SYSTEMATIC AND COMBINATORIAL APPROACHES FOR METABOLIC ENGINEERING OF OPTIMAL YEAST STRAINS TO PRODUCE FUELS AND CHEMICALS

BY

JOSHUA C. QUARTERMAN

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Food Science and Human Nutrition with a concentration in Food Science in the Graduate College of the University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Doctoral Committee:

Associate Professor Michael J. Miller, Chair
Associate Professor Yong-Su Jin, Director of Research
Associate Professor Christopher Rao
Professor Huimin Zhao
ABSTRACT

The overall goal of my thesis study is to use a metabolic engineering approach for developing optimal yeast cell factories capable of efficiently fermenting various sugars abundant in renewable biomass. The research has broad applications for sustainable biological production of value-added fuels and chemicals. Initially, a systematic approach based on global reaction stoichiometry was applied to select gene knockout targets in *Saccharomyces cerevisiae* for enhancing bioethanol production from glucose with minimal impact on biomass yield. Due to the limited scope of stoichiometric models, genetic screening was combined with the model-based technique, and this led to identification of three knockout mutants (Δcox9, Δqcr6, and Δqcr9) with superlative characteristics for enhancing fermentation of glucose and other hexoses (sucrose, fructose, and mannose). Although deletion of a respiratory enzyme subunit (Cox9) identified by the model-based approach severely inhibited galactose metabolism, the deletion was a necessary intermediate step for the respiration-deficient yeast to reach efficient galactose fermentation rapidly through serial subculture in galactose media. The combination of systematic and combinatorial methods led to an optimal phenotype on galactose that could not have been achieved by either method independently and demonstrates a promising approach for directing adaptive evolution toward fermentative metabolism. To understand the genetic basis of the improved phenotype, genome sequencing was conducted and used to identify a loss of function mutation in a repressor of galactose metabolic genes (Gal80p); this mutation was found to act synergistically with inhibition of respiration for efficient galactose fermentation by *S. cerevisiae*. A similar ‘fermentative evolution’ approach, involving deletion of COX9 and serial subculture in xylose minimal media, was applied to improve ethanol fermentation by a xylose-
fermenting yeast strain previously developed in our group (SR8). Genome sequencing and a yeast mating experiment led to identification of a frameshift mutation in a transcriptional regulatory complex subunit (Spt3p) for improving xylose fermentation in engineered respiration-deficient yeast. Ethanol production was significantly improved in the xylose-evolved mutant, but excessive xylitol production demonstrated a redox imbalance problem due to NAD$^+$ shortage without respiration. Two separate strategies were effective for alleviating the redox imbalance problem and further improving ethanol fermentation from xylose in the respiration-deficient evolved strain: i) expression of a NADH-consuming acetate reduction pathway, or ii) expression of a mutant NADH-preferring xylose reductase.

In summary, this metabolic engineering study utilizes model-based and evolutionary tools for development of yeast cell factories that can rapidly ferment sugars abundant in non-food plant biomass. Furthermore, we demonstrate a novel strategy for redirecting sugar metabolism toward the fermentation pathway by systematic deletion of a respiration-related gene and adaptive evolution in selective conditions. The work in this study advances knowledge of limiting factors for sugar fermentation under anaerobic conditions and describes metabolic engineering strategies for overcoming these limitations. Efficient sugar-fermenting strains that can function in the absence of oxygen (i.e. without respiration), such as the ones developed in this study, are desirable because they prohibit respiratory utilization of sugars and oxidative metabolism of alcohol or other fermentation products, both of which can reduce product yield. Thus, engineering and characterization of respiration-deficient yeast strains provides valuable knowledge and understanding that may have applications for industrial fermentation processes.
ACKNOWLEDGEMENTS

I am very thankful to my thesis advisor, Dr. Yong-Su Jin, for his support and guidance throughout my graduate research. Also, I would like to express my appreciation to the other members of my advisory and thesis committees for their valuable input and direction: Drs. Michael Miller, Hans Blaschek, Huimin Zhao, Chris Rao, Hao Feng, and Nathan Price.

I consider it a privilege to have worked with many great scientists and friends in Dr. Jin’s research group. I would like to thank my former and current colleagues Soo Rin Kim, Pan-Jun Kim, Suk-Jin Ha, Jin-Ho Choi, Hyo-Jin Kim, Won-Heong Lee, Na Wei, Guochang Zhang, Ching-Sung Tsai, Eun Joong Oh, Heejin Kim, In Iok Kong, Tim Turner, Panchalee Panthibul, Seong-Oh Seo, Suryang Kwak, Haiqing Xu, Hong Nan, Nurik Kuanyshev, Qiaosi Wei, Stephan Lane, Clarissa Koga, Anastashia Lesmana, Ian Liu, Matt Au, Dan Kim, and Clarissa Florencia. Also, I am thankful for great collaborators in various places such as Xueyang Feng (VT), Jeff Skerker (UCB), Eun-Ju Yun (KU), and Jin-Ho Seo (SNU). This study would not have been possible without generous funding from the Energy Biosciences Institute (EBI) and the Jonathan Baldwin Turner (JBT) fellowship.

Finally, I thank my wife Katie, my parents Clay and Mary, my sister Kelly, and all my friends and family members for their love, support, and prayers through the years.
# TABLE OF CONTENTS

**CHAPTER I**  INTRODUCTION TO METABOLIC ENGINEERING FOR PRODUCTION OF FUELS AND CHEMICALS  .............................................................................................................................1

**CHAPTER II**  ENHANCED HEXOSE FERMENTATION BY *SACCHAROMYCES CEREVISIAE* THROUGH INTEGRATION OF STOICHIOMETRIC MODELING AND GENETIC SCREENING  ..................................................................................................................26

**CHAPTER III**  RAPID AND EFFICIENT GALACTOSE FERMENTATION BY ENGINEERED *SACCHAROMYCES CEREVISIAE* THROUGH AN INVERSE METABOLIC ENGINEERING APPROACH  ......................................................................................................................56

**CHAPTER IV**  INCREASING XYLOSE FERMENTATION EFFICIENCY BY ENGINEERED RESPIRATION-DEFICIENT *SACCHAROMYCES CEREVISIAE*  ........................................................................94

**CHAPTER V**  TWO REDOX BALANCING STRATEGIES FOR IMPROVING XYLOSE FERMENTATION BY ENGINEERED *SACCHAROMYCES CEREVISIAE*  ..........................................114

**CHAPTER VI**  SUMMARY AND FUTURE STUDIES  .................................................................................................131

**CHAPTER VII**  REFERENCES  ..............................................................................................................................137
CHAPTER I  INTRODUCTION TO METABOLIC ENGINEERING FOR PRODUCTION OF FUELS AND CHEMICALS

The content of this chapter will be submitted to *Applied Microbiology and Biotechnology* as a review article. I am the first author of the paper and Yong-Su Jin is the co-author.
1.1 Needs and motivation: biological production of fuels and chemicals

Over the last few decades, advances in metabolic engineering and synthetic biology have contributed to improvement in the cellular properties of a variety of microorganisms for biosynthesis of value-added products. Furthermore, consumers and industries are increasingly aware of environmental issues and unsustainability surrounding the traditional chemical synthesis of fuels, bulk chemicals, and other products from petroleum-derived hydrocarbons (Saxena et al., 2009). For the development of a sustainable industrial society with controlled greenhouse gas emissions, a shift from traditional petroleum resources to renewable biomass feedstocks is generally viewed as an important and positive step. Some visionaries expect that the growing ‘carbohydrate economy’ will one day replace much of the old ‘hydrocarbon economy’ that has prevailed since the Industrial Revolution (Ragauskas et al., 2006). In this context, global interest is being directed towards research and commercialization of more bio-based renewable products to address current and future societal needs. The rise of the biorefinery concept provides vast opportunity for scientists and engineers to apply an ever-growing set of biotechnology tools in the context of metabolic engineering for developing microbial cell factories that can produce the next generation of bioproducts.

Recent advances in microbial production of bioethanol and other biofuels offer a prominent example of the paradigm shift towards biological processes utilizing renewable raw materials. The Deepwater Horizon oil spill in the Gulf of Mexico, political tensions in the Middle East, and unstable oil prices have directed particular attention at the need for sustainable, efficient, cost-effective, and cleaner-burning domestic energy sources. In general, the overall life-cycle analysis for biofuels showed an improvement over traditional fossil fuel resources as the greenhouse gas emitted during combustion is mitigated by the inorganic carbon fixed from
atmospheric carbon dioxide during growth of the biomass feedstock (Davis et al., 2009). Furthermore, biofuels such as ethanol are less toxic, biodegradable, and generate fewer pollutants than petroleum fuels (John et al., 2011). However, the status quo of using food crops for biofuel conversion is a limited solution and cannot satisfy the full demand for renewable fuels. Corn and sugarcane are the most widely utilized raw materials for biofuel synthesis because of well-established farming practices and simple, cheap processes for release of fermentable sugars. Despite the economic and technological feasibility of first generation bioethanol from these food crops, its production is controversial and greatly limited due to unavailability of more arable land for growing the feedstocks and concerns about food scarcity and increasing prices of food commodities (John et al., 2011).

For a sustainable future, bioethanol and other biofuels should be generated from any of a wide variety of alternative, renewable non-food organic sources. For example, one plausible alternative that has been a focus of extensive research is use of terrestrial non-food biomass, such as agricultural residues, wood waste, or energy grasses, as a raw material for biofuel production (Himmel et al., 2007). These lignocellulosic biomass sources are advantageous because of low cost, minimal land use change, and avoidance of the competition between food and fuel. However, current chemical and biological technologies have yet to overcome the significant obstacles to releasing sugars from recalcitrant lignocellulose and efficiently converting pentoses (e.g. xylose) to target fuel molecules with high yields and productivities (Jones and Mayfield, 2012). Marine plant biomass is another promising feedstock that can be used for producing fuels and chemicals and has some advantages over terrestrial biomass. For example, macroalgae does not require arable land, can grow in salt water or municipal waste water, synthesizes biomass very rapidly, and can be depolymerized easily due to lack of hemicellulose and lignin (Wei et al.,
Galactose is a major sugar obtained from hydrolysis of some marine macroalgae, but industrial microorganisms do not utilize this sugar efficiently due to tight regulation of the galactose metabolic pathway, called the Leloir pathway (Bro et al., 2005; Holden et al., 2003). In light of these technological challenges, engineered microbial strains for converting renewable biomass to biofuels should possess at least two specific characteristics: (i) capable of efficiently metabolizing various sugars abundant in biomass sources, and (ii) rapid (productivity) and efficient (yield) production of fuel or chemical molecules. To overcome the deficiencies of naturally existing microorganisms, concepts and methodologies from metabolic engineering have been employed to develop optimal strains with these characteristics.

1.2 Overview of metabolic engineering: basic concepts and methodologies

Before the advent of techniques for DNA recombination, microbial strain improvement relied heavily on the use of mutagens, such as ultraviolet (UV) irradiation or chemical agents, and clever selection techniques to isolate improved strains for overproduction of some target molecules (Thoma, 1971). While this approach was not unsuccessful, the mutant strains were poorly understood and their development depended solely on random, time-consuming processes with little science or engineering involved. In the 1970s, the development of recombinant DNA technology introduced a new dimension to strain improvement by providing the skills for modification of specific enzymatic reactions in a cell’s metabolic pathways. The control of pathway modification is a key element that distinguishes metabolic engineering from previous strain improvement approaches. Thus, metabolic engineering has been defined as “the directed improvement of product formation or cellular properties through the modification of specific biochemical reaction(s) or the introduction of new one(s) with the use of recombinant DNA
technology” (Stephanopoulos et al., 1998). This multidisciplinary field draws from the knowledge base and skill set provided by biochemistry, genetics, molecular biology, cellular physiology, chemical engineering, and more.

Since metabolic engineering is a strategic and directed approach for strain improvement, it requires a solid understanding of the system being studied and not just individual reactions or pathways (Ostergaard et al., 2000c). A holistic approach is important for studying cellular metabolism because there are often some unexpected or uncharacterized interactions between pathways or reactions in a cell. Thus, one should make some distinction between genetic engineering, with its focus on an individual enzyme or pathway, and metabolic engineering, with its focus on integrated metabolic pathways and regulatory networks. Some pioneers of metabolic engineering suggest that the field preceded and paved the way for systems biology by creating a definite need for an overall or system-wide view of metabolism (Woolston et al., 2013).

Two important and interrelated components make up the foundation of metabolic engineering: a careful analysis of the metabolic system and precise synthesis of the recombinant strain (Ostergaard et al., 2000c). Initially, when techniques for DNA recombination were still young, more attention was focused on the synthetic side of the field, which includes the expression of new genes in a host organism, deletion or amplification of existing genes, modulation of enzymatic activity, or transcriptional deregulation (Stephanopoulos et al., 1998). Generally, the synthesis step is fairly straightforward in cases where the gene to be expressed is available and the host organism is genetically well-characterized. However, it is the analysis step that is often limiting in metabolic engineering studies (Ostergaard et al., 2000c). It is important to realize that the analytical step distinguishes metabolic engineering, a field with a strong engineering or problem-solving component, from an applied biology. There are at least
three major tasks from a network analysis perspective, and each of these tasks presents some significant challenges. First, analysis involves a detailed evaluation of cellular physiology. Systems biology tools including metabolic flux determination and metabolite profiling as well as measurement of gene expression, cell growth, or metabolite exchange rate may be helpful for characterizing cellular function in a microbial strain (Stephanopoulos and Sinskey, 1993). The key question here is: how can we identify the most important parameters for defining the physiological state of the cell? Second, the physiological (or phenotypic) information should be used to elucidate control points in the network and thus identify specific genetic, enzymatic, or regulatory modifications towards the accomplishment of the objective. A set of genetic manipulations may be chosen and ranked in ascending order of difficulty (Stephanopoulos, 2002). Third, after synthesis of the recombinant strain, the actual impact of the genetic and enzymatic modifications should be assessed for designing further genetic manipulations (Stephanopoulos et al., 1998). We emphasize here that analysis and synthesis are not two separate or linear steps, but actually these two steps are closely related and often proceed as a cycle through several rounds of engineering until an optimal strain emerges (Ostergaard et al., 2000c).

Over the last two decades, metabolic engineering has grown quickly and found a wide variety of applications, such as increasing the range of substrate utilization, improving product yield or production rate, diversifying the range of products synthesized, reducing byproduct accumulation, or improvement of other cellular properties (Stephanopoulos, 1999). Some interesting properties that have been conferred to microorganisms include the ability to thrive in stressful environments (Hong et al., 2010) or balancing the redox potential within the cell (Berrios-Rivera et al., 2002). Metabolic engineering has also found applications for
biodegradation of harmful pollutants in the environment (Haro and de Lorenzo, 2001),
production of chiral compounds for pharmaceutical synthesis (Lee et al., 2009), and production
of biopolymer molecules with desirable structures and properties (Aldor and Keasling, 2003). In
general, metabolic engineering has a distinct focus on designing microbial factories for industrial
applications, such as large-scale production of fuels, chemicals, and pharmaceuticals, in an
efficient and economical way. A cost-effective process for industrial use cannot be realized until
the three factors of titer, productivity, and yield have all been optimized (Woolston et al., 2013).
Therefore, metabolic engineers are concerned with much more than just piecing genes together
to construct a new functioning metabolic pathway. In just a few months, synthetic biologists
may be able to construct a new working pathway for synthesis of a desired product on the order
of milligrams, but it can take years for metabolic engineers to optimize the strain so that it can
support an efficient large-scale process. This distinguished metabolic engineering from the
overlapping field of synthetic biology (Woolston et al., 2013).

The identification of novel gene targets impacting the production phenotype of
recombinant strains is a limiting step of metabolic engineering and a critical aspect of industrial
strain construction (Stephanopoulos et al., 1998). Two general approaches have been described
to locate interesting gene targets for metabolic engineering: systematic approaches (Fig. 1.1) and
combinatorial approaches (Fig. 1.2) (Alper et al., 2005b). The following sections will describe
these general approaches and then review some applications in the context of engineering
Saccharomyces cerevisiae for production of fuels and chemicals.
1.3 Systematic approaches for metabolic engineering

Many early metabolic engineering studies relied primarily on ad hoc target selection for strain improvement (Stephanopoulos et al., 1998). This means that some logical genetic changes were identified for increasing flux through a particular target pathway by reason and visual inspection of a metabolic network. The ad hoc pathway manipulation approach is quite limited because of unknown interactions in the system and lack of any rigorous analysis or justification. The growth of metabolic engineering has been fueled by development of more comprehensive tools for analysis.

1.3.1 Stoichiometric modeling

The availability of genome sequences for many organisms has facilitated the development of metabolic models for cellular metabolism that permit systematic approaches of gene target identification (Alper, 2006). These models enable the mathematical description of metabolic function in terms of distribution of metabolic flux, which is a fundamental indicator of cellular physiology and the most important factor for metabolic engineering of a cell (Stephanopoulos et al., 1998). Over the last two decades, a particular emphasis has been placed on using stoichiometric models for predicting metabolic flux distribution and identifying gene targets in a systematic way. Stoichiometric models are advantageous because they only require information about stoichiometry and reversibility of enzymatic reactions without the need for any of the difficult-to-measure kinetic parameters that are required for more theoretical approaches (Orth et al., 2010; Varma and Palsson, 1994). The application of stoichiometric models depends on a mass balance around each of the internal metabolites in a network and results in a system of differential equations representing the dynamic mass balance (Llaneras and
Subsequently, a steady-state assumption can be applied, which is based on the idea that changes in intracellular metabolism are very fast compared to extracellular changes in resultant phenotype (Lee et al., 2006). By assuming a quasi steady-state condition for internal metabolites, the sum of all fluxes involving formation, degradation and transport of a metabolite must equal zero. This basic assumption is critical because it allows the dynamic mass balance equation to be simplified into a system of linear equations. In matrix form, the steady-state flux balance equation can be written as:

\[
S \cdot \nu = 0 \quad \text{(Equation 1)}
\]

where \( S \) is the stoichiometric matrix and \( \nu \) is a vector of reaction fluxes. Additional thermodynamic information should also be considered to account for all irreversible reactions \( i \) in the flux vector \( \nu \):

\[
\nu_i \geq 0 \quad \text{(Equation 2)}
\]

In most cases, Equation 1 describes an underdetermined system with infinite solutions because there are often many more reaction fluxes (unknowns) than metabolites (equations) in a network (Schilling et al., 1999). Thus, Equation 1 defines what a metabolic network cannot do and describes a solution space where every possible flux distribution \( \nu \) lies, but it cannot predict a single flux solution. This is quite logical as cells can show different phenotypes depending on carbon source, oxygen availability, and other environmental conditions (Llaneras and Picó, 2008). In this review, we will focus on two fundamental stoichiometric model-based approaches to solve Equation 1 with the constraints in Equation 2 in a meaningful way: flux balance analysis and metabolic pathway analysis. These tools are related in that they are both based on the stoichiometric matrix and a steady-state assumption, but they also have clear differences in flux calculation and applications for metabolic engineering.
1.3.2 Flux balance analysis

Flux balance analysis (FBA) has been defined as “a mathematical approach for analyzing the flow of metabolites through a metabolic network” (Orth et al., 2010). This method provides an explicit prediction of metabolic behavior based on some defined reaction constraints and an objective function. Linear programming can be applied to solve for the flux balances (Equation 1) together with the thermodynamic constraints (Equation 2) and other reaction constraints of the metabolic network. Therefore, the in silico flux distribution, which represents the strain’s predicted phenotype under the defined conditions, is highly dependent on a suitable choice for objective function. Maximization of biomass has been demonstrated as a useful objective function in many studies with microorganisms, such as *Escherichia coli* or *Saccharomyces cerevisiae* (Edwards et al., 2001; Famili et al., 2003). Biomass production is considered by including a biomass reaction in the stoichiometric matrix, which converts precursor metabolites to biomass at appropriate stoichiometric ratios based on biomass composition (Orth et al., 2010). For a *S. cerevisiae* genome-scale model, the maximization of biomass objective led to in silico predictions consistent with experimental results in ~70-80% of the cases considered (Famili et al., 2003). While this objective function has been used in a vast majority of FBA applications, it should be noted that others have also been effective, such as minimizing ATP generation, minimizing nutrient consumption, or maximizing synthesis of a specific metabolite (Llaneras and Picó, 2008). Some significant limitations of the FBA approach include its assumption of a cellular objective and general lack of kinetic or regulatory information.
1.3.3 Metabolic pathway analysis

In contrast to the above, metabolic pathway analysis provides no explicit prediction of phenotype but rather calculates all the possible flux solutions in a metabolic network without need for fixed flux rates or an objective function. This gives a complete understanding of an organism’s metabolic capabilities (Trinh et al., 2009). Two closely related metabolic pathway analysis approaches are elementary flux mode (EFM) analysis and extreme pathway (EP) analysis; here we will focus on the former. An EFM has been defined as “a minimal set of enzymes that can operate at steady state with all irreversible reactions proceeding in the appropriate direction” (Schuster and Hilgetag, 1994; Schuster et al., 2002). For a given network, a set of EFMs define and describe all the independent metabolic routes that are both stoichiometrically and thermodynamically feasible. Thus, EFM analysis is a powerful mathematical tool for decomposing a complex metabolic network of many interconnected reactions into unique, organized pathways that support steady state metabolism. The modes are calculated by solving Equation 1 with Equation 2 and a non-decomposability constraint, which ensures a finite number of solutions and that each solution is unique (Trinh et al., 2009). EFM analysis can provide determination of the most efficient physiological state of a cell, analysis of metabolic network properties such as robustness and regulation, and a tool for rational design of the host strain’s metabolism (Llaneras and Picó, 2008; Pfeiffer et al., 1999; Schuster et al., 2000). The major limitation of this approach is computationally expensive calculation, which grows exponentially with the size of the metabolic network.
1.3.4 Rational strain design by EFM analysis

The calculation of all unique and feasible pathways in a metabolic network permits the rational design of strains for accomplishing a specific metabolic objective. By careful in silico manipulation, a host strain can be designed such that undesirable pathways are deactivated and cell growth is coupled to the operation of more desirable pathways. Previously, Trinh et al. described a rational approach for strain design that involves calculation of EFMs in a metabolic network and sequential deletion of reactions to eliminate inefficient pathways (Trinh et al., 2009). This strategy was used to improve ethanol production from hexoses and pentoses by engineered E. coli (Trinh et al., 2008). Here, we describe a distinct but related approach for systematic identification of gene deletion targets that may be beneficial for yeast metabolic networks.

The in silico rational design strategy is illustrated in Fig. 1.3 with the analysis of a 58-reaction metabolic network for glucose fermentation by S. cerevisiae. The network includes glycolysis, pentose phosphate pathway, TCA cycle, fermentation pathway, a biomass equation, and transport fluxes, as well as 47 internal metabolites at steady-state and seven external metabolites as sources or sinks (glucose, oxygen, carbon dioxide, ethanol, acetate, glycerol, and biomass). In this example, our objective was to find reaction deletions that will constrain the cell to use the most efficient pathway for converting glucose to ethanol while maintaining sufficient biomass for cell growth. The first step was to calculate the elementary modes for the given metabolic network with METATOOL or other appropriate software (Fig. 3a). Second, the 1,918 total modes were evaluated by calculating the ethanol yield and biomass yield for each, and then the modes were plotted to show their distribution in relation to these parameters (Fig. 3b). Due to high ethanol yield and moderate biomass yield, 4 of the modes were classified as “good”
EFMs (i.e. desirable pathways) and the remaining 1,914 were classified as “bad” EFMs (i.e. undesirable pathways) for the objective. Third, we identified a pool of 23 “bad” reactions that do not participate in the “good” EFMs and thus only contribute to the inefficient or undesirable pathways. The key of the third step is to select a reaction target from the pool of “bad” reactions that is included in the most elementary modes in order to deactivate as many of the inefficient pathways as possible (Fig. 3c). This step should be repeated to identify multiple deletion targets for minimizing undesirable modes. In this case study, succinate dehydrogenase (SDH) was the best deletion target for accomplishing the objective and allowed elimination of almost 80% of the “bad” EFMs. The second round of reaction deletion identified the acetate transporter (ACETt) as the best target and sequential deletion of the first two reactions reduced the number of undesirable modes to less than 5% of the original value. In the third round, glucose-6-phosphate dehydrogenase (G6PDH) was the best of the remaining targets and elimination of the first three reaction targets reduced the number of inefficient pathways to less than 1% of the original value (Fig. 3d). Thus, the sequential deletion of three reactions (SDH, ACETt, G6PDH) eliminates a maximal number of undesirable modes and supports the operation of metabolic pathways for efficient glucose fermentation with sufficient biomass for cell growth.

1.4 Combinatorial approaches for metabolic engineering

Previously, we emphasized that metabolic engineering is a directed, or rational, approach for strain improvement. However, one should not understand this to mean that random processes have no role in metabolic engineering of cells. In fact, random mutations can be used to select superior strains which can provide a rich source of information about metabolic pathways and metabolic control (Stephanopoulos et al., 1998). Inverse metabolic engineering is an important
and useful approach, which uses random genetic perturbations in combination with high throughput screening for identifying interesting metabolic engineering targets (Alper, 2006).

1.4.1 Inverse metabolic engineering

There are three basic steps involved in inverse metabolic engineering, which have been used together in many ways for elucidating strategies towards the directed engineering of useful phenotypes (Bailey et al., 1996). The first step of this approach is the identification of some beneficial phenotype in a mutant strain. The desired phenotype may occur naturally in the organism or through well-designed evolutionary engineering experiments. Evolutionary engineering involves some spontaneous or induced mutations in a host strain and appropriate selection pressure for enriching fitter variants during cultivation in liquid or solid media (Çakar et al., 2012; Sauer, 2001). Another option for generating mutant strains with beneficial phenotypes is the introduction of random genetic perturbations through a genomic DNA library (Hong et al., 2010; Lee et al., 2011) or randomized gene knockout library (Badarinarayana et al., 2001). Also, directed evolution has been very useful to generate diverse mutants with beneficial traits for inverse metabolic engineering; this approach involves random mutagenesis of a target gene, protein, or organism followed by gene recombination and screening or selection (Arnold, 1998). While screening of mutants can be tedious by conventional methods, tools such as flow cytometry or microfluidic devices have aided high-throughput screening and selection of interesting mutant strains (Alper, 2006; Wang et al., 2014). The second step of inverse metabolic engineering involves determining the underlying genetic or molecular basis for a desired phenotype. One of the great experimental advances towards this end is the significant increase in throughput and reduction in cost for genome sequencing (Gill, 2003). Genome-wide
transcript analysis has also been used as a tool for uncovering changes responsible for superior cellular properties of a mutant strain (Bro et al., 2005). As improvements in genomics technologies continue, it is expected that locating relevant gene sequences will no longer be a rate-limiting step of inverse metabolic engineering (Gill, 2003). Finally, the third step of inverse metabolic engineering is to transfer the identified genetic perturbation(s) to another strain or organism for confirming the genotype-phenotype relationship. Subsequently, the engineered strain may be further evolved or the strategy may be executed multiple times until arriving at the desired phenotype.

1.5 Metabolic engineering of *S. cerevisiae* for production of fuels and chemicals

For many centuries, *S. cerevisiae* has been widely employed for alcohol production, especially for human consumption. This yeast is well-known to be one of the best ethanol-producing organisms in nature. Furthermore, *S. cerevisiae* is an attractive host for industrial bioprocesses due to high resistance to ethanol and other inhibitors and status as a GRAS organism (generally recognized as safe). Based on its traditional importance in the food industry and recent applications in the field of biotechnology, a detailed knowledge of genetics, physiology, and biochemistry, as well as process technologies for large-scale fermentation, has accumulated over time. *S. cerevisiae* was the first eukaryotic organism whose genome was completely sequenced (Goffeau et al., 1996) and now several databases exist containing a wealth of information about genes, open reading frames, and gene products (Nevoigt, 2008). The availability of efficient transformation methods (Gietz and Schiestl, 2007a) and many genetic tools, such as specialized expression vectors and selectable markers, has facilitated advanced genetic engineering techniques (Nevoigt, 2008). In addition, an extensive collection of systems
biology tools and metabolic reconstructions makes *S. cerevisiae* a desirable organism for *in silico* metabolic engineering. A total of nine predictive genome-scale metabolic models have been created for *S. cerevisiae* in the last ten to fifteen years to provide a mathematical framework for analyzing the metabolic potential of this organism (Österlund et al., 2012), and the predictive power of these models has been validated against experimental data (Feist et al., 2008). These are some of the major reasons that this host has been heavily utilized for various metabolic engineering studies.

Several extensive and thorough review articles have been written on metabolic engineering of *S. cerevisiae* (Hong and Nielsen, 2012; Nevoigt, 2008; Ostergaard et al., 2000c). It is not the purpose of this section to give another exhaustive overview of the subject. Rather, we intend to focus on the most impactful studies utilizing the systematic and/or combinatorial approaches discussed previously for enhancing bioconversion of sugars in renewable biomass.

### 1.5.1 Bioethanol production from sugars in renewable biomass

*S. cerevisiae* strains have long been selected for their ability to convert sugar to ethanol in order to make wine, beer, and other alcoholic beverages. Thus, the yeast has evolved to efficiently consume glucose, which is the most abundant sugar in nature and a component of starch, fruits such as grapes, sucrose sources such as sugarcane, and cellulose and hemicellulose. However, it is known that glucose fermentation to ethanol is still not optimal, even under anaerobic conditions, because some carbon flux is directed towards glycerol, biomass, CO$_2$, and other minor by-products (Bro et al., 2006; Nissen et al., 2000). For example, an *in silico* model was used to identify strategies for reducing glycerol production in *S. cerevisiae* by preventing excess NADH through biomass synthesis. The best strategy involved expression of a NADP$^+$-
dependent glyceraldehyde-3-phosphate dehydrogenase and resulted in a 40% lower glycerol yield but only 3% improvement in ethanol yield (Bro et al., 2006). In another study, deletion of two key enzymes for glycerol synthesis, \textit{GPD1} and \textit{GPD2}, eliminated glycerol production and increased ethanol yield from glucose about 12% during aerobic batch cultivations; however, the double knockout mutant also showed severe growth reductions (and no growth anaerobically) (Nissen et al., 2000). Other research has shown that deletion of the glycerol facilitator \textit{FPS1}, with or without glutamate synthase overexpression, also significantly reduces glycerol and increases ethanol production in \textit{S. cerevisiae} (Kong et al., 2006). Finally, petite mutants of yeast have also demonstrated usefulness for improving ethanol production from glucose (Bacila and Horii, 1979; Hutter and Oliver, 1998). Petite mutants include any yeast strain with a DNA mutation resulting in inability to utilize non-fermentable carbon sources such as ethanol, glycerol, or acetate. These respiration-deficient mutants can still grow on glucose or other fermentable carbon sources but form smaller colonies than wild type cells (Tzagoloff and Dieckmann, 1990).

Xylose, which makes up a substantial portion of hemicellulose, is the second most abundant sugar in nature but is not able to be metabolized by native \textit{S. cerevisiae} strains (Ostergaard et al., 2000c). Over the last several decades, two main strategies have been employed for engineering xylose-utilizing mutants; both strategies require introduction of a heterologous pathway for converting xylose to xylulose, which can enter the pentose phosphate pathway. The first strategy involves introduction of genes encoding xylose reductase (XR) and xylitol dehydrogenase (XDH) from the xylose-fermenting yeast \textit{Scheffersomyces stipitis} and in some cases overexpression of an endogenous xylulokinase (XK) from \textit{S. cerevisiae} (Jin et al., 2000; Toivari et al., 2001). The major challenges associated with this pathway are generally
redox imbalance and xylitol production, which are especially prevalent in anaerobic conditions. The first study to overcome the hurdle of anaerobic xylose utilization with this pathway made use of an evolutionary engineering strategy with selection of evolved strains by step-wise decreases in oxygen availability (Sonderegger and Sauer, 2003). A more recent study engineered an efficient xylose-fermenting strain with minimal xylitol production by combining strong and balanced expression levels of the heterologous S. stipitis xylose utilization pathway and serial subcultures on xylose (Kim et al., 2013a). Evolutionary engineering has also demonstrated utility for improving co-fermentation of xylose and arabinose to ethanol by recombinant S. cerevisiae expressing the XR/XDH pathway and a bacterial arabinose utilization pathway (Sanchez et al., 2010b). Finally, a genomic DNA fragment library from S. stipitis was used to identify a beneficial overexpression target with high degree of homology to native S. cerevisiae TAL1 for a 70% increase in ethanol production from xylose (Jin et al., 2005). The second general strategy for conversion of xylose to xylulose by yeast involves introduction of the xylose isomerase gene (xylA) from Piromyces or bacteria. This is a one-step, redox-neutral pathway but it is not energetically favorable compared to the XR/XDH pathway (Karhumaa et al., 2007). For recombinant S. cerevisiae strains expressing fungal xylose isomerase, evolutionary engineering has been effective in numerous studies for improving fermentation parameters during growth on xylose (Kuyper et al., 2004; Shen et al., 2012; Zhou et al., 2012) or a mixture of xylose and other sugars (Kuyper et al., 2005; Wisselink et al., 2009). In another study, directed evolution generated a mutant xylose isomerase with a 77% increase in enzymatic activity and the mutant gene led to an 8-fold increase in xylose uptake rate and ethanol productivity (Lee et al., 2012).
Galactose is abundant in some marine plant biomass, but utilization of this sugar by *S. cerevisiae* is relatively slow and inefficient. Although capable of growing on galactose, its galactose uptake rate is approximately three times lower than that on glucose and ethanol yield is also significantly lower (Ostergaard et al., 2000d). To overcome this deficiency, one study attempted to modify the GAL gene regulatory network in *S. cerevisiae*, which is a tightly regulated system. They found that deletion of three negative regulators of the GAL system (Gal6, Gal80, Mig1) led to a 41% increase in galactose pathway flux, which was the best among the mutants tested (Ostergaard et al., 2000a). A follow-up study used transcript analysis to identify phosphoglucomutase (*PGM2*) as an overexpression target resulting in a 70% increase in galactose uptake rate as compared to the reference strain (Bro et al., 2005). Another group employed a genome-wide perturbation library to discover *SNR84* (coding for a small nuclear RNA) and truncated *TUP1* (coding for a general repressor of transcription) as overexpression targets for improving galactose fermentation as much as *PGM2* overexpression (Lee et al., 2011). Finally, a recent article reported that adaptive evolution with a laboratory strain over many generations was useful for increasing galactose uptake rate and growth rate in *S. cerevisiae*; furthermore, this approach led to identification of unknown and unexpected genetic targets (Hong et al., 2011a).

### 1.5.2 Production of other fuels and chemicals from sugars in renewable biomass

Besides ethanol, metabolic engineering approaches have also been useful for developing *S. cerevisiae* cell factories with increased carbon flux toward one of a wide variety of other fuels or chemicals. For example, a systematic model-based approach was employed to identify a key deletion target (*GDH1*) in yeast for enhancing NADPH availability and improving sesquiterpene
synthesis from galactose by 85%. Although the knockout strain (Δgdh1) showed a significant reduction in growth rate, subsequent overexpression of GDH2 maintained product titer with less growth penalty (Asadollahi et al., 2009). In another study, an in silico gene deletion simulation was effective for selecting two knockout targets (PDC1 and GDH1) involved in vanillin production, and the best knockout strain (Δpdc1) improved vanillin productivity by a factor of 2 in glucose-limited continuous culture (Brochado et al., 2010b). More recently, flux balance analysis led to a metabolic engineering strategy for overproduction of succinate which involved deletion of SDH3 (encoding the primary succinate consuming reaction in the TCA cycle) and SER3/SER33 (encoding 3-phosphoglycerate dehydrogenase). When combined with adaptive evolution, the resulting S. cerevisiae strain produced 30-fold more succinate than the control strain with a 2.8-fold decrease in specific growth rate (Otero et al., 2013). Also, EFM analysis was applied for evaluating the effect of network modifications and changing culture conditions on poly-β-hydroxybutyrate (PHB) production in recombinant yeast including a PHB synthesis pathway. The results of the study showed that adding an ATP citrate-lyase reaction and a transhydrogenase reaction to the metabolic network increased maximum theoretical PHB yield from 0.67 to 0.84 (Carlson et al., 2002). In addition to the model-based techniques, various combinatorial methods have also been successful for improving biosynthesis of fuel and chemicals by engineered S. cerevisiae. For example, after deletion of four genes (TPII, NDE1, NDE2, and GUT2) identified by rational means, evolutionary engineering led to a glycerol-producing strain with almost maximum product yield from glucose (Overkamp et al., 2002). Another group constructed a pyruvate decarboxylase-negative strain and used a two-step evolutionary engineering approach to develop glucose-tolerant, pyruvate-producing yeast with product yield of 0.54 g/g glucose (van Maris et al., 2004). Finally, Ishida et al. utilized an
evolutionary engineering approach with a starting strain expressing six copies of lactate dehydrogenase and thereby achieved 122 g/L of lactate production (Ishida et al., 2006).

1.6 Concluding remarks and future perspectives

The systematic model-based approaches, which are based on network stoichiometry, provide a valuable tool for system-wide analysis of metabolism, gene target identification, and phenotype prediction under specified conditions. A major advantage of these systematic approaches is in providing a cheaper and faster way to narrow an extensive set of candidate gene targets without costly and time-consuming initial wet experiment screening. However, it is important to note that stoichiometric models usually do not account for kinetic and regulatory effects that may significantly impact in vivo results, and they require a significant accumulation of knowledge about an organism’s metabolic capabilities and interactions within its complex biomolecular network. Thus, the model-based methods are somewhat limited and should generally be used in combination with other approaches. On the other hand, inverse metabolic engineering is a powerful combinatorial approach for identifying novel gene targets responsible for a desired phenotype. A particular advantage of this technique is unique access to still poorly understood cellular subsystems, unknown regulatory genes, or non-coding regions that are critical for cellular function (Alper, 2006; Sauer and Schlattner, 2004); perturbation targets in these categories could not be identified by a systematic approach like stoichiometric modelling.

In the past, the main challenge associated with many combinatorial engineering studies was in discerning the genetic basis responsible for a desirable phenotype in some mutant strain (Bailey et al., 1996). However, rapid and on-going advances in genomics technologies have greatly improved the ability to gain understanding about novel or interesting mutant yeast strains. The
various applications discussed in this review demonstrate the capability for systematic and
combinatorial methods to reveal previously unknown perturbation targets for strain development
and improvement. Due to continued efforts in these areas, we expect to see significant advances
in research and definite progress toward commercialization of yeast-derived biofuels and
biochemicals (Kim et al., 2012).
Fig. 1.1 Systematic approaches for metabolic engineering of microorganisms. (a) A stoichiometric model is based on a metabolic network of reactions and a mass balance around each of the intracellular metabolites. Under steady state conditions, there is no net accumulation of metabolites and the model can be simplified to a system of algebraic expressions ($S \cdot v = 0$). (b) Flux balance analysis allows an explicit prediction of metabolic phenotype by defining some reaction constraints and applying an objective function to solve the balance equation. (c) Metabolic pathway analysis calculates all feasible solutions that exist in a metabolic network by solving the balance equation with an additional non-decomposability constraint.
Fig. 1.2 Combinatorial approaches for metabolic engineering of microorganisms. Inverse metabolic engineering involves random genetic perturbations in combination with high throughput screening for identifying interesting genetic targets. Spontaneous or induced mutations are commonly introduced to a host strain or gene fragment by (a) evolutionary engineering, (b) genomic DNA libraries, or (c) directed evolution.
Fig. 1.3 Rational *in silico* design strategy for identifying deletion targets by EFM analysis. (a) Calculate EFMs for the metabolic network. (b) Evaluate the modes based on the metabolic objective and then classify modes into categories (e.g. “good” vs. “bad”). (c) Identify the reaction that includes the most undesirable or “bad” EFMs. (d) Sequentially determine multiple knockout targets in order to minimize the inefficient pathways.
CHAPTER II  ENHANCED HEXOSE FERMENTATION BY SACCHAROMYCES CEREVISIAE THROUGH INTEGRATION OF STOICHIOMETRIC MODELING AND GENETIC SCREENING

The content of this chapter has been accepted for publication by Journal of Biotechnology. I was the first author of the paper and Soo Rin Kim, Pan-Jun Kim, and Yong-Su Jin were co-authors. I performed the research with help from the co-authors and Dr. Yong-Su Jin was the director of the research.
2.1 Introduction

_Saccharomyces cerevisiae_, or baker’s yeast, is the preferred microorganism for industrial ethanol fermentation since this species has long been employed by the food industry to convert sugar into alcohol (Nevoigt, 2008). In addition to its industrial importance, bioethanol production in yeast is also an excellent model system for probing of genotype-phenotype relationships and discovering genetic perturbations for optimization of metabolic flux through central carbon metabolism (Dikicioglu et al., 2008). _S. cerevisiae_ and many other organisms first catabolize hexoses (e.g. glucose, fructose, mannose) through the glycolytic pathway with pyruvate as the end product. In turn, pyruvate can be reduced to ethanol via a two-step fermentative pathway. Alternatively, in the presence of oxygen, pyruvate may be completely oxidized to carbon dioxide via tricarboxylic acid (TCA) cycle enzymes, which results in transfer of electrons to oxygen and ATP generation (Sonnleitner and Käppeli, 1986). It is well-known that _S. cerevisiae_ can grow and remain viable in anaerobic conditions without oxidative degradation of pyruvate because some ATP is generated during glycolysis. However, when even small amounts of oxygen are present, some of the carbon source will be metabolized oxidatively (cellular respiration). In Crabtree positive yeasts such as _S. cerevisiae_, a low external glucose concentration is more conducive to the respiratory metabolism than a very high glucose concentration. The cellular respiration process is more efficient than fermentation in terms of net energy payback to the cell but is counterproductive for the goal of optimal ethanol production. Similarly, _S. cerevisiae_ can also direct carbon flux toward synthesis of other fermentation by-products such as glycerol or acetate in order to maintain redox balance within the cells (Costenoble et al., 2000; Remize et al., 2000). Therefore, we hypothesized that elimination of some unknown reactions may shift carbon flux away from the undesirable pathways (i.e. TCA
cycle, glycerol/acetate synthesis, etc.) and increase the ethanol-producing capability of the organism. It is important that any deleted reaction(s) should not lead to undesirable side effects caused by unknown interactions in the system. For example, a previous metabolic engineering study described yeast knockout strains with increased ethanol yield on glucose, but the strains were accompanied by a substantial reduction in specific growth rate (Nissen et al., 2000). The objective of this study was to engineer *S. cerevisiae* towards optimal ethanol fermentation with minimal impact on growth phenotype by using a genome-scale metabolic model to systematically guide the identification of novel gene deletion targets.

The yeast *S. cerevisiae* was the first eukaryotic organism whose genome was completely sequenced and now several databases exist containing a wealth of information about genes, open reading frames, and gene products (Goffeau et al., 1996; Nevoigt, 2008). The availability of very efficient transformation methods and many genetic tools, such as specialized expression vectors and selectable markers, has facilitated advanced genetic engineering techniques (Gietz and Schiestl, 2007b). In addition, an extensive collection of systems biology tools and metabolic reconstructions makes *S. cerevisiae* an amenable microorganism for *in silico* metabolic engineering. A total of six predictive genome-scale metabolic models have been created in the last fifteen years to provide a mathematical framework for analyzing the metabolic potential of this organism, and the predictive power of these models has been validated against experimental data (Feist et al., 2008; Nookaew et al., 2011). Constraint-based flux analysis (or flux balance analysis) is a proven mathematical approach for application of the genome-scale models in order to predict phenotypic results and identify gene targets for metabolic engineering. This method only requires information about reaction stoichiometry, metabolic requirements for growth, and a few other strain-specific parameters without any of the difficult-to-measure kinetic parameters
that are required for more theoretical approaches (Orth et al., 2010; Varma and Palsson, 1994). An optimal flux distribution representing the strain’s predicted phenotype under defined conditions can be calculated based on specified reaction constraints and an objective function, which is generally accepted to be biomass maximization in microorganisms (Edwards et al., 2001; Famili et al., 2003). The main advantage of using stoichiometric modeling techniques is for providing an efficient way to narrow large sets of candidate genes without costly and time-consuming initial wet experiment screening of extensive gene sets. Inspired by the fact that stoichiometric models have been successfully applied to study and improve production of a variety of value-added products in *S. cerevisiae* (Asadollahi et al., 2009; Bro et al., 2006; Brochado et al., 2010a; Ng et al., 2012), we hypothesized that a model-based approach could be used to generate new metabolic engineering strategies for enhancing ethanol fermentation in this yeast.

For accomplishing the objective, we used the genome-scale stoichiometric model *iND750* (Duarte et al., 2004a) to evaluate the effect of gene deletion on ethanol production in yeast. A single gene deletion simulation was conducted at various oxygen uptake rates, and the *in silico* knockout phenotypes were scanned for improvement in ethanol production while maintaining sufficient biomass for cell growth. Twenty potential gene targets linked to two enzymes in the electron transport chain (cytochrome c oxidase, ubiquinol cytochrome c reductase) were identified for further evaluation. Though stoichiometric modeling is a powerful tool for *in silico* metabolic engineering, it is important to note that neither constraint-based flux analysis nor the *iND750* model accounts for kinetic and regulatory effects that may significantly impact *in vivo* results. It is possible that genetic regulation may outweigh stoichiometric effects in certain situations and thus lead to unexpected results (Alper et al., 2005a). To overcome this
limitation and compare all knockout targets identified by the gene deletion simulation, laboratory screening was conducted in shake flask experiments with minimal media. The combination of a model-based gene targeting strategy and experimental screening led to identification of three outstanding single knockout strains for improving ethanol production on glucose in microaerobic conditions. The beneficial effects of these mutant strains were demonstrated on two other common hexoses (fructose and mannose) and a disaccharide (sucrose). Furthermore, this study describes the significant variation in phenotype among respiration-deficient mutants of *S. cerevisiae* and provides knowledge of specific genetic targets for optimizing ethanol production coupled with cell growth.

2.2 Materials and methods

2.2.1 *In silico* design

The *S. cerevisiae* genome-scale metabolic model iND750 was used in this study (Duarte et al., 2004a). The model contains 750 genes and 1,266 associated reactions, including 1,149 intracellular reactions, 116 membrane exchange fluxes, and a biomass equation. Compartmentalization of 646 unique metabolites in the cell, which have been assigned to one of eight cellular compartments (cytosol, mitochondrion, peroxisome, nucleus, endoplasmic reticulum, Golgi apparatus, and vacuole), results in 1,059 total species in the model that are all stoichiometrically balanced. The dimensions of the stoichiometric matrix are therefore 1,059 metabolites by 1,266 reactions. The exchange fluxes for the following metabolites were unconstrained in the model to assume excess for simulating minimal media conditions: NH$_4^+$, SO$_4^{2-}$, K$^+$, Na$^+$, PO$_4^{3-}$, and H$_2$O. Glucose was the sole carbon source with an uptake rate of 5 mmol/gDW·hr, and the non-growth associated ATP maintenance requirement was constrained to
1 mmol/gDW·hr (Stouthamer, 1979). Oxygen uptake rates were constrained in the model to appropriate values for yeast fermentation in oxygen-limited conditions (see Results section). The COBRA (constraint-based reconstruction analysis) toolbox is a software package and valuable systems biology tool allowing for predictive computations of cellular metabolism using constraint-based models (Becker et al., 2007). All gene deletion simulations in this study were conducted using the singleGeneDeletion function in the COBRA toolbox. This function constrains the flux to zero for the reaction(s) corresponding to a deleted gene based on gene-reaction associations in the model; then, the algorithm calculates a metabolic flux distribution by applying an objective function and defined reaction constraints. The constraint-based flux analysis was performed using linear programming library GLPK (http://ftp.gnu.org/gnu/glpk/) in MATLAB (2010b, Mathworks, Natick, MA) with maximization of biomass as the objective function.

2.2.2 Strains and knockout strain construction

*S. cerevisiae* strains BY4742 (*MATalpha, leu2, his3, ura3, lys2*), CEN.PK2-1D (*MATalpha, leu2, his3, ura3, trp1*), and JAY291 (*MATa*) (Argueso et al., 2009) were used in this study as control strains and for engineering of knockout strains. Single knockout mutants with BY4742 background were obtained from Yeast Knockout MATalpha Collection (OpenBiosystems, Lafayette, CO). The COX9 deletion cassette with KanMX marker was cloned from BY4742 Δcox9 by PCR with primers COX9-f and COX9-r. The QCR6 deletion cassette with KanMX marker was cloned from BY4742 Δqcr6 by PCR with primers QCR6-f and QCR6-r. The QCR9 deletion cassette with KanMX marker was cloned from BY4742 Δqcr9 by PCR with primers QCR9-f and QCR9-r. All primers were obtained from Integrated DNA
Technologies (Coralville, IA) and are listed in Table 2.1. The deletion cassette was then integrated into the genome of CEN.PK2-1D or JAY291 strain by transformation using the EZ-Transformation kit (BIO 101, Vista, CA) or a high efficiency lithium acetate protocol (Gietz and Schiestl, 2007b). Positive transformants were selected by G418 resistance on YPD agar plates. Gene deletion was confirmed by colony PCR with confirmation primers Cd-f and KanMX-r for COX9 deletion, primers Q6d-f and KanMX-r for QCR6 deletion, or primers Q9d-f and KanMX-r for QCR9 deletion (Table 2.1).

2.2.3 Media and culture conditions

To prepare yeast strains for inoculation, cells were cultivated overnight at 30°C and 250 RPM in YPD medium (10 g/L yeast extract, 20 g/L Bacto peptone, 20 g/L glucose). Batch fermentation was carried out at 30°C and 100 RPM in yeast synthetic complete (YSC) medium (1.7 g/L yeast nitrogen base, 5 g/L ammonium sulfate, and 0.79 g/L Complete Supplement Mixture (MP Biomedicals, Solon, OH) for supplying amino acids and nucleobases) with 40 g/L of the appropriate sugar (glucose, fructose, sucrose, or mannose). To select yeast transformants with the KanMX marker, a YPD agar plate was used with 200 μg/mL of G418.

2.2.4 Fermentation experiments

Yeast pre-cultures were grown with 5 mL of YPD medium in an orbital shaker at 30°C to prepare inoculums for fermentation experiments. Cells in exponential phase were harvested and inoculated after removing used YPD. Flask fermentation experiments were performed using 25 mL (or 50 mL) of YSC medium with 40 g/L of the appropriate sugar in a well-controlled shaking incubator (Thermoscientific, MaxQ4000, Dubuque, IA) under oxygen-limited conditions. Initial
cell densities were adjusted to OD<sub>600</sub> of ~0.1. During the course of the fermentation, the
temperature and agitation rate were kept constant at 30°C and 100 RPM, respectively.

2.2.5 Analytical methods

Cell concentration was measured by optical density (OD) at 600 nm using a UV-visible
spectrophotometer (Biomate 3, ThermoScientific, Madison, WI). The concentration of sugar,
ethanol, glycerol, and acetate in batch fermentations was determined by a high-performance
liquid chromatography system (Agilent Technologies 1200 series) with a refractive index
detector (RID) and a Rezex ROA-Organic Acid H<sup>+</sup> column (Phenomenex Inc., Torrance, CA).
The column temperature was kept constant at 50°C and the elution was performed with 0.005 N
H<sub>2</sub>SO<sub>4</sub> at a flow rate of 0.6 ml/min.

2.3 Results

2.3.1 In silico gene deletion simulations

The S. cerevisiae iND750 genome-scale model accounts for 750 genes associated with
1266 metabolic reactions (including membrane exchange fluxes) and 1061 internal and external
metabolites (eight cellular compartments). In order to obtain reliable predictions, it is important
to consider which metabolic state(s) should be used to most accurately constrain the model. For
example, the yeast S. cerevisiae is well-known to adjust to external oxygen availability by using
respiratory and/or fermentative metabolic modes (Jouhten et al., 2008); therefore, varying the
oxygen uptake rate (OUR) in silico can have a significant effect on model predictions. Based on
previous studies with a yeast genome-scale metabolic model (Duarte et al., 2004b), the oxygen
uptake rate was varied from 0 to 20 mmol/gDW·hr to determine the effect of this variable on cellular metabolism and identify an appropriate range for model simulations. For all model simulations, glucose was the sole carbon source (uptake rate = 5 mmol/gDW·hr) and flux constraints were adjusted for minimal media. Under complete anaerobic conditions (OUR = 0), the iND750 model predicted a low maximum specific growth rate (0.0853 hr\(^{-1}\)) and high ethanol yield and productivity (0.424 g ethanol/g glucose; 8.288 mmol/gDW·hr), which is in accordance with fermentative metabolism. As the oxygen uptake rate increased to a saturation point (OUR = 12.5 mmol/gDW·hr), the model predicted a shift towards respiratory metabolism with the maximum specific growth rate increasing to its maximum value (0.478 hr\(^{-1}\)) and ethanol production rate decreasing to zero. Four different oxygen uptake rates at <1 mmol/gDW·hr were selected to mimic microaerobic conditions for all subsequent gene deletion simulations; these values were chosen after comparing in silico ethanol yields with typical flask fermentation results in oxygen-limited conditions.

The gene deletion simulation was conducted at the oxygen uptake rate values of interest using the COBRA toolbox, and the in silico yields of ethanol, glycerol, and biomass were calculated as criteria for screening the 750 knockout mutants. In most cases, biomass yield and ethanol yield are inversely related, so knockout target selection must strike a balance between improving ethanol production and maintaining sufficient biomass synthesis for cell growth. We searched for deletion targets predicting an increase in ethanol yield while maintaining at least 80% of the wild type cell’s biomass yield. Twenty such genes were identified and are listed in Table 2.2 with the gene name, related enzyme, and enzyme classification. All 20 genes are linked to two oxidative phosphorylation enzymes, cytochrome c oxidase (COX) and ubiquinol cytochrome c reductase (QCR), which are both critical for the cell’s respiratory metabolism. The
model predicts an identical flux distribution for each of the twenty knockout strains and thus an equal degree of improvement towards the objective. Fig. 2.1a shows a plot of the in silico ethanol and biomass specific productivity versus oxygen uptake rate for the wild-type strain and all twenty mutants from Table 2.2. When comparing the mutant phenotypes to the wild type, the model predictions show an increasing improvement in ethanol yield and increasing reduction in biomass yield at higher oxygen uptake rates (Fig. 2.1b). Under anaerobic conditions, the wild type and mutant strains have the same flux distribution. At an oxygen uptake rate of 1 mmol/gDW/hr, the in silico ethanol yield and biomass yield of the knockout strains are 4% higher and 20% lower than the wild type, respectively. The results from the model-based simulations seem to indicate that COX-related genes and QCR-related genes may be good deletion targets for improving glucose fermentation to ethanol in oxygen-limited or fully aerobic conditions.

2.3.2 Experimental screening of knockout mutants

In order to validate the in silico results and determine the effect of various subunits on overall enzyme function, a total of eighteen cytochrome c oxidase (COX) single knockout mutants and nine ubiquinol cytochrome c reductase (QCR) single knockout mutants were screened by flask fermentation in glucose minimal media. Fig. 2.2a and Fig. 2.2b show the average percent change in ethanol production rate and biomass yield as compared to control strain BY4742 for each of the COX knockout strains and QCR knockout strains, respectively. The eighteen COX mutants that were evaluated are deficient in one of the following genes: any of six out of the ten COX genes identified by the in silico gene deletion study (Table 2.2) or any of twelve other COX genes not included in the iND750 model. Four of the ten COX genes
identified by the *in silico* simulation were either mitochondrial genes (*COX1, COX2, COX3*) or resulted in a lethal knockout phenotype (*COX4*) and therefore could not be deleted *in vivo*. The nine QCR mutants that were evaluated are deficient in any one of the QCR genes predicted by the model-based simulation (Table 2.2) except for *COB*, which is a mitochondrial gene and thus could not be deleted.

Fig. 2.2 clearly illustrates that all deletion mutants did not have a similar phenotype. However, in qualitative agreement with the model-based predictions, twelve of the eighteen COX mutants (Δ*cox6*, Δ*cox8*, Δ*cox9*, Δ*cox10*, Δ*cox12*, Δ*cox14*, Δ*cox16*, Δ*cox17*, Δ*cox18*, Δ*cox20*, Δ*cox23*, Δ*cox25*) and seven of the nine QCR mutants (Δ*cor1*, Δ*cyt1*, Δ*qcr2*, Δ*qcr6*, Δ*qcr7*, Δ*qcr9*, Δ*qcr10*) did show an improvement in ethanol production rate and a reduction in biomass yield as compared to the control strain. Among the eighteen COX mutants, BY4742 Δ*cox9* stood out as the best strain for our objective with the largest increase in ethanol production rate; the ethanol-producing performance of this strain exceeded the second best COX deletion strain BY4742 Δ*cox12* by a factor of approximately two. The BY4742 Δ*cox9* strain consumed glucose at a 22% faster rate and showed a 12% increase in ethanol yield over the control strain; however, the Δ*cox9* mutant also showed a significant loss in final cell density (P < 0.05). The complete fermentation profiles for BY4742 and BY4742 Δ*cox9* are displayed in Fig. 2.3a and Fig. 2.3b, respectively. Among the nine QCR mutants, BY4742 Δ*qcr9* was the best strain in terms of ethanol production rate. The BY4742 Δ*qcr9* strain had an 18% faster glucose consumption rate and a 7% higher ethanol yield when compared to the control strain. In addition, we also observed that BY4742 Δ*qcr6* was the best of the QCR mutants in terms of ethanol production rate coupled with biomass yield. This was especially interesting because deletion of either *COX9* or *QCR9* seemed to have some negative effect on cell growth. The
BY4742 $\Delta$qcr6 mutant showed no significant reduction in maximum specific growth rate or final cell density but still consumed glucose 17% faster and produced ethanol with a 6% higher yield than BY4742.

2.3.3 Evaluation of $\Delta$cox9, $\Delta$qcr6, and $\Delta$qcr9 mutations in CEN.PK2-1D background

To determine whether or not the observed improvement in ethanol yield and productivity by BY4742 $\Delta$cox9, BY4742 $\Delta$qcr6, and BY4742 $\Delta$qcr9 is strain-specific, we deleted COX9, QCR6, and QCR9 in the laboratory strain CEN.PK2-1D. Geneticin (G418) resistance and colony PCR were used to confirm successful gene deletion. In flask fermentations with glucose minimal media, the engineered strains CEN.PK2-1D $\Delta$cox9, CEN.PK2-1D $\Delta$qcr6, and CEN.PK2-1D $\Delta$qcr9 showed improvement in ethanol production rate as compared to the control strain by 33%, 36%, and 38%, respectively. Furthermore, deletion of COX9, QCR6, and QCR9 in CEN.PK2-1D background improved ethanol yield by 10%, 13%, and 15%, respectively. Interestingly, the CEN.PK2-1D $\Delta$qcr6 strain also showed no significant reduction in maximum specific growth rate or final cell density versus the control strain. While the $\Delta$cox9 and $\Delta$qcr9 mutants grew slower than the control during the lag phase and early exponential phase of the fermentation, the $\Delta$qcr6 strain maintained a similar or even slightly higher cell density. Fig. 2.3c and Fig. 2.3d display the related fermentation profiles for the $\Delta$qcr6 and $\Delta$qcr9 mutants in CEN.PK2-1D background. Also, Table 2.3 presents a summary of the glucose consumption rate, ethanol yield, and ethanol production rate by the wild type, $\Delta$cox9 mutant, $\Delta$qcr6 mutant, and $\Delta$qcr9 mutant in both BY4742 and CEN.PK2-1D strains. The results demonstrate that the beneficial effects of these three deletions during glucose fermentation are general to S. cerevisiae and not specific to any particular strain.
2.3.4 Evaluation of Δcox9, Δqcr6, and Δqcr9 mutants on other hexoses

Fructose, sucrose, and mannose are three other hexoses abundant in nature that can be fermented to ethanol by *S. cerevisiae*. We were interested to evaluate the identified single knockout mutants on these three hexoses to determine if the improvement in ethanol production rate and ethanol yield on glucose would extend to other sugars with a similar metabolism as glucose. Therefore, we conducted another series of flask fermentations with the knockout strains BY4742 Δcox9, BY4742 Δqcr6, and BY4742 Δqcr9 in minimal media with 40 g/L of fructose, sucrose, or mannose as the sole carbon source. The ethanol production rate, ethanol yield, and final cell density for the wild type and all three mutant strains is summarized in Fig. 2.4a, Fig. 2.4b, and Fig. 2.4c, respectively; the figure includes results for all four hexoses used in this study including glucose results discussed previously. Fig. 2.4a shows that each of the three knockout strains showed a significant improvement (P < 0.05) in ethanol production rate on glucose, fructose, mannose, or sucrose. Furthermore, all of the knockout strains had a higher average ethanol yield than the control on each carbon source (Fig. 2.4b); the BY4742 Δqcr6 strain on glucose was the only case where this improvement in ethanol yield was not statistically significant (P = 0.087). Finally, the BY4742 Δcox9 strain had a significant reduction (P < 0.05) in final cell density for each of the four hexoses while BY4742 Δqcr6 had no significant change in final cell density for any hexose (Fig. 2.4c). The results clearly demonstrate that the outstanding characteristics of the three knockout strains identified in this study may be extended to a variety of fermentable sugars in microaerobic conditions.
2.3.5 Growth on non-fermentable carbon source

The \( \Delta \text{cox9}, \Delta \text{qcr6}, \) and \( \Delta \text{qcr9} \) mutants were streaked onto YP glycerol plates to compare growth on a non-fermentable carbon source. Fig. 2.5 shows the results in both BY4742 and CEN.PK2-1D background. In both cases, the \( \Delta \text{cox9} \) and \( \Delta \text{qcr9} \) mutants were not able to grow at all, which indicates a complete respiration deficiency by these knockout strains. However, the \( \Delta \text{qcr6} \) mutant did show some growth on glycerol, although the colony size was considerably smaller than for the wild-type strain. This indicates only a partial respiration deficiency by deleting \( QCR6 \) and explains why the \( \Delta \text{qcr6} \) mutant was able to grow better than either \( \Delta \text{cox9} \) or \( \Delta \text{qcr9} \) during hexose fermentation.

2.3.6 Industrial strain fermentation with high initial glucose concentration

In order to evaluate the knockout strains in a condition where more fermentation takes place, we conducted another experiment using the industrial \( S. \text{cerevisiae} \) strain JAY291, which consumes glucose much faster (2.93 g/L·hr) and also produces ethanol much faster (1.21 g/L·hr) than either of the laboratory strains used previously (BY4742 or CEN.PK2-1D). We proceeded to delete \( QCR6 \) and \( QCR9 \) in the JAY291 background and then evaluated the mutants (\( \Delta \text{qcr6} \) and \( \Delta \text{qcr9} \)) in minimal media with a high initial glucose concentration (100 g/L). We chose to omit \( \text{COX9} \) deletion in JAY291 since deletion of \( QCR9 \) and deletion of \( \text{COX9} \) resulted in a similar growth phenotype for the laboratory strains. Fig. 2.6 shows the results for the high-sugar industrial strain fermentation. Although the improvement by our knockout mutants is reduced in these highly fermentative conditions, we still see a higher average ethanol production rate (9-12%) and higher glucose consumption rate (9-10%) by JAY291 \( \Delta \text{qcr6} \) or JAY291 \( \Delta \text{qcr9} \) as compared to the control. For both mutants, the improvement in glucose consumption rate was
statistically significant (P < 0.05). However, the increase in average ethanol production rate was not statistically significant at the 95% confidence interval for JAY291 Δqcr6 (P = 0.109) or JAY291 Δqcr9 (P = 0.055). There was almost no improvement in ethanol yield by the knockout strains under these conditions. The JAY291 Δqcr6 and JAY291 Δqcr9 strains both grew as well as the control throughout the fermentation and actually showed a slightly higher final cell density.

2.4 Discussion

Rapid and efficient hexose fermentation by *S. cerevisiae* with maximum yield and productivity of ethanol is important for many food and fuel applications. In addition, the ethanol production pathway is also an excellent model system for optimization of metabolic flux through central carbon metabolism. In this study, we employed an *in silico* genome-scale metabolic model to identify and select knockout targets for improving ethanol production in yeast while maintaining acceptable overall growth rate. The model-based strategy allowed us to narrow an extensive gene set and systematically guided us to consider genes encoding two oxidative phosphorylation enzymes, which catalyze the last two reactions in the respiratory electron transport chain. Ubiquinol cytochrome c reductase (EC 1.10.2.2) and cytochrome c oxidase (EC 1.9.3.1), also known as complex III and complex IV in the electron transport chain, are integral membrane proteins in the mitochondrial membrane which are critical for aerobic life due to their ability to maintain a proton gradient and thus synthesize ATP (Brunori et al., 2005). In *S. cerevisiae*, it is known that the ubiquinol cytochrome c reductase (QCR) enzyme contains ten non-identical subunits and is responsible for transferring electrons from reduced ubiquinone to ferricytochrome c (Shi et al., 2001). Subsequently, the cytochrome c oxidase (COX) enzyme,
which possesses eleven distinct subunits and depends on a number of other related proteins for assembly and activity, transfers electrons from ferrocytochrome c to oxygen and thus catalyzes the reduction of oxygen to water (Geier et al., 1995). The *Saccharomyces* Genome Database (Cherry et al., 2012) contains a wealth of information about genes that are directly or indirectly associated with these two multi-subunit enzymes. In this study, the stoichiometric model-based simulation predicted that deletion of any one of the multiple genes encoding the COX enzyme or QCR enzyme would have an identical effect on phenotype (Fig. 2.1). While stoichiometric modeling was useful for narrowing a large set of gene targets, this approach was not sufficient to distinguish between multiple genes linked to a single enzyme because the metabolic model does not account for genetic regulation and other unknown potential cellular interactions.

Distinguishing between the twenty gene targets in Table 2.2 is where the utility of the *in silico* model ends and the need for experimental screening begins. To our knowledge, no previous work has been done to evaluate and compare the phenotypic effect in *S. cerevisiae* by *in vivo* deletion of each one of the QCR-related genes or COX-related genes identified by our model-based gene targeting strategy.

Eighteen COX single knockout mutants and nine QCR single knockout mutants were screened by glucose fermentation in oxygen-limited conditions to evaluate the model predictions. Over two-thirds of the knockout mutants identified by the *in silico* study exhibited experimental behavior qualitatively consistent with model predictions; however, the notable exceptions illustrate the danger of using a purely stoichiometric model-based approach. The experimental screening results demonstrate a high degree of variability among the phenotypes of QCR and COX single knockout mutants (Fig. 2.2) and indicate that the individual subunits contribute to the overall function of the enzyme complexes in very different ways. For the nine QCR single
deletion mutants, the average change in ethanol production rate was a 16% increase and the average change in biomass yield was a 9% decrease as compared to the control strain. This result indicates a qualitative fit of experimental data with model predictions. However, the best QCR-related ethanol-producing strain ($Δqcr9$) increased ethanol production rate by 27% and reduced biomass yield by 20%; the worst QCR-related ethanol-producing strain ($Δrip1$) decreased ethanol production rate by 15% and increased biomass yield by 15%. The model-based gene targeting strategy was unable to predict these variations among QCR-related gene deletion mutants or to identify the most outstanding single knockout mutants. The data in Fig. 2.2, showing the variable effect of QCR and COX gene deletions on ethanol production rate and biomass yield, could be used to build a more accurate genome-scale metabolic model of S. cerevisiae.

Many previous studies have reported that utilizing respiration-deficient yeast strains, or petite mutants, can improve ethanol production efficiency (Bacila and Horii, 1979; Hutter and Oliver, 1998; Kim et al., 2010). Petite mutants include any yeast strain with a nuclear DNA mutation resulting in inability to utilize non-fermentable carbon sources such as ethanol, glycerol, or acetate. Petite mutants can still grow on glucose or other fermentable carbon sources but form smaller colonies than wild type cells (Tzagoloff and Dieckmann, 1990). For industrial ethanol production, two of the major advantages of respiration deficiency include: (i) elimination of oxidative metabolism of sugars and other fermentation products, and (ii) more leniency in fermentation conditions since a strict anaerobic environment is no longer required. Though it is well known in a general sense that respiration inhibits fermentation performance, knowledge is lacking of phenotype variation among respiration-deficient mutants and specific genetic targets for maximizing ethanol production with minimal impact on growth phenotype. There are
numerous approaches (genetic or chemical) to generate respiration-deficient mutants of *S. cerevisiae*, and all of the resulting strains are not equivalent. For example, some respiration-deficient mutants may have better fermentation performance than others, as illustrated by the large variation in ethanol production among 20 knockout strains identified and screened in this study (Fig. 2.2). Also, many respiration-deficient mutants are accompanied by a significant penalty in terms of cell growth, but the growth defect may be quite different among these strains (Fig. 2.2). In this study, the model-based gene deletion simulation indicated that COX-related or QCR-related deletion targets are better than other components of the respiratory system (e.g. NADH dehydrogenase, succinate dehydrogenase, fumarase, etc.) for improving ethanol production and maintaining sufficient biomass production. By experimental screening of the COX- and QCR-related mutants, we identified three specific knockout strains with maximum ethanol production and little or no growth penalty (Δcox9, Δqcr6, and Δqcr9). Previously, it was reported that *COX9* exists as a single copy in haploid strains of *S. cerevisiae*, produces one major transcript, and encodes subunit VIIa of the COX enzyme complex, which is essential for respiratory function (Wright et al., 1986). Similarly, QC9p, the smallest subunit of the QCR complex, is critical for correct structure of the enzyme and thus necessary for respiratory growth; its deletion led to elimination of electron transfer activity at the ubiquinol oxidase site (Graham et al., 1992; Phillips et al., 1990). In another study, QCR6p was demonstrated necessary for activity of the QCR complex only at high temperature (37°C), but the deletion mutant showed normal respiratory growth at 30°C (Yang and Trumpower, 1994). The present study is the first to describe the Δcox9, Δqcr6, and Δqcr9 strains in terms of ethanol-producing capability and to demonstrate the superlative characteristics of these three knockout mutants among many others for improving hexose fermentation in *S. cerevisiae* under microaerobic conditions. The Δqcr6
mutant was particularly interesting with a 24% improvement in ethanol production rate and no growth defect. Future work will be needed for elucidating the molecular mechanism to explain why eliminating specific protein subunits (Qcr6p, Qcr9p, or Cox9p) from the QCR or COX enzyme complex has a more beneficial influence on ethanol production and biomass yield than removing other subunits or related proteins.

We have demonstrated that the benefits of deleting COX9, QCR6, or QCR9 in S. cerevisiae for ethanol production in limited-oxygen conditions are consistently observed in different strain backgrounds and extend to common hexoses such as glucose, fructose, mannose, or sucrose (Fig. 2.4). Even the industrial strain fermentation with high initial glucose concentration showed an increase in average ethanol production rate and glucose consumption rate by the knockout strains tested; however, it is also important to note that improvements were significantly reduced in these highly fermentative conditions (Fig. 2.6). We admit that extreme sugar concentration or strict anaerobic conditions will lessen the beneficial effects by the knockout strains described in this study. It is well known that excess glucose represses respiration and contributes to highly fermentative metabolism in S. cerevisiae even in the presence of oxygen (i.e. Crabtree effect). Also, strict anaerobic conditions will certainly minimize respiration and maximize fermentation by yeast. In these situations, S. cerevisiae performs fermentation very well, and there is often little or no room for improvement by metabolic engineering. Still, some advantage could come from engineering strains that assure optimal fermentation in various conditions and does so regardless of changing process parameters such as dissolved oxygen or sugar concentration. For example, some applications may benefit from improved fermentation at lower sugar concentrations under continuous culture conditions. Also, many large-scale fermentations do not begin with a strict anaerobic
environment (i.e. no reactor purging, no nitrogen charging, no tight sealing of the vessel). Thus, the fermentation will behave anaerobically only after all initial dissolved oxygen is consumed by the cells, which means that substantial process time in a microaerobic state is possible. In changing or non-ideal fermentation conditions, the Δqcr6 mutant may offer particular advantages due to faster ethanol production rate, higher yield, and no biomass penalty.

In conclusion, we emphasize that the main significance of the present study lies in the methodology used. The single knockout strains described in this work are quite beneficial in certain conditions (microaerobic, lower external sugar concentration), but we cannot claim to have the best ever ethanol-producing *S. cerevisiae* strains. Still, the combination of genome-scale metabolic modelling with genetic screening using a systematic knockout collection could prove to be valuable in many other cases for reducing a large set of gene targets and identifying specific targets of interest.
2.5 Figures and tables

Fig. 2.1 (a) The *in silico* ethanol and biomass specific productivity for the wild type strain and twenty knockout mutants and (b) the percent change in ethanol and biomass yield by gene deletion as a function of oxygen uptake rate. Symbols: wild-type ethanol (closed blue triangle), knockout ethanol (open blue triangle), wild-type biomass (closed green circle), and knockout biomass (open green circle).
Fig. 2.2 A total of (a) eighteen cytochrome c oxidase single knockout mutants and (b) nine cytochrome c reductase single knockout mutants were screened by glucose fermentation in minimal media. The average percent change in ethanol production rate and biomass yield is displayed for each knockout mutant as compared to control strain BY4742. All results are the average of duplicate experiments and the error bars represent one standard deviation.
Fig. 2.3 Fermentation profiles for glucose consumption and ethanol production during flask fermentation by *S. cerevisiae* strains: (a, b) BY4742 and BY4742 Δcox9, (c, d) CEN.PK2-1D, CEN.PK2-1D Δqcr6, and CEN.PK2-1D Δqcr9. All results are the average of duplicate experiments. The error bars represent one standard deviation and are not visible when smaller than the symbol size.
Fig. 2.4 Comparison of (a) ethanol production rate, (b) ethanol yield, and (c) final cell density by laboratory strain BY4742 and knockout strains BY4742 Δcox9, BY4742 Δqcr6, and BY4742 Δqcr9 during fermentation in minimal media with four different carbon sources: glucose, mannose, fructose, and sucrose. All results are the average of duplicate experiments and the error bars represent one standard deviation. The student t test was used to establish significant differences between fermentations conducted with the knockout strains and the control strain.

*, P < 0.05 (95% confidence). †, P < 0.10 (90% confidence).
Fig. 2.4 (cont.)
Fig. 2.5 Wild-type *S. cerevisiae* with CEN.PK2-1D background or BY4742 background and single knockout mutants of cytochrome c oxidase (∆cox9) and cytochrome c reductase (∆qcr6, ∆qcr9) were grown on YP glycerol agar plates. Plates were incubated at 30°C for 3 days (CEN.PK2-1D) or 6 days (BY4742).
Fig. 2.6 Comparison of ethanol yield, ethanol production rate, and glucose consumption rate by industrial strain JAY291 and knockout strains JAY291 Δqcr6 and JAY291 Δqcr9 during fermentation in minimal media with high initial glucose concentration. All results are the average of duplicate experiments and the error bars represent one standard deviation. The student t test was used to establish significant differences between fermentations conducted with the knockout strains and the control strain. *, P < 0.05 (95% confidence). †, P < 0.10 (90% confidence).
Table 2.1 Primers used in this study

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>COX9-f</td>
<td>AGTTTTGTGGTTGAGCAGTCG</td>
<td>COX9 deletion cassette using a KanMX marker</td>
</tr>
<tr>
<td>COX9-r</td>
<td>GGCAAAATGGCAGGTATTCG</td>
<td></td>
</tr>
<tr>
<td>Cd-f</td>
<td>CCTTCGATGGATTCGTACGT</td>
<td>confirm COX9 deletion</td>
</tr>
<tr>
<td>KanMX-r</td>
<td>CTTTTCCCTACCATGGTTGT</td>
<td></td>
</tr>
<tr>
<td>QCR6-f</td>
<td>GCTACAATCAAGCCCAGCTGCTAT</td>
<td>QCR6 deletion cassette using a KanMX marker</td>
</tr>
<tr>
<td>QCR6-r</td>
<td>CTACCTGCATTCCAATGGGCG</td>
<td></td>
</tr>
<tr>
<td>Q6d-f</td>
<td>GGTGGAAGTTGATGATGGA</td>
<td>confirm QCR6 deletion</td>
</tr>
<tr>
<td>QCR9-f</td>
<td>GGGTGACGAATTTCAGTGACT</td>
<td>QCR9 deletion cassette using a KanMX marker</td>
</tr>
<tr>
<td>QCR9-r</td>
<td>GCAAGACACACCCCAAACCTA</td>
<td></td>
</tr>
<tr>
<td>Q9d-f</td>
<td>TGCCATGAGAAGAGGTATTTA</td>
<td>confirm QCR9 deletion</td>
</tr>
</tbody>
</table>
Table 2.2 *In silico* gene deletion targets for improving ethanol production while maintaining adequate biomass for cell growth

<table>
<thead>
<tr>
<th>Standard Name</th>
<th>Systematic Name</th>
<th>Enzyme Name</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>COX1</td>
<td>Q0045</td>
<td>cytochrome c oxidase (complex IV)</td>
<td>Oxidative Phosphorylation</td>
</tr>
<tr>
<td>COX2</td>
<td>Q0250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COX3</td>
<td>Q0275</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COX4</td>
<td>YGL187C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COX6</td>
<td>YHR051W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COX7</td>
<td>YMR256C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COX8</td>
<td>YLR395C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COX9</td>
<td>YDL067C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COX12</td>
<td>YLR038C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COX13</td>
<td>YGL191W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COB</td>
<td>Q0105</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COR1</td>
<td>YBL045C</td>
<td>ubiquinol 6 cytochrome c reductase (complex III)</td>
<td></td>
</tr>
<tr>
<td>CYT1</td>
<td>YOR065W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QCR2</td>
<td>YPR191W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QCR6</td>
<td>YFR033C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QCR7</td>
<td>YDR529C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QCR8</td>
<td>YIL166W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QCR9</td>
<td>YGR183C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>QCR10</td>
<td>YHR001W-A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RIP1</td>
<td>YEL024W</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.3 Fermentation parameters\(^{a}\) of \textit{S. cerevisiae} wild type and Δ\textit{cox}9, Δ\textit{qcr}6, and Δ\textit{qcr}9 mutants in glucose with various strain backgrounds

<table>
<thead>
<tr>
<th>Strains</th>
<th>Glucose</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(r_{\text{Glc}}) ± SD</td>
<td>(P_{\text{EtOH}}) ± SD</td>
<td>(Y_{\text{EtOH}}) ± SD</td>
</tr>
<tr>
<td>BY4742</td>
<td>0.811 ± 0.009</td>
<td>0.275 ± 0.002</td>
<td>0.339 ± 0.002</td>
</tr>
<tr>
<td>BY4742 Δ\textit{cox}9</td>
<td>0.991 ± 0.002</td>
<td>0.376 ± 0.003</td>
<td>0.380 ± 0.004</td>
</tr>
<tr>
<td>BY4742 Δ\textit{qcr}6</td>
<td>0.952 ± 0.006</td>
<td>0.342 ± 0.010</td>
<td>0.359 ± 0.009</td>
</tr>
<tr>
<td>BY4742 Δ\textit{qcr}9</td>
<td>0.959 ± 0.009</td>
<td>0.349 ± 0.007</td>
<td>0.364 ± 0.004</td>
</tr>
<tr>
<td>CEN.PK2-1D</td>
<td>1.312 ± 0.004</td>
<td>0.448 ± 0.008</td>
<td>0.342 ± 0.006</td>
</tr>
<tr>
<td>CEN.PK2-1D Δ\textit{cox}9</td>
<td>1.586 ± 0.001</td>
<td>0.594 ± 0.006</td>
<td>0.375 ± 0.004</td>
</tr>
<tr>
<td>CEN.PK2-1D Δ\textit{qcr}6</td>
<td>1.573 ± 0.002</td>
<td>0.608 ± 0.002</td>
<td>0.387 ± 0.001</td>
</tr>
<tr>
<td>CEN.PK2-1D Δ\textit{qcr}9</td>
<td>1.567 ± 0.002</td>
<td>0.619 ± 0.003</td>
<td>0.395 ± 0.003</td>
</tr>
</tbody>
</table>

\(^{a}\)Fermentation parameters were measured at 36.5 hrs (BY4742) or 22 hrs (CEN.PK2-1D) of incubation. \(r_{\text{Glc}}\), glucose consumption rate (g glucose/L·hr); \(P_{\text{EtOH}}\), ethanol productivity (g ethanol/L·hr); \(Y_{\text{EtOH}}\), ethanol yield (g ethanol/g glucose); SD means standard deviation
CHAPTER III  RAPID AND EFFICIENT GALACTOSE FERMENTATION BY ENGINEERED SACCHAROMYCES CEREVISIAE THROUGH AN INVERSE METABOLIC ENGINEERING APPROACH

The content of this chapter is in preparation for submission. I was the first author of the paper and Xueyang Feng, Jeffrey Skerker, Ian Liu, Adam Arkin, Huimin Zhao, and Yong-Su Jin were co-authors. I performed the research with help from the co-authors and Dr. Yong-Su Jin was the director of the research.
3.1 Introduction

Biological production of fuels and chemicals by native or engineered microorganisms has received much attention in recent years as a sustainable alternative to chemical synthesis from limited petroleum-derived hydrocarbons. In order to become an economically viable option with significant environmental benefits, fuels and chemicals should be synthesized from renewable non-food plant sources, such as lignocellulosic biomass or marine plant biomass (Ragauskas et al., 2006). These feedstocks are renewable, cheap, and abundant; however, they also contain significant portions of sugars that cannot be readily fermented by industrial microorganisms. Slow and inefficient conversion of nonfavored carbon sources by fermenting microorganisms is an intrinsic problem delaying the sustainable production of biofuels, biochemicals, and other biomaterials (Stephanopoulos, 2007). The identification of novel gene targets for improving utilization of non-glucose sugars which are abundant in plant cell wall hydrolyzates is a critical aspect of industrial strain construction and has been a focus of numerous metabolic engineering studies.

Galactose is the major sugar obtained from hydrolysis of some marine biomass, such as red seaweed (Yoon et al., 2010), and is also found in other industrial sources such as cheese whey (Siso, 1996) or molasses (Rosen, 1987). The metabolic pathway for conversion of galactose to the glycolytic intermediate glucose-6-phosphate, called the Leloir pathway, is under strict genetic control and has long been of interest as a model system for studying gene expression and transcriptional control in eukaryotes (Timson, 2007). The yeast Saccharomyces cerevisiae, which is a common microorganism for industrial biotechnology, possesses the complete Leloir pathway and is capable of growing on galactose, but the specific uptake rates are approximately 2.5-3.3-fold lower than those on glucose. Moreover, ethanol yield and
productivity are known to be considerably lower during galactose fermentation as compared to glucose fermentation because of slower respiro-fermentative metabolism caused by a lack of the Crabtree effect (Ostergaard et al., 2000e). In previous studies, two distinct metabolic engineering approaches have been applied to improve galactose utilization by *S. cerevisiae*. Ostergaard et al. used a rational approach to achieve balanced overexpression of galactose metabolic genes by modifying the regulatory network of the Leloir pathway. They found that deletion of three genes encoding negative regulators of GAL gene expression (*GAL6, GAL80,* and *MIG1*) led to a 41% increase in galactose uptake rate, and overexpression of one gene encoding a positive activator (*GAL4*) led to a 26% increase in uptake rate (Ostergaard et al., 2000b). However, the *ad hoc* pathway manipulation approach is limited because of the lack of rigorous system-wide analysis and thus potential for unknown or unexpected interactions between reactions or pathways. Therefore, combinatorial approaches, which involve random genetic perturbations and appropriate screening or selection techniques, have proven to be more valuable for identifying novel genetic targets to improve galactose fermentation in yeast. A particular advantage of combinatorial engineering is unique access to still poorly understood cellular subsystems, unknown regulatory genes, or non-coding regions that are critical for cellular function and could not be identified by a more rational approach (Alper, 2006; Sauer and Schlattner, 2004). For example, Lee et al. employed a genome-wide perturbation library and careful selection techniques to engineer superior galactose-fermenting yeast strains, which led to discovery of overexpression targets *SNR84* (encoding a small nuclear RNA) and truncated *TUP1* (encoding a general repressor of transcription) (Lee et al., 2011). Also, Hong et al. utilized an evolutionary engineering approach with a laboratory strain over many generations to increase
galactose uptake rate and growth rate; this led to identification of an unexpected mutation in RAS2 (encoding a GTP-binding protein) (Hong et al., 2011b).

The aforementioned metabolic engineering studies have demonstrated the utility of rational or combinatorial approaches for improving galactose utilization by S. cerevisiae, but the resulting engineered or evolved strains show enhanced galactose utilization only under aerobic conditions because of an inherent linkage between galactose metabolism and respiration (Hong et al., 2011b). However, aerobic conditions are undesirable for the production of biofuels and chemicals because respiratory utilization of sugars can lead to lower yields (Zhang et al., 2010). Furthermore, it is not economical or convenient to maintain aeration during industrial-scale fermentation (Peng et al., 2012). Thus, respiration-deficient yeast strains may be beneficial to ensure optimal fermentation in various conditions by eliminating oxidative metabolism of sugars and other fermentation products. Respiration-deficient strains have been employed for improving ethanol productivity and/or yield by S. cerevisiae from glucose (Bacila and Horii, 1979; Hutter and Oliver, 1998; Oner et al., 2005) or xylose (Jin et al., 2004; Peng et al., 2012). However, development of a respiration-deficient galactose-fermenting strain has not been attempted and is especially challenging because galactose metabolism is associated with heightened respiratory activity (Fendt and Sauer, 2010).

In this study, we aimed to construct optimal yeast strains that maximize carbon flux from galactose through the fermentation pathway rather than the respiration pathway. For this purpose, a respiration-deficient mutant was used for strain development; specifically, we employed a laboratory strain with deletion of COX9 (encoding a subunit of cytochrome c oxidase). The BY4742 Δcox9 mutant was shown to have higher fermentation rate and efficiency than the parental strain with only a minor growth disadvantage on glucose, fructose, sucrose, or
mannose. However, this Cox9-negative strain was unable to grow with galactose as the sole carbon source. Interestingly, we found that spontaneous genetic mutations during prolonged incubation allowed the respiration-deficient strain to use galactose fermentatively, and the best mutants were enriched by serial subculture. As follows, we present the results of an integrated systems biology analysis of the evolved galactose-fermenting strains developed in this study. Furthermore, we describe a new paradigm for evolutionary engineering, termed “fermentative evolution,” whereby a metabolic “death valley” may be a beneficial intermediate step to provide pressure for adaptive evolution of microbial strains toward optimal production of ethanol or other fermentation products.

3.2 Materials and methods

3.2.1 Yeast strains and evolutionary engineering

*S. cerevisiae* strain BY4742 (*MATalpha, leu2, his3, ura3, lys2*) was used in this study as a control strain. Single knockout mutant BY4742 Δ*cox9* was obtained from Yeast Knockout MATalpha Collection (OpenBiosystems, Lafayette, CO) and was the starting strain for adaptive evolution. After almost 100 hours, the Δ*cox9* mutant began to grow in galactose (40 g/L) minimal media and evolved mutants were generated by serial subculture over ≈15 generations. Two single colony isolates were obtained from the last shake flasks and named JQ-G1 and JQ-G2. All strains and plasmids used in this work are summarized in Table 3.1. The primers used for amplification of gene deletion cassettes and mutant *YHP1* insert for cloning are shown in Table 3.2, along with confirmation primers.
3.2.2 Medium and culture conditions

To prepare yeast strains for inoculation, cells were cultivated overnight at 30°C and 300 RPM in YPD medium (10 g/L yeast extract, 20 g/L Bacto peptone, 20 g/L glucose). Batch fermentation was carried out at 30°C and 100 RPM in yeast synthetic complete (SC) medium composed of 1.7 g/L yeast nitrogen base, 5 g/L ammonium sulfate, and 0.79 g/L CSM (MP Biomedicals, Solon, OH) for supplying amino acids and nucleobases. Galactose (or other sugar) was added to provide the carbon source. For anaerobic fermentations, ergosterol and Tween80 were also included in the medium at a final concentration of 10 mg/L and 420 mg/L, respectively. To select yeast transformants with an amino acid auxotrophic marker, a SC agar plate prepared with CSM-His-Leu-Ura (MP Biomedicals, Solon, OH) was used.

3.2.3 Fermentation experiments

Yeast pre-cultures were grown with 5 mL of YPD medium in an orbital shaker at 30°C to prepare inoculums for fermentation experiments. Cells in exponential phase were harvested and inoculated after removing used YPD. Flask fermentation experiments were performed using 25 mL (or 50 mL) of SC medium with 40 g/L of the appropriate sugar in a well-controlled shaking incubator (Thermoscientific, MaxQ4000, Dubuque, IA) under oxygen-limited conditions. Initial cell densities were adjusted to OD\textsubscript{600} of ~0.1 unless otherwise noted. During the course of the fermentation, the temperature and agitation rate were kept constant at 30°C and 100 RPM, respectively. Anaerobic fermentations were conducted in 20 mL of SC medium with the appropriate sugar in a sealed and purged 100 mL vial.
3.2.4 Analytical methods

Cell growth was measured by optical density (OD) at 600 nm using a UV-visible spectrophotometer (Biomate 3, Thermoscientific, Madison, WI). The concentration of extracellular metabolites was determined by a high-performance liquid chromatography system (Agilent Technologies 1200 series) with a Rezex ROA-Organic Acid H+ column (Phenomenex Inc., Torrance, CA). Column temperature was kept constant at 50°C and the elution was performed with 0.005 N H$_2$SO$_4$ at a flow rate of 0.6 ml/min. Galactose, ethanol, acetate, and glycerol were detected with a refractive index detector (RID) and oxaloacetate was detected with a diode-array detector (DAD) at 210 nm.

3.2.5 Genome sequencing and SNP discovery

The genomic DNA of BY4742, BY4742 Δcox9, and evolved strains JQ-G1 and JQ-G2 was prepared with the YeaStar Genomic DNA Kit (Zymo Research, Irvine, CA) and quality was confirmed in a 1% agarose gel. Sequencing (SE100) was done using an Illumina HiSeq2000 machine at the W.M. Keck Center for Comparative and Functional Genomics at the University of Illinois at Urbana-Champaign. The barcoded shotgun genomic DNA libraries were constructed with the TruSeq Sample Prep Kit (Illumina, San Diego, CA). The raw data was processed with Casava 1.8. Approximately 11,000,000 reads were obtained for each sample, which corresponds to >50X sequence coverage. Variant analysis (either SNPs or INDELs) was performed using CLC Genomics Workbench version 7.0.4. Reads were trimmed based on quality scores with default program settings and the trimmed reads were mapped to an S288C reference sequence (obtained from Genbank; accession numbers for 16 chromosomes and the mitochondrial genome: NC_001133 through NC_001148 and NC_001224) using the following settings:
Masking mode = No masking; Mismatch cost = 2; Insertion cost = 3; Deletion cost = 3; Length fraction = 0.5; Similarity fraction = 0.8; Global alignment = No; Non-specific match handling = Ignore. The BY4742, BY4742 Δcox9, JQ-G1, and JQ-G2 SNPs or INDELs were identified using the probabilistic variant detection algorithm in CLC Genomics Workbench and the following settings: Ignore non-specific matches = YES; Ignore broken pairs = Yes; Minimum coverage = 30; Variant probability = 90.0; Required variant count = 30; Require presence in both forward and reverse reads = No; Ignore variants in non-specific regions = Yes; Filter 454/Ion homopolymer indels = No; Maximum expected variants = 1. Each strain yielded approximately 150 variants compared to S288C. The SNPs unique to JQ-G1 and JQ-G2 relative to BY4742 Δcox9 were then determined by comparing the SNP files using the "Filter against Known Variants" function in CLC Genomics Workbench. Of the 144 variants common to JQ-G1 and JQ-G2, only 9 variants were unique when compared to the parental strain BY4742 Δcox9 (Table 3.3).

3.2.6 Metabolic flux analysis

The BY4742 and JQ-G1 strains were inoculated into defined minimal media as described previously (Christen and Sauer, 2011) with 10 g/L [1-13C] galactose (Cambridge Isotope Laboratories). Cells were grown in duplicates and harvested during mid-exponential phase (Fig. 3.10). Hydrolysis of biomass, derivatization of amino acids, GC-MS analysis, measurement of isotope labeling patterns, and metabolic flux calculations were carried out as discussed thoroughly in a previous study (Feng and Zhao, 2013). In general, extracellular fluxes were calculated from HPLC data using two samples from exponential growth phase. A stoichiometric model developed by Feng and Zhao (Feng and Zhao, 2013) was modified to include the Leloir
pathway and used in this study to fit the metabolic flux data. The stoichiometric reactions included in the metabolic model and the standard deviation of calculated fluxes is shown in Fig. 3.11. The measured and simulated isotopomer labeling patterns of amino acids are compared in Fig. 3.12.

3.2.7 Metabolite profiling

The BY4742 and JQ-G1 strains were inoculated into defined minimal media as described previously (Christen and Sauer, 2011) with 10 g/L galactose. Cells were grown in triplicates and harvested during mid-exponential phase (Fig. 3.10). Quenching was done quickly in cold methanol (≤-40°C) with a 5:1 ratio of solvent to culture broth. Metabolite extraction, derivatization, and GC-MS analysis were carried out as described previously (Kim et al., 2013b). For processing the GC-MS data, we used AMDIS software version 2.7.1 (Davies, 1998) for peak deconvolution and automatic detection and then uploaded the data to SpectConnect (Styczynski et al., 2007) for peak alignment and generating the data matrix. The Golm Metabolome Database mass spectral reference library (Hummel et al., 2007) was used for metabolite identification.

3.3 Results

3.3.1 Physiological characterization of evolved mutants

While it has been known that galactose cannot be utilized by native S. cerevisiae strains under anaerobic conditions (i.e. without metabolic fluxes towards the respiratory pathway), we found that a respiration-deficient mutant (BY4742 Δcox9) was able to grow and use galactose
fermentatively after serial subcultures in minimal media with galactose as the sole carbon source. The change in galactose uptake rate and ethanol yield during the serial subcultures was quite abrupt and dramatic (Fig 3.6a), with most of the phenotype improvement already observed in the second subculture. Two evolved strains, JQ-G1 and JQ-G2, were isolated from the last shake flask after enrichment over a relatively short period of time (≈6 days). The evolved strains not only consumed galactose faster than the wild-type (2.5-fold improvement) but also were far superior in terms of ethanol yield (1.9-fold improvement) and volumetric productivity (4.8-fold improvement). As expected, the phenotype advantages of the evolved strain were accompanied by a 20% reduction in biomass production. Fig. 3.1 shows the complete fermentation profile on galactose by wild-type BY4742 (a), knockout BY4742 ∆cox9 (b), and evolved knockout JQ-G1 (c). The intermediate step of COX9 deletion between the parental strain and evolved strains strongly inhibited galactose utilization, but the deletion was also crucial for achieving optimal galactose fermentation rapidly through adaptive evolution. In contrast, serial subculture with the respiration-competent wild-type strain resulted in relatively little increase in galactose uptake rate or ethanol yield during the same time period required to generate our respiration-deficient evolved strains (JQ-G1 and JQ-G2). Also, it is noteworthy that COX9 deletion offered a great advantage for enhancing galactose utilization in a condition where the deletion greatly inhibited growth, but serial subculture of the ∆cox9 mutant with fructose or mannose as the sole carbon source did not lead to improvement in sugar consumption or ethanol production (Fig. 3.6b-c). As reported previously, BY4742 ∆cox9 grows well in minimal media with fructose or mannose and produces ethanol even better than the parental strain.

To evaluate stability of the galactose-evolved mutant phenotype, JQ-G1 and JQ-G2 were grown in minimal media with glucose for an extended period (3 serial subcultures or ≈15
generations) and then immediately transferred to galactose minimal media. The extended growth in glucose did not have any significant effect on galactose uptake rate, ethanol yield or ethanol productivity by the evolved strains (P > 0.05). Also, we tested detailed metabolic phenotype of the evolved strains on other hexoses. When growing the galactose-evolved mutant on glucose, we observed that JQ-G1 produced 22% more ethanol and consumed glucose 21% faster than the wild-type (BY4742), although with an 8% reduction in biomass production. Similar results were observed during fermentation experiments with mannose, fructose, or sucrose as the sole carbon source, except that the JQ-G1 strain did not show a significant reduction in biomass production on the latter two sugars as compared to the wild-type (P > 0.05). Fig 3.2 shows the specific sugar uptake rate (a) and specific ethanol production rate (b) by JQ-G1 versus BY4742 on all sugars evaluated. For each of the five carbon sources, the evolved strain showed significant improvement in sugar uptake rate and ethanol productivity (P < 0.05), with the largest advantages on galactose and the smallest advantage on sucrose.

The previous fermentations were performed with fairly low initial cell density (0.03 g/L), but we wanted to get an idea of the fermentation parameters of the evolved strain with a larger inoculum. By increasing initial cell density to ~1.5 g/L, we were able to improve galactose uptake rate and ethanol productivity by JQ-G1 in minimal media to 1.84 g/L·hr and 0.68 g/L·hr, respectively, as compared to 1.01 g/L·hr and 0.34 g/L·hr by BY4742. In these conditions, JQ-G1 grew as well as, or slightly better than, the wild-type strain and also generated about one-fourth the acetate. Ethanol yield by the evolved strain was similar in the low cell density (0.38 g/g) and high cell density (0.37 g/g) fermentations. Fig. 3.7 shows the fermentation profile of these strains in the high cell density experiment.
3.3.2 Changes in genotype and reverse engineering of galactose-fermenting strains

The entire genome of the evolved mutants JQ-G1 and JQ-G2, as well as control strains BY4742 and BY4742 Δcox9, was sequenced and compared in order to identify genetic perturbations associated with improved galactose metabolism without respiration. Any SNPs observed in the evolved strains but not present in either control strain are real and may contribute to the beneficial phenotype. After eliminating silent substitutions or genetic mutations in non-coding regions, there were only two remaining SNPs shared in both evolved strains: a truncation in \textit{GAL80} (Glu348*) and a point mutation in \textit{YHP1} (Glu67Gln). \textit{GAL80} is known to be a transcriptional repressor involved in repression of \textit{GAL} genes (Johnston, 1987). \textit{YHP1} is another transcriptional repressor encoding a homeoprotein (Kunoh et al., 2000), but this gene has not been reported to be linked with galactose metabolism.

Initially, we hypothesized that the observed mutations in \textit{GAL80} and \textit{YHP1} might cause loss of function of the related genes in evolved strain JQ-G1. In order to test the effect of these deletions, we constructed single knockout strains BY4742 Δgal80 and BY4742 Δyhp1, as well as double and triple knockout strains with deletion of every possible combination of \textit{GAL80}, \textit{YHP1}, and \textit{COX9} (since the evolved strains are \textit{COX9}-negative). The knockout strains were evaluated in minimal media with 40 g/L of galactose, and then specific sugar uptake rate and specific ethanol production rate were calculated for each mutant (Fig. 3.3). First, the results clearly indicate the negative impact of transcriptional repressor GAL80p on galactose fermentation parameters, as deletion of the associated gene (\textit{GAL80}) increased specific galactose uptake rate and specific ethanol production rate in the wild-type strain by a factor of 1.6 and 2.3, respectively. The BY4742 Δgal80 mutant had a shorter lag time on galactose and grew faster than the control during the early part of the fermentation, but final cell density was similar for
both strains. Furthermore, deletion of \textit{GAL80} in a respiration-deficient strain (BY4742 \textit{Δ}cox9) had an even more drastically positive effect on galactose fermentation, resulting in a double knockout mutant with 5.4-fold improvement in specific ethanol productivity over the wild-type and a similar ethanol yield as JQ-G1 ($0.389 \pm 0.008$ g/g versus $0.384 \pm 0.003$ g/g, respectively). The BY4742 \textit{Δ}cox9 \textit{Δ}gal80 strain was better than all other knockout strains in this study in terms of galactose fermentation rate and efficiency, although specific ethanol production rate was still 19\% lower than that of evolved strain JQ-G1. Not surprisingly, the \textit{Δ}cox9 \textit{Δ}gal80 mutant also had a longer lag time and generated 27\% less biomass than the wild-type strain. Second, the results demonstrate that \textit{YHP1} deletion did not have a significant effect on growth or other galactose fermentation parameters in any of the strains tested. Also, expression of mutant \textit{YHP1} (Glu67Gln) on a single-copy plasmid in BY4742 \textit{Δ}cox9 \textit{Δ}gal80 \textit{Δ}yhp1 resulted in very marginal improvement in ethanol yield as compared to a control strain harboring the null vector (Fig. 3.8).

Third, our results show that \textit{COX9} deletion was beneficial in all \textit{GAL80}-negative strains for increasing ethanol yield and productivity from galactose, but the same deletion (\textit{Δ}cox9) was inhibitory for growth and fermentation in the \textit{GAL80}-positive strains. To confirm the effect of deleting \textit{COX9} in the evolved strain, we introduced a \textit{COX9} cassette into the JQ-G1 strain and selected transformants on a glycerol plate. The JQ-G1-\textit{COX9} strain showed a 33\% reduction in specific galactose uptake rate and a 49\% reduction in specific ethanol production rate as compared to JQ-G1, although the \textit{COX9}-expressing strain had a shorter lag time. Table 3.4 provides a summary of important fermentation parameters for engineered strains in this study.

### 3.3.3 Integration of fluxomic and metabolomic data

In order to elucidate metabolic changes related to efficient galactose fermentation by a respiration-deficient strain, we sought quantification of metabolic flux and metabolite abundance
in the evolved strain as compared to the wild-type strain. Extracellular fluxes and isotope labeling patterns of proteinogenic amino acids were used to calculate the distribution of intracellular metabolic fluxes in the wild-type and evolved strains during steady-state growth on [1-$^{13}$C] galactose (Fig. 3.4). The galactose uptake fluxes were normalized to a relative value of 100 in both strains. Based on the $^{13}$C-MFA results, the BY4742 and JQ-G1 strains were similar in that glycolysis was the main catabolic pathway for breakdown of glucose-6-phosphate generated by the Leloir pathway. In both strains, at least 95% of carbon flux was processed by glycolysis (relative flux value of 94.5 – 97.7) and thus only a small percentage was shuttled toward the pentose phosphate pathway (relative flux value of 2.9 – 5.3). After glycolysis, the metabolic fluxes split at the node of pyruvate to enter either the ethanol fermentation pathway or the TCA cycle. Comparison of TCA cycle fluxes in the two strains revealed major differences, with the evolved strain demonstrating almost 3-fold reduction as compared to the wild-type (relative flux value of 23.4 versus 63.6). The higher TCA cycle flux in the parental strain is consistent with previous studies that have shown the importance of oxygen supply and high cellular energy charge for galactose metabolism (Sanchez et al., 2010a; van den Brink et al., 2009). However, results in the current study demonstrate that much of the excess TCA cycle flux has been redirected to the ethanol fermentation pathway in evolved strain JQ-G1, while still maintaining enough ATP and biomass precursors for cell growth. Almost 73% of total pyruvate flux was converted to ethanol in the evolved strain, while only about 52% of pyruvate flux passed through the two-step fermentation pathway in the wild-type strain. Conversely, much less of the pyruvate flux was directed to acetyl-CoA by the evolved strain as compared to the wild-type, either in the cytosol (relative flux value of 8.1 versus 38.4) or the mitochondria (relative flux value of 2.0 versus 6.6). The futile cycle including pyruvate carboxylase and
phosphoenolpyruvate carboxykinase was active in both strains but with similar flux values. Differences between the calculated flux results in this study and two previous $^{13}$C-MFA studies with \textit{S. cerevisiae} growing on galactose may be due to the use of different strain backgrounds: CEN.PK 113-7D by Ostergaard et al. (Ostergaard et al., 2001), FY4 by Fendt and Sauer (Fendt and Sauer, 2010), and BY4742 in this study. The host dependence could lead to large variations in metabolic flux, as was found in pathway switch experiments for xylose utilization (Feng and Zhao, 2013).

To further investigate the metabolic flux distributions during growth of evolved strain JQ-G1 on galactose and to identify potential rate-limiting reactions in central carbon metabolism, we also measured intracellular metabolite concentrations by GC-MS analysis. The results showed 38 metabolites whose intracellular concentration was significantly changed ($P < 0.05$) in the evolved mutant as compared to the control strain (Fig. 3.5). Pyruvate, a central node for many high flux metabolic pathways, had a higher concentration in the evolved strain, and this result matches the $^{13}$C-MFA data showing higher fermentation pathway flux from pyruvate to ethanol in this strain. Also, it was not surprising that some TCA cycle intermediates, such as fumarate and malate, had a lower concentration in the JQ-G1 strain since its TCA cycle fluxes were also significantly lower than the control strain. However, citrate, which is a longer six-carbon intermediate early in the TCA cycle, was an exception in terms of the TCA cycle metabolites and had a higher concentration in the JQ-G1 strain. In general, the level of free amino acids was higher in the evolved strain, with aspartate and glutamate as the only individual amino acids not consistent with that trend. Interestingly, the evolved strain had reduced levels of galactitol, which has been shown to play a role in galactose-induced toxicity (de Jongh et al., 2008). Finally, two other interesting metabolites, trehalose and glyceric acid-3-phosphate, had
reduced concentrations in the evolved strain and were included in Fig. 3.5, although the
difference was not quite significant at the 95% confidence interval (P = 0.057 and P = 0.075,
respectively). In summary, the evolved mutant showed a clear separation from the control in the
concentration of many key intracellular metabolites, such as amino acids, TCA cycle
intermediates, and galactitol.

3.4 Discussion

Due to the limitations of rational metabolic engineering approaches (Çakar et al., 2012),
evolutionary engineering has been widely used as a combinatorial tool in order to identify
unknown or unexpected genetic targets for S. cerevisiae strain improvement. The approach
depends on some spontaneous or induced mutations in a starting strain and appropriate selection
pressure for enriching fitter variants during serial or prolonged cultivation in liquid or solid
media (Çakar et al., 2012; Sauer, 2001). In numerous studies, evolutionary engineering has been
used to improve strain performance by growing a sub-optimal wild-type microorganism in
aerobic conditions for many generations. While this conventional approach is perfectly suited to
select for improved strains in the context of sugar consumption or cell growth, it is more
challenging to find suitable selective conditions for evolved strains with maximum product
formation. Genetic mutations leading to faster growth by the host strain are important for out-
competing rival mutants and thus are selected for; however, these mutations may not result in the
best product yield or production rate. On the other hand, mutations leading to better product
formation are usually not critical to the strain’s survival or performance, and therefore those
mutants may not be enriched. To overcome these challenges, we demonstrated that a metabolic
“death valley” involving some stress or adverse genetic perturbations may be advantageous to
increase the rate of spontaneous mutagenesis in a population and provide pressure for rapid
evolution toward a more desirable phenotype. Though somewhat counterintuitive, we developed
an efficient galactose-fermenting mutant by employing a starting strain with a mutation \((\Delta \text{cox9})\)
that was detrimental to growth on galactose (i.e. a metabolic “death valley”). The respiration
deficiency of the starting strain had two very positive effects for evolutionary engineering: (i)
forcing the cells to evolve quickly in order to survive in galactose minimal media, and (ii)
redirecting metabolic flux toward the fermentation pathway to improve product yield (rather than
directing flux toward respiratory metabolism). The entire adaptive evolution process required
only about 15 generations (≈6 days) and resulted in evolved strain JQ-G1 with a 590%
improvement in specific ethanol productivity as compared to BY4742. In contrast, Hong et al.
evolved a wild-type strain of yeast for a much longer time period (400 generations or ≈62 days)
and increased specific ethanol production rate by about 170% in the best evolved mutant 62C
(Hong et al., 2011b). Furthermore, adaptive evolution led to some evolutionary trade-offs when
comparing the performance of 62C in glucose and galactose (Hong and Nielsen, 2013).
However, mutations for aiding galactose fermentation by evolved strain JQ-G1 did not have a
negative effect when fermenting glucose or other hexoses (Fig. 3.2). Based on the results in this
study, we suggest that our fermentative evolution approach offers an alternative to conventional
evolutionary engineering and may be advantageous for maximizing production of ethanol or
other fermentation products in a host strain.

A previous study by Fendt and Sauer reported that several TCA cycle and electron
transport chain enzymes were significantly up-regulated in \(S. \text{cerevisiae}\) during growth on
galactose. Similarly, \(^{13}\text{C-MFA}\) calculations showed a heightened TCA cycle flux by the wild-
type yeast strain during growth on galactose as compared to glucose (Fendt and Sauer, 2010).
The importance of respiratory metabolism for galactose utilization was further demonstrated by the inability of respiration-deficient mutant Δrip1 to consume galactose in aerobic substrate-switch experiments due to low energy charge (van den Brink et al., 2009). In the present study, we also observed no growth by respiration-deficient mutant Δcox9 during batch fermentation in galactose minimal media. Although heightened respiration is generally associated with galactose utilization, this mode of metabolism is certainly not ideal for attaining optimal conversion of galactose to ethanol or other fermentation products as discussed previously. In this study, we demonstrated the possibility of efficient galactose utilization without respiration by deletion of COX9, encoding an electron transport chain enzyme subunit, and adaptive evolution, which led to a loss of function mutation in a negative regulator of the Leloir pathway (Gal80p). The metabolic flux results clearly demonstrate the shift in carbon flux from the TCA cycle to the ethanol fermentation pathway in the evolved strain JQ-G1 (Fig. 3.4). Also, re-introduction of a COX9 cassette back into the evolved strain significantly reduced its galactose fermentation performance in terms of sugar uptake rate and ethanol production (Fig. 3.3; Table 3.4); this result shows the value of inhibiting respiration for optimizing fermentation pathway flux from galactose (with some minor penalty in terms of cell growth). Although Gal80p has been known to strongly inhibit galactose utilization in yeast (Ostergaard et al., 2000b), we demonstrate here the synergy between inhibition of respiration and elimination of Gal80p function for attaining high fermentation rate and yield from galactose. Specifically, we reported that double knockout mutant Δcox9Δgal80 significantly out-performed single knockout strain Δgal80 in galactose minimal media under microaerobic conditions (Fig. 3.3; Table 3.4). When respiration was inhibited by anaerobic conditions, deletion of GAL80 alone was sufficient for rapid and efficient galactose fermentation with no additional improvements by COX9 deletion (Fig. 3.9).
To our knowledge, this is the first study demonstrating efficient galactose fermentation by respiration-deficient *S. cerevisiae*. The evolved strain JQ-G1, which was obtained through both systematic and combinatorial engineering, was quite impressive when compared to other galactose fermentation results by engineered yeast in the literature, although a laboratory strain background and minimal media conditions were used in this study. We expect that key fermentation parameters can be further improved by employing an industrial strain as a host and optimizing process conditions. The utility of the evolved strain could be extended to produce other fermentation products such as lactic acid or 2,3-butanediol from galactose by introducing the required pathways with additional metabolic engineering. The point mutation in *YHP1*, which was identified by genome sequencing, did not have a significant effect on the phenotype of the evolved strain. We could not completely reproduce the phenotype of JQ-G1 by reverse engineering (although the BY4742 Δcox9 Δgal80 double knockout strain was quite similar) so additional work will be needed to determine other genetic or epigenetic factors that may play a role in this strain.

### 3.5 Figures and tables
Fig. 3.1 Strategy for fermentative evolution involves systematic deletion of a respiratory enzyme subunit and subsequent evolution of the respiration-deficient knockout strain by serial subculture in selective conditions. The strategy was applied for engineering *S. cerevisiae* towards optimal galactose fermentation. Galactose fermentation profiles by the wild-type strain BY4742 (a), knockout strain BY4742 ∆*cox9* (b), and evolved knockout strain JQ-G1 (c). Symbols: galactose (blue triangle), ethanol (red square), cell growth (open circle), glycerol (orange diamond), and acetate (green triangle).
Fig. 3.2 Fermentation physiology of the evolved mutant strain JQ-G1 compared with reference BY4742 in glucose, fructose, mannose, sucrose, and galactose. The data are shown as: (a) specific sugar uptake rate (mmol/g DW/hr) and (b) specific ethanol production rate (mmol/g DW/hr). *, P < 0.05.
Fig. 3.3 Plot of specific ethanol production rate versus specific galactose uptake rate for evolved mutant strain JQ-G1 (light green) compared with the parental strain BY4742 (dark blue) and all engineered strains prepared in this study – BY4742 Δcox9 (red), BY4742 Δyhp1 (light blue), BY4742 Δgal80 (dark green), BY4742 Δgal80Δyhp1 (brown), BY4742 Δcox9Δgal80 (orange), BY4742 Δcox9Δgal80Δyhp1 (purple), and JQ-G1 with re-integration of COX9 (yellow). Dashed red arrows indicate steps for fermentative evolution.
Fig. 3.4 Metabolic flux calculation by $^{13}$C-based flux analysis approach for BY4742 (blue) and JQ-G1 (red) strains during exponential growth on galactose. The galactose uptake flux is normalized to 100 units for both wild-type and mutant strains. The green arrows indicate metabolites that are used for cell growth in the biomass equation.
Fig. 3.5 Intracellular metabolite profile from wild-type strain BY4742 and mutant strain JQ-G1 during exponential growth on galactose in triplicates. Red color indicates higher metabolite concentration and green color indicates lower metabolite concentration as compared to the average across all samples.
Fig. 3.6 Evolution of the BY4742 Δcox9 strain by serial subculture on various carbon sources. Changes in the ethanol yield (blue) and sugar uptake rate (green) are shown during enrichment on galactose (a), fructose (b), or mannose (c). Fermentation parameters were calculated at approximately 48 hrs (galactose) or 43 hrs (fructose and mannose) from inoculation in fresh media. All subcultures were performed in oxygen-limited conditions (100 rpm) with minimal media and 40 g/L sugar. The initial cell density was adjusted to 0.03 g/L. For comparison, the fermentation performance of the parental strain BY4742 is shown at the left of each graph under similar conditions.
Fig. 3.6 (cont.)
Fig. 3.7 High initial cell density ($\text{OD}_{600} = 5$) fermentation by parental strain BY4742 (a) and evolved strain JQ-G1 (b) in minimal media with 40 g/L galactose.
Fig. 3.8 Effect of mutant *YHP1* on galactose fermentation by *S. cerevisiae*. (a) Strategy for evaluating effect of mutant *YHP1* from evolved JQ-G1. Plasmids were sequenced after construction to confirm presence of point mutation (199G>C) in pRS413-m*YHP1*. (b) Ethanol yield (g/g) during flask fermentation by engineered *S. cerevisiae* strain BY4742 ∆*cox9*∆*gal80*∆*yhp1* expressing one of the following: (i) empty vector pRS413, (ii) pRS413 with wild-type *YHP1* insert, or (iii) pRS413 with mutant *YHP1* insert from JQ-G1. Results are the average of three biological replicates.
Fig. 3.9 Results of anaerobic galactose fermentation by wild-type strain BY4742, knockout strains BY4742 Δgal80 and BY4742 Δcox9Δgal80, and evolved strain JQ-G1. Cells were inoculated at initial OD=1 into a sealed 100 ml vial containing 20 ml SC medium and 40 g/L galactose as the sole carbon source. The vials were purged with nitrogen and incubated at 30°C and 100 rpm. After 27 hours cultivation, samples of the culture broth were taken for HPLC analysis and measurement of cell growth. Bars show consumed galactose and ethanol production (a) and glycerol production and final cell density (b) for all strains.
Fig. 3.10 Fermentation profiles for BY4742 (blue diamonds) and JQ-G1 (red triangles) in minimal media with 10 g/L galactose and very low initial cell density (OD$_{600}$ ~ 0.03). For both metabolite profiling and $^{13}$C-metabolic flux analysis experiments, cells were grown in these conditions and harvested during mid-exponential phase. Data is shown for: (a) galactose consumption, (b) ethanol production, and (c) cell growth.
Fig. 3.10 (cont.)
Fig. 3.11 Metabolic flux calculation by $^{13}$C-MFA and the standard deviation associated with each flux value. The galactose uptake flux is normalized to 100 units for both wild-type and mutant strains.
<table>
<thead>
<tr>
<th>Reaction</th>
<th>BY4742</th>
<th>JQ-G1</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAA_mit(abcd)+ACCOA_mit(ef) ==&gt; Cit_mit(dcbfea)</td>
<td>63.7 0.9</td>
<td>23.6 0.3</td>
</tr>
<tr>
<td>Cit_mit(abcd)(def) ==&gt; ICTT_mitabcdef</td>
<td>63.7 0.9</td>
<td>23.6 0.3</td>
</tr>
<tr>
<td>ICTT_mitabcdef ==&gt; AKG_mitabcde + CO2(f) + NADPH</td>
<td>63.7 0.9</td>
<td>23.6 0.3</td>
</tr>
<tr>
<td>AKG_mitabcde ==&gt; SucCoA_mitbcde + CO2(a) + NADH</td>
<td>63.6 1.0</td>
<td>23.4 0.3</td>
</tr>
<tr>
<td>SucCoA_mitabcd ==&gt; Suc_mit[0.5]abcd + [0.5]dcba + ATP</td>
<td>63.6 1.0</td>
<td>23.4 0.3</td>
</tr>
<tr>
<td>Suc_mit[0.5]abcd + [0.5]dcba ==&gt; Fum_mit[0.5]abcd + [0.5]dcba + FADH2</td>
<td>63.6 1.0</td>
<td>23.4 0.3</td>
</tr>
<tr>
<td>Fum_mit[0.5]abcd + [0.5]dcba ==&gt; Mal_mitabcd</td>
<td>62.9 1.4</td>
<td>20.7 0.5</td>
</tr>
<tr>
<td>Mal_mitabcd ==&gt; OAA_mitabcd + NADH</td>
<td>62.9 1.4</td>
<td>20.7 0.5</td>
</tr>
<tr>
<td>Pyr_mitabc + CO2(d) + NADPH</td>
<td>0.7 0.5</td>
<td>2.7 0.3</td>
</tr>
<tr>
<td>Acetaldehyde(ab) + NADH ==&gt; EtOH(ab)</td>
<td>130.2 0.4</td>
<td>152.5 1.3</td>
</tr>
<tr>
<td>G3PG(abc) + NADH ==&gt; Glycerol(abc)</td>
<td>4.4 0.2</td>
<td>21.8 0.0</td>
</tr>
<tr>
<td>EOH(ab) ==&gt; EtOH_ext(ab)</td>
<td>102 0.0</td>
<td>134.9 0.0</td>
</tr>
<tr>
<td>Acetate(ab) ==&gt; Acetate_ext(ab)</td>
<td>9.5 0.0</td>
<td>4.0 0.0</td>
</tr>
<tr>
<td>Glycerol(abc) ==&gt; Glycerol_ext(abc)</td>
<td>4.4 0.0</td>
<td>21.8 0.0</td>
</tr>
<tr>
<td>R5P(abcd) + Cl(f) ==&gt; His(edcbaf)</td>
<td>1.3 0.2</td>
<td>4.2 0.5</td>
</tr>
<tr>
<td>PEP(abc) + PEP(def) + E4P(ghij) ==&gt; Phe(abcdefgij)</td>
<td>0.5 0.0</td>
<td>0.3 0.0</td>
</tr>
<tr>
<td>AKG_mitabcde ==&gt; Glu(abcd)</td>
<td>0.1 0.1</td>
<td>0.1 0.1</td>
</tr>
<tr>
<td>OAA_cytabcd ==&gt; Asp(abcd)</td>
<td>0.5 0.1</td>
<td>0.2 0.0</td>
</tr>
<tr>
<td>G3PG(abc) ==&gt; Ser(abc) + NADH</td>
<td>4.3 0.2</td>
<td>0.7 0.3</td>
</tr>
<tr>
<td>Ser(abc) + NADH ==&gt; GLY(ab) + Cl(c)</td>
<td>3.9 0.2</td>
<td>0.4 0.3</td>
</tr>
<tr>
<td>GLY(ab) ==&gt; Cl(b) + CO2(a) + NADH</td>
<td>3.1 0.2</td>
<td>-0.3 0.3</td>
</tr>
<tr>
<td>Pyr_mitabc ==&gt; Ala(abc)</td>
<td>0.4 0.1</td>
<td>0.5 0.1</td>
</tr>
<tr>
<td>Pyr_mitabc + Pyr_mitdef + NADPH ==&gt; Val(abcdef)</td>
<td>0.3 0.1</td>
<td>0.9 0.1</td>
</tr>
<tr>
<td>OAA_cytabcd ==&gt; OAA_ext(ab)</td>
<td>9.8 0.0</td>
<td>3.5 0.0</td>
</tr>
</tbody>
</table>

Fig. 3.11 (cont.)
Fig. 3.12 Measured and simulated isotopomer labeling patterns of protein-bound amino acids in wild-type strain BY4742 (a) and evolved strain JQ-G1 (b) during exponential growth on galactose.
Table 3.1 Strains and plasmids used in this study.

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
</tr>
<tr>
<td>BY4742</td>
<td>MATalpha, ura3, his3, leu2, lys2</td>
</tr>
<tr>
<td>BY4742 Δyhp1</td>
<td>BY4742 Δyhp1::LEU2</td>
</tr>
<tr>
<td>BY4742 Δcox9</td>
<td>BY4742 Δcox9::KanMX4</td>
</tr>
<tr>
<td>BY4742 Δgal80</td>
<td>BY4742 Δgal80::URA3</td>
</tr>
<tr>
<td>BY4742 Δgal80Δyhp1</td>
<td>BY4742 Δgal80::URA3 Δyhp1::LEU2</td>
</tr>
<tr>
<td>BY4742 Δcox9Δgal80</td>
<td>BY4742 Δcox9::KanMX4 Δgal80::URA3</td>
</tr>
<tr>
<td>BY4742 Δcox9Δgal80Δyhp1</td>
<td>BY4742 Δcox9::KanMX4 Δgal80::URA3 Δyhp1::LEU2</td>
</tr>
<tr>
<td>JQ-G1</td>
<td>BY4742 Δcox9 evolved</td>
</tr>
<tr>
<td>JQ-G1-COX9</td>
<td>JQ-G1 cox9::COX9</td>
</tr>
<tr>
<td>B-CGY control</td>
<td>BY4742 Δcox9Δgal80Δyhp1 pRS413</td>
</tr>
<tr>
<td>B-CGY-YHP1</td>
<td>BY4742 Δcox9Δgal80Δyhp1 pRS413-YHP1</td>
</tr>
<tr>
<td>B-CGY-mYHP1</td>
<td>BY4742 Δcox9Δgal80Δyhp1 pRS413mYHP1</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
</tr>
<tr>
<td>pRS413</td>
<td>HIS3 marker, a single copy plasmid</td>
</tr>
<tr>
<td>pRS413-YHP1</td>
<td>Wild-type YHP1 inserted in pRS413</td>
</tr>
<tr>
<td>pRS413-mYHP1</td>
<td>Mutant YHP1 from JQ-G1 inserted in pRS413</td>
</tr>
</tbody>
</table>
Table 3.2 Primers used in this study.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAL80-f</td>
<td>TCTCGATAGTTGGTTCCCGTTTTTCCACTCCCGT CATGAGATTGA CAGAGAGTG CAC</td>
<td>GAL80 deletion cassette - forward</td>
</tr>
<tr>
<td>GAL80-r</td>
<td>TCGCTGCACTGGGGGCAACAGGCAAGGCAAGAT</td>
<td>GAL80 deletion cassette - reverse</td>
</tr>
<tr>
<td>G1</td>
<td>CGATGTCCTCTTTAAAATGCGCAAG</td>
<td>GAL80 upstream (confirmation)</td>
</tr>
<tr>
<td>G2</td>
<td>TTAGTTTTGTGGCCGCATCCTTCTTG</td>
<td>URA3 reverse (confirmation)</td>
</tr>
<tr>
<td>YHP1-f</td>
<td>GAAATAACCAATAACAAATAATATCGTATATATATAAC ATGAGATTGTACTGAGAGTG CAC</td>
<td>YHP1 deletion cassette – forward</td>
</tr>
<tr>
<td>YHP1-r</td>
<td>TAGCATCAGTGCTCAAGAAAATAACACTGAGAAA GATTACTGTCGGGTATTCACACCG</td>
<td>YHP1 deletion cassette - reverse</td>
</tr>
<tr>
<td>Y1</td>
<td>CCGGAAATGGTAAACATCGAACGC</td>
<td>YHP1 upstream (confirmation)</td>
</tr>
<tr>
<td>Y2</td>
<td>CCAACCATTGCTATTGGTCTTGG</td>
<td>LEU2 reverse (confirmation)</td>
</tr>
<tr>
<td>Y3</td>
<td>CCAAGACCATTGCTAGAAGAAGAGC</td>
<td>YHP1 reverse (confirmation)</td>
</tr>
<tr>
<td>mYHP1-f</td>
<td>GCCggatccAAAACCGTTAATATGAAAGAACC GTCC</td>
<td>Cloning mutant YHP1 – forward</td>
</tr>
<tr>
<td>mYHP1-r</td>
<td>GCCetcgcAGCTTTGTGTTTCACACCGTC</td>
<td>Cloning mutant YHP1 - reverse</td>
</tr>
</tbody>
</table>

Note: nucleotide sequences in lowercase letters represent restriction enzyme recognition sites.
Table 3.3 List of variants common to evolved strains JQ-G1 and JQ-G2 that were not present in starting strain BY4742 Δcox9.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Region</th>
<th>Type</th>
<th>Reference</th>
<th>Allele</th>
<th>Coding region</th>
<th>Amino acid change</th>
<th>Non-synonymous</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>715509</td>
<td>SNV</td>
<td>G</td>
<td>A</td>
<td>YBR248C</td>
<td></td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>1361983</td>
<td>SNV</td>
<td>C</td>
<td>G</td>
<td>YDR451C</td>
<td>Glu67Gln</td>
<td>Yes</td>
<td>YHP1</td>
</tr>
<tr>
<td>X</td>
<td>156712</td>
<td>SNV</td>
<td>G</td>
<td>A</td>
<td>YJL136C</td>
<td></td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>531404</td>
<td>SNV</td>
<td>C</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XII</td>
<td>52619</td>
<td>Del</td>
<td>T</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XII</td>
<td>799073</td>
<td>SNV</td>
<td>G</td>
<td>C</td>
<td>YLR335W</td>
<td></td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>XIII</td>
<td>172635</td>
<td>SNV</td>
<td>G</td>
<td>T</td>
<td>YML051W</td>
<td>Glu348*</td>
<td>Yes</td>
<td>GAL80</td>
</tr>
<tr>
<td>XIV</td>
<td>6637</td>
<td>Del</td>
<td>A</td>
<td>-</td>
<td>YNL338W</td>
<td>His26fs</td>
<td>Yes</td>
<td>Dubious ORF</td>
</tr>
<tr>
<td>XVI</td>
<td>260986</td>
<td>Ins</td>
<td>-</td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.4 Comparison of galactose fermentation parameters by *S. cerevisiae* strains used in this study.

<table>
<thead>
<tr>
<th>Strains</th>
<th>$Y_{\text{EtOH}} \pm \text{SD}$</th>
<th>$P_{\text{EtOH}} \pm \text{SD}$</th>
<th>$X_{\text{Final}} \pm \text{SD}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BY4742</td>
<td>0.197 ± 0.021</td>
<td>1.349 ± 0.245</td>
<td>1.037 ± 0.006</td>
</tr>
<tr>
<td>BY4742 ∆yhp1</td>
<td>0.213 ± 0.017</td>
<td>1.426 ± 0.168</td>
<td>1.011 ± 0.034</td>
</tr>
<tr>
<td>BY4742 ∆cox9</td>
<td>0</td>
<td>0</td>
<td>0.03</td>
</tr>
<tr>
<td>BY4742 ∆gal80</td>
<td>0.289 ± 0.000</td>
<td>3.143 ± 0.082</td>
<td>1.017 ± 0.013</td>
</tr>
<tr>
<td>BY4742 ∆gal80∆yhp1</td>
<td>0.301 ± 0.003</td>
<td>3.227 ± 0.122</td>
<td>1.065 ± 0.021</td>
</tr>
<tr>
<td>BY4742 ∆cox9∆gal80</td>
<td>0.389 ± 0.007</td>
<td>7.548 ± 0.199</td>
<td>0.759 ± 0.004</td>
</tr>
<tr>
<td>BY4742 ∆cox9∆gal80∆yhp1</td>
<td>0.390 ± 0.014</td>
<td>7.457 ± 0.469</td>
<td>0.774 ± 0.008</td>
</tr>
<tr>
<td>JQ-G1</td>
<td>0.384 ± 0.003</td>
<td>9.290 ± 0.153</td>
<td>0.830 ± 0.011</td>
</tr>
<tr>
<td>JQ-G1-COX9</td>
<td>0.293 ± 0.000</td>
<td>4.758 ± 0.072</td>
<td>0.843 ± 0.025</td>
</tr>
</tbody>
</table>

*Fermentation parameters: $Y_{\text{EtOH}}$, ethanol yield (g ethanol/g galactose); $P_{\text{EtOH}}$, specific ethanol productivity (g ethanol/g cells∙hr); $X_{\text{Final}}$, final cell density (g cells/L); SD means standard deviation.
CHAPTER IV  INCREASING XYLOSE FERMENTATION EFFICIENCY BY ENGINEERED RESPIRATION-DEFICIENT SACCHAROMYCES CEREVISIAE
4.1 Introduction

Most eukaryotic organisms that naturally assimilate xylose, such as the yeast *Scheffersomyces stipitis*, make use of the xylose reductase (XR) and xylitol dehydrogenase (XDH) pathway. The XR enzyme converts xylose to intermediate xylitol, which requires NADH or NADPH as a cofactor. Then, XDH converts xylitol to xylulose, and this step regenerates NADH only (Yablochkova et al., 2003). When the XR enzyme uses NADPH as the cofactor, the redox imbalance involved with xylose conversion through the pathway creates a demand for respiration to supply needed NAD$^+$. Therefore, many yeasts cannot ferment xylose in the absence of oxygen because of the redox limitation (Bruinenberg et al., 1983). The widely used industrial yeast *Saccharomyces cerevisiae*, which is well-suited for metabolic engineering studies, cannot metabolize xylose but can assimilate xylulose through the pentose phosphate pathway. Numerous studies have introduced the heterologous XR/XDH pathway into *S. cerevisiae* for developing recombinant xylose-fermenting strains. Still, oxygen availability and respiration are generally important for ethanol production from xylose by engineered *S. cerevisiae* (Jin et al., 2004). Evolutionary engineering has been used to generate mutant *S. cerevisiae* strains that can grow anaerobically on xylose, but the evolved strains still showed a fairly low ethanol yield of $\leq 0.25$ g/g in these conditions and the mutations were not well characterized (Sonderegger and Sauer, 2003).

As mentioned in previous chapters, there are many advantages for engineering ethanol-fermenting yeast strains that can function in the absence of oxygen (i.e. without respiration). Anaerobic conditions are desirable because they prohibit ethanol consumption and oxidative metabolism of sugars, both of which can significantly reduce product yield. Also, it is more economical to operate industrial-scale fermentations without aeration. Respiration-deficient
strains mimic an anaerobic fermentation in any environment and thus have advantages for industrial fermentation processes. In this study, the objective was to engineer an efficient xylose-fermenting yeast strain without respiration, despite the typical challenge of NAD$^+$ shortage associated with the XR/XDH pathway in anaerobic conditions. Improved xylose-fermenting mutants were generated by (i) deletion of a respiratory enzyme subunit (COX9) in an engineered \textit{S. cerevisiae} strain (SR8) expressing the XR/XDH pathway genes, and (ii) adaptive evolution of SR8 $\Delta$cox9 in xylose minimal media. Genome sequencing and a yeast mating experiment were combined to understand the underlying genetic basis (or bases) of the mutant phenotype.

4.2 Materials and methods

4.2.1 Strains and plasmids

The xylose-fermenting \textit{S. cerevisiae} strain SR8 was constructed previously in our laboratory (Kim et al., 2013a) by (i) balanced heterologous expression of \textit{XYL1} (encoding xylose reductase), \textit{XYL2} (encoding xylose dehydrogenase), and \textit{XYL3} (encoding xylulokinase), (ii) laboratory evolution on xylose, and (iii) deletion of \textit{ALD6} (encoding acetaldehyde dehydrogenase). For constructing the SR8 $\Delta$cox9 mutant, a \textit{COX9} deletion cassette with \textit{KanMX} marker was amplified from BY4742 $\Delta$cox9 (Yeast Knockout MATalpha Collection, OpenBiosystems, Lafayette, CO) by PCR with primers COX9-f and COX9-r (Table 2.1). The deletion cassette was integrated into the genome of SR8 by transformation using a high efficiency lithium acetate protocol (Gietz and Schiestl, 2007b) and then positive transformants were selected by G418 resistance on YPD agar plates. For constructing the SR8-m\textit{SPT3} and SR8 $\Delta$cox9-m\textit{SPT3} strains, the CRISPR-Cas system (DiCarlo et al., 2013) was used to introduce a
frameshift mutation (Glu223fs) in SPT3 of the SR8 and SR8 Δcox9 strains, respectively. The guide RNA, which was designed to cut the genomic DNA in the SPT3 region, was cloned into vector pRS42H using restriction enzymes KpnI and SacI. The 90 bp donor DNA was designed to leave a single nucleotide deletion (669delG) in SPT3 and also included silent mutations in the 20-nt guide sequence to avoid repeated cutting by Cas9. Sequences for the guide RNA, donor DNA, and all related primers are listed in Table 4.1. A vector containing a Cas9-encoding gene and clonNAT resistance marker was transformed into the host strains (SR8 and SR8 Δcox9) followed by co-transformation with the donor DNA and the guide RNA vector. Positive transformants with mutant SPT3 were selected on a YPD agar plate with clonNAT and hygromycin. The targeted frameshift mutation in SPT3 (Glu223fs) was confirmed by Sanger sequencing for engineered strains SR8-mSPT3 and SR8 Δcox9-mSPT3; the same mutation was not observed in the SPT3 sequence of control strains SR8 or SR8 Δcox9.

4.2.2 Medium and culture conditions

To prepare yeast strains for inoculation, cells were cultivated overnight at 30°C and 300 RPM in YPD medium (10 g/L yeast extract, 20 g/L Bacto peptone, 20 g/L glucose). Batch fermentation was carried out by resuspending YPD-grown cells in 25 ml (or 10 ml) yeast synthetic complete (SC) medium composed of 1.7 g/L yeast nitrogen base, 5 g/L ammonium sulfate, and 0.79 g/L CSM (MP Biomedicals, Solon, OH) for supplying amino acids and nucleobases. Xylose (or other sugar) was included in the SC media to provide the carbon source. Precultured cells were inoculated to fresh media at initial cell density of 0.03 g/L, unless noted otherwise. Fermentation flasks were incubated at 30°C and 100 RPM in a well-controlled shaking incubator to maintain an oxygen-limited condition during the course of the experiment. At regular intervals, samples of 200 μl were taken to measure the OD600 and metabolite
concentrations. For cloning, *E. coli* was grown overnight in Luria-Bertani medium at 37°C with 50 μg/ml of ampicillin added when necessary. To select yeast transformants with an antibiotic resistance gene marker, YPD agar plates were used containing one or more of the following as appropriate: 200 μg/ml G418, 120 μg/ml clonNAT, and/or 300 μg/ml hygromycin.

4.2.3 Adaptive evolution on xylose

The SR8 Δcox9 mutant was the starting strain for adaptive evolution in xylose minimal media (SCX40). After preculture in YPD, the mutant was able to grow in SCX with much slower initial xylose uptake rate as compared to SR8. When the stationary phase was reached, the cells were transferred to fresh SCX media for adaptive evolution but no growth or xylose consumption was observed in the second subculture for an extended time period. After approximately 300 hours of regular monitoring, the SR8 Δcox9 mutant began to grow and consume xylose more rapidly than during the first subculture with heightened ethanol yield. The culture was sampled and transferred iteratively to fresh media during the stationary phase but further improvements in xylose fermentation parameters were not observed over the course of three subsequent subcultures. Before each serial transfer, multiple xylose-evolved mutants were isolated from the last fermentation flask on agar plates. The SR8 Δcox9 evolved mutants were screened and the best xylose-fermenting strains were named JQ-X1, JQ-X2, and JQ-X3.

4.2.4 Genome sequencing and SNP discovery

The genomic DNA of SR8 Δcox9 and evolved strains JQ-X1, JQ-X2, and JQ-X3 was prepared with the YeaStar Genomic DNA Kit (Zymo Research, Irvine, CA) and quality was confirmed in a 1% agarose gel. Sequencing (SE100) was done using an Illumina HiSeq2000
machine at the W.M. Keck Center for Comparative and Functional Genomics at the University of Illinois at Urbana-Champaign. The barcoded shotgun genomic DNA libraries were constructed with the TruSeq Sample Prep Kit following the manufacturer’s instructions (Illumina, San Diego, CA). The raw data was processed with Casava 1.8, and over 15,000,000 reads were obtained for each sample. Variant analysis (either SNPs or INDELs) was performed using CLC Genomics Workbench version 7.0.4. Reads were trimmed based on quality scores with default program settings and the trimmed reads were mapped to an S288C reference sequence, as described previously (section 3.2.5). The SNPs unique to evolved strains JQ-X1, JQ-X2, and JQ-X3 relative to SR8 Δcox9 were then determined by comparing the SNP files using the "Filter against Known Variants" function in CLC Genomics Workbench. Of the variants common to the evolved mutants, only 2 variants resulting in an amino acid change were unique when compared to the parental strain (Table 4.3).

4.2.5 Yeast mating and tetrad dissection

The haploid strains JQ-X1 (MATalpha) and SR8 (MATa) were mated by crossing on a YPD agar plate and incubating at 30°C for two days. For identifying diploids from the mating plate, fresh cells from the intersection of the two haploid strains were collected and diluted so as to isolate single colonies when spread on a second YPD plate. A halo assay test was used to confirm the presence of diploids from the mating experiment. First, the halo assay involves spreading each of two mating type tester strains (DBY7730 or DBY7442) on a separate YPD plate. Then, the potential diploids are spotted on the lawn of each tester strain, and the plates are incubated overnight. For this test, the lack of a halo around a colony indicates a diploid since it is not mating competent with either of the two tester strains. After identifying diploids by the halo test, the diploid colonies were precultured and the cells were washed with water before
inoculating to sporulation medium (10 g/L potassium acetate, 1 g/L yeast extract, 0.5 g/L glucose) for ~7 days to generate tetrads. When tetrads were observed under a microscope, the cells were treated with Zymolase and β-glucuronidase to break down the ascus coat and then streaked on a YPD plate. Four ascospores were separated from each tetrad using a microneedle and the spores were placed at least 5 mm apart on the plate. Plates were incubated at 30°C and all spores were confirmed to be haploid by repeating the halo assay test.

4.2.6 Analytical methods

Cell concentration was measured by optical density (OD) at 600 nm using a UV-visible spectrophotometer (Biomate 3, Thermoscientific, Madison, WI). The concentration of sugar, ethanol, glycerol, and acetate in batch fermentations was determined by a high-performance liquid chromatography system (Agilent Technologies 1200 series) with a refractive index detector (RID) and a Rezex ROA-Organic Acid H⁺ column (Phenomenex Inc., Torrance, CA). The column temperature was kept constant at 50°C and the elution was performed with 0.005 N H₂SO₄ at a flow rate of 0.6 ml/min.

4.3 Results and discussion

4.3.1 Characterization of evolved mutants

Despite the inherent connection between heightened respiratory activity and efficient xylose utilization by engineered S. cerevisiae (Jin et al., 2004), we found that an engineered xylose-fermenting strain with respiration deficiency (SR8 ∆cox9) was able to consume xylose more rapidly with a substantial improvement in ethanol yield after serial subcultures in xylose minimal media. Similar to the adaptive evolution on galactose by a respiration-deficient strain
(Section 3.3.1), the change in xylose uptake rate and ethanol yield by SR8 \( \Delta \text{cox9} \) during the serial subculture was quite abrupt with little or no further improvement after the second transfer. During the evolution process, many xylose-evolved mutants were isolated and screened from each fermentation flask, and the best ethanol-producing strains were named JQ-X1, JQ-X2, and JQ-X3. The evolved strains were superior to SR8 in terms of both ethanol yield (32% improvement) and volumetric productivity (37% improvement) during batch fermentation in xylose minimal media. Fig. 4.1 shows a complete fermentation profile by SR8 (a), knockout SR8 \( \Delta \text{cox9} \) (b), and evolved knockout JQ-X1 (c) in minimal media with 40 g/L xylose. One interesting characteristic of the evolved JQ-X1 strain is a much higher xylitol yield (0.18 ± 0.01 g/g) than SR8 (0.01 ± 0.00 g/g), which may be related to a shortage of NAD\(^+\) for driving the xylitol dehydrogenase reaction; this is likely caused by lack of respiration in the JQ-X1 strain. While xylitol production is often considered to be an undesirable characteristic of xylose-fermenting strains, we expected that the NADH surplus in the JQ-X1 strain would be useful for further metabolic engineering (see Chapter V). Despite conversion of more xylose to xylitol, the evolved mutant still fermented ethanol faster and more efficiently than either SR8 or SR8 \( \Delta \text{cox9} \). When calculated on the basis of xylose assimilated, the ethanol yield by JQ-X1 (0.35 ± 0.02 g/g xylose assimilated) far exceeded that of the parental strain SR8 (0.22 ± 0.01 g/g xylose assimilated). Table 4.2 describes some other relevant fermentation parameters for the engineered and evolved strains in this study.

4.3.2 Changes in genotype of the evolved xylose-fermenting strains

The entire genome of the evolved mutants JQ-X1, JQ-X2, and JQ-X3, as well as parental strain SR8 \( \Delta \text{cox9} \), was sequenced and compared in order to identify genetic perturbations
associated with improved xylose metabolism without respiration. Any SNP observed in the evolved strains but not present in the control strain may contribute to our beneficial phenotype. After eliminating silent substitutions or genetic mutations in non-coding regions, there were only two remaining SNPs shared in all evolved strains: a point mutation in MYO4 (Leu37Phe) and a frameshift mutation in SPT3 (Glu223fs). MYO4 encodes a myosin motor, which is required for transport of mRNA and other cargo to polarized regions of the cell (Haarer et al., 1994; Pruyne et al., 2004). SPT3 is known to encode a subunit of SAGA, which is a transcriptional regulatory complex in yeast (Grant et al., 1997). It is noteworthy that Spt3p interacts with Spt15p for activating transcription, and a previous study found that three separate mutations in SPT15 improved ethanol production from glucose and ethanol tolerance by S. cerevisiae (Alper et al., 2006).

4.3.3 Evaluation of spores from yeast mating and tetrad dissection

In order to study the relationship between genotype and phenotype in the evolved JQ-X1 strain, a yeast mating experiment was conducted (Fig. 4.2a). After mating of the evolved strain with parental strain SR8, diploids were isolated and inoculated into sporulation media to generate tetrads. Twenty tetrads were dissected under microscope using a microneedle and plated on an agar plate. The plates were incubated for 2-3 days and thirteen of the twenty tetrads grew a complete family of four spores. To identify the spores in each family with or without COX9, the thirteen families (52 total spores) were plated on YPD with antibiotic G418. Since COX9 was deleted by replacement with a kanMX marker, only the COX9-negative spores should be able to grow on G418. After incubation of the plates, we found that two of the four spores in each family were COX9-negative and thus half of the spores in a family are respiration-deficient. Also, the mating type of each spore was determined by plating all 52 spores on YPD spread with
one of the mating type tester strains (A or alpha). As expected, each of the thirteen families contained two mating type A and two mating type alpha spores (Fig. 4.2b). The phenotype of three families (12 total spores) was evaluated by fermentation in xylose minimal media, and the ethanol yield was calculated at 102 hours (Fig. 4.2c). The fermentation results led to two very interesting observations. First, the average ethanol yield for the COX9-negative spores (0.36 ± 0.03 g/g) was significantly higher than that for the COX9-positive spores (0.20 ± 0.04 g/g) in the three families tested. Second, the two COX9-negative spores (or the two COX9-positive spores) within a family did not show the same phenotype but in every case one outperformed the other in terms of ethanol yield. Based on the genome sequencing results for JQ-X1, we expected that a mutation in SPT3 and/or MYO4 (Table 4.3) may play a role in the improved fermentation performance by certain spores. Although sequencing of MYO4 from the spores did not reveal any interesting associations, we found the SPT3 frameshift mutation (Glu223fs) in all of the respiration-deficient spores with better phenotype in three families. Of the respiration-competent spores, two out of three with better phenotype in a family had the SPT3 mutation (Fig. 4.3). The results led to the hypothesis that the SPT3 frameshift mutation was indeed an important genetic perturbation for improving xylose fermentation by a respiration-deficient xylose-utilizing strain.

4.3.4 Effect of mutant SPT3 on SR8 and SR8 Δcox9 phenotype

In order to clearly determine the effect of mutant SPT3 on xylose fermentation with or without respiration, we constructed the engineered strains SR8-mSPT3 and SR8 Δcox9-mSPT3 by using the CRISPR-Cas9 genome editing system. After strain construction, the phenotypes were evaluated during flask fermentation in xylose minimal media. For the respiration-deficient strains (SR8 Δcox9 and SR8 Δcox9-mSPT3), xylose fermentation results are shown in Fig. 4.4.
along with results for respiration-deficient evolved strain JQ-X1. In the COX9-negative background, mutant SPT3 led to a 25% improvement in xylose uptake rate and a 22% improvement in ethanol productivity as compared to parental strain SR8 Δcox9. In comparison, evolved strain JQ-X1 showed a similar 24% improvement in xylose uptake rate and a 31% improvement in ethanol productivity as compared to the parental strain (SR8 Δcox9) in the same conditions. The SR8 Δcox9-mSPT3 strain also demonstrated a 3.9-fold increase in xylitol production and a 54% reduction in glycerol production as compared to SR8 Δcox9; this result was similar to the 3.3-fold increase in xylitol production and 46% reduction in glycerol production by JQ-X1 versus the SR8 Δcox9 strain. In summary, the results demonstrate that the frameshift mutation in SPT3 (Glu223fs), which was identified by genome sequencing of evolved strain JQ-X1, is an important factor for improving ethanol fermentation from xylose in a respiration-deficient XR/XDH-expressing strain; however, this mutation is also associated with heightened xylitol production. For the COX9-positive strains (SR8 and SR8-mSPT3), the mutant SPT3 led to a substantial decrease in xylose uptake rate, cell growth, and ethanol productivity as compared to the control (SR8). The SPT3 mutation was clearly detrimental for xylose fermentation when respiration was able to proceed normally.
Fig. 4.1 Comparison of fermentation profiles in xylose (40 g/L) minimal media by engineered and evolved *S. cerevisiae* strains (a) SR8, (b) SR8 Δcox9, and (c) JQ-X1 (SR8 Δcox9 evolved).
Fig. 4.2 (a) Xylose-fermenting strains SR8 and JQ-X1 were mated to obtain daughter cells for studying genotype-phenotype relationships in the parental strains. (b) In order to discern mating type and genotypic information about the daughter cells, the spores were plated on YPD plates with (clockwise from top left) no addition, antibiotic G418, mating type A tester, or mating type alpha tester. (c) The phenotype of the spores was evaluated by fermentation in xylose minimal media and ethanol yield was calculated for each spore.
Fig. 4.2 (cont.)
Fig. 4.2 (cont.)
Fig. 4.3 For three families of spores generated by mating of SR8 and JQ-X1 followed by tetrad dissection, the *SPT3* region was amplified by PCR and then sequenced by Sanger sequencing. The table organizes the spores by genotype (*COX9*-positive or *COX9*-negative) and phenotype (better or worse in terms of ethanol yield) and also identifies where the frameshift mutation in *SPT3* was found (green box).

<table>
<thead>
<tr>
<th></th>
<th>High Ethanol Yield</th>
<th>Low Ethanol Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>COX9-</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JQ-X1</td>
<td>9a</td>
<td>SR8 Δcox9</td>
</tr>
<tr>
<td></td>
<td>3a</td>
<td>9b</td>
</tr>
<tr>
<td></td>
<td>2d</td>
<td>3c</td>
</tr>
<tr>
<td><strong>COX9+</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9c</td>
<td>9d</td>
</tr>
<tr>
<td></td>
<td>3b</td>
<td>3d</td>
</tr>
<tr>
<td></td>
<td>2c</td>
<td>2b</td>
</tr>
</tbody>
</table>

= *SPT3* frameshift mutation (669delG)
Fig. 4.4 Effect of mutant SPT3 in XR/XDH-expressing *S. cerevisiae* strains with respiration deficiency (SR8 Δ*cox9* m*SPT3* and JQ-X1) as compared to the control (SR8 Δ*cox9*). Fermentation was conducted in 25 ml of xylose minimal media under oxygen-limited conditions with initial cell density ~0.3 g/L (OD<sub>600</sub>=1.0). Bars show consumed xylose (a) and production of ethanol, xylitol, and glycerol (b) for all strains after 23 hours of cultivation.
Table 4.1 Primers used in this study.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guide RNA</td>
<td>TCTACAGCGGCCGCGagctcTCTTTGAAAGATAAT GTAAGTTAGCTTTTCACCTCATATTATACAGAAACCTTGAT GTTTTCTTTCGAGTATATAACGTTGATTACATGTA GTTTGAAAGTAAAACTCTAGATTTTGTAGTGCCCTC TT GGGCTAGCGTAAAGGTGGCGATGTTGCGCCGCTCG AAACCTTCCCGAGTAGAAATAGATCGAGATTTT CGCAATTAACCTGAGTTTTTAGAGCTAGAATAGC A GTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCAGTCTCGGTGTTGCTTTTTTTGTCTTTTATGTCTtgaccGCCGCGGCTCTATA</td>
<td>Designed for cutting genomic DNA in the SPT3 region</td>
</tr>
<tr>
<td>Guide-f</td>
<td>TCTACAGCGGCCGCGAGCTCTCT</td>
<td>Primers for amplifying guide RNA</td>
</tr>
<tr>
<td>Guide-r</td>
<td>TATAGACGCCGCGGGTACCAAG</td>
<td></td>
</tr>
<tr>
<td>Donor_F</td>
<td>GAAATAAAAGGTCTTTCAAGGACTGGTCTGGAATTTCGCAATTGACAGAGGAAAGCCCATGA</td>
<td>60 bp primers with 30 bp homology for amplifying donor DNA</td>
</tr>
<tr>
<td>Donor_R</td>
<td>GGGCAGCTTTTCTCTGACCTACATCATCGGGGTTTCCCTCTGACATTTGCGAAGT</td>
<td></td>
</tr>
<tr>
<td>SPT3-f</td>
<td>TGGGCAGTGGCAGCTGGAATG</td>
<td>Primers for Sanger sequencing of SPT3</td>
</tr>
<tr>
<td>SPT3-r</td>
<td>GTCAAAGGCCCTATGCTCATGTC</td>
<td></td>
</tr>
</tbody>
</table>

Note: nucleotide sequences in lowercase letters represent restriction enzyme recognition sites; red letters denote the 20 nt guide sequence for Cas9
Table 4.2 Comparison of xylose fermentation parameters by engineered and evolved *S. cerevisiae* strains in minimal media with 40 g/L xylose. Cells were inoculated at initial cell density ~0.03 g/L (OD$_{600}$ = 0.1).

<table>
<thead>
<tr>
<th>Fermentation parameters</th>
<th>Strains</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SR8</td>
</tr>
<tr>
<td>Consumed xylose (g/L)</td>
<td>32.23 ± 0.49</td>
</tr>
<tr>
<td>Ethanol titer (g/L)</td>
<td>7.04 ± 0.60</td>
</tr>
<tr>
<td>( Y_{\text{Ethanol}}^* )</td>
<td>0.22 ± 0.01</td>
</tr>
<tr>
<td>Xylitol titer (g/L)</td>
<td>0.19 ± 0.07</td>
</tr>
<tr>
<td>( Y_{\text{Xylitol}} )</td>
<td>0.01 ± 0.00</td>
</tr>
<tr>
<td>Glycerol titer (g/L)</td>
<td>2.42 ± 0.42</td>
</tr>
<tr>
<td>( X_{\text{Final}} )</td>
<td>1.63 ± 0.00</td>
</tr>
</tbody>
</table>

*Fermentation parameters: \( Y_{\text{Ethanol}}^* \), ethanol yield (g ethanol/g xylose assimilated); \( Y_{\text{Xylitol}} \), xylitol yield (g xylitol/g xylose); \( X_{\text{Final}} \), final cell density (g cells/L); all parameters were calculated at 167 hours.
Table 4.3 List of variants common to evolved strains JQ-X1, JQ-X2, and JQ-X3 that cause an amino acid change and were not present in starting strain SR8 Δcox9.

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Region</th>
<th>Type</th>
<th>Reference</th>
<th>Allele</th>
<th>Coding region</th>
<th>Amino acid change</th>
<th>Non-synonymous</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>92160</td>
<td>SNV</td>
<td>G</td>
<td>T</td>
<td>YAL029C</td>
<td>Leu37Phe</td>
<td>Yes</td>
<td>MYO4</td>
</tr>
<tr>
<td>IV</td>
<td>1259364</td>
<td>Del</td>
<td>G</td>
<td>-</td>
<td>YDR392W</td>
<td>Glu223fs</td>
<td>Yes</td>
<td>SPT3</td>
</tr>
</tbody>
</table>
CHAPTER V TWO REDOX BALANCING STRATEGIES FOR IMPROVING XYLOSE FERMENTATION BY ENGINEERED SACCHAROMYCES CEREVISIAE
5.1 Introduction

Redox imbalance is known to be a major challenge for efficient xylose fermentation in recombinant *Saccharomyces cerevisiae* transformed with the xylose reductase (XR) and xylitol dehydrogenase (XDH) genes from *Scheffersomyces stipitis*. The heterologous pathway is generally associated with xylitol production, especially in anaerobic or microaerobic conditions, due to different coenzyme specificity between XR and XDH. Still, some recent reports have noted advantages of the XR/XDH pathway as compared to the redox-neutral xylose isomerase (XI) pathway from *Piromyces* (Bettiga et al., 2008; Karhumaa et al., 2007). In order to alleviate the redox imbalance problem and improve ethanol production from xylose by engineered *S. cerevisiae*, expression of a mutant XR with preference for cofactor NADH (rather than NADPH) has been demonstrated as a useful strategy. For example, one study used protein engineering to generate a mutant XR with a single amino acid change (Arg276His) for improving the ratio of specific activity with NADH to specific activity with NADPH by 46-fold. Expression of the mutant XR with a wild-type XDH gene from *S. stipitis* led to a 20% increase in ethanol production and 52% reduction in xylitol as compared to a similar strain with wild-type XR (Watanabe et al., 2007b). Alternatively, expression of a mutant XDH with preference for cofactor NADP⁺ (rather than NAD⁺) has also been shown to be effective for improving ethanol production by engineered xylose-fermenting *S. cerevisiae* (Matsushika et al., 2008; Watanabe et al., 2007a).

More recently, a rational metabolic engineering strategy led to an acetate and xylose co-consuming *S. cerevisiae* strain that makes use of the NADH-producing xylose utilization pathway by introducing an NADH-consuming acetate reduction pathway (Wei et al., 2013). Acetic acid is a necessary but problematic component of cellulosic hydrolysates that results from
acetylation of hemicellulose and lignin in the plant cell wall (Klinke et al., 2004). Acetate assimilation can enhance microbial fermentation in two specific ways: (i) by improving fermentation yields due to conversion of an additional carbon source, and (ii) by detoxifying the media through lowering the concentration of an inhibitory compound. Native S. cerevisiae strains possess an acetyl coenzyme A (acetyl-CoA) synthetase for conversion of acetic acid to acetyl-CoA. Expression of NADH-dependent acetylating acetaldehyde dehydrogenase (AADH), which is not present in wild-type S. cerevisiae, was shown to complete the acetate reduction pathway by bioconversion of acetyl-CoA to acetaldehyde. Subsequently, the native alcohol dehydrogenase catalyzes the reduction of acetaldehyde to ethanol. Operation of this pathway by engineered yeast led to a 6% increase in ethanol yield and 11% reduction in by-product yield during anaerobic fermentation of mixed sugars (Wei et al., 2013). Although a recent report claimed superiority of the redox-neutral xylose isomerase pathway (Lee et al., 2014), the potential for acetate reduction due to surplus NADH generated in the XR/XDH pathway may offer advantages for cellulosic biofuel production. The acetate and xylose co-consumption strategy has turned a significant challenge of the XR/XDH pathway into a valuable asset.

Previously, we engineered a respiration-deficient xylose-fermenting yeast strain by deletion of an electron transport chain enzyme subunit and evolutionary engineering (see Chapter IV). While the evolved JQ-X1 strain showed significant improvement in ethanol production as compared to parental strain SR8 in xylose minimal media, we also reported high xylitol yield by the mutant due to redox imbalance. Excess xylitol production in the JQ-X1 strain indicates a shortage of NAD$^+$ for driving the xylitol dehydrogenase reaction, and this is likely related to respiration deficiency. Thus, we hypothesized that two aforementioned redox balancing strategies, namely introduction of an NADH-dependent acetate reduction pathway or
introduction of an NADH-preferring xylose reductase, could supply more NAD³⁺ for reducing xylitol production and further improving ethanol production in the respiration-deficient strain. In this study, a heterologous gene (adhE) encoding acetylating acetaldehyde dehydrogenase (AADH) was transformed into JQ-X1; the results demonstrate operation of the acetate assimilation pathway and improved ethanol production by the adhE-expressing strain during anaerobic fermentation of xylose and acetate. Separately, the mutant S. stipitis gene encoding an XR with preference for cofactor NADH (Arg276His) was also transformed into the JQ-X1 strain on a multi-copy plasmid. However, the results did not show improvement in ethanol production or reduction in xylitol by the transformants.

5.2 Materials and methods

5.2.1 Strains and plasmids

The evolved respiration-deficient S. cerevisiae strain JQ-X1 (see chapter IV) was used in this study for expressing mutant xylose reductase and the acetate reduction pathway. To introduce an acetylating acetaldehyde dehydrogenase (AADH) into the JQ-X1 strain, we used an E. coli adhE (NCBI Gene ID: 945837) that was codon-optimized for S. cerevisiae. The codon-optimized adhE was cloned into integration plasmid pITy3 with PGK promoter and terminator; the gene for resistance to antibiotic hygromycin was also cloned into the plasmid. After cutting with restriction enzyme XhoI to linearize, the pITy3-co-adhE plasmid was then transformed into JQ-X1 by a high efficiency lithium acetate procedure (Gietz and Schiestl, 2007a) for integration at the delta sequences. Positive transformants were selected on a YPD agar plate with hygromycin and colony PCR was used to confirm the presence of adhE. To increase enzymatic
activity in the JQ-X1-\textit{adhE} strain, the CRISPR-Cas system (DiCarlo et al., 2013) was used to introduce another copy of codon-optimized \textit{adhE}, resulting in strain JQ-X1-\textit{adhE}-\textit{adhE}. The guide RNA, which was designed to cut the genomic DNA in the \textit{PHO13} region, was cloned into Blasticidin resistance vector OG539 (Oxford Genetics) using restriction enzymes KpnI and SacI. A vector containing clonNAT resistance marker and Cas9-encoding gene was transformed into the host strain (JQ-X1-\textit{adhE}) first followed by co-transformation with donor DNA and the guide RNA vector. Positive transformants were selected on a YPD agar plate containing clonNAT and blasticidin and then colony PCR was used to confirm the presence of \textit{adhE} in the \textit{PHO13} region.

For expressing an NADH-preferring xylose reductase in the JQ-X1 strain, a mutant \textit{S. stipitis XYL1} (Arg276His) was cloned into vector pRS42H with GPD promoter and CYC terminator. The pRS42H-mutXR plasmid was transformed into JQ-X1 by a high-efficiency lithium acetate procedure and positive transformants were selected on a YPD agar plate with hygromycin.

5.2.2 Medium and culture conditions

To prepare yeast strains for inoculation, cells were cultivated overnight at 30°C and 300 RPM in YPD medium (10 g/L yeast extract, 20 g/L Bacto peptone, 20 g/L glucose). Batch fermentation was carried out by resuspending YPD-grown cells in yeast synthetic complete (SC) medium composed of 1.7 g/L yeast nitrogen base, 5 g/L ammonium sulfate, and 0.79 g/L CSM (MP Biomedicals, Solon, OH) for supplying amino acids and nucleobases. Xylose was included in the SC media to provide the carbon source (SCX), and 3 M sodium acetate was also added to the appropriate concentration for xylose/acetate co-consumption experiments (SCXA). Precultured cells were harvested in exponential phase and inoculated to fresh media after
removing used YPD. Fermentation flasks were incubated at 30°C and 100 RPM in a well-controlled shaking incubator (Thermoscientific, MaxQ4000, Dubuque, IA) during the course of the experiment. At regular intervals, samples of 200 μl were taken to measure the OD_{600} and metabolite concentrations. For anaerobic fermentations, ergosterol and Tween80 were included in SCXA medium at a final concentration of 10 mg/L and 420 mg/L, respectively. Anaerobic fermentations were conducted in 20 mL of SCXA (or YPXA) media in a sealed and nitrogen-purged 100 mL vial. To select yeast transformants with an antibiotic resistance gene marker, YPD agar plates were used containing one or more of the following as appropriate: 300 μg/ml hygromycin, 120 μg/ml clonNAT, and/or 100 μg/ml blasticidin. For cloning, *E. coli* was grown overnight in Luria-Bertani medium at 37°C with 50 μg/ml of ampicillin added when necessary.

5.2.3 Enzymatic activity assays

For measuring activity of AADH in engineered yeast strains, cell extracts were obtained from anaerobic xylose fermentation during exponential phase growth. Cells were washed with a cocktail of 1 mM dithiothreitol and protease inhibitor tablet (Roche, catalog #04693116001) and then vortexed with sterile glass beads for 10 min to break the cell walls and release cellular contents. The enzymatic reaction was conducted in a 96-well plate (total volume = 200 μL) by adding 24 μL of acetyl coenzyme A (10.29 mM), 20 μL of NADH (2.82 mM), and 5 μL of protein extract into 151 μL of potassium phosphate buffer (50 mM, pH 7.5). Enzymatic activity was evaluated in triplicates using a microplate reader (BioTek, Winooski, VT) by monitoring NADH oxidation as measured from the change in absorbance at 340 nm. The AADH activity was normalized by total protein concentration in the extract as determined from a Pierce BCA Protein Assay Kit (Thermo Scientific).
To measure XR activity in engineered yeast strains, cell extracts were obtained during the exponential phase of aerobic growth in glucose media (YPD). The media was separated by centrifugation and the cells were resuspended in Y-PER solution (Pierce, Rockford, IL) for 20 min at room temperature to break the cell walls and release cellular contents. After a 10 min centrifugation at 4°C, the crude enzyme extract (supernatant) was obtained and stored on ice. The enzymatic reaction was conducted in a 96-well plate (total volume = 200 μL) by adding 20 μl of xylose (1 M), 20 μl of NADH or NADPH (2 mM), and 20 μl of protein extract into 140 μl of sodium phosphate buffer (50 mM, pH 6.5). Enzymatic activity was evaluated in triplicates using a microplate reader (BioTek, Winooski, VT) by monitoring NADH (or NADPH) oxidation as measured from the change in absorbance at 340 nm. The XR activity was normalized by total protein concentration in the extract as determined from a Pierce BCA Protein Assay Kit (Thermo Scientific).

5.2.4 Analytical methods

Cell concentration was measured by optical density (OD) at 600 nm using a UV-visible spectrophotometer (Biomate 3, ThermoScientific, Madison, WI). The concentration of sugar, ethanol, glycerol, and acetate in batch fermentations was determined by a high-performance liquid chromatography system (Agilent Technologies 1200 series) with a refractive index detector (RID) and a Rezex ROA-Organic Acid H⁺ column (Phenomenex Inc., Torrance, CA). The column temperature was kept constant at 50°C and the elution was performed with 0.005 N H₂SO₄ at a flow rate of 0.6 ml/min.
5.3 Results and discussion

5.3.1 Expression of acetate reduction pathway in JQ-X1 strain

In order to create a demand for surplus NADH generated by the XR/XDH pathway, we introduced a heterologous gene (adhE) encoding NADH-consuming acetylating acetaldehyde dehydrogenase (AADH) to evolved strain JQ-X1 and thus completed the acetate reduction pathway in the resulting JQ-X1-ahdE strain. Due to balancing of redox cofactors and assimilation of a second carbon source, we expected that co-fermentation of xylose and acetate by the adhE-expressing strain would show reduced xylitol production and improved ethanol production as compared to parental strain JQ-X1. To test the hypothesis, the JQ-X1-adhE strain was evaluated during anaerobic fermentation of xylose and acetate in various conditions. In complex media with 2 g/L of acetate and 40 g/L of xylose, JQ-X1-adhE was able to consume 1.39 g/L of acetate in the 222 hours required to deplete greater than 90% of initial xylose (Fig. 5.1b). The control JQ-X1 did not consume acetate and used xylose at a 5.5% lower rate than JQ-X1-adhE (Fig. 5.1a). The relatively slow xylose utilization rate by the engineered strains was due to low initial cell density (~0.03 g/L) in this experiment. Still, we were able to validate our hypothesis since the adhE-expressing strain showed a 15% improvement in ethanol yield and 55% reduction in xylitol yield as compared to the non-acetate consuming strain. To improve sugar uptake rate and fermentation rate, we conducted a similar experiment in complex media (YPXA) with a ten-fold higher initial cell density of 0.3 g/L (Fig. 5.2). With the addition of more cells, JQ-X1-adhE showed a 3.3-fold improvement in xylose uptake rate as compared to the same strain in the low cell density experiment and also consumed 1.15 g/L acetate in 67 hours (Fig. 5.2b). As expected, JQ-X1 was still unable to consume acetate in the higher cell
density experiment (Fig. 5.2a). The acetate-consuming strain produced 49% more ethanol and had a 27% lower xylitol yield versus the control, although final xylitol titer was similar in the two strains. Table 5.1 shows the key fermentation parameters by JQ-X1-adhE and the control during the xylose/acetate co-fermentation experiment. In minimal media with xylose and acetate, the engineered S. cerevisiae strains JQ-X1 and JQ-X1-adhE were able to grow anaerobically but only with a lower initial acetate concentration (~1 g/L) and addition of 50 mM sodium phthalate buffer. In this condition, JQ-X1-adhE consumed all acetate in 78 hours and demonstrated 15% higher ethanol yield than the control with a 20% lower xylitol yield. Clearly, the results indicate that the NADH-consuming acetate reduction pathway can restore redox balance and improve ethanol production in a respiration-deficient xylose-fermenting yeast strain.

To increase activity of AADH, the CRISPR/Cas system was used to introduce another copy of adhE into the JQ-X1-adhE strain. Colony PCR confirmed success of the transformation and enzymatic activity assay showed a significant 38% increase in AADH activity by JQ-X1-adhE-adhE as compared to JQ-X1-adhE (Fig. 5.3). However, the transformants with an additional adhE copy used acetate more slowly than the original JQ-X1-adhE strain and also had lower xylose uptake rate. Metabolic burden and/or some detrimental mutation(s) during transformation may play a role in the unexpected phenotype of the JQ-X1-adhE-adhE strain.

5.3.2 Expression of NADH-preferring xylose reductase in JQ-X1 strain

To alleviate the redox imbalance problem in strain JQ-X1, we also attempted another strategy that involved expression of a mutant XR previously shown to have a preference for cofactor NADH (rather than the NADPH preference of the wild-type XR). Twelve transformants bearing mutant XR on multi-copy plasmid were selected and screened in xylose
minimal media (SCX). The calculated ethanol yield and xylitol yield for all transformants and the control strain (JQ-X1) is shown in Fig. 5.4. Contrary to our hypothesis, the strains with mutant XR had lower ethanol yield in this condition and 11 of 12 also had higher xylitol yield than the control strain. However, in complex media with xylose as the sole carbon source, the strains expressing mutant XR showed more promising results (Fig. 5.5). The main changes in this experiment were use of complex media (YPX), comparison against a proper control strain with empty vector (JQ-X1-pRS42H), and addition of hygromycin to all fermentation media to prevent plasmid loss. In this condition, the results show a 44% improvement in ethanol production and a 22% reduction in xylitol production for the JQ-X1-mXR strain as compared to the control (JQ-X1-pRS42H). Also, the mutant XR strain consumed xylose faster (16% improvement) and generated more biomass (23% improvement) than the control. A xylose reductase enzymatic activity assay confirmed that expression of the mutant XR significantly increased the ratio of NADH-specific activity to NADPH-specific activity in the engineered xylose-fermenting strains (Fig. 5.6). We hypothesize that higher NADH-specific XR activity is the major cause of improved xylose utilization and ethanol production by the JQ-X1-mXR strain.

5.4 Figures and tables
Fig. 5.1 Comparison of xylose/acetate fermentation profiles by engineered *S. cerevisiae* strains

(a) JQ-X1 and (b) JQ-X1-adhE with low initial cell density (OD$_{600}$ = 0.1).
Fig. 5.2 Comparison of xylose/acetate fermentation profiles by engineered *S. cerevisiae* strains (a) JQ-X1 and (b) JQ-X1-adhE with high initial cell density (OD$_{600}$ = 1.0). All results are the average of duplicate experiments. The error bars represent one standard deviation and are not visible when smaller than the symbol size.
Fig. 5.3 Enzymatic activity assay of acetylating acetaldehyde dehydrogenase reaction in two engineered yeast strains during acetate assimilation as compared with control strain JQ-X1. Results are the average of triplicate assays and error bars indicate standard deviation. The student t test was used to establish significant differences between the adhE-expressing strains and the control. *, P < 0.05 (95% confidence)
Fig. 5.4 Calculated ethanol yield and xylitol yield during fermentation in xylose minimal media by twelve transformants harboring a vector with mutant NADH-preferring xylose reductase (red diamonds) as compared to control strain JQ-X1 (blue square).
Fig. 5.5 Xylose fermentation profiles by the JQ-X1 strain harboring either empty vector pRS42H (a) or the same vector with a mutant NADH-preferring xylose reductase (b). Flask fermentations were conducted in complex media with 40 g/L of xylose at initial cell density ~0.15 g/L (OD$_{600}$ = 0.5). All results are the average of duplicate experiments and the error bars represent standard deviation.
Fig. 5.6 Ratio of NADH-specific xylose reductase activity to NADPH-specific xylose reductase activity in JQ-X1 strains harboring either empty vector pRS42H (p1 and p2) or the same vector with a mutant NADH-preferring xylose reductase (mXR1 and mXR2). Results are the average of triplicate assays and error bars indicate standard deviation. The student t test was used to establish significant differences between the mutant XR strains and the controls. *, P < 0.05 (95% confidence)
Table 5.1 Comparison of fermentation parameters by JQ-X1 and JQ-X1-adhE strains in YP media with 40 g/L xylose and 2 g/L acetate. Cells were inoculated at initial cell density ~0.3 g/L (OD$_{600} = 1$).

<table>
<thead>
<tr>
<th>Fermentation parameters</th>
<th>JQ-X1</th>
<th>JQ-X1-adhE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumed xylose (g/L)</td>
<td>28.24 ± 0.23</td>
<td>37.06 ± 0.14</td>
</tr>
<tr>
<td>Consumed acetate (g/L)</td>
<td>0.07 ± 0.00</td>
<td>1.15 ± 0.03</td>
</tr>
<tr>
<td>Ethanol titer (g/L)</td>
<td>9.39 ± 0.10</td>
<td>14.02 ± 0.09</td>
</tr>
<tr>
<td>$Y_{\text{Ethanol}}^*$</td>
<td>0.371 ± 0.00</td>
<td>0.409 ± 0.00</td>
</tr>
<tr>
<td>Xylitol titer (g/L)</td>
<td>2.93 ± 0.02</td>
<td>2.80 ± 0.01</td>
</tr>
<tr>
<td>$Y_{\text{Xylitol}}$</td>
<td>0.10 ± 0.00</td>
<td>0.08 ± 0.00</td>
</tr>
<tr>
<td>Glycerol titer (g/L)</td>
<td>1.50 ± 0.04</td>
<td>0.97 ± 0.02</td>
</tr>
<tr>
<td>$X_{\text{Final}}$</td>
<td>1.08 ± 0.01</td>
<td>1.82 ± 0.03</td>
</tr>
</tbody>
</table>

*Fermentation parameters: $Y_{\text{Ethanol}}^*$, ethanol yield (g ethanol/g xylose assimilated); $Y_{\text{Xylitol}}$, xylitol yield (g xylitol/g xylose); $X_{\text{Final}}$, final cell density (g cells/L); all parameters were calculated at 66.5 hours.
6.1 Summary

Systematic and combinatorial methods were studied and applied for uncovering genetic targets in yeast to improve cellular phenotype. Specifically, perturbations were identified for rapid utilization of various sugars in renewable biomass and efficient conversion to fermentation products with high yield and productivity. Throughout the four research chapters in the thesis, strain improvement was driven by metabolic engineering concepts and methodologies, such as flux balance analysis, evolutionary engineering, rational strain design, and inverse metabolic engineering. Furthermore, systems biology approaches involving collection of genomic, fluxomic, and metabolomic data were applied to characterize engineered strains and accumulate valuable understanding about improved phenotypes at both systems and molecular levels. The most important findings from each chapter are summarized below.

First, I performed an in silico gene deletion experiment based on a genome-scale metabolic model in order to determine beneficial deletions for ethanol production by *S. cerevisiae*. Genes coding for two oxidative phosphorylation reactions (cytochrome c oxidase and ubiquinol cytochrome c reductase) were identified by the model-based simulation as potential deletion targets for enhancing ethanol production and maintaining acceptable overall growth rate in oxygen-limited conditions. Since the two target enzymes are composed of multiple subunits, we conducted a genetic screening study to evaluate the in silico results and compare the effect of deleting various portions of the respiratory enzyme complexes. Over two-thirds of the knockout mutants identified by the in silico study did exhibit experimental behavior in qualitative agreement with model predictions, but the exceptions illustrate the limitation of using a purely stoichiometric model-based approach. Furthermore, there was a substantial quantitative variation in phenotype among the various respiration-deficient mutants that were screened in this study,
and three genes encoding respiratory enzyme subunits were identified as the best knockout targets for improving hexose fermentation in microaerobic conditions. Specifically, deletion of either \textit{COX9} or \textit{QCR9} resulted in higher ethanol production rates than the parental strain by 37\% and 27\%, respectively, with slight growth disadvantages. Also, deletion of \textit{QCR6} led to improved ethanol production rate by 24\% with no growth disadvantage. The beneficial effects of these gene deletions were consistently demonstrated in different strain backgrounds and with four common hexoses. The combination of stoichiometric modeling and genetic screening using a systematic knockout collection was useful for narrowing a large set of gene targets and identifying targets of interest.

Next, I found that deletion of respiratory enzyme subunit Cox9p, which was identified by the previous model-based approach, led to severe inhibition of galactose fermentation. However, the metabolic “death valley” (i.e. no growth on galactose) was a necessary intermediate phenotype for the respiration-deficient yeast to reach optimal galactose fermentation rapidly through serial subcultures in galactose media. The resulting strain JQ-G1 was able to produce ethanol with a 94\% increase in yield and 6.9-fold improvement in specific productivity as compared to the wild-type strain. \textsuperscript{13}C-metabolic flux analysis demonstrated a three-fold reduction in carbon flux through the TCA cycle of the evolved mutant with redirection of flux toward the fermentation pathway. Also, the intracellular concentration of galactitol and most TCA cycle intermediates was significantly reduced in JQ-G1 while pyruvate concentration was elevated as compared to the wild-type strain. For understanding the genetic basis of the evolved phenotype, genome sequencing was conducted and used to identify a loss of function mutation in a master negative regulator of the Leloir pathway (Gal80p); this mutation was found to act synergistically with inhibition of respiration for efficient galactose fermentation. Thus, deletion
of GAL80 alone was sufficient for rapid galactose utilization (1.30 g/L·h) with high ethanol yield (0.383 g/g) in strict anaerobic conditions. This is especially interesting since galactose metabolism in S. cerevisiae typically requires most or all cellular energy production by respiration; therefore, the wild-type yeast cannot metabolize galactose under strict anaerobic conditions. The results in this section demonstrate a promising approach for directing adaptive evolution toward fermentative metabolism and for generating evolved yeast strains with improved phenotypes under anaerobic conditions.

Finally, I applied a similar ‘fermentative evolution’ approach for improving xylose fermentation in an engineered strain (SR8) expressing the heterologous XR/XDH pathway. The strategy involved deletion of COX9 and evolutionary engineering in xylose minimal media. The resulting evolved strain JQ-X1 showed a substantial 32% improvement in ethanol yield and 37% improvement in ethanol productivity as compared to SR8 in xylose media. Genome sequencing identified two mutations that were common to all isolated evolved mutants and also not found in the reference strain. Yeast mating and tetrad dissection focused the search on one of the two mutations – a frameshift deletion in SPT3 (encoding a subunit of the transcriptional regulatory complex SAGA) which was discovered in all respiration-deficient spores with better phenotype on xylose. One apparent disadvantage of JQ-X1 was a significantly higher xylitol yield (0.18 g/g) as compared with the parental strain (0.01 g/g) due to imbalance of redox cofactors without sufficient respiration to replenish NAD⁺. Expression of an NADH-consuming acetate reduction pathway was effective for restoring redox balance and improving ethanol production in the respiration-deficient xylose-fermenting yeast strain. The acetate-consuming strain JQ-X1-adhE produced 49% more ethanol and had a 27% lower xylitol yield in xylose/acetate-containing media as compared to JQ-X1. Also, expression of a mutant NADH-preferring xylose reductase
was promising for restoring redox balance and improving fermentation parameters by the evolved JQ-X1 strain. The engineered JQ-X1-mXR strain produced 44% more ethanol with a 22% reduction in xylitol as compared to the control (JQ-X1-pRS42H) during fermentation in xylose media.

6.2 Future research directions

In chapter III, I noted that deletion of *GAL80* permits galactose fermentation in strict anaerobic conditions, which was not possible in wild-type laboratory strain BY4742. More experiments are needed to confirm whether this is a strain-specific effect or it can be generalized to all *S. cerevisiae* strains. I propose to delete *GAL80* in several other strain backgrounds, such as CEN.PK and D452-2, and then evaluate the phenotype of the knockouts and control strains in a sealed and purged anaerobic vial with 20 mL of minimal media and 40 g/L galactose.

In chapter IV, I found that a mutation in *SPT3* seemed to contribute to the improved phenotype of xylose-fermenting strain JQ-X1. To confirm this, I introduced the mutation into the SR8 and SR8 Δ*cox9* strains using the CRISPR/Cas system to evaluate the effect in xylose. More work is needed to evaluate the effect of the frameshift mutation in *SPT3* in various strains and during growth on various sugars. I propose to introduce the mutation into strains BY4742 and BY4742 Δ*cox9* using the same donor DNA and guide RNA designed by me. Using these strains, the effect of the mutation can be evaluated in a respiration-deficient or respiration-sufficient laboratory strain during growth on glucose, galactose, or other sugars.

In chapter V, I introduced another copy of *adhE* to the JQ-X1-adhE strain in order to increase capacity for acetate assimilation. However, all transformants showed a reduction in acetate consumption after introducing the extra *adhE* copy in the *PHO13* region as described
previously. I propose to clone a different, unused antibiotic resistance marker (e.g. clonNAT marker) into the pITy3-co-adhE integration plasmid in place of the used hygromycin marker. Then, the plasmid can be linearized and transformed to strain JQ-X1-adhE for inserting additional copies of adhE at the delta sequences. This is a simpler alternate method to increase AADH activity in the xylose/acetate co-consuming strain (JQ-X1-adhE) for further improving redox balance and ethanol fermentation. The resulting transformants should be tested in YPXA under anaerobic conditions to evaluate the phenotype.

In chapter V, a mutant NADH-preferring xylose reductase was transformed into JQ-X1 on a multi-copy plasmid. In complex media with 40 g/L of xylose, the transformants did show higher ethanol production and lower xylitol production as compared to the control strain with empty vector pRS42H. In order to increase stability and expression of this NADH-preferring xylose reductase, I propose to engineer a mutant XR construct for integration into the genome of JQ-X1. The mutant XR can be cloned into the pITy3 plasmid with a strong promoter and then the plasmid can be digested with XhoI and transformed to JQ-X1. The resulting transformants should be tested in YPX and SCX to evaluate the phenotype.
CHAPTER VII  REFERENCES


Hong, K., Nielsen, J., 2012. Metabolic engineering of Saccharomyces cerevisiae: a key cell factory platform for future biorefineries. Cellular and Molecular Life Sciences. 69, 2671-2690.


Yang, M., Trumpower, B. L., 1994. Deletion of \textit{QCR6}, the gene encoding subunit six of the mitochondrial cytochrome bc1 complex, blocks maturation of cytochrome c1, and causes
temperature-sensitive petite growth in *Saccharomyces cerevisiae*. J. Biol. Chem. 269, 1270-1275.

