, and L. Guarente, *Saccharomyces cerevisiae contains two functional citrate synthase genes.* Mol Cell Biol, 1986. **6**(6): p. 1936-42.

110.	Albertyn, J., et al., *GPD1, which encodes glycerol-3-phosphate dehydrogenase, is essential for growth under osmotic stress in Saccharomyces cerevisiae, and its expression is regulated by the high-osmolarity glycerol response pathway.* Mol Cell Biol, 1994. **14**(6): p. 4135-44.

111.	Brewster, N.K., et al., *Regulation of pyruvate carboxylase isozyme (PYC1, PYC2) gene expression in Saccharomyces cerevisiae during fermentative and nonfermentative growth.* Arch Biochem Biophys, 1994. **311**(1): p. 62-71.

112.	Gonzalez, E., et al., *Characterization of a (2R,3R)-2,3-butanediol dehydrogenase as the Saccharomyces cerevisiae YAL060W gene product. Disruption and induction of the gene.* J Biol Chem, 2000. **275**(46): p. 35876-85.

113.	Vuralhan, Z., et al., *Identification and characterization of phenylpyruvate decarboxylase genes in Saccharomyces cerevisiae.* Appl Environ Microbiol, 2003. **69**(8): p. 4534-41.

114.	McAlister-Henn, L. and L.M. Thompson, *Isolation and expression of the gene encoding yeast mitochondrial malate dehydrogenase.* J Bacteriol, 1987. **169**(11): p. 5157-66.

115.	Oyedotun, K.S. and B.D. Lemire, *The quaternary structure of the Saccharomyces cerevisiae succinate dehydrogenase. Homology modeling, cofactor docking, and molecular dynamics simulation studies.* J Biol Chem, 2004. **279**(10): p. 9424-31.

116.	Enomoto, K., R. Ohki, and H. Muratsubaki, *Cloning and sequencing of the gene encoding the soluble fumarate reductase from Saccharomyces cerevisiae.* DNA Res, 1996. **3**(4): p. 263-7.

117.	Arthington, B.A., et al., *Cloning, disruption and sequence of the gene encoding yeast C-5 sterol desaturase.* Gene, 1991. **102**(1): p. 39-44.

118.	Fernandez, E., F. Moreno, and R. Rodicio, *The ICL1 gene from Saccharomyces cerevisiae.* Eur J Biochem, 1992. **204**(3): p. 983-90.

119.   San Filippo, J., P. Sung, and H. Klein, *Mechanism of eukaryotic homologous recombination.* Annu Rev Biochem, 2008. **77**: p. 229-57.

120.   Orr-Weaver, T.L., J.W. Szostak, and R.J. Rothstein, *Yeast transformation: a model system for the study of recombination.* Proc Natl Acad Sci U S A, 1981. **78**(10): p. 6354-8.

121.   Gardner, J.M. and S.L. Jaspersen, *Manipulating the yeast genome: deletion, mutation, and tagging by PCR.* Methods Mol Biol, 2014. **1205**: p. 45-78.

122.   Barrangou, R., et al., *CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes.* Science, 2007. **315**(5819): p. 1709-1712.

123.   Brouns, S.J.J., et al., *Small CRISPR RNAs Guide Antiviral Defense in Prokaryotes.* Science, 2008. **321**(5891): p. 960-964.

124.   Garneau, J.E., et al., *The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA.* Nature, 2010. **468**: p. 67.

125.   Deltcheva, E., et al., *CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.* Nature, 2011. **471**: p. 602.

126.   Goldstein, A.L. and J.H. McCusker, *Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae.* Yeast, 1999. **15**(14): p. 1541-53.

127.   Bao, Z., et al., *Homology-Integrated CRISPR–Cas (HI-CRISPR) System for One-Step Multigene Disruption in Saccharomyces cerevisiae.* ACS Synthetic Biology, 2015. **4**(5): p. 585-594.

128.   Dong, X., et al., *PlasMapper: a web server for drawing and auto-annotating plasmid maps.* Nucleic Acids Research, 2004. **32**(Web Server issue): p. W660-W664.

APPENDIX

Table 8: 50X TAE buffer recipe.

Component	Amount	Units
Tris Base	242	g/L
0.5 M EDTA (pH 8.0)	100	mL
Glacial Acetic Acid	57.1	mL

Table 9: SCM medium recipe.

Component	Amount	Units
YNB (without ammonium sulfate and amino acids)	1.7	g/L
Ammonium Sulfate	5	g/L
Dextrose	20	g/L
CSM-URA-ADE Dropout Mix	0.77	g/L
Uracil (if making SCM medium without adenine)	0.02	g/L
Adenine (if making SCM medium without uracil)	0.02	g/L
Filter sterilize.		

Table 10: SOC medium recipe.

Component	Amount	Units
Tryptone	20	g/L
Yeast Extract	5	g/L
NaCl	0.5	g/L
250 mM KCl	10	mL
NaOH	Variable	--
4M $MgCl_2$	2.5	mL/L
1M Glucose	80	mL/L
Tryptone, yeast extract, NaCl, and KCl should be combined in MilliQ water, and NaOH should be added as needed to adjust the pH to 7.0. The volume of this portion of the medium should be brought to 920 mL and autoclaved. The $MgCl_2$ and glucose should be added directly before use.		

Table 11: SOB plate medium recipe.

Component	Amount	Units
Tryptone	20	g/L
Yeast Extract	5	g/L
NaCl	0.5	g/L
Agar	14	g/L
250 mM KCl	10	mL
NaOH	Variable	--
4M $MgCl_2$	2.5	mL/L
100 mg/mL Ampicillin (AMP)	100	µg/mL
Tryptone, yeast extract, NaCl, KCl, and agar should be combined in MilliQ water, and NaOH should be added as needed to adjust the pH to 7.0. The volume of this portion of the medium should be brought to 1000 mL and autoclaved. $MgCl_2$ and AMP should be added aseptically before pouring plates.		

Table 12: YP(A)D medium recipe.

Component	Amount	Units
Yeast Extract	10	g/L
Peptone	20	g/L
Dextrose	20	g/L
Adenine Sulfate (if making YPAD medium)	40	mg/L
Agar (if making solid plate medium)	20	g/L
Autoclave to sterilize.		

Table 13: YNB medium recipe.

Component	Amount	Units
Dextrose	20	g/L
YNB (without ammonium sulfate and amino acids)	1.7	g/L
Ammonium Sulfate	5	g/L
Agar (if making solid plate medium)	20	g/L
Filter sterilize liquid medium. If making plate medium, the agar should be added to 500 mL MilliQ water and autoclaved. The rest of the components should be combined in 500 mL MilliQ water, filter sterilized, and added to the autoclaved portion before pouring plates.		

Table 14: Optimal estimated parameter values for the batch kinetic model.

Parameter	Value	Units
$Y_{X/G}$	0.1605	$\dfrac{g\,Dry\,Cell\,Weight}{g\,Glucose}$
$Y_{X/E}$	0.4478	$\dfrac{g\,Dry\,Cell\,Weight}{g\,Ethanol}$
$Y_{X/A}$	1.000	$\dfrac{g\,Dry\,Cell\,Weight}{g\,Acetate}$
k_1	2.958	$\dfrac{g\,Ethanol}{g\,Dry\,Cell\,Weight}$
k_2	0	$\dfrac{g\,Acetate}{g\,Dry\,Cell\,Weight}$
k_3	0.8098	$\dfrac{g\,Acetate}{g\,Dry\,Cell\,Weight}$
α_1	0.2836	$\dfrac{mg\,Product}{g\,Glucose}$
α_2	0.3710	$\dfrac{mg\,Product}{g\,Ethanol}$
α_3	0.3806	$\dfrac{mg\,Product}{g\,Acetate}$
β	0.2125	$\dfrac{mg\,Product}{g\,Dry\,Cell\,Weight\,\cdot\,hr}$
$\mu_{max,G}$	0.1821	h^{-1}
$\mu_{max,E}$	0.5000	h^{-1}
$\mu_{max,A}$	0.0004	h^{-1}
K_{SG}	0.0000	$\dfrac{g\,Glucose}{L}$
K_{SE}	20.3859	$\dfrac{g\,Ethanol}{L}$
K_{SA}	0.350948	$\dfrac{g\,Acetate}{L}$
a_{ge}	1	--
a_{ga}	1	--
a_{eg}	1	--
a_{ea}	34.6166	--
a_{ag}	1.0845	--
a_{ae}	1.0254	--

Table 15: gBlock DNA used in this work.

Name	Description	Sequence
MD_GF5	contains ADE2 homology donor, ADE2 guide sequence, and flanking spacers and BsaI restriction sites; flanked upstream and downstream by sequences for PCR amplification	GAAGCTCAGGAGTTAACTCACCTTTG GTCTCACCAAAACGATTTGGGTTTTC CATTCGTCTTGAAGTCGAGGACTTTG GCATACGATGGACTTCGTTGTAAAGA ATAAGGAAATGATTCCGGAAGCTTTG GAAGTACTGACTTTGGCATACGATGG AAGGTTTTAGAGAGAGACCTTTCACA TTGGAGCTAGCTAGGACG
MD_GF7	contains ADE2 homology donor, ADE2 guide sequence, DR sequence, BsaI restriction sites, and junction to connect with hygromycin gene; flanked upstream and downstream by sequences for PCR amplification	GCATCCGCTCTTTGGTCTCACCAAAA CGATTTGGGTTTTCCATTCGTCTTGA AGTCGAGGACTTTGGCATACGATGG ACTTCGTTGTAAAGAATAAGGAAAT GATTCCGGAAGCTTTGGAAGTACTGA CTTTGGCATACGATGGAAGGTTTTAG AGCTATGCTGTTTTGAATGGTCCCAA AACAGGCCTGAGACCTTTCTGCGCAT GC
MD_GF8	contains CAN1 homology donor, CAN1 guide sequence, DR sequence, BsaI restriction sites, and junction to connect with hygromycin gene; glanked upstream and downstream by sequences for PCR amplification	GCATCCGCTCTTTGGTCTCACCAAAA CGCAGTTTTCGTTACTGCTGCATTTG GCGCTTTGGCTTACATGGAGACATCG TGACAAAGTTTTCGAATGGCTATTAA ATATCACTGGTGTTGCAGGCTTTGCT TACATGGAGACATCTACGTTTTAGAG CTATGCTGTTTTGAATGGTCCCAAAA CAGGCCTGAGACCTTTCTGCGCATGC

Table 16: Primer sequences used in the CRISPR-Cas9 work.

Name	Description	Sequence
MD_1_f	Forward primer; amplifies MD_GF5	GAGTTAACTCACCTTTGGTCT
MD_1_r	Reverse primer; amplifies MD_GF5	TAGCTCCAATGTGAAAGGT
hygro_MD_1_f	Forward primer; amplifies hygromycin gene and T3 promoter from pAG26; flanked by BsaI site, junction to connect with gBlock, and 5' DR sequence	CTTTGGTCTCAGGCCTCCATGATTACGCCAAGCTCG
hygro_MD_1_r	Reverse primer; amplifies hygromycin gene and T3 promoter from pAG26; flanked by BsaI site, junction to connect with gBlock, and 5' DR sequence	GAAAGGTCTCTCTCTAAAACTTATTCCTTTGCCCTCGGACG
MD_2_f	Forward primer; amplifies MD_GF7	TCCGCTCTTTGGTCTCACCA
MD_2_r	Reverse primer; amplifies MD_GF7	TGCGCAGAAAGGTCTCAGG

Name	Description	Sequence
CIT3_hygroB_70_f	Forward primer; amplifies hygromycin resistance gene from plasmid pAG26; contains CIT3 homology on 5' end	TTCCTGCCAGAATCAATGTTATGGCTGTTGATGACAGGTGGTGTGCCAACTTTCCAGCAAGCTGCTTCTTGTTTAGCTTGCCTTGTCCCC
CIT3_hygroB_70_r	Reverse primer; amplifies hygromycin resistance gene from plasmid pAG26; contains CIT3 homology on 3' end	AAATTCAAGCATTGCTGTAAATCGAGGATCCGGTTTACGCAAAACTGCGTGCCCATAACCGGGGATCACACGAATCGACAGCAGTATAGCG
LPD1_hygroB_70_f	Forward primer; amplifies hygromycin resistance gene from plasmid pAG26; contains LPD1 homology on 5' end	AGAGGTATTGACGTCAACGGTGATATCAAAATTAACGTAGCAAACTTCCAAAAGGCTAAGGATGACGCTGGTTTAGCTTGCCTTGTCCCC
LPD1_hygroB_70_r	Reverse primer; amplifies hygromycin resistance gene from plasmid pAG26; contains LPD1 homology on 3' end	ATTTTATAGTCAATGCCGGCTTCTTTCAATTGCTCTTCGGTTTTACCAACCCATGCTACTTCTGGGTGAGCGAATCGACAGCAGTATAGCG
LRO1_hygroB_70_f	Forward primer; amplifies hygromycin resistance gene from plasmid pAG26; contains LRO1 homology on 5' end	TCCAGAAGACTGATTTTCATTCTTGGTGCATTCTTAGGTGTACTTTTGCCGTTTAGCTTTGGCGCTTATCGTTTAGCTTGCCTTGTCCCC
LRO1_hygroB_70_r	Reverse primer; amplifies hygromycin resistance gene from plasmid pAG26; contains LRO1 homology on 3' end	GTTCCGTCCCCCTCGGTGAGGAATACAGGTTGCTTGCTTTCGTAGTCGATGGTCAAATTCAGAGCAGAGGCGAATCGACAGCAGTATAGCG
CIT3_U_f	Forward primer; amplifies 5' homology region of CIT3	GATGACTGGCTTGTCAAGCATTC
CIT3_U_r	Reverse primer; amplifies 5' homology region of CIT3; contains hygromycin gene homology on 3' end	CGCTGGCCGGGTGACCCGGCGGGGACAAGGCAAGCTAAACAAGAAGCAGCTTGCTGGAAAG
hygroB_CIT3_f	Forward primer; amplifies hygromycin resistance gene	ATGACAGGTGGTGTGCCAACTTTCCAGCAAGCTGCTT

	from plasmid pAG26; contains CIT3 homology on 5' end	CTTGTTTAGCTTGCCTTGT CCCC
hygroB_CIT 3_r	Reverse primer; amplifies hygromycin resistance gene from plasmid pAG26; contains CIT3 homology on 3' end	CGGTTTACGCAAAACTGC GTGCCCATAACCGGGGAT CACACGAATCGACAGCAG TATAGCG
CIT3_D_f	Forward primer; amplifies 3' homology region of CIT3; contains hygromycin gene homology on 5' end	ATCGTATGTGAATGCTGGT CGCTATACTGCTGTCGATT CGTGTGATCCCCGGTTATG GG
CIT3_D_r	Reverse primer; amplifies 3' homology region of CIT3	GCTTGCTTTGGTAAGTGCT TCC
LPD1_U_f	Forward primer; amplifies 5' homology region of LPD1	CGTACCCTTTGACTCACCT CG
LPD1_U_r	Reverse primer; amplifies 5' homology region of LPD1; contains hygromycin gene homology on 3' end	CGCTGGCCGGGTGACCCG GCGGGGACAAGGCAAGCT AAACCAGCGTCATCCTTA GCCTTTTG
hygroB_LPD 1_f	Forward primer; amplifies hygromycin resistance gene from plasmid pAG26; contains LPD1 homology on 5' end	ATTAACGTAGCAAACTTC CAAAAGGCTAAGGATGAC GCTGGTTTAGCTTGCCTTG TCCCC
hygroB_LPD 1_r	Reverse primer; amplifies hygromycin resistance gene from plasmid pAG26; contains LPD1 homology on 3' end	TGCTCTTCGGTTTTACCAA CCCATGCTACTTCTGGGTG AGCGAATCGACAGCAGTA TAGCG
LPD1_D_f	Forward primer; amplifies 3' homology region of LPD1; contains hygromycin gene homology on 5' end	ATCGTATGTGAATGCTGGT CGCTATACTGCTGTCGATT CGCTCACCCAGAAGTAGC ATGGG
LPD1_D_r	Reverse primer; amplifies 3' homology region of LPD1	CCGGCATTTGGACCGATA ATG
LRO1_U_f	Forward primer; amplifies 5' homology region of LRO1	CCGTGTACAAAATTGAAT GCCCAA
LRO1_U_r	Reverse primer; amplifies 5' homology region of LRO1; contains hygromycin gene homology on 3' end	TGTCAAGGAGGGTATTCT GGGCCTCCATGTCGCTGG CCGGGTGACCCGGCGGGG ACAAGGCAAGCTAAACGA TAAGCGCCAAAGCTAAAC GG
hygroB_LRO 1_f	Forward primer; amplifies hygromycin resistance gene	TTCTTAGGTGTACTTTTGC CGTTTAGCTTTGGCGCTTA

	from plasmid pAG26; contains LRO1 homology on 5' end	TCGTTTAGCTTGCCTTGTC CCC
hygroB_LRO 1_r	Reverse primer; amplifies hygromycin resistance gene from plasmid pAG26; contains LRO1 homology on 3' end	TGCTTGCTTTCGTAGTCGA TGGTCAAATTCAGAGCAG AGGCGAATCGACAGCAGT ATAGCG
LRO1_D_f	Forward primer; amplifies 3' homology region of LRO1; contains hygromycin gene homology on 5' end	ATCGTATGTGAATGCTGGT CGCTATACTGCTGTCGATT CGCCTCTGCTCTGAATTTG ACCATCG
LRO1_D_r	Reverse primer; amplifies 3' homology region of LRO1	CTAGCCTAAAAAGGCGCG TCG
CIT3_conf_1	Confirmation primer	CAGGTGGTGTGCCAACTTT C
CIT3_conf_2	Confirmation primer	CTTCCAACCAAGTGGGTG GT
CIT3_conf_3	Confirmation primer	GGCTGTTGATGACAGGTG GTG
CIT3_conf_4	Confirmation primer	CAATTCAACGCGTCTGTG AGG
LPD1_conf_ 1	Confirmation primer	GGAAGGCACTGTCAAGGA AGA
LPD1_conf_ 2	Confirmation primer	CTTTTGGGTGGCTTTGGCA AC
LPD1_conf_ 3	Confirmation primer	CCAAATGCATACGGAAGC GCA
LPD1_conf_ 4	Confirmation primer	GGCTCTCGCTGAATTCCCC AA

Table 18: Expected overlap PCR products in the traditional yeast gene knockout work.

CIT3 deletion module
TATTTAAGATGACTGGCTTGTCAAGCATTCTAATCATAAAAAAAAGATCGTATTTGATCAAGAATTTA
TACATAGACGCCGCTAAATAATTGAATACAAAATGGTACAAAGGCTTCTACCGGGCGCACATATATG
CAGAAGGTCCTTCAATTCGTCTGCAATTATAAAGTCTTCTGCATTGACTCTCAAGGAAGCATTAGAAA
ACGTGATACCTAAGAAAAGAGATGCTGTGAAGAAATTGAAGGCCTGTTATGGCAGCACGTTTGTCGG
ACCGATTACCATTTCATCAGTTCTAGGTGGGATGAGAGGTAATCAGTCAATGTTTTGGCAAGGAACAT
CATTAGATCCCGAACATGGCATTAAATTTCAAGGTTTAACGATTGAAGAATGTCAAAATAGATTACCT
AATACAGGTATTGATGGCGATAATTTCCTGCCAGAATCAATGTTATGGCTGTTGATGACAGGTGGTGT
GCCAACTTTCCAGCAAGCTGCTTCTTGTTTAGCTTGCCTTGTCCCCGCCGGGTCACCCGGCCAGCGAC
ATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTC
GCCCGTACATTTAGCCCATACATCCCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCACG
GCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACG
CGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTC
TTCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAA
CCATGGGTAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAG
CGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGC
GTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCAC
TTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTA
TTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTC
TGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGG
CCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATC
CCCATGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGAT
GAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCA
ACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGA
TTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGC
GCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCAT
TGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGGCGCAGGGTC
GATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGC
GGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGT
CCGAGGGCAAAGGAATAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTTGTCATTTGTA
TAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGATTTATATTTTTTTTCGCCTCG
ACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAAT
GCTGGTCGCTATACTGCTGTCGATTCGTGTGATCCCCGGTTATGGGCACGCAGTTTTGCGTAAACCGG
ATCCTCGATTTACAGCAATGCTTGAATTTGCGCAAAAGAGGCCTATTGAATTTGAGAACGACAAGAA
CGTTTTGTTGATGCAAAAATTGGCAGAAATAGCGCCTAAGGTTTTGTTGGAACACGGAAAGAGTAAG
AATCCATTTCCCAATGTTGACTCTGCATCGGGGATTTTGTTTTATCATTATGGAATCAGGGAATTATTA
TTCTTTACCGTCATTTTTGGGTGTTCAAGGGCCATGGGACCCTTGACACAACTTGTTTGGGATCGCATT
CTAGGTTTACCAATTGAAAGGCCCAAGAGTTTGAACTTGGAGGGTCTGGAAGCACTTACCAAAGCAA
GCAATGTTAACAAGTTGTAACGCAGTTCCAATTTACAAGAATGCTTCGTTTGCTATTACAATATTGAA
ATATAAATAAAATTCCATACAGCATGTCTAATCATAGCTAATTTATACATAT

LPD1 deletion module
ATGCTAATAAACAATTTGATGATAATTCGTACCCTTTGACTCACCTCGAATATATATAGATATATATA
TACATATAACGTATATTTATATATATACGGTTTTGTTCGATTGTCTCTGTCGTACCATCAAAGAACATA
CTAACAGTTCACAATGTTAAGAATCAGATCACTCCTAAATAATAAGCGTGCCTTTTCGTCCACAGTCA
GGACATTGACCATTAACCAAGTCACATGATGTAGTCATCATCGGTGGTGGCCCTGCTGGTTACGTGGCT
GCTATCAAAGCTGCTCAATTGGGATTTAACACTGCATGTGTAGAAAAAAGAGGCAAATTAGGCGGTA
CCTGTCTTAACGTTGGATGTATCCCCTCCAAAGCACTTCTAAATAATTCTCATTTATTCCACCAAATGC
ATACGGAAGCGCAAAAGAGAGGTATTGACGTCAACGGTGATATCAAAATTAACGTAGCAAACTTCCA
AAAGGCTAAGGATGACGCTGGTTTAGCTTGCCTTGTCCCCGCCGGGTCACCCGGCCAGCGACATGGA
GGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATGTGACTGTCGCCCGT
ACATTTAGCCCATACATCCCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCACGGCGCGA
AGCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGCTCCCCTCACAGACGCGTTGA
ATTGTCCCCACGCCGCGCCCCTGTAGAGAAATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTT
CATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAACAACCATG
GGTAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCT
CCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGG
ATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATCGGCACTTTG

143

CATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGC
ATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCA
GCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCA
TTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCA
TGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAG
CTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCAACA
ATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATTC
CCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCT
ACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGG
TCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGAT
GCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGCGGC
CGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCG
AGGGCAAAGGAATAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTTGTCATTTGTATAG
TTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACA
TCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCT
GGTCGCTATACTGCTGTCGATTCGCTCACCCAGAAGTAGCATGGGTTGGTAAAACCGAAGAGCAATT
GAAAGAAGCCGGCATTGACTATAAAATTGGTAAGTTCCCCTTTGCGGCCAATTCAAGAGCCAAGACC
AACCAAGACACTGAAGGTTTCGTGAAGATTTTGATCGATTCCAAGACCGAGCGTATTTTGGGGGCTC
ACATTATCGGTCCAAATGCCGGTGAAATGATTGCTGAAGCTGGCTTAGCCTTAGAATATGGCGCTTCC
GCAGAAGATGTTGCTAGGGTCTGCCATGCTCATCCTACTTTGTCCGAAGCATTTAAGGAAGCTAACAT
GGCTGCCTATGATAAAGCTATTCATTGTTGAAAACAGGAAATAATAAACAGTATAGTATATATATTTA
TGAAGAACCGCTTAGTATTGAGTAAGTAAAAAAAAATTGCACAAGGAATTAATATATGAGGAAACTA
CTAGTCGATCAGATAAGATATTATGACCAGATATGCAAAAAAAGG

LRO1 deletion module

TAATAATAATGTCCGTGTACAAAATTGAATGCCCAAGAAGGTGAAATTTGAAAAATAAGAAATCTAC
TAAATACCGATACGAAGAAGCGTATAGTAACAGCCATTACAAAAGGTTCTCTACCAACGAATTCGGC
GACAATCGAGTAAAAAATGGGCACACTGTTTCGAAGAAATGTCCAGAACCAAAAGAGTGATTCTGAT
GAAAACAATAAAGGGGGGTTCTGTTCATAACAAGCGAGAGAGCAGAAACCACATTCATCATCAACAG
GGATTAGGCCATAAGAGAAGAAGGGGTATTAGTGGCAGTGCAAAAAGAAATGAGCGTGGCAAAGAT
TTCGACAGGAAAAGAGACGGGAACGGTAGAAAACGTTGGAGAGATTCCAGAAGACTGATTTTCATTC
TTGGTGCATTCTTAGGTGTACTTTTGCCGTTTAGCTTTGGCGCTTATCGTTTAGCTTGCCTTGTCCCCGC
CGGGTCACCCGGCCAGCGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTC
AGGGGCATGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATGTATAATCATTTGCATCCA
TACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGG
AAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGAAATATAAAAGGT
TAGGATTTGCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACAGTTCTCAC
ATCACATCCGAACATAAACAACCATGGGTAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTT
TCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTT
TCAGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAA
AGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGG
AATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCCT
GAAACCGAACTGCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATC
TTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGA
TTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAGTG
CGTCCGTCGCGCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTC
GTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACT
GGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTT
GGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCCGCGG
CTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGACGGCAATTTCGA
TGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGT
ACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTG
GAAACCGACGCCCCAGCACTCGTCCGAGGGCAAAGGAATAATCAGTACTGACAATAAAAAGATTCTT
GTTTTCAAGAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCG
TGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGTAATA
TCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGCTGTCGATTCGCCTCTGCTCTGAATTTGAC
CATCGACTACGAAAGCAAGCAACCTGTATTCCTCACCGAGGGGGACGGAACCGTTCCGCTCGTGGCG
CATTCAATGTGTCACAAATGGGCCCAGGGTGCTTCACCGTACAACCCTGCCGGAATTAACGTTACTAT
TGTGGAAATGAAACACCAGCCAGATCGATTTGATATACGTGGTGGAGCAAAAAGCGCCGAACACGTA

144

```
GACATCCTCGGCAGCGCGGAGTTGAACGATTACATCTTGAAAATTGCAAGCGGTAATGGCGATCTCG
TCGAGCCACGCCAATTGTCTAATTTGAGCCAGTGGGTTTCTCAGATGCCCTTCCCAATGTAAATGACC
GACATTGACTCACTATCCATCCGTGTATTATTTCAAAGAGCGAAAAGAAGGCGCGTCGCGTCGACGC
GCCTTTTTAGGCTAGAAAATAAACAGAAAACAAAAACAAAAACAAAAAAAGGCGAAAAAACAAACG
AAAAAACAAAC
```

Table 19: PCR Master Mix components for CRISPR-Cas9 experiments.

Component	Amount	Units
10X Polymerase Buffer	12.5	μL
10 mM dNTP Mix	3.75	μL
50 mM $MgCl_2$	5.00	μL
5 μM stock of forward primer	12.75	μL
5 μM stock of reverse primer	12.75	μL
Taq Polymerase	1.25	μL
Sterile H_2O	71.25	μL

Table 20: PCR components for the traditional yeast gene knockout work.

Component	Amount	Units
Phusion or Q5 master mix	25	μL
10 μM stock of forward primer	2.5	μL
10 μM stock of reverse primer	2.5	μL
Template DNA	variable	μL
Sterile H_2O	variable	μL

Table 21: Thermocycler program used in the CRISPR-Cas9 work.

Cycles	Temperature	Time
1	94°C	2 minutes
35	94°C	30 seconds
	$T_{melting}$-5°C	30 seconds
	74°C	1-2 minutes/kilobase product
1	74°C	5 minutes
1	4°C	hold

Table 22: Thermocycler program used in the traditional yeast gene knockout work.

Cycles	Temperature	Time
1	98ºC	30 seconds
35	98ºC	10 seconds
	45ºC-72ºC (calculated by NEB online calculator)	30 seconds
	72ºC	30 seconds/kilobase product
1	72ºC	10 minutes
1	4ºC	hold

Table 23: Lithium Acetate Mix recipe.

Component	Amount	Units
1M Lithium Acetate	3	mL
10X TE	3	mL
Sterile H_2O	24	mL

Table 24: PEG Mix recipe.

Component	Amount	Units
50% PEG	2	mL
10X TE	250	μL
1M Lithium Acetate	250	μL

Table 25: The full list of possible knock out reactions used in Optknock analysis.

Reaction Model ID	Reaction Name	Associated Genes
r_0003	(R,R)-butanediol dehydrogenase	YAL060W
r_0016	2-aceto-2-hydroxybutanoate synthase	YCL009C and YMR108W
r_0028	2-methylcitrate synthase	YPR001W
r_0062	3-methyl-2-oxobutanoate decarboxylase	YGR807C
r_0064	3-methyl-2-oxopentanoate decarboxylyase	YDL080C
r_0095	Acetaldehyde condensation	YGR087C
r_0097	Acetolactate synthase	YCL009C and YMR108W
r_0113	Acetyl-CoA synthetase	YAL054C
r_0114	Acetyl-CoA synthetase nuclear	YLR153C
r_0231	C-14 sterol reductase	YNL280C
r_0233	C-22 sterol desaturase (NADP)	YMR015C
r_0234	C-3 sterol dehydrogenase	YGL001C
r_0235	C-3 sterol dehydrogenase (4-methylzymosterol)	YGL001C
r_0236	C-3 sterol keto reductase (4-methylzymosterol)	YLR100W
r_0237	C-3 sterol keto reductase (zymosterol)	YLR100W
r_0238	C-4 methyl sterol oxidase	YGR060W
r_0239	C-4 methyl sterol oxidase	YGR060W
r_0240	C-4 methyl sterol oxidase	YGR060W
r_0241	C-4 sterol methyl oxidase (4,4-dimethylzymosterol)	YGR060W
r_0242	C-5 sterol desaturase	YLR056W
r_0243	C-8 sterol isomerase	YMR202W
r_0244	C-s24 sterol reductase	YGL012W
r_0300	citrate synthase	YNR001C
r_0301	citrate synthase, peroxisomal	YCR005C
r_0302	citrate to cis-aconitate(3-)	YLR304C
r_0303	citrate to cis-aconitate(3-), cytoplasmic	YLR304C
r_0355	Dimethylallyltranstransferase	YJL167W
r_0451	Fumarase	YPL262W
r_0452	Fumarase, cytoplasmic	YPL262W

Reaction Model ID	Reaction Name	Associated Genes
r_0454	Fumarate reductase	YEL047C
r_0462	Geranyltranstransferase	YJL167W
r_0489	Glycerol-3-phosphatase	YER062C
r_0491	Glycerol-3-phosphate dehydrogenase (NAD)	YDL022W
r_0492	Glycerol-3-phosphate dehydrogenase (NAD)	YOL059W
r_0501	Glycine cleavage system	YAL044C and YDR019C and YFL018C and YMR189W
r_0504	glycine-cleavage complex (lipoamide)	YAL044C and YDR019C and YFL018C and YMR189W
r_0505	glycine-cleavage complex (lipoamide)	YAL044C and YDR019C and YFL018C and YMR189W
r_0506	glycine-cleavage complex (lipoylprotein)	YAL044C and YDR019C and YFL018C and YMR189W
r_0507	glycine-cleavage complex (lipoylprotein)	YAL044C and YDR019C AND YFL018C and YMR189W
r_0508	glycine-cleavage complex (lipoylprotein)	YAL044C and YDR019C and YFL018C and YMR189W
r_0509	glycine-cleavage system (lipoamide)	YAL044C and YDR019C and YFL018C and YMR189W
r_0567	Indole-3-pyruvate decarboxylase	YGR087C
r_0658	Isocitrate dehydrogenase (NAD+)	YNL037C and YOR136W
r_0662	Isocitrate lyase	YER065C
r_0668	Itaconate-CoA ligase (ADP-forming)	YGR244C and YOR142W
r_0698	Lanosterol synthase	YHR072W
r_0713	Malate dehydrogenase	YKL085W
r_0714	malate dehydrogenase, cytoplasmic	YOL126C
r_0715	malate dehydrogenase, peroxisomal	YDL078C
r_0716	malate synthase	YIR031C

Reaction Model ID	Reaction Name	Associated Genes
r_0717	malate synthase	YNL117W
r_0831	oxoglutarate dehydrogenase (dihydrolipoamide S-succinyltransferase)	YDR148C and YFL018C and YIL125W
r_0832	oxoglutarate dehydrogenase (lipoamide)	YDR148C and YFL018C and YIL125W
r_0854	Phenylpyruvate decarboxylase	YDR380W
r_0958	pyruvate carboxylase	YBR218C
r_0959	pyruvate decarboxylase	YGR087C
r_0960	pyruvate decarboxylase (acetoin-forming)	YGR087C
r_0961	pyruvate dehydrogenase	YBR221C and YER178W and YFL018C and YGR193C AND YNL071W
r_0986	S-adenosyl-methionine delta-24-sterol-c-methyltransferase	YML008C
r_1010	squalene epoxidase (NAD)	YGR175C
r_1011	squalene epoxidase (NADP)	YGR175C
r_1012	squalene synthase	YHR190W
r_1021	succinate dehydrogenase (ubiquinone-6)	YDR178W and YJL045W and YKL141W and YLL041C
r_1023	Succinate-semialdehyde dehydrogenase (NADP)	YBR006W
r_1030	Tetrahydrofolate aminomethyltransferase	YAL044C and YDR019C and YFL018C and YMR189W